Malat1 is a determinant of $CD4^+$ T cell function

Katie Alice West

PhD

University of York

Biology

December 2022

Abstract

In this thesis, we show that *Malat1* is a critical regulator of one of the most fundamental transitions in the immune system, the differentiation of T helper (Th) cells from naïve $CD4^+$ T cells. We determine that *Malat1* down-regulation is a hallmark of Th cell differentiation, yet its deletion results in enhanced immunity in both *in vitro* and in mixed bone marrow chimera *in vivo* infection models. This is because *Malat1*$^{-/-}$ Th1 and Th2 cells produce lower levels of the anti-inflammatory cytokine IL-10. Mechanistically, we determine that *Malat1* regulates the expression of IL-10 through the key transcription factor MAF. Notably, we show that *Malat1* dependent regulation of IL-10 is female specific, as loss of *Malat1* does not alter IL-10 expression in male $CD4^+$ T cells in *in vitro* polarised Th2 cells or in an *in vivo Schistosoma mansoni* egg induced model of lung inflammation. Additionally, we find that *Malat1* loss alters gene expression levels, which is reflected through impaired induction of the Th cell differentiation programme. . Next, through RAP-MS analysis, we determine the *Malat1*-RBP interactome in both EL4 cells (a T cell line) and primary naive $CD4^+$ T cells. We determine that *Malat1* interacts with multiple RBPs which include members of the hnRNP and SR protein families. Some of these interactions are unique to either EL4 cells or primary naïve $CD4^+$ T cells with some core interaction partners identified in both cell types. Collectively, our results reveal that *Malat1* is an essential orchestrator of CD4+ T cell function with relevance to immune health and pathology.

Contents

Table of figures

Table of tables

Acknowledgements

It is a pleasure to reflect on the support from friends, family and colleagues throughout my PhD. Funds from the BBSRC white rose DTP, HYMS, and the Wellcome Trust have kindly supported my research. Thank you to the COVID 19 PhD Scholarship fund for extending my PhD to mitigate the impacts of COVID on my research.

I would like to thank my two supervisors Prof Dimitris Lagos and Prof Dawn Coverley for the opportunity to work in their labs and their support and guidance to develop my scientific skills. Thank you to Dimitris for taking the time to help improve my confidence in presenting and for his support during difficult times. Thank you to Dawn for her enthusiasm and encouragement throughout the project.

I would like to express my gratitude to my TAP members, Dr Michael Plevin and Prof Paul Kaye for their helpful insights and valuable comments in TAP meetings.

I would also like to thank the technology facility for their training in microscopy and flow cytometry – with particular thanks to Dr Karen Hogg and Dr Graeme Park for teaching me compensation. I am particularly grateful to Dr Adam Dowle for his mass spectrometry expertise.

I am grateful for the support from our collaborators at the Sanger Institute, in particular Dr Kylie James for her support with RNA-seq analysis and interpretation. Our colleagues at the Francis Crick Institute for their time and iCLIP expertise. I would like to thank Prof Jernej Ule, Dr Miha Modic and Urska Janjos for their great enthusiasm, ideas and support with iCLIP.

I would like to thank past and present members of the Hewitson laboratory. With particular thanks to Dr James Hewitson, Joanna Greenman and Sofia Hain

Porter, for their time, help with infection experiments and for sharing their expert flow cytometry knowledge.

Thank you to Dr Dave Boucher for his constant supply of sweets and to Dr Alyssa Cull for her kindness, positive outlook and advice.

I would also like to thank all of the past and present members of the Lagos lab group. Thank you to Joshua Lee for his support and friendship throughout our time in the lab. Thank you to Magnus Gwynne for their help with my last few months of laboratory work, their enthusiasm and for carrying on making interesting findings on this project.

I would like to thank my family, Helen, Daniel, and Joe for their constant encouragement and unwavering love. Wethers for her support at my online poster sessions, and company when writing. Finally, a heart-felt thank you to Josh for his friendship, sense of humour and always listening. Your great love and support have been a huge help throughout these last few years.

Thank you all,

Katie

Author's declaration

I declare that this thesis is a presentation of original work, and I am the sole author. This work has not previously been presented for an award at the University of York, or elsewhere for a degree or award. Any collaborators on this project have been acknowledged where appropriate, and any work mentioned that is not my own is included in the reference list.

1. Introduction

Statement

Work I have previously produced as part of the following review is also included in this chapter.

West and Lagos, 2019. Long Non-Coding RNA Function in $CD4^+$ T Cells: What We Know and What Next? Non-Coding RNA, 5(3), p.43.

1.1 The non-coding transcriptome

1.1.1 Non-coding RNA

Mammalian cells are capable of transcribing approximately 80% of their genome. Historically, protein-coding genes have been the focus of extensive study, however, the fully spliced, protein-coding transcript isoforms of these genes comprise just above 2% of the transcribed genome (Djebali et al., 2012). The remaining untranslated genes generate non-coding RNAs (ncRNAs) which previously have been regarded as transcriptional noise. In more recent years, some ncRNAs have been shown to play a vital role in a myriad of cellular functions including gene regulation and nuclear organisation, with several ncRNAs now implicated in multiple disease pathologies (Kopp & Mendell, 2018; X. Zhang et al., 2017).

ncRNAs are broadly classified based on their post-transcriptional length. Those which are less than 200bp are known as small non-coding RNAs, examples include microRNAs (miRNAs), transfer ribonucleic acids (tRNAs), and small nuclear RNAs (snRNAs). Conversely, transcripts longer than 200bp are known as long non-coding RNAs (lncRNAs) (X. Zhang et al., 2017). Here, we will focus on the role and regulation of lncRNAs.

Mammalian genomes encode numerous lncRNAs, with some thought to comprise up to 100,000 lncRNA genes (Harrow et al., 2012; Y. Zhao et al., 2016). In humans, a conservative estimation has annotated 16,000 lncRNA genes which generate over 28,000 unique transcripts that lack protein-coding capacity (Marchese et al., 2017). The structural features of lncRNAs varies, many lncRNAs are transcribed by RNA polymerase II, and consequently, some transcripts comprise a 5' cap and poly-A tail. Other lncRNA transcripts lack a 3' poly-A tail and undergo alternative 3' processing. Typically, lncRNA transcripts are localised in cellular compartments such as the nucleus or cytoplasm, they are expressed at low copy numbers and can be poorly evolutionary conserved. Given the archetypal low expression and poor sequence conservation of lncRNAs, this draws the cellular relevance of lncRNAs into question. The fraction of annotated lncRNAs which have a cellular function remains unknown, however, there is a lack of evidence to support the functionality of most lncRNAs which indicates they are a product of pervasive transcription (Ulitsky, 2016). However, many lncRNAs are expressed in a cell-type and tissue-specific manner strongly suggesting that lncRNAs may play a cell-specific role. Although the exact function or lack thereof is not known for all these lncRNA genes, those which have been studied can be classified in numerous ways.

One method of categorising lncRNAs is based on their genomic location. Long intergenic non-coding RNAs (lincRNAs) are a type of lncRNA that are located at least 1kb away from any other gene in the genome (coding or non-coding) - meaning there is no overlap (***Figure 1.1*****A**) (Tsagakis et al., 2020). Interestingly, lincRNAs represent over half the lncRNAs found in humans, they are typically transcribed by RNA Pol II and so are generally polyadenylated, spliced, and are an average length of 1kb (Kung et al., 2013; Ransohoff et al., 2018). Classic examples of lincRNAs include X-inactive

specific transcript (*Xist*) which is essential for X chromosome inactivation (XCI) and HOX transcript antisense RNA (*Hotair*) which suppresses the transcription of HOX genes.

Many other lncRNAs overlap in some form with other genes, these can be categorised based on their location and orientation. Distinct from lincRNAs, a lncRNA that is transcribed entirely from the intronic region of a protein-coding gene is known as an intronic lncRNA (***Figure 1.1*B**). Contrasting to lincRNAs, only a small fraction of intronic lncRNAs have been studied regarding function. LncRNAs that are categorised based on orientation are known as sense, anti-sense or bi-directional lncRNAs (Tsagakis et al., 2020). Sense lncRNAs are located and transcribed from the sense strand of a coding gene (***Figure 1.1*C**). These lncRNAs can overlap with part or all of the sequence of the protein-coding gene. Contrastingly, antisense lncRNAs are transcribed from the antisense strand of a coding gene, like with sense lncRNAs, antisense lncRNAs can overlap with part or all the protein-coding transcript (***Figure 1.1*D**). Antisense lncRNAs are the second most common type of lncRNA transcript, which characteristically bind to other RNAs – typically the protein-coding gene it is antisense to. Typically, both sense and antisense transcripts contain mRNA-like features such as a 5' cap and polyA tail (Ma et al., 2013). The last category of lncRNA genes is bi-directional lncRNAs which are located near a protein-coding gene but are transcribed in the opposite direction (***Figure 1.1*E**). Of these categories of lncRNAs, lincRNAs and antisense lncRNAs are the most abundant and thus the most studied transcripts.

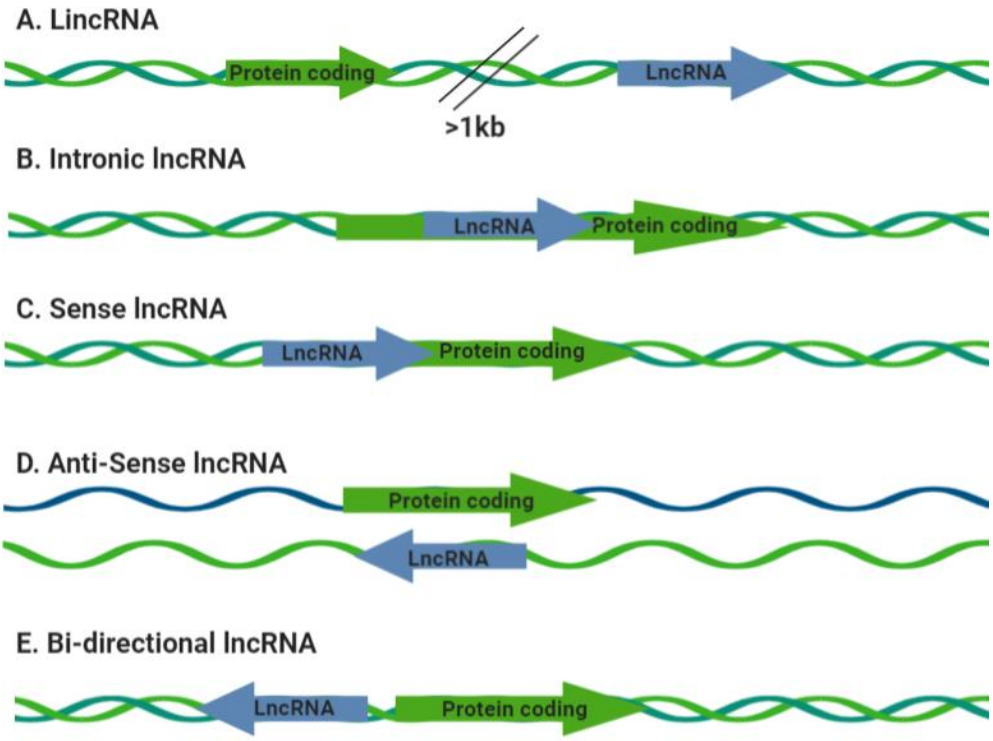


Figure 1.1: Schematic representation of the location and orientation based lncRNA classification system.

Different categories of lncRNAs based on location and orientation. **A)** LincRNA **B)** Intronic lncRNA **C)** Sense lncRNA **D)** Antisense lncRNA **E)** Bi-directional lncRNA

Another method of lncRNA classification is based on functionality. LncRNAs can be described as either c*is* and/or *trans-acting*, through modular interactions with RNA, DNA, and proteins (Y. G. Chen et al., 2017; Kopp & Mendell, 2018). LncRNAs that act *in-trans* exert their function at a different site to which they were transcribed. Examples of *in-trans* lncRNA functions include binding and sequestering proteins in cellular compartments, sponging other RNAs, and regulating gene expression in different genomic locations to the site of transcription. Contrastingly, *cis*-acting lncRNAs function by interacting with genes neighbouring its site of transcription and can help localise epigenetic modifiers to these locations (Garitano-Trojaola et al., 2013). However, there are many exceptions and nuances to this rule. An example of this is the lncRNA functional intergenic repeating RNA element (*Firre*). *Firre* originates from the X chromosome and interacts with the nuclear-matrix factor hnRNPU to contact regions located far away on the same chromosome (Hacisuleyman et al., 2014). In cases such as this, it is particularly difficult to assign an *in-cis* or *in-trans* mode of action. Nevertheless, it is clear that lncRNAs play a role in a wide range of cellular functions both locally and distally (***Figure 1.2).***

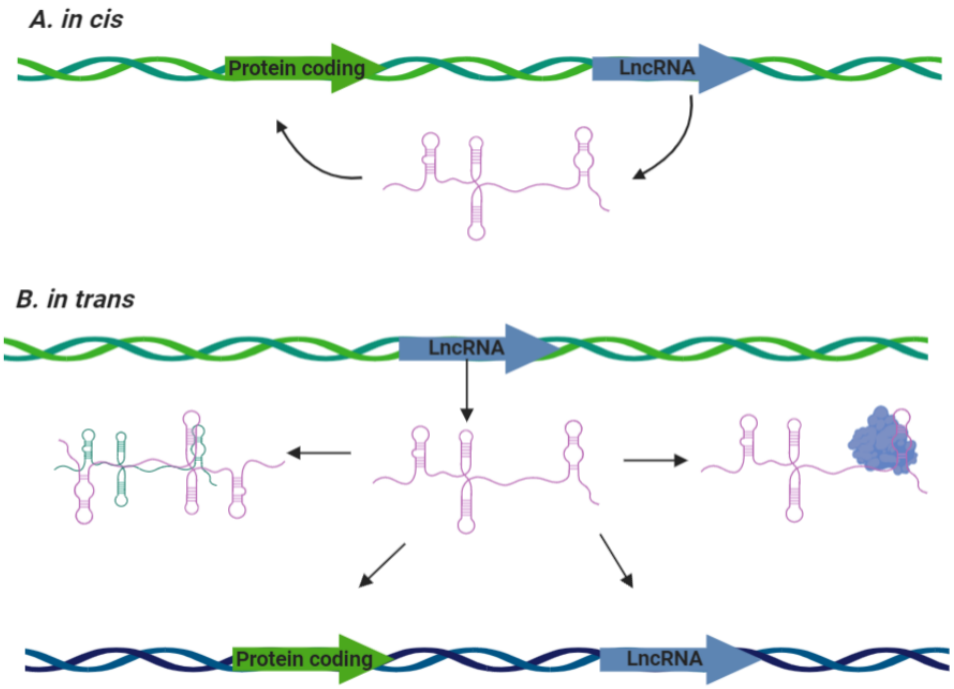


Figure 1.2: Schematic representation of LncRNA *in-cis* vs *in-trans* mode of action.

Representation of lncRNA acting **A)** *In-cis* by affecting local gene regulation or **B)** *in-trans* acting away from the site of transcription by interacting with other RNAs, proteins or distal DNA.

1.1.2 Long non-coding RNA conservation

If a gene or protein is well conserved between species, this often suggests functionality. This is because, if a gene conferred an advantageous function, it would be positively selected within the population. Concequently, if the gene was to mutate in a way that weakened its advantageous nature it would be purged from the population by purifying/negative selection (Haerty & Ponting, 2013). Yet, in stark contrast to protein-coding genes, lncRNA sequences are typically poorly conserved and mutate at a much faster rate than protein-coding genes and mRNA 3'UTRs, suggesting that lncRNAs may lack functional capacity and could be a product of pervasive transcription (Haerty & Ponting, 2013). However, when compared to other neutral DNA sequences or small introns found in *Drosophila*, lncRNAs are under greater selective pressure, which indicates in general lncRNAs may have some functional capacity (Haerty & Ponting, 2013).

There are several ways to explain the low conservation of lncRNAs. One is by examining features of lncRNAs that are known to be functional. Some lncRNAs functionally interact with proteins through a small fraction of the total sequence ~ 10nt. This suggests that only a small portion of the total lncRNA sequence needs to be conserved to be functional (Diederichs, 2014). An alternative explanation is that the act of transcription across the lncRNA locus is conserved between species. In these instances, it is the act of transcription which affects the expression of neighbouring genes, not the lncRNA product. This process is known as syntenic transcription and would require little lncRNA sequence conservation (Diederichs, 2014). Secondary and tertiary structures are often required for lncRNA function (Smith et al, 2013). In these cases, sequence mutations are commonly tolerated unless they interfere with lncRNA structural motifs. This suggests that lncRNAs could be more tolerant of mutations than

protein-coding genes. Additionally, lncRNAs with similar k-mer content have been shown to have related functions, despite a lack of linear sequence homology (Kirk et al., 2018). As such despite their low conservation, it is possible that more lncRNAs could be functional than initially anticipated.

1.1.3 Long non-coding RNA function

Despite the notion that some lncRNAs lack functional capabilities, several lncRNAs have been shown to play a central role in a myriad of cellular processes. Early studies uncovered the essential role of the lncRNA *Xist* in XCI. *Xist* is crucial for the recruitment of chromatin-modifying enzymes to the inactive X chromosome helping it form a heritable repressed state (Brockdorff & Turner, 2015). It is now well established that lncRNAs function in a wide range of processes such as cell differentiation, cell cycle progression, metabolism, and response to viral infections (Marchese et al., 2017). Mechanistically, lncRNAs affect these essential cellular functions in a variety of ways. Several lncRNAs have been shown to enhance or impair transcription by altering the structure of nuclear domains, interacting with enhancers, or interfering with transcriptional enzymes. Many more lncRNAs have been shown to act post-transcriptionally for example by regulating splicing, altering mRNA stability, affecting translation, and interfering with protein stability. This growing body of evidence demonstrates that lncRNAs are functionally diverse, this is likely due to their ability to form different structures and molecular interactions. The importance of lncRNAs in these crucial cellular processes is further highlighted in numerous diseases which have been linked to lncRNA dysregulation (Marchese et al., 2017).

lncRNA localisation is often helpful for determining functionality. lncRNAs are frequently enriched in the nucleus, with many located in specific nuclear compartments, further suggesting they are not a result of transcriptional noise (Statello

et al., 2020). One example is the lncRNA nuclear paraspeckle assembly transcript 1 (*Neat1*) which is located in nuclear paraspeckles. Nuclear paraspeckles are a nuclear sub-structure found in the interchromatin area and play an important role in regulating gene expression. *Neat1* is a key component of paraspeckles and is essential for their formation and maintenance (Bond & Fox, 2009). Several lncRNAs are located in the cytoplasm, as exporting RNA from the nucleus poses an energetic burden to the cell this suggests that lncRNAs are being exported so they are able to carry out a beneficial function. Interestingly, some cytoplasmic lncRNAs affect the function of specific organelles and cytoplasmic structures. The lncRNA cytoplasmic endogenous regulator of oxidative phosphorylation 1 (*Cerox1*) is important for mitochondrial function and regulates mitochondrial complex I catalytic activity (Sirey et al., 2019). Thus, determining lncRNA location can improve our understanding of potential lncRNA functions.

Nuclear lncRNAs can alter transcription in numerous ways. For example, lncRNAs can alter chromatin architecture through interactions with chromatin-modifying enzymes. LncRNAs are known to interact with several chromatin-modifying enzymes such as the polycomb repressive complex (PRC), switch/sucrose nonfermentable (SWI/SNF) complex and the histone demethylase lysine specific demthylase 1 (LSD1) (Sun et al., 2018). LncRNAs can directly bind to these complexes and guide them to specific genomic regions, or act as a decoy to prevent these complexes from reaching their targets. One study showed that the lncRNA TCF21 antisense RNA inducing demethylation (*Tarid*) binds to the DNA demethylation regulator growth arrest and DNA damage inducible alpha (GADD45α) and targets it to specific genomic locations to regulate gene expression (Arab et al., 2014). LncRNAs can also affect chromatin structure indirectly through binding to non-

chromatin modifying protein intermediates as reviewed by (Sun et al., 2018) LncRNAs are also known to influence transcription more directly by affecting transcription factors. The emerging theme of these studies is that lncRNAs typically interact with proteins to indirectly bind DNA and alter transcription.

Cytoplasmic lncRNAs interact with RNA and proteins. Interestingly, some lncRNAs contain specific miRNA binding sites and bind/sponge specific miRNAs for example to prevent degradation. In *Arabidopsis,* the lncRNA induced by phosphate starvation 1 (*Ips1*) sponges the miR-399 to control a gene which is essential for regulating root phosphate content (Franco-Zorrilla et al., 2007). Other cytoplasmic lncRNAs interact with mRNAs during translation. Studies suggest that the trafficking of lncRNAs to ribosomes aids bound mRNA degradation (Carlevaro-Fita et al., 2016). Yet, many other studies suggest that lncRNA interaction with polysomes is not random and has a potential translational regulatory role (Pereira et al., 2019). Of note, lncRNAs have been shown to be enriched in exosomes, a type of extracellular vesicle (EV) which play a key role in cell-cell communication. Several exosomal lncRNAs have been identified as new cancer biomarkers such as lncRNA breast cancer anti-oestrogen resistance 4 (*Bcra4*) (Dragomir et al., 2018). This collectively demonstrates that lncRNAs have key roles in a wide range of cellular functions and exert their function through direct or indirect interactions with DNA, RNA and proteins throughout the cell.

1.1.4 Long non-coding RNA therapeutics

Despite 98% of the human genome lacking protein-coding capacity, nearly all of the drugs currently available to treat disease act on one of approximately 700 disease-linked proteins (Santos et al., 2017). LncRNAs are dysregulated in a wide range of diseases such as Alzheimer's, diabetes, and many cancers. As several

lncRNAs are also expressed in a highly cell-specific manner this highlights the potential of lncRNAs as beneficial therapeutic targets (Arun et al., 2018). Additionally, lncRNAs could potentially be more advantageous targets than some proteins as they are not translated, have a fast turnover and typically low expression, which may enable quicker effects of drugs at a lower dose. Furthermore, current lncRNA targeting approaches have low toxicity (Statello et al., 2020).

Several lncRNA targeting techniques are being developed, including small interfering RNA (siRNA) mediated degradation, chemically modified antisense oligonucleotides (ASOs), sterically blocking lncRNA promoters and, steric inhibition of RNA protein interactions via ASO or small molecules (Arun et al., 2018). The most advanced therapeutic targeting of lncRNAs is currently based on the use of ASOs. These are single-stranded DNA oligos which bind to target RNAs via Watson and Crick base pairing. This induces RNase H cleavage of the target RNA which impairs transcription and reduces the levels of the RNA target (Fluiter et al., 2009). To improve their efficacy ASOs are typically chemically modified to increase their stability for example by creating an RNA-DNA-RNA ASO with a modified sugar backbone. Fusing these ASOs to aptamers can enable targeted delivery of these ASOs (Dassie & Giangrande, 2013)

siRNA / ASO technology has successfully diminished lncRNA levels in cell lines, however, only a handful have been tested *in vivo*. One example where ASOs have been used successfully *in vivo* is in a mouse mammary tumour virus (MMTV) PyMT model. ASOs that target the lncRNA metastasis-associated lung adenocarcinoma transcript 1 (*Malat1*) were subcutaneously delivered, this reduced *Malat1* expression which caused primary tumour differentiation and a significant reduction in metastases (Arun et al., 2016). The FDA has recently approved the use of

antisense drugs for Duchenne muscular dystrophy which has led to the development of multiple pre-clinical models reviewed by (Arun et al., 2018). An interesting pre-clinical technique to study lncRNAs which are restricted to humans is the patient-derived xenograft models (PDXs) in which PDX tumours are implanted into nude mice. The lncRNAs *Hotair* and Survival Associated Mitochondrial Melanoma Specific Oncogenic Non-Coding RNA (*Sammson)* have been targeted by ASOs/siRNAs in these models (Gupta et al., 2010; Leucci et al., 2016).

The use of small molecules to target lncRNAs is less well explored. Generating small molecules which bind lncRNAs at high affinity and specificity requires knowledge of lncRNA structural motifs which at present is only available for a handful of lncRNAs (Brown et al., 2012). Small molecules which target lncRNA-protein interactions or act as a decoy to compete with the lncRNA for protein binding could be a beneficial therapeutic approach. Recently, small molecules which bind *Xist* have been identified. The X1 compound was found to have drug-like properties and specifically bind to the RepA domain of *Xist*. This binding disrupted *Xist* interactions with the PRC2 complex which consequently suppressed histone trimethylation and impaired XCI in a female-specific manner (Aguilar et al., 2022). This study highlighted the potential of generating small molecules that target lncRNAs for therapeutic purposes. Thus, targeting pathogenic lncRNAs is an exciting new area of therapeutics with some RNA-targeting therapeutics already being clinically approved to treat disease.

1.1.5 The non-coding transcriptome and mammalian complexity

Remarkably, in eukaryotes as the complexity of an organism increases the ratio of non-coding to protein-coding DNA rises. Approximately 98% of the human genome is comprised of non-coding DNA, contrasting to 25-50% in simple eukaryotes

and >50% non-coding DNA in complex fungi and plants (Mattick, 2004). Consequently, some scientists have postulated that non-coding transcripts may account for some of the greater complexity seen in mammals (Mattick, 2004).

The immune system contains some of the most complex cellular networks observed in mammals. It provides essential protection against an unpredictable and vast number of pathogens. Cells derived from both the innate and adaptive arms of the immune system undergo complex development, translocate throughout the body, and act in a highly coordinated manner to remove threats. The specificity of this response results in a highly sophisticated defence mechanism enabling protection against infection and reinfection throughout life. Thus, mammalian immunity is an ideal platform for the study of lncRNA function.

1.2 The immune system

1.2.1 Innate and adaptive immunity

The immune system identifies threats by recognising structural features of pathogens or allergenic molecules by one of two broad mechanisms. The first is the recognition of common molecular features of microbes and toxins that are not present in the host, this recognition response is encoded by germline genes and constitutes the basis of the innate immune system. The other recognition method is to identify unique structural features of invading threats, this type of recognition requires somatic gene recombination to enable appropriate detection and is the basis of the adaptive immune response. The innate immune system is known as the first line of defence, once a pathogen is encountered a rapid response is initiated as the recognition molecules it uses are expressed broadly on many cells. The adaptive immune response is typically initiated after the innate immune response as it is comprised of a smaller number of

cells known as lymphocytes. B and T cells can rearrange their receptors to identify an antigen, then subsequently increase in numbers to mount an appropriate response. A unique feature of the adaptive immune system is that it can create memory cells, meaning a more effective and rapid adaptive response can be produced against the specific antigen if it is encountered again (Chaplin, 2010).

1.2.2 Cells of the innate immune system

Granulocytes are the most common type of white blood cell; they are characterised by their ability to release cytosolic granules. There are four main granulocyte subsets neutrophils, eosinophils, basophils and mast cells (Lin & Loré, 2017). Neutrophils are the most common subset of granulocytes and are the most abundant white blood cell in both mice and humans. The main antimicrobial functions of neutrophils are phagocytosis, degranulation and formation of neutrophil extracellular traps (NETs) (Rosales, 2018). Neutrophils interact with a pathogen directly through the recognition of pathogen-associated molecular patterns (PAMPs) by pattern recognition receptors (PRRs) on neutrophils, they can also interact with a pathogen if they have been opsonised by microbes or complement proteins. Neutrophils can degrade pathogens through phagocytosis which involves the ingestion of the microorganism into a phagolysosome which destroys the pathogen through enzymes and exposure to low pH. NETs are extracellular structures comprised of granular proteins that assemble on chromatin, these also contribute to pathogen destruction when they are too large to be ingested (Lin & Loré, 2017; Rosales, 2018)

Monocytes are critical cells of the innate immune response to inflammatory signals. In mice, there are two main populations of monocytes which can be discriminated by their expression of Ly6C. Ly6C^{hi} monocytes are capable of tissue infiltration, they are pro-inflammatory they express high levels of C-C chemokine

receptor 2 (CCR2). They transport antigens to the lymph node and can differentiate into macrophages or dendritic cells depending on local signalling. Ly6C^{low} monocytes are also known as patrolling monocytes and are involved in early responses to inflammatory signals, they express low levels of CCR2 (Kratofil et al., 2017).

Macrophages are a heterogeneous population of innate immune cells that originate from either monocyte or embryonic progenitors (Zhang et al., 2021). They are able to phagocytose pathogens, recruit other immune populations and have some antigen-presenting abilities. Broadly speaking macrophages can be activated into two distinct subsets. Classically activated or M1-activated macrophages can be polarised by type 1 cytokines such as interferon-gamma (IFNγ) or tumour necrosis factor alpha (TNFα) or lipopolysaccharide (LPS) from bacteria. M1 macrophages are pro-inflammatory, they upregulate inducible nitric oxide synthase (iNOS) and express cytokines such as interleukin 1 beta (IL-1β), IL-6, IL12, IL-23 and TNFα (Martinez & Gordon, 2014). Alternatively activated macrophages or M2 macrophages are produced in response to allergens or helminth parasites and arise when exposed to IL-4. They are characterised by their production of YM1 and RElMα. They produce anti-inflammatory cytokines such as IL-10 and have been linked to wound healing and tissue repair (Martinez & Gordon, 2014).

Dendritic cells (DCs) act as a bridge between the innate and adaptive immune system they are the most efficient, potent and professional antigen-presenting cells (APCs) in the immune system (Patente et al., 2019). DCs express high levels of major histocompatibility complex (MHC) class II and cluster of differentiation 11 c (CD11c), they are found in nearly all tissues where they sample the local environment for antigen presentation. DCs are found in two different functional states, immature and mature. The maturation of dendritic cells is initiated by PAMP or damage-

associated molecular patterns (DAMP) recognition. This transition stimulates a switch in metabolism and gene expression which enables DCs to migrate from peripheral tissues to secondary lymphoid organs where antigen presentation may take place (Patente et al., 2019).

1.2.3 Cells of the adaptive immune system

B cells are central to the adaptive immune response, their main role is to produce antibodies against invading pathogens (Hoffman et al., 2016). Antibodies facilitate the neutralisation of infections by inactivating pathogen proteins, activating macrophages by binding to Fc receptors (FcRs) and activating the classical pathway of the complement system. There are three main classes of B cells in mice and humans, B1 and B2 lymphocytes which comprise marginal zone and follicular B cells. B1 cells arise from the foetal liver, and reside in the peritoneal and pleural cavities, they produce IgM antibodies and do not require T-cell help to produce antibodies. B2 cells originate from the bone marrow and can differentiate into marginal zone or follicular lineages in the spleen. The strength of B cell receptor (BCR) signalling determines if a marginal zone or follicular zone B cell is generated (Hoffman et al., 2016)

Cytotoxic T cells (CTLs) play an essential role in the immune system's defence against intracellular pathogens, viruses, and bacteria and are important in tumour cell surveillance (Zhang & Bevan, 2011). Cytotoxic T cells express the T cell receptor (TCR) and interact with peptides presented by MHC Class I molecules which are found on all nucleated cells. $CD8^+$ T cells have three major mechanisms by which they kill infected or malignant cells. They produce TNFα and IFNγ which have anti-tumour and anti-microbial properties. Cytotoxic T cells release cytotoxic granules – perforin and granzymes. Perforin creates a hole in the membrane of an infected or malignant cell which enables granzymes to enter the infected cell. Granzymes cleave proteins

inside the cell ultimately resulting in apoptosis. Activated $CD8^+$ T cells also express FasL these interact with FAS on the surface of a target cell which activates a caspase signalling cascade which induces apoptosis in the target cell (Zhang & Bevan, 2011).

This thesis will explore $CD4^+$ T cell functions. Below $CD4^+$ T cell development, differentiation and functions will be discussed in greater detail.

1.2.4 $CD4^+$ T lymphocytes

1.2.4.1 $CD4^+$ T lymphocyte development

T cells play a central role in the adaptive immune system. The different classes of T cells can be broadly defined by their surface receptors, for example, αβ or γδ TCR. The major class of T cells in humans express the αβ TCR which is important for recognising MHC-antigen complexes. Once αβ T cells have matured they express either CD4 or CD8 molecules on their surface. $CD8^+$ T cells are also known as cytotoxic T lymphocytes they are important for killing cells that are infected with intracellular microbes or abnormal cancerous cells and recognise MHC class I molecules which are expressed on all nucleated cells. $CD4^+$ T cells also known as T helper (Th) cells recognise MHC class II molecules expressed on APCs such as dendritic cells. As suggested by their name $CD4^+$ T cells help the immune system remove pathogens, largely by secreting cytokines to drive the function of other important cells in the immune system (Pennington et al., 2005). Here, we will focus on the role and regulation of $CD4^+$ T cells.

T cells are derived from haematopoietic stem cells (HSCs) produced in the bone marrow. The thymus is populated with these bone marrow-derived precursors and can become committed to the T cell lineage due to recombination of the TCR loci. Each $CD4^+$ T cell expresses a unique TCR that can recognise one specific antigen

(Spits, 2002). To ensure that T cells can detect a diverse range of antigens, during their development T cells rearrange their TCRs. These receptors are assembled by somatic rearrangement of variable (V), diversity (D) - (TCRβ locus only), joining (J) and constant (C) elements of the TCR α and β chains in a process known as V(D)J recombination (***Figure 1.)***. This process is mediated by lymphoid-specific recombinase activating genes (RAG) 1 and 2. These enzymes cleave the TCR gene near the V, D and J segments which are then repaired by a variety of enzymes including terminal deoxynucleotidyl transferase (TdT) which adds random nucleotides to the VDJ regions to further increase the diversity of the TCR repertoire. However, in some cases, this can produce a non-functional or autoreactive TCR (Chaplin, 2010).

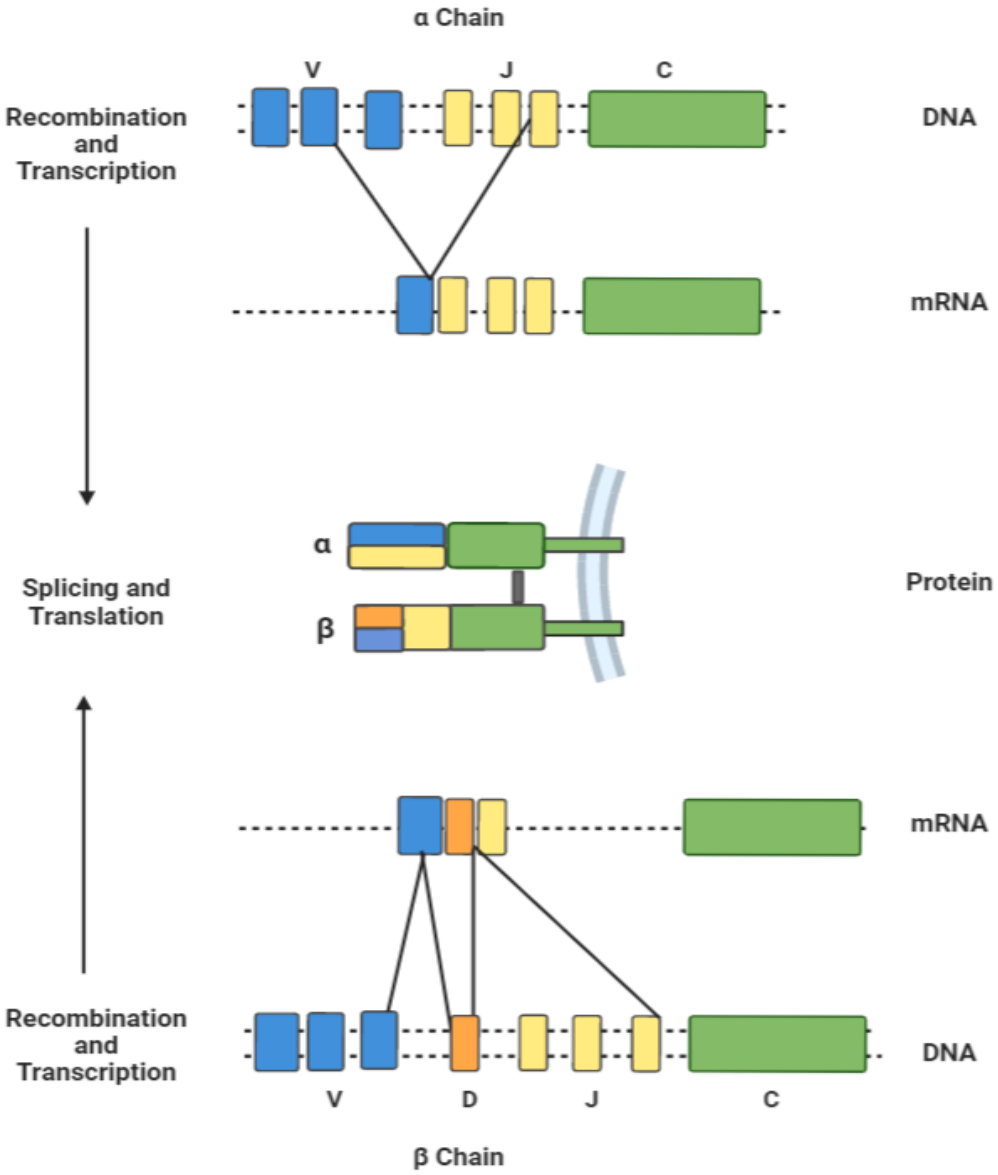


Figure 1.3: Schematic representation of VDJ recombination at the TCR locus.

The germline TCR α and β locus are comprised of several V and J segments with additional D segments at the β locus. An example of only a few segments is shown for illustration purposes. At the α locus, a V to J recombination occurs. At the β locus, there is an initial D to J segment rearrangement followed by a D-J to V segment recombination. The VDJ product is then transcribed and spliced to a c region. These are then translated and expressed on the cell surface.

To ensure T cells can respond to pathogens but are tolerant to self-antigen, $CD4^+$ T cells undergo a selection process in the thymus. Once a T cell precursor has begun to differentiate and rearrange the TCRβ chains it moves to the cortex of the thymus where it creates a potentially functional αβ TCR. Next, the developing lymphocyte tests its TCR with MHC molecules expressed by specialised cortical epithelium. This interaction's strength determines whether the cell will progress through its development or be cleared by apoptosis. If there is no interaction this indicates a non-functional T cell and so it undergoes apoptosis. If the cells are positively selected, they migrate to the thymus medulla and are tested for potential auto reactivity. In this process, the cells are presented with various tissue-specific proteins by thymic medullary epithelial cells, genes encoding these proteins are mediated by the gene called autoimmune regulator (AIRE). If the T cell can recognise these self-antigens they are purged from the population. (Chaplin, 2010). Approximately 5% of $CD4^+$ T cells leave the thymus and enter the bloodstream (Surh & Sprent, 1994). As these newly developed cells have yet to encounter a target antigen these are referred to as recent thymic emigrants as they have restricted function.

1.2.4.2 $CD4^+$ T lymphocyte activation and differentiation

To become activated the naïve $CD4^+$ T cells must receive two signals. Signal 1 comprises the binding of TCR with antigen complexed to MHC II-antigen on an APC such as a dendritic cell. Signal 2 also known as co-stimulation involves the binding of CD80 and CD86 on an APC to CD28 on the $CD4^+$ T cell. If a T cell only receives signal 1, this can cause the T cell to become unresponsive and leads to an anergic state. The combination of signals 1 and 2 generates a cascading response which induces transcription factors such as nuclear factor of activated T cells (NFAT) and nuclear factor kappa light chain enhancer of activated B cells (NF-ĸB) stimulating

transcription of IL-2 which ultimately results in $CD4^+$ T cell proliferation (Luckheeram et al., 2012). The T cell activation process is tremendously complex and alters the expression of thousands of genes.

Following activation, T cells differentiate into effector T cells when provided with a third signal - cytokines. Initially, two distinct populations of effector $CD4^+$ T cells were identified each with unique cytokine expression profiles; Th1 cells which secrete IFNγ and Th2 cells which have IL-4 as their signature cytokine (Mosmann et al., 1986; Mosmann & Coffman, 1989). New subsets of $CD4^+$ T cells have been discovered each with a unique cytokine profile (***Figure 1.4***). These are broadly separated into two subsets, effector $CD4^+$ T cells which are key in host defence against pathogens (Th1, Th2, Th9, Th17) and a regulatory $CD4^+$ subset which dampens the immune response to aid in the prevention of autoimmunity (Tr1, nTreg, iTreg) (Jäger & Kuchroo, 2010). $CD4^+$ T cells undergo remarkable and finely regulated gene expression changes upon antigenic stimulation, which can be studied and manipulated in both *in vivo* and *in vitro* models.

Key transcription factors dictate the differentiation of different sub populations of T cells in response to cytokines. When $CD4^+$ cells are stimulated by antigens and APCs the cells produce IL-2, as such are designated Th0 cells. IL-12 produced by macrophages or natural killer (NK) cells stimulates the activated T cells and results in the expression of the T-box transcription factor (T-bet) and signal transducer and activator of transcription 4 (STAT4), this creates a Th1 cell which is characterised by the expression of IL-2, IFN-γ and lymphotoxin. Th1 cells play a crucial role in cell-mediated immune responses characteristically involved in combating intracellular bacteria and protozoa – typically by activating macrophages and in addition to $CD8^+$ T cells and stimulating B cells.

Contrastingly, when activated T cells are stimulated with IL-4 produced by basophils or mast cells this generates high expression of signal transducer and activator of transcription (STAT6) and GATA Binding Protein 3 (GATA3) and a Th2 cell (Luckheeram et al., 2012). Th2 cells express IL-4, IL-5, IL-9, IL-13 and granulocyte macrophage colony-stimulating factor (GM-CSF), and they are important in supporting humoral immune responses and combating helminth parasites. Th2 cells are important for driving B cell proliferation and IgE class switching, stimulating eosinophils and stimulating macrophages to an M2 phenotype (Walker & McKenzie, 2018).

In the case of Th17 polarisation, this is induced by IL-6, transforming growth factor beta (TGF-β) and IL-23, this stimulates signal transducer and activator of transcription 3 (STAT3) and RAR-related orphan receptor gamma t (RORγT). Th17 cells produce IL-17A, IL-17F, IL-22 and IL-21 which stimulate neutrophil production, as such Th17 cells are important in many autoimmune and inflammatory diseases (Kolls & Sandquist, 2018).

Treg cells are a subset of T cells that help regulate the immune response. Treg cells can be classed into multiple categories, peripherally induced pTreg which are developed in peripheral lymphoid organs after antigen presentation. Thymic-derived Treg cells n(Treg) are released from the thymus and already express FOXP3. TGFβ and IL-10 induce FOXP3 expression in pTreg cells (Shevyrev & Tereshchenko, 2020).

In more recent years, studies have revealed the plastic capabilities of $CD4^+$ T cells. This can be defined as the ability of a specific subset of $CD4^+$ T cells to take on the characteristics of a different $CD4^+$ T cell subset. This is regulated by extracellular cues such as cytokines (Dupage & Bluestone, 2016). It has become evident that

distinct polarised Th cell populations are not uniform and there is a large degree of heterogeneity in these subsets. Th1-Th3 and Th2-Th17 hybrid cells have been reported highlighting a potential continuum of Th cell phenotypes rather than distinct phenotypes.

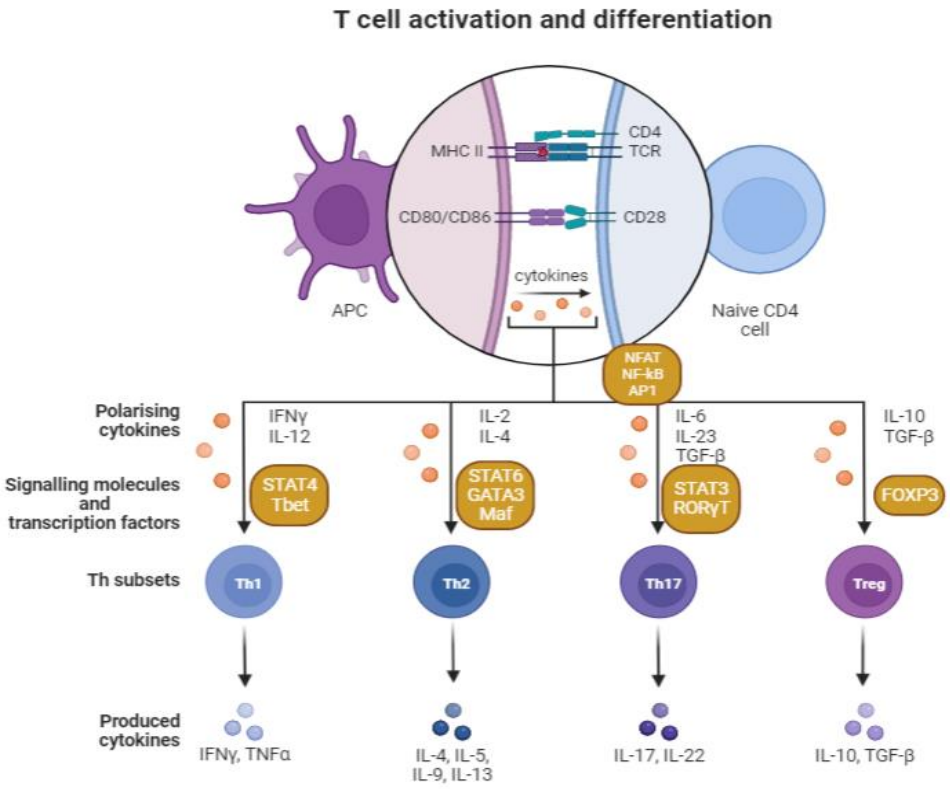


Figure 1.4: Schematic representation of the CD4⁺ T cell differentiation process.

Schematic representation of CD4⁺ T cell differentiation including key transcription factors yellow boxes and secreted cytokines.

1.2.4.3 $CD4^+$ T cell effector functions

Th1 cells are prominently known for their role in the elimination of intracellular pathogens and effective defences against viral infections, with uncontrolled Th1 responses implicated in autoimmunity (Spellberg & Edwards, 2001). IFNϒ produced by Th1 cells is essential to activate macrophages and microglial cells and enhance their phagocytic capabilities (Szulc & Piasecki, 1988). IFNϒ also upregulates the expression of MHC class I and class II molecules on numerous cell types which enhances antigen presentation to T cells (Volk et al., 1986). Additionally, IFNϒ stimulates endothelial cells and keratinocytes to secrete proinflammatory cytokines such as TNF. Secretion of cytokines by Th1 cells stimulates adhesion molecule expression on endothelial cells, which drives the accumulation of leukocytes in areas of inflammation (Luckheeram et al., 2012).

In contrast, Th2 cells are known to exert their effect on immunologic responses to extracellular pathogens and helminth parasites (Spellberg & Edwards, 2001). Aberrant Th2 responses are linked with allergy and asthma development. Th2 cells stimulate the production of high titres of antibodies by B cells. IL-4 and IL-13 stimulate B cell proliferation, antibody production and Ig class switching particularly from IgG to IgE (Punnonen et al., 1993; Punnonen & de Vries, 1994). IL-4 produced by Th2 cells acts in a positive feedback loop to further promote Th2 cell differentiation in naïve $CD4^+$ T cells as they encounter antigens (Nakayama et al., 2017). IL-5 produced by Th2 cells is a potent haematopoietic cytokine which stimulates the production of eosinophils in the bone marrow in addition to dictating chemotaxis of eosinophils and basophils at the site of inflammation (Warringa et al., 1992). Consequently, inflammation as a result of Th2-mediated pathology can be characterised by the presence of eosinophils and basophils in addition to mast cell

degranulation (release of pro-inflammatory molecules) which is dependent on cross-linking surface-bound IgE. This coordinated action can result in enhanced immunity against for example helminthic parasites, which can be coated in IgE antibodies and destroyed by the granular contents of eosinophils (Nakayama et al., 2017).

Th17 cells, have a pro-inflammatory bias and have been shown to be essential for the defence of the immune system against extracellular bacteria, and fungi and the development of autoimmune diseases. Th17 cells produce IL-17 which stimulates epithelial cells, endothelial cells and fibroblasts to secrete TNFα, IL-1β, IL-6, IL-8, CXCL1, and CXCL8 which drive neutrophil recruitment and inflammatory responses (Liang et al., 2007). IL-21 promotes $CD8^+$ T cell function, NK cells and B cell maturation (Monteleone et al., 2005; Ozaki et al., 2002).

1.2.4.4 Disease models for studying $CD4^+$ T cells

Two mouse strains BALB/c and C57BL/6 mice are commonly used to study immune responses. BALB/c mice are genetically prone to develop type 2 immune responses and C57BL/6 mice are prone to develop type 1 immune responses. This is because BALB/c mice hypersecrete IL-4 and have a defect in IL-12 secretion. When BALB/c and C57BL/6 mice were infected with *Leishmania major* they found that $CD4^+$ T cells derived from BALB/c mice had a Th2 type phenotype, these mice were not protected from infection. In contrast, C57BL/6 mice generated a robust Th1 response and showed enhanced immunity to the infection (Locksley et al., 1987). Similarly, if the same mouse strains were exposed to a helminth parasite BALB/c mouse are resistant to infection and displayed decreased worm burdens than C57BL/6 mice which displayed chronic infection and scarring (Else & Grencis, 1991).

1.3 *Ma1at1*

1.3.1 *Malat1* overview

Malat1 also known as nuclear enriched abundant transcript 2 (Neat2) is one of the most extensively analysed lncRNAs (Zhang et al., 2017). It was first identified in a differential gene expression study on early-stage non-small cell lung cancer (NSCLC) patients, which demonstrated that high expression of *Malat1* was significantly associated with metastasis (Ji et al., 2003). Later work went on to link abnormally high *Malat1* expression with a multitude of cancers, including gastric cancer (Okugawa et al., 2014), ovarian cancer (Chen et al., 2016), bladder cancer (Li et al., 2017), breast cancer (Xu et al., 2015), in addition to many others. Abnormally high expression of *Malat1* can be an indicator of poor prognosis in several types of cancers, suggesting that *Malat1* may be involved in carcinogenesis (Li et al., 2018).

Malat1 is encoded on chromosome 11q13.1 in humans and 19qA in mice. The human transcript is ~8.7 kb in length and ~6.7 kb in mice. Interestingly, *Ma1at1* is processed into two transcripts by RNaseP and RNaseZ. The large 6.7kb transcript is localised in nuclear speckles (nuclear sites enriched with splicing factors), and the smaller 61nt transcript known as the mascRNA is exported to the cytoplasm (Nakagawa et al., 2012a). The larger processed *Malat1* transcript lacks a traditional poly A tail instead, its 3’ A rich tract pairs with an upstream U-rich region to form a unique triple helix structure which protects the transcript from degradation by exonucleases and contributes to its stable nature (Brown et al., 2012; X. Zhang et al., 2017). This triple helical structure was first identified in the polyadenylated nuclear (PAN) RNA produced by the Kaposi sarcoma-associated γ-herpesvirus (KSHV), this

triple helix structure is rare, but is also found in the lncRNA *Neat1* (Brown et al., 2012; Wilusz et al., 2012).

Malat1 is expressed ubiquitously throughout multiple human and mouse tissues and has been observed at high levels comparable to that of housekeeper genes i.e genes which are required for the maintenance of basic cellular functions and are expressed across tissues (*Malat1* is observed at levels between 5,000 -10,000 copies per cell). In some cases, transcripts of *Malat1* are expressed at higher RNA levels than transcripts of housekeeping genes (Eißmann et al., 2012). In contrast to many lncRNAs, *Malat1* has a high degree of sequence conservation across 33 mammalian species including humans and mice (Eißmann et al., 2012). (***Figure 1.***5), with over 50% total conservation and 80% conservation at the 3' end of the transcript (Arun et al., 2020). Collectively, the dysregulation of *Malat1* in cancer, its specific nuclear localisation and its high degree of conservation are strongly suggestive that *Malat1* has a crucial biological function.

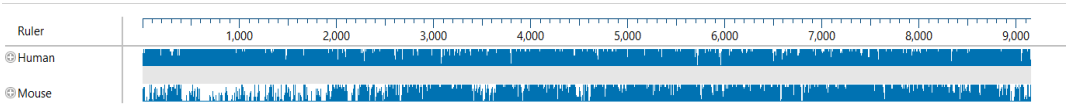


Figure 1.5: Schematic representation of the conservation between human and mouse *Malat1* RNA transcripts.

Pairwise alignment was used to compare conservation of human and mouse *Malat1* RNA FASTA sequences using laser gene software.

1.3.2 *Malat1* functions

Numerous functions of *Malat1* have been reported (***Figure 1.***6), for example *Malat1* is associated with the regulation of alternative splicing (AS). The role of *Malat1* in splicing is directly linked to its nuclear speckle localisation – a site enriched in splicing factors. AS is important for increasing proteomic complexity. Numerous proteins are crucial in the regulation of AS in particular serine/arginine-rich (SR) proteins, heterogeneous nuclear ribonuclear proteins (hnRNPs) and small nuclear ribonuclear proteins (snRNPS). Immunoprecipitation of splicing proteins in HeLa cells has indicated that *Malat1* binds to SR proteins such as serine and arginine-rich splicing factor 1 (SRSF1), serine and arginine-rich splicing factor 2 (SRSF2) and serine and arginine-rich splicing factor 3 (SRSF3), and that depletion of *Malat1* using antisense oligonucleotides alters AS of endogenous pre-mRNAs (Tripathi et al., 2010b). *Malat1* has also been linked to enhancing the progression of ovarian cancers, by regulating AS through the splicing factor RNA binding protein fox -1, homolog 2 (RBFOX2) (Gordon et al., 2019).

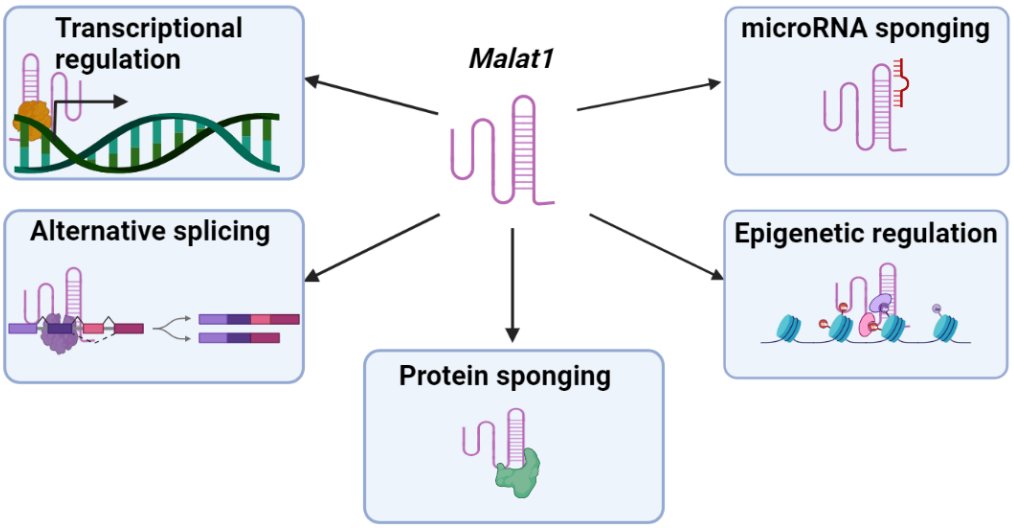


Figure 1.6: Schematic representation of reported functions of *Malat1.*

Representation of *Malat1* functions including regulation of splicing, transcription, protein localisation, miRNA localisation and epigenetic regulation.

Interestingly, *Malat1* and its neighbouring gene *Neat1* have both been shown to bind active chromatin sites (West et al., 2014b). Using capture hybridization analysis of RNA targets (CHART)-seq hundreds of *trans*-binding sites for both *Neat1* and *Malat1* have been identified. These binding sites largely overlap with active genes. *Neat1* was shown to bind near transcription start sites (TSS) and transcriptional termination sites (TTS). Contrastingly, *Malat1* binds near TSS and across the gene bodies. This suggests that they have complementary roles. Many proteins which reside in nuclear bodies were also shown to bind to both lncRNAs lending further support to their complementary function. This collectively suggests that *Malat1* and *Neat1* are important in linking active genes to nuclear subdomains.

Other work has shown that *Malat1* interacts with several RNA-binding proteins (RBPs). For example, multiple studies have indicated that *Malat1* interacts with PRC2 – a protein complex involved in epigenetic repression of transcription. A recent study investigating HIV-1 replication showed that *Ma1at1* enables HIV-1 transcription and infections via PRC2, as *Malat1* detaches enhancer of zeste homolog 2 (EZH2) a PRC2 component from binding the HIV-1 LTR promoter, thus relieving epigenetic silencing marks (Qu et al., 2019). In addition, *Malat1* has been shown to bind PRC2 components EZH2 and suppressor of zeste 12 protein homolog (SUZ12) in a T cell lymphoma line (Kim et al., 2017). Other proteins often emerge as common *Malat1* protein interaction partners, these include TAR DNA-binding protein 43, (TDP-43) and Stauffen1 (STAU1), however, 127 potential *Malat1* interaction partners have been identified using RNA pull-down approaches followed by proteomic analysis in a liver cancer cell line (Chen et al., 2017a).

A growing body of work has linked *Malat1*-dependent regulation of its targets through directly sponging miRNAs/ acting as a competing endogenous RNA (ceRNA)

(Zhou et al., 2021). To name a few examples, *Malat1* has been reported to bind miR-125b which inhibits bladder cancer progression, or miR-1914-3p which promotes metastasis in NSCLC (Han et al., 2013; Jin et al., 2019). However, as miRNAs are predominantly cytoplasmic and *Malat1* is located in nuclear speckles it is unclear how these interactions take place. It is possible miRNAs are sequestered by *Malat1* in the early stage of miRNA processing. Further investigation into how and where these interactions take place would add confidence to *Malat1s* role in miRNA regulation.

1.3.3 *Malat1* knockout models

To begin unravelling the function and molecular relevance of *Malat1* three independent groups generated *Malat1* knockout (KO) mouse models (Eißmann et al., 2012; Nakagawa et al., 2012; Zhang et al., 2012). Each group abolished the expression of *Malat1* by altering the genomic locus of *Malat1* (***Figure 1.***7). Eißmann et al deleted the entire 7kb *Malat1* gene including 250 bp upstream of the TSS and 321 bp downstream of the *Malat1* gene (***Figure 1.7A*)**. qPCR analysis using primers which span the length of *Malat1* confirmed loss of *Malat1* expression in multiple tissues, with no residual *Malat1* transcript expression observed. Zhang et al generated a 3kb deletion in the 5' end of the *Malat1* genomic locus including its promoter (***Figure 1.7B*)**. Both southern and northern blot analyses confirmed the loss of *Malat1* in multiple tissue types. However, residual *Malat1* expression was observed in some regions of the brain. The authors postulate that the neural-specific RNA is produced from the 3' end of *Malat1* and is likely generated from an internal promoter that is activated in specific neural tissues. Nakagawa et al, prevented transcript expression of *Malat1* by introducing a transcriptional termination site in the form of the LacZ gene followed by a polyadenylation sequence 69 bp downstream of the TSS (***Figure 1.7C).*** Complete loss of full-length *Malat1* and mascRNA was confirmed in multiple tissue

types by northern blotting and situ hybridisation. Yet, as with the Zhang model, the Nakagawa group also noted residual *Malat1* expression in brain tissue, adding weight to the argument of a neural-specific 3' promoter. Interestingly, it was consistently observed that loss of *Malat1* resulted in no obvious phenotype. The knockout mice were fertile and show no developmental defects. However, the Zhang model noted that some of the genes which are adjacent to *Malat1* in the genome were dysregulated in adult mice which suggested an *in-cis* function. Nakagawa and colleagues also observed dysregulation of the neighbouring lncRNA gene *Neat1* in specific tissues upon loss of *Malat1*. Although there is no overt phenotype when *Malat1* is knocked-out, there are, however, other possible functions of *Malat1* and its function may only be revealed under certain conditions in specific cell types. Interestingly, AS is seemingly not affected in *Malat1* knockout mouse models (Nakagawa et al., 2012), this could be explained by potential redundant compensatory mechanisms during the development of *Malat1* knockout mice which do not occur in cultured cell lines (Zhang et al., 2017).

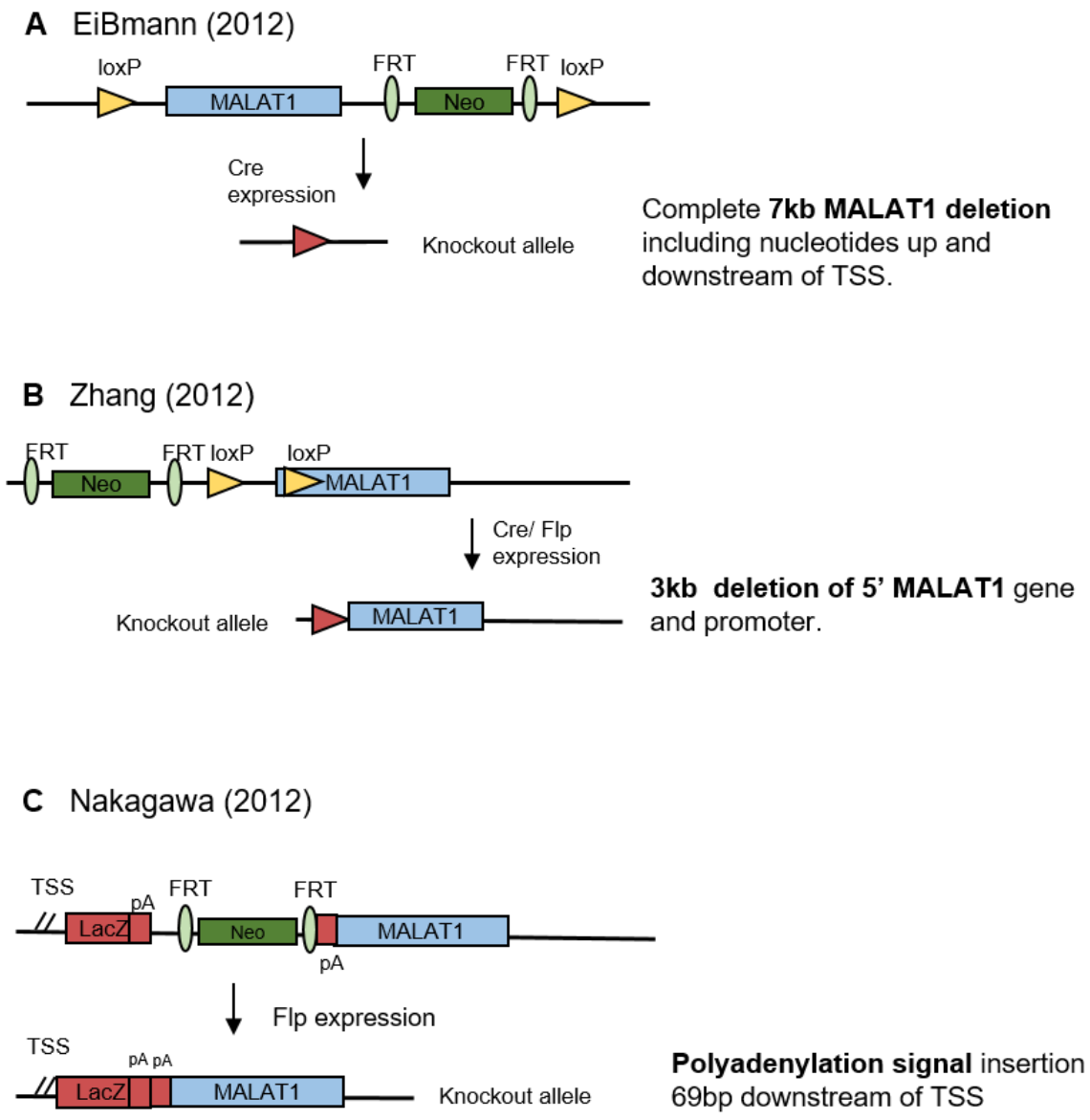


Figure 1.7: Schematic representation of the generation of *Malat1* KO mouse models.

A) EiBmann model in which the full length of *Malat1* was deleted including 250 bp upstream of the TSS and 321 bp downstream of the *Malat1* gene. **B)** Zhang model a 3kb deletion of the 5' end of *Malat1* including its promoter was deleted. **C)** Polyadenylation signals and the LacZ gene were inserted.

1.3.4 *Malat1* in the immune system

This thesis aims to explore the role of lncRNAs in the immune system. Given the high abundance and conservation between mammalian species, *Malat1* is an ideal lncRNA to probe functional relevance in the immune system. A growing body of evidence has described the functional relevance of *Malat1* in immune cells. A summary of reported functions for *Malat1* in immune contexts is described below.

1.3.5 The role of *Malat1* in macrophages

Several studies have examined the role of *Malat1* in macrophages. One study found that treatment of monocytes with LPS upregulated *Malat1* (M1 phenotype), whereas upon treatment with IL-4 *Malat1* was downregulated (M2 phenotypes). The knockdown of *Malat1* impaired LPS induced M1 activation and enhanced IL-4 M2 activated macrophages. (Cui et al., 2019). In support of this finding, treatment of MPC-83 cells with extracellular vesicles encapsulating *Malat1* induced and M1 phenotype (IL-6 and TNFα expression) (J. Liu et al., 2021)Yet, contrasting studies suggest that siRNA reduction of *Malat1* in bone marrow-derived macrophages, enhanced both an M1 and M2 phenotype, (Masoumi et al., 2019). The discrepancies in results between studies could be explained by differences in experimental design and the starting cell population.

1.3.6 The role of *Malat1* in neutrophils

Additional studies have indicated that *Malat1* may regulate neutrophil maturation and function. One study suggested that *MALAT1* secreted by human umbilical vein endothelial cells (HUVEC) cells in EVs promoted NET formation in co-culture experiments. (Gao et al., 2020). Wei and colleagues went on to investigate the role of *Malat1* in a rat model of ischemia-reperfusion injury (a common

complication after lung transplantation). Silencing expression of *Malat1* alleviated inflammation in the lungs following transplantation, partially through downregulating IL-8 which in turn inhibited neutrophil chemotaxis (Wei et al., 2019).

1.3.7 The role of *Malat1* in $CD8^+$ T cells

The role of lncRNAs in $CD8^+$ T cell differentiation remains poorly understood. However, recent research has examined the role of *Malat1* in the regulation of this process. During microbial infections, $CD8^+$ T cells give rise to short-lived effector cells and memory cells. Notably, *Malat1* was found to be associated with repressive H3K27me3 chromatin marks at several memory cell-associated genes. Consequently, the knockdown of *Malat1* reduced effector $CD8^+$ T cell formation and resulted in increased expression of several memory cell-associated genes (Kanbar et al., 2022).

1.3.8 The role of *Malat1* in B cells

Malat1 has also been shown to regulate B cell function. *MALAT1* expression was found to be higher in diffuse large B-cell lymphoma (DLBCL) compared to human normal B lymphocytes (IM-9I). siRNA-mediated knockdown of *MALAT1* decreased cell survival and increased the proportion of cells that were in the G0/G1 phase of the cell cycle. Additionally, in a subcutaneous tumour xenograft model, siRNA-mediated knockdown of *Malat1* decreased tumour volumes (Li et al., 2017). Additional studies using DLBCL cells found also found *MALAT1* to be upregulated in patient samples. Programmed death-ligand 1 (PD-L1) is an immune checkpoint protein which plays an essential role in suppressing the adaptive immune response through binding to PD-1 expressed on $CD4^+$ and $CD8^+$ T cells, this process reduces the proliferation of antigen-specific-T cells and apoptosis of Treg cells. PD-L1 expression is often enhanced during tumorigenesis and was found to be upregulated

in DLBCL tissues. miRNA-195 was found to be negatively correlated with *MALAT1* and *CD274* transcript abundance. The authors found that *MALAT1* sponged miRNA-195. Consequently, the knockdown of *MALAT1* using shRNAs increased miRNA-195 levels and decreased PD-L1 expression, which in turn decreased cell proliferation and immune escape (Wang et al., 2019). Mantle cell lymphoma (MCL) results from the malignant transition of B cells at the mantle zone of a lymph node. Similar to DLBCL, *MALAT1* expression was elevated in MCL tumours compared to healthy controls with high *MALAT1* expression significantly correlated with reduced patient survival. Knockdown of *MALAT1* impaired MCL cell proliferation and enhanced apoptosis (X. Wang et al., 2016). In summary, these studies provide compelling evidence for the role of *MALAT1* in promoting B cell proliferation, particularly in the context of B cell cancers.

1.3.9 *Malat1* in T helper cells

Interestingly, *Malat1* has been reported to play a functional role in T helper cells. However, the literature describes conflicting results (***Figure 1.8***). A recent *in vivo* study assessed the role of *Malat1* in $CD4^+$ T cell responses to acute infection with lymphocytic choriomeningitis virus (LCMV) (Yao et al., 2018). Flow cytometry analysis of T cell subsets revealed no significant difference in the proportion of different $CD4^+$ T cell populations between *Malat1* KO and WT peripheral cells after 8 days of infection. Contrastingly, siRNA-mediated knockdown of *Malat1* in primary naïve T cells pushes activated T cells towards a Th1/Th17 phenotype and inhibits Treg differentiation *in vitro* (Masoumi et al., 2019). Xue et al report that the knockdown of *Malat1* impairs Th17 differentiation and reduces the pathology of mouse models of acute viral myocarditis (AVMC) and *in vitro* (Xue et al., 2022). Other work has examined the role of *Malat1* in the regulation of the Th1/Th2 balance in asthma. They

suggest that overexpression of *Malat1* using pcDNA3.1-*Malat1* reduced IFN-γ, IL-2 and T-bet levels (Th1 markers) and increased IL-4, IL-10 and GATA3 levels. The authors postulate this regulation occurs through a miR-155/CTLA-4 axis (Liang & Tang, 2020). To gain a cohesive understanding of *Malat1* function in $CD4^+$ T cells further investigation is required.

Although the role of *Malat1* in immunity is only beginning to be uncovered, there is compelling evidence to suggest that *Malat1* plays a context-specific role in numerous cells of the immune system.

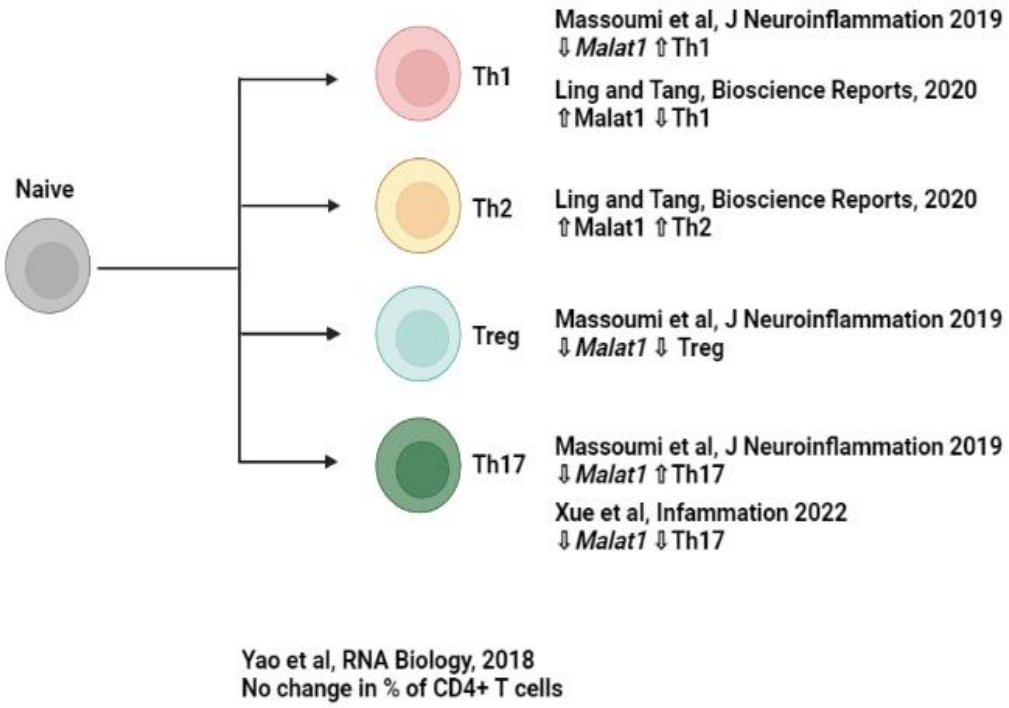


Figure 1.8: Schematic representation of the reported functions of *Malat1* in CD4$^+$ T cells.

Representation of findings from papers examining the role of *Malat1* in CD4$^+$ T cells.

1.4 Hypothesis and aims

A major challenge in the lncRNA field surrounds the question of functionality - are lncRNAs functionally relevant or a result of pervasive transcription. Here we aim to determine the role of the lncRNA *Malat1* in the immune system, specifically $CD4^+$ T cells. Several functions for *Malat1* have already been reported in an immune context. $CD4^+$ T cells offer the ideal platform for probing functionality given their centrality in immune responses and highly sophisticated transcriptional responses in response to antigen and differentiation stimuli. The central hypothesis tested in this thesis is that *Malat1* has a non-redundant role in the adaptive immune system through specific $CD4^+$ T cell functions.

In this thesis, I aim to characterise any potential roles of *Malat1* in $CD4^+$ T cells and determine how this impacts the immune response to pathogens. Below are the aims for the project and the associated results chapters:

1) To determine the role of *Malat1* in $CD4^+$ T cells – Chapter 3.

2) To characterise transcriptomic changes when comparing WT and *Malat1*$^{-/-}$ $CD4^+$ T cells in in naïve and effector cell populations – Chapter 4.

3) To investigate the *Malat1* protein interactome in $CD4^+$ T cells – Chapter 5.

4) To characterise RBP functions in the presence and absence of *Malat1* in $CD4^+$ T cells – Chapter 6.

5) To compare functions of *Malat1* in male and female $CD4^+$ T cells – Chapter 7.

2. Materials and methods

2.1 Animals

2.1.1 Ethical approval

Ethical approval was obtained from the University of York Animal Welfare and Ethical Review Body. Schedule 1 was performed under the Lagos non-PPL approved 6 Sept 2021. All procedures were performed under the authority of United Kingdom Home Office Project License PFB579996 and P49487014.

2.1.2 Mice

C57BL/6 CD45.2 mice were obtained from Charles River (UK). *Malat1* $^{-/-}$ mice were obtained from the Riken Institute (Nakagawa et al., 2012). All mice were housed under specific pathogen free conditions and bred internally in the BSF facility as a homozygous line. In all *in vitro* $CD4^+$ T cell polarisation experiments mice were between 6 and 12 weeks old. The mice were euthanised by rising CO_2 concentrations followed by cervical dislocation.

2.2 Cell line culture methods

EL4 cells are a murine T lymphocyte cell line. Cells were cultured in Roswell park memorial institute medium (RPMI 1640) (Gibco) supplemented with 10% heat inactivated FCS (Gibco), 1x glutamax (Gibco) and 1x penicillin Streptomycin (Gibco) under sterile conditions. Cells were cultured at 37 ^oC, 5% CO_2 and were passaged twice a week at a 1 in 20 split.

NIH 3T3 cells are a murine fibroblast cell line. Cells were cultured in dulbeccos modified eagle media (DMEM) (Gibco) supplemented with 10% heat inactivated FBS, 1 x glutamax and 1 x penicillin streptomycin (Gibco). To passage cells, any medium was first removed from the cell culture vessel and the cells washed with 1 x Phosphate buffered saline (PBS) $^{-Ca2+\ -Mg2+}$ (Invitrogen). Pre-warmed trypsin

(Invitrogen) was subsequently added to the vessel of cells to cleave the adhesion molecules. Cells were placed in a 37 ^{o}C, 5% CO_2 incubator for ~ 5 minutes or until the cells had detached. Trypsin was then diluted with media and the cells split 1:3. Cells were cultured at 37 ^{o}C, 5% CO_2 incubator and were passaged three times a week.

2.2.1 Enumeration of cells

The number of viable cells was typically determined by diluting 10 µl of cell suspension with 10 µl 0.4% trypan blue (Cytiva). 10 µl was loaded onto a haemocytometer chamber and all unstained, round, and glossy cells (those that are viable) were counted within a 16 square set using a bench top microscope. The total number of cells counted was multiplied by 2 (to account for the dilution) and then multiplied by 1 x 10^4. This gave a viable number of cells per ml which was used for experimental calculations.

2.2.2 Freezing cells

For long term storage cells were cryo-preserved and placed in liquid nitrogen. Cell number was determined then cells were pelleted by centrifugation at 250 g and re-suspended in 0.5-1 ml of freezing medium heat inactivated FCS supplemented with 10% dimethyl sulphoxide (DMSO) (Sigma). The cells were transferred to a cryovial, tightly wrapped in several layers of blue roll and placed at -80 ^{o}C for up to one week. Cells were placed into liquid nitrogen for long term storage.

2.2.3 Cell recovery

Cells were recovered from the liquid nitrogen and rapidly thawed in hand. Pre-warmed media was gently pipetted into the tube and the cells transferred to a flask or plate containing pre-warmed media. Cells were then cultured at 37 ^{o}C, 5% CO_2 until required.

2.3 *In vitro* T cell polarisation

2.3.1 Tissue isolation for $CD4^+$ T cell culture

The spleen, mesenteric, brachial, inguinal, and auxiliary lymph nodes were isolated from WT or *Malat1*$^{-/-}$ mice. A single cell suspension was generated by gently pressing the spleen and lymph nodes through a 70 µM cell strainer (Falcon) using the thumb side of a 2 ml syringe plunger, in to a 50 ml falcon tube. The strainer was washed with media to flush through any remaining cells. Subsequently, the cells were pelleted by centrifugation at 475 g for 5 minutes, the supernatant discarded, and the cells re-suspended in ~4 ml of ACK buffer (Gibco) to lyse the red blood cells. Cells were incubated in ACK for 5-10 minutes then 5 ml of media added, and the cells centrifuged at 475g for 5 minutes. The colour of the pellet was assessed to ensure red blood cells had lysed.

2.3.2 Fluorescent activated cell sorting (FACs) isolation of naïve $CD4^+$ T cells

The cells were counted using a haemocytometer and trypan blue staining. Cells were re-suspended in 97 µl of RPMI per 1 x 10^7 cells, and 3 µl of a $CD4^+$ micro beads (Miltenyi Biotech) were added to the cell suspension per 1 x 10^7 cells. Cells were left to incubate at 4 ^oC for 30 minutes. A LS column (Miltenyi Biotech) and was placed into a MACs separator (Miltenyi Biotech). The MACs column was rinsed with 3 ml of RPMI then the cell suspension was loaded onto the column. The column was rinsed with 3 x 3 ml of RPMI, and cells collected in a 15 ml falcon (negative fraction). The column was removed from the MACs separator and placed into a new 15 ml falcon. 5 ml of media was added to the column and cells were immediately flushed out using

the plunger. Purity of sorted cells was compared to the negative fractions by flow cytometry.

As the $CD4^+$ cell suspension will contain a mixture of both activated and naïve cells, to increase the percentage of naïve $CD4^+$ T cells in the population the cells were further sorted. Thus, cells were counted, pelleted, and re-suspended in 500 µl of RPMI containing the following fluorescently labelled antibodies (***Table 2.1***)

Table 2.1 Antibodies used for FACs sorting

Antibody	**Antibody dilution**	**Brand**	**Clone**
CD4 Per CP 5.5	1:250	Bio legend	R445
CD44 FITC	1:200	Bio legend	IM7
CD62L PE	1:200	Bio legend	MEL-14
CD8 APC	1:200	Bio legend	53-6.7
CD11b APC	1:200	Bio legend	M1170
MHC II APC	1:200	Bio legend	M5/114.15.2

Single-stained and un-stained conditions were used throughout. Cells were incubated for 20-30 minutes on wet ice protected from light. Cells were washed twice in RPMI and re-suspended at 1 x 10^7 cells / ml of RPMI. Cells were sorted using the MoFlo Astrios (Beckman Coulter) by a member of the York Biology Technology Facility based on $CD4^+CD62L^+CD44^-CD8^-CD11b^-MHC\ II^-$ expression into a 15 ml falcon tube (***Figure 2.1***).

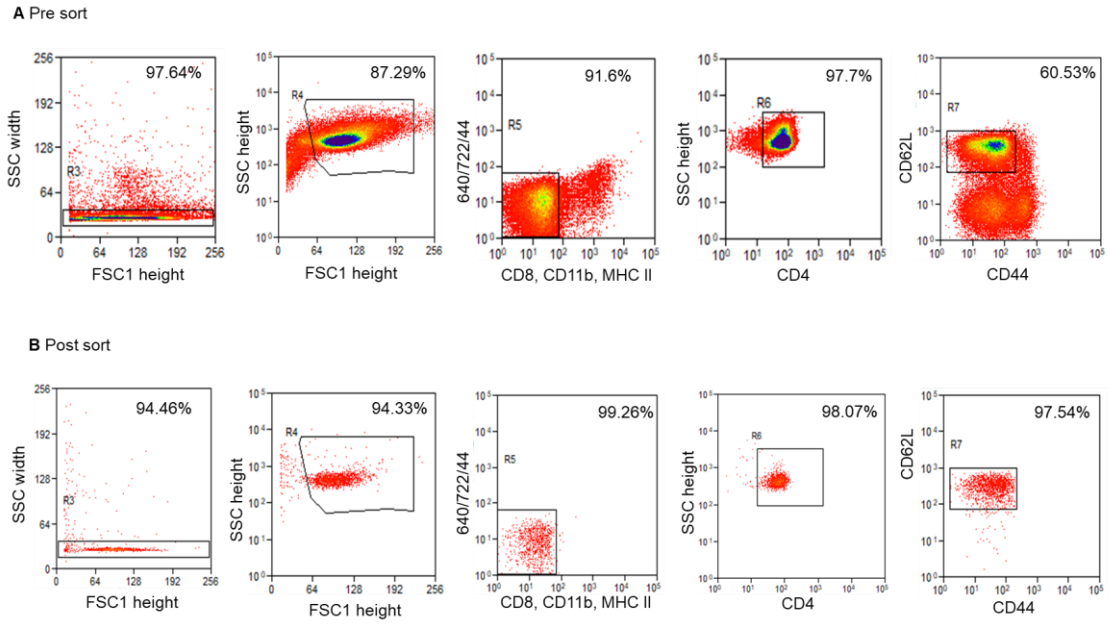


Figure 2.1: FACs sorting of naive CD4⁺ T cells.

Representative flow cytometry plots showing the gating strategy used to isolate naïve $CD4^+$ T cells by FACs sorting from pooled spleen and lymph nodes. Cells were sorted as single, live, $CD4^+CD62L^+CD44^-CD8^-CD11b^-MHC\ II^-$ **A)** Example of pre-sort populations **B)** Example of post-sort population.

2.3.3 Naïve $CD4^+$ T cell enrichment with miltenyi beads

As an alternative strategy to isolate naïve $CD4^+$ T cells, a naïve mouse $CD4^+$ T cell isolation kit was used (Miltenyi Biotech). Once the red blood cells had been lysed as per 2.3.1, the cells were counted using a haemocytometer and pelleted by centrifugation at 450 x g for 5 minutes. The cell pellet was re-suspended in 40 µl of MACs buffer per 10^7 total cells (2 mM EDTA, 0.5% BSA, 1x PBS, in house). 3 µl of biotin antibody cocktail was added to the cell suspension per 10^7 total cells. The cocktail of biotin-conjugated monoclonal antibodies included antibodies against CD8a, CD11b, CD11c, CD19, CD25, CD45R (B220), CD49b (DX5), CD105, MHC II, Ter-119, and TCRγ/δ. The antibody cocktail was mixed well with the cell sample by pipetting then incubated for 10 minutes in a fridge. Next, 20 µl of MACs buffer per 10^7 cells was added to the cell suspension along with 6 µl of anti-biotin microbeads per 10^7 cells and 3 µl of CD44 Micro beads per 10^7 cells. The sample was mixed well by pipetting and returned to the fridge for a further 15 minutes. The cells were then washed in 5 ml of MACs buffer and re-suspended in 1 ml of MACs buffer in preparation for magnetic separation.

A LS column (Miltenyi Biotech) was used for selection. The column was placed into a MACs Separator (Miltenyi Biotech). The MACs column was rinsed with 3 ml of MACs buffer then the cell suspension was loaded onto the column. The column was rinsed once with 3 ml of MACs buffer and cells collected in a 15 ml falcon (negative fraction representing naïve $CD4^+$ T cells). The column was removed from the MACs Separator and placed into a new 15 ml falcon. 5 ml of MACs buffer was added to the column and cells were immediately flushed out using the plunger to isolate the labelled cells (positive fraction). Successful isolation was confirmed by flow cytometry staining for CD4, CD62L and CD44 (***Figure 2.2***).

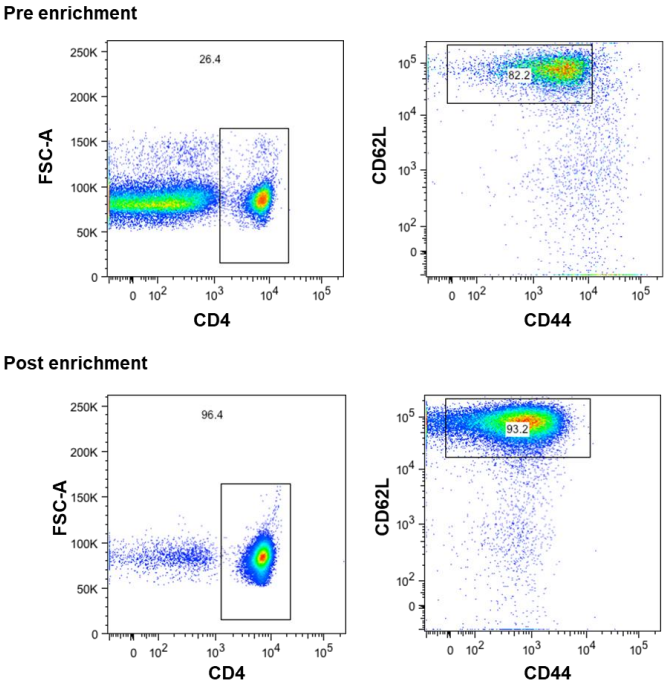


Figure 2.2: Bead enrichment of naive CD4⁺ T cells.

Representative flow cytometry plots of naïve CD4⁺ T cells pre and post bead enrichment. Cells were isolated from pooled spleen and lymph nodes.

2.3.4 T cell activation and polarisation

Flat bottomed 96 well plates were coated with 10 µg/ ml of anti-CD3 (Biolegend). Naïve $CD4^+$ T cells were plated in 200 µl of RPMI at density of 5 x 10^5 cells per well. To activate the cells, 4 µg/ ml of soluble anti-CD28 (Biolegend) was added to the media. To push cells towards a Th1 phenotype cells were cultured with 25 ng/ ml of rIL-12 (PeproTech) and 5 µg/ ml of anti IL-4 (Biolegend). To create Th2 polarised cells, 25 ng/ ml of rIL-4 (PeproTech) and 5 µg/ ml of anti IFNY (Biolegend) was added to the wells. Cells were cultured for 4 days at 37 ^oC, 5% CO_2. After 4 days cells were counted, washed with media and placed into a new 96 well plate with 10 U/ ml of human rIL-2 (PeproTech). Cells were cultured for a further 2 days at 37 ^oC, 5% CO_2 in rIL-2 before harvesting for experimentation. To induce Th17 differentiation, naive $CD4^+$ T cells were stimulated with 10 µg/ ml plate-bound anti-CD3 and 4 mg/ ml soluble anti-CD28 (37.51), and 1 ng/ ml of rTGF-b (PeproTech), 37.5 ng/ ml rIL-6 (PeproTech), 5 mg/ ml anti–IFNγ, and 5 mg/ ml anti–IL-4. After 3 days of stimulation, cells were transferred to a new 96-well plate in the presence of half the concentration of recombinant cytokines and inhibiting antibodies. Cells were harvested and analysed by flow cytometry at day 5. Antibodies used for T cell polarisation are listed in (***Table 2.2***).

Table 2.2. Antibodies used in T cell activation and polarisation experiments.

Antibody	Concentration	Clone
Anti-mouse IFNY	5 μg/ ml	XMG1.2
Anti-mouse IL4	5 μg/ ml	11B11
Anti-mouse CD28	4 μg/ ml	37.51
Anti-mouse CD3	10 μg/ ml	145-2C11

2.3.5 Recombinant Cytokines

Recombinant cytokines were purchased as lyophilised powders from PeproTech. Powders were reconstituted at 20 μg/ ml. rIL-12 p70 stock solutions were reconstituted in sterile 1x PBS pH 7.2-7.4 and 0.1% bovine serum albumin (BSA) (Thermo). rIL-4 was reconstituted in sterile water containing 0.1% BSA. Samples were reconstituted under sterile conditions and stored at -80^{o}C until required. Once thawed cytokines were not re-used.

2.4 Genotyping *Malat1* knockout mice

All *Malat1* knockout mice used in experimentation were confirmed as knockouts using the following genotyping protocol or qRT-PCR. Mouse ear clips were first heated to 95^{o}C submerged in 40 μl of a 25 mM NaOH solution for 15 minutes. NaOH was neutralised by adding 40 μl of 42 mM Tris-HCL Ph 7.5. Samples were cooled on ice and prepared for Phusion PCR screening. In order to amplify the desired region of the *Malat1* gene the reaction mix consisted of 10 μl 5 HF x Buffer (Thermofisher), 1 μl dNTP (10 mM stock), 2.5 μl of each primer 10 mM (***Table 2.3***), 0.25 μl Phusion Hot start II polymerase (Thermofisher), 29.25 μl nuclease free water

(Ambion) and 2 µl template per reaction. The PCR reaction took place at 98°C 30 seconds, 98°C 10 seconds, 60°C 30 seconds, 72°C 60 seconds (repeat previous 3 steps 40 cycles), 72°C10 minutes, 4°C ∞. The C1000 touch thermocycler (Biorad) was used. The PCR products were separated by size using a 2% TAE gel containing SYBRsafe (ThermoFisher), gel was run at 120 v for approximately 30 minutes. The gel was imaged on the INGENIUS gel doc (Syngene) using the automatic settings. Genotype was determined by comparing PCR product sizes 260 base pair (bp) band corresponding to WT and 900 bp band corresponding to KO.

Table 2.3. Forward and reverse primers used for genotyping

Target	**Primer names**	**Primer sequences 5′-3′**
Malat1	NEAT2 F NEAT2 R	AGC GTA GAG CAG CAC AGC TG GCT CTG GTC AGC CTC CAT TA
LacZ	LACZ R	GCA CAT CTG AAC TTC AGC

2.5 Measurement of RNA levels by qRT-PCR

Gene expression changes were measured by quantitative real-time reverse transcription polymerase chain reaction (qRT-PCR).

2.5.1 RNA extraction

Initially, cells were pelleted by centrifugation at 250 g and the supernatant discarded. The cells were re-suspended by disturbance in 1 ml of PBS and pelleted again by centrifugation. The supernatant was removed and cells were lysed by adding 700 µl of Quiazol (Quiagen) to the cell pellet. After thorough mixing the lysates were

stored at -80 ^{o}C until required. Frozen samples were thawed at room temperature then 140 µl of chloroform was added. Samples were vigorously shaken and left at room temperature for 5 minutes. The samples were then centrifuged at 12,000 g, 4 ^{o}C for 15 minutes. The upper aqueous layer was removed ensuring no white precipitate (protein/DNA) was taken with the sample. 525 µl of 100% ethanol was mixed with the upper aqueous phase. Up to 700 µl of sample was transferred to a RNeasy mini column (Quiagen). Samples were centrifuged at 8000 g for 15 seconds at room temperature and the flow through discarded. The same process was repeated until all the sample was passed through the column. 700 µl of RWT buffer (Quiagen) was added to the RNeasy column, samples were again centrifuged at 8000 g and the flow through discarded. The column was washed twice with RPE buffer (Quiagen) by centrifuging at 8000 g for 15s and discarding the flow through. The column was then placed in to a new Eppendorf and centrifuged again for 1 minute at 8000 g to remove any residual buffer. The column was transferred to a new RNase free Eppendorf and 30 µl of RNAse free water was pipetted directly onto the membrane. The samples were centrifuged at 8000 g for 1 minute to elute the RNA. RNA yield and quality was assessed using the Nanodrop 2000 (ThermoFisher).

2.5.2 Generation of cDNA

qRT-PCR requires a DNA template, the isolated RNA was converted into complementary DNA (cDNA) using the reverse transcriptase enzyme Superscript II (Invitrogren). 5 µl of RNA was added to a PCR tube in addition to 1 µl random hexamer (50 ng/ µl) (Thermofisher), 1 µl of dNTP mix (10mM) (Quiagen), and the final reaction volume made up to 10 µl with DEPEC H_2O. To denature the RNA samples were heated to 70 ^{o}C for 6 minutes using the thermocycler heated lid. Samples were then cooled on wet ice and a second master mix added which was comprised of

4 µl of 5x first strand buffer (Invitrogen), 2 µl of DTT, 1 µl of RNase OUT (Invitrogen) and 1 µl of superscript II reverse transcriptase (Invitrogen). Reverse transcription was carried out using the C1000 touch thermocycler (Biorad). Samples were heated to 25 ^{o}C for 10 minutes, followed by a 50 minute incubation at 50 ^{o}C and a 5 minute incubation at 85 ^{o}C to terminate the reaction. Samples were cooled on ice and stored at -20 ^{o}C until required.

2.5.3 qRT-PCR (SYBR green)

Fast SYBR green (Invitrogen) is a fluorescent dye which binds double stranded DNA formed by PCR and was used to quantify the PCR amplified DNA. Reaction mixes comprised 10 µl of fast SYBR green mix, 0.6 µl forward primer (10 mM) 0.6 µl reverse primer (10 mM), 7.8 µl of H_2O and 1 µl of cDNA (***Table 2.4***). Alternatively, if quantitect primers were used reaction mix was as follows, 10 µl of fast SYBR green mix, 2 µl of quantitect primer mix, 7 µl of H_2O and 1 µl of cDNA (***Table 2.5***).

Where possible, all reactions were performed in duplicate. The qRT-PCR was carried out using the StepOnePlus Real-Time PCR system (Applied biosystems). Standard fast thermocycling conditions were used. Where possible no template controls (NTC) were incorporated.

2.5.4 PCR primers

Table 2.4. Forward and reverse primers used for RT-qPCR

Target	**Primer names**	**Primer sequences 5′-3′**
Neat1	*Neat1 F*	CCT AGG TTC CGT GCT TCC TC
		CAT CCT CCA CAG GCT TAC

	Neat1 R	
Malat1	*Malat1 F* *Malat1 R*	GCA GTG TGC CAA TGT TTC GT AGT CTG CTG TTT CCT GC TCC
U6	*U6 F* *U6 R*	CGC TTC GGC AGC ACA TAT AC TTC ACG AAT TTG CGT GTC AT
Hotair	*Hotair F* *Hotair R*	CCT TAT AAG CTC ATC GGA GCA CAT TTC TGG GTG GTT CCT TT
Dynlt1b	*Dynlt1b F* *Dynlt1b R*	CGA AGA CTT CCA GGC CTC A GGC GCT TTC TAT AGC CTC CT
Il21	*Il21 F* *Il21 R*	GCC TCC TGA TTA GAC TTC GTC AC CAG GCA AAA GCT GCA TGC TCA C
Eef1g	*Eef1g F* *Eef1g R*	CAG TGA CAT CGT TCC TCC AGC T CGA GTC TTC AAG TGA GTG TCC AG
Ago4	*Ago4 F* *Ago 4 R*	CAC ACG CAT CAT CTA CT ACCGC GCC GAT AGT CTT CCT CCA CCA
Emilin2	*Emilin2 F* *Emilin2 R*	CGC TCA AAT CGC ACT CCA GAG A TCG GTT GCT TCT GAG GGT TCC T

Tnfsf9	*Tnfsf9 F* *Tnfsf9 R*	CCA AGT ACC TTC TCC AGC ATA GG GCF TTG TGG GTA GAG GAG CAA A
Sox5	*Sox5 F* *Sox5 R*	CGC CAG ATG AAA GAG CAA CTC AG TGA GTC AGG CTC TCC AGT GTT
Eif4a1	*Eif4a1 F* *Eif4a1 R*	TAC ATC GGT GCC TCT TGT CA GGG TAC CCA CGA TGA TAT GG
18s	*18s F* *18s R*	TGC CAG AGT CTC GTT CGT TA GGT GCA TGG CCG TTC TTA
Cd69	*Cd69 F* *Cd69 R*	CCC TTG GGC TGT GTT AAT AGT G AAC TTC TCG TAC AAG CCT G
Hnrnpa1	*Hnrnpa1 F* *Hnrnpa1 R*	TGG AAG CAA TTT TGG AGG TGG CGT TCC GTG GTT TAG CAA AGT

Additionally, several quantitect primers were used throughout experimentation (**Table 2.5**). Quantitect primers are a pool of forward and reverse primers for the desired target that have previously been validated and shown to have 100% PCR efficiency and are suitable for SYBR Green based RT-PCR.

Table 2.5. Quantitect primers used for RT-qPCR

Target	Assay name
STAT6	Mm-STAT6-1-SG
IL-10	Mm-IL-10-1-SG
MAF	Mm-MAF-1-SG
STAT4	Mm-STAT4-1-SG
TigD3	Mm-Tigd3-1-SG
Tbet	Mm-Tbx21-1-SG
GATA3	Mm-GATA3-1-SG
BLIMP1	Mm-Prdm1-1-SG
Bhlhe40	Mm-Bhlhe40-1-SG
IL-4	Mm_IL4_va.1_SG
IFNg	MM_IFng_1_SG

2.5.5 Primer efficiency

In a PCR reaction the quantity of cDNA doubles with each cycle. However, the efficiency of amplification can vary between primers. Thus, to enable comparison between RNA levels, different primers must amplify at a similar efficiency ~90-110%. To determine primer efficiency a 6-point standard curve was generated by serially diluting a cDNA template. The qPCR reaction was carried out as normal with samples in duplicate. Using Excel (Microsoft 2016) Ct values were averaged and logged. Subsequently a scatter plot was created plotting average Ct values vs log values and

the slope of the trend line determined. To calculate efficiency the following equation was used:

Efficiency (%) = $(\frac{-1}{10^{slope}-1})$ x 100

All in house PCR primers have been tested to ensure appropriate PCR efficiency. Some primers had previously been optimised by previous laboratory members.

2.5.6 qRT-PCR analysis

The average Ct value (PCR cycle at which fluorescence is above background) was used for subsequent analysis. Examination of target gene expression was analysed using the $2^{-\Delta\Delta Ct}$ method using Excel 2016 (Microsoft).

2.6. *Malat1* knockdown approaches

2.6.1 siRNA knockdown

$1.5x10^6$ cells were pelleted by centrifugation at 250 g. The supernatant was removed, and the cells re-suspended in 1 ml of PBS. Cells were pelleted once more and the supernatant removed, care was taken to ensure a dry pellet was obtained. The pellet was re-suspended in 150 µl of R buffer (Invitrogen) containing 100 nM of *Malat1* Lincode SMART pool siRNA (Dharmacon) or 100 nM of Lincode SMART pool NTC (Dharmacon) (***Table 2.6)***. EL4 cells were electroporated using the Neon ® transfection system (Invitrogen) in a 100 µl tip (Invitrogen) at a pulse voltage of 1,080 (v), a pulse width of 50 (ms) and 1 pulse number. Cells were plated into a pre-warmed 6 well plate containing 2 ml of RPMI, 10% FBS 1x L Glu (without antibiotics). Cells were cultured at 37 ºC, 5% CO_2 for 48 hours. Cells were harvested at this time point for analysis. *Malat1* knockdown was assessed by qRT-PCR.

Table 2.6. siRNA sequences used for knockdown experiments.

Target name	**Sequence 5′-3′**
Lincode SMARTpool siRNA N-174310-01 *Malat1*	GUUUAAAUGCUUACGAUCA
Lincode SMARTpool siRNA N-174310-02 *Malat1*	AUAGAGUAGCUUAUCGAAA
Lincode SMARTpool siRNA N-174310-03 *Malat1*	GGUUAGAGAAGGCGUGUAC
Lincode SMARTpool siRNA N-174310-04 *Malat1*	GAGAGCAUGCGGUGCGGUA
Lincode non-targeting pool sequence 1	UGGUUUACAUGUCGACUAA
Lincode non-targeting pool sequence 2	UGGUUUACAUGUUGUGUGA
Lincode non-targeting pool sequence 3	UGGUUUACAUGUUUUCUGA
Lincode non-targeting pool sequence 4	UGGUUUACAUGUUUUCCUA

2.6.2 siRNA reconstitution

siRNAs used throughout experiments were ordered as lyophilised powders (Dharmacon). Tubes containing siRNA were briefly centrifuged to ensure the siRNA pellet was at the bottom of the tube. The pellet was re-suspended in 500 µl of RNase-free 1x siRNA buffer (Dharmacon) to create a 20 µM stock. After gentle mixing by pipetting the solution was further mixed for 30 minutes on a tube rotator. siRNA was aliquoted and stored at -20 ^{o}C until required.

2.6.3 GapmeR knockdown procedure

Cells of interest were plated into a flat bottomed 96 well tissue culture plate with 200 µl of cell culture medium. 100 nM of targeted or non-targeting GapmeRs were added to the plate wells and left to enter the cells by gymnosis (***Table 2.7).*** Cells were harvested at various time points post GapmeR introduction and analysed by qRT-PCR to assess the level of knockdown.

Table 2.7: GapmeR sequences used in knockdown experiments.

Target name	**Sequence 5′-3′**
Antisense LNA GapmeR Control: Negative control A	A*A*C*A*C*G*T*C*T*A*T*A*C*G*C
Antisense LNA GapmeR control: *Malat1* (mouse) positive control	G*T*C*A*C*A*A*T*G*C*A*T*T*C*T*A

2.6.4 GapmeR reconstitution

GapmeRs were purchased as lyophilised powders from Quiagen. Upon arrival, GapmeRs were briefly centrifuged to ensure that all of the material was collected at the bottom of the tube. GapmeRs were re-suspended at 50 µM in nuclease free water. After gentle mixing, samples were aliquoted under sterile conditions and stored at -20^{o}C until required. Freeze thaw cycles were kept to a minimum, no GapmeR was freeze thawed any more than five times.

2.7 Flow cytometry

2.7.1 Flow cytometry buffers

FACs buffer*:* PBS supplemented with 0.5% BSA and 0.05% sodium azide (Sigma)

Intracellular cytokine Fix buffer: BD fix cytofix/cytoperm (BD Biosciences 554722)

Intracellular cytokine1x Perm buffer*:* Diluted from 10x BD perm/wash 10x solution (BD Biosciences) with deionised water.

Intranuclear staining buffer: Fix/Perm buffer (eBiosciences)

2.7.2 Flow cytometry antibodies

Combinations of the following antibodies were used for flow cytometry to investigate protein expression in specific cell populations (**Table 2.8).**

Table 2.8. Antibodies used for flow cytometry experiments.

Target	Fluorophore/ Conjugation	Manufacturer	Clone	Dilution factor
IL-10	PE	Bio legend	JE55-16E3	1 in 100
IL-4	APC	Bio legend	11B11	1 in 100
IFNγ	FITC	Bio legend	XMG1.2	1 in 200
CD4	PerCPcy5.5	Bio legend	R445	1 in 125
TCRβ	PE Cy7	Bio legend	H57.597	1 in 200
CD44	FITC	Bio legend	IM7	1 in 250
CD62L	PE	Bio legend	MEL-14	1 in 250

CD8	APC	Bio legend	53-6.7	1 in 250
CD11b	APC	Bio legend	M1170	1 in 250
MHC II	APC	Bio legend	M5/114.15.2	1 in 250
CD45.2	BV786	Biolegend	104	1 in 250
CD45.1	Alexaflour700	Biolegend	A20	1 in 125
SiglecF	PerCP710	Biolgend	IRNM44N	1 in 250
F4/80	APC	Biolegend	BM8	1 in 125
CD19	APC Cy7	Biolegend	6D5	1 in 250
Ly6g	FITC	Biolegend	IA8	1 in 250
CD11b	Pacblue	Biolegend	M1/70	1 in 250
CD4	BV650	Biolegend	RM4-5	1 in 250
IL-13	E450	eBiosciences	eBio13A	1 in 100
CD62L	PE	Biolegend	MEL-14	1 in 250
CD41	APC	Biolegend	MEReg30	1 in 250
IL-4	PE Dazzle	Biolegend	11B11	1 in 250
Ly6G	APC Cy7	Biolegend	IA8	1 in 250
CD41	FITC	Biolegend	MWReg30	1 in 250
Ly6C	BV605	Biolegend	HK1.4	1 in 250
SiglecF	E710	Biolegened	IRNM44N	1 in 250
CD64	PE	Biolegend	X54-5/7.1	1 in 250
MHCII	Alexaflour700	Biolegend	M5/114.15.2	1 in 250

iNOS	PE e610	Biolegend	CxNF7	1 in 125
RELM*α*	Unconjugated	Peprotech	H1717	1 in 200
Ym1	Biotin	R & D Biosystems	BAF2446	1 in 200
Goat anti-Rabbit	Alexaflour A647	Invitrogen	A27040	1:400
Streptavadin	PeCy7	Biolegend	405206	1:400

2.7.3 Flow cytometry surface staining

For flow cytometry staining, cells were harvested and washed twice in 1x phosphate buffered saline (PBS $^{-Ca2+-Mg2+}$) (Sigma) by centrifugation at 450 g. Dead cells were stained by re-suspending cells in 100 µl of PBS$^{-Ca2+-Mg2+}$ containing 0.1 µl of live dead zombie aqua (Biolegend) and incubated for 10 minutes protected from light on wet ice. Cells were washed in 1 ml of FACs buffer. Subsequently, Fc receptors were blocked by resuspending the pellet in 3 µl of 2 mg/ ml rat IgG (Sigma) and incubating the sample for 5 minutes on wet ice protected by light. To stain for surface markers an appropriate antibody cocktail was added to the cell suspension and left to stain of wet ice for 20 minutes protected from light. Cells were washed twice in FACs buffer by centrifugation at 450 g. If appropriate cells were fixed in 4% paraformaldehyde (PFA) for 20 minutes on wet ice protected from light (Fisher bioreagents) Cells were washed twice in FACs buffer and re-suspended in 450 µl of FACs buffer and analysed by flow cytometry.

2.7.4 Intracellular cytokine staining for flow cytometry

To assess cytokine production, cells were incubated in RPMI for 4 hours in the presence of 10 µg/ ml brefeldin, 0.5 µg/ ml PMA, and 1 µg/ ml ionomycin at 37 ^{o}C, 5% CO_2. Surface staining was performed as described in 2.1.3. However, cells were fixed in 150 µl of BD fix and incubated on wet ice for 20 minutes. Cells were washed twice in 1x perm buffer and re-suspended in 50 µl of 1 x BD perm with the desired antibody cocktail. Cells were stained with the desired antibody cocktails on wet ice for 20 minutes protected from light. Cells were washed twice in 500 µl of 1 x BD perm and finally re-suspended in 450 µl of FACs buffer and analysed by flow cytometry.

2.7.5 Flow cytometry data analysis

The LSR Fortessa-X-20 (BD) or LX375 CytoFLEX (Beckman Coulter) was used for all flow cytometry. Compensation was carried out using single stained cells. All samples were gated on single, live cells, and analysis carried out using FlowJoV10.6.1® (FlowJo®).

2.7.6 Flow cytometry gating strategy

To examine specific cell populations, different gating strategies and antibody panels were used. An example gating strategies to identify CD45.1 and CD45.2 cells in a mixed bone marrow chimera model is shown in (***Figure 2.3***). An example of identification of different lymphoid populations is shown in (***Figure 2.4***). Similarly, an example gating strategy is shown that was used to identify intracellular cytokines in $CD4^+$ T cells (***Figure 2.5***). Finally, a gating strategy is shown that was used to identify different myeloid subpopulations (***Figure 2.6***). Of note, Magnus Gwynne analysed the myeloid panels in this thesis.

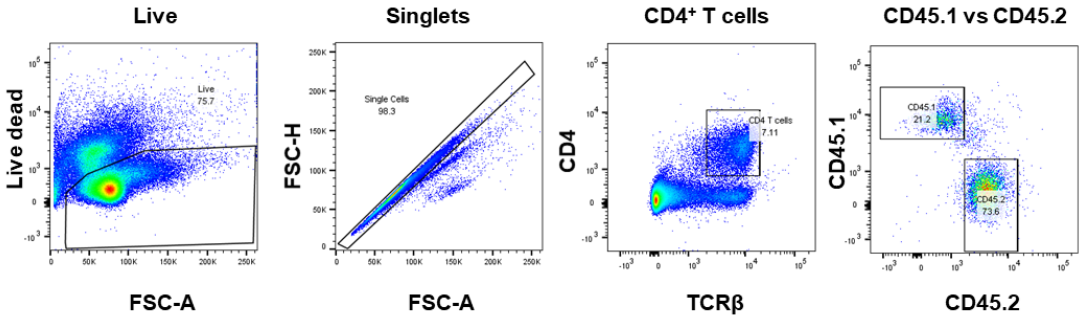


Figure 2.3: CD45.1 vs CD45.2 gating strategy example.

Example gating strategy used to determine CD45.1 and CD45.2 $CD4^+$ T cells using flow cytometry. Example is from spleen tissues derived a mixed bone marrow chimera of WT and *Malat1*$^{-/-}$ mice infected with *S.mansoni* worms.

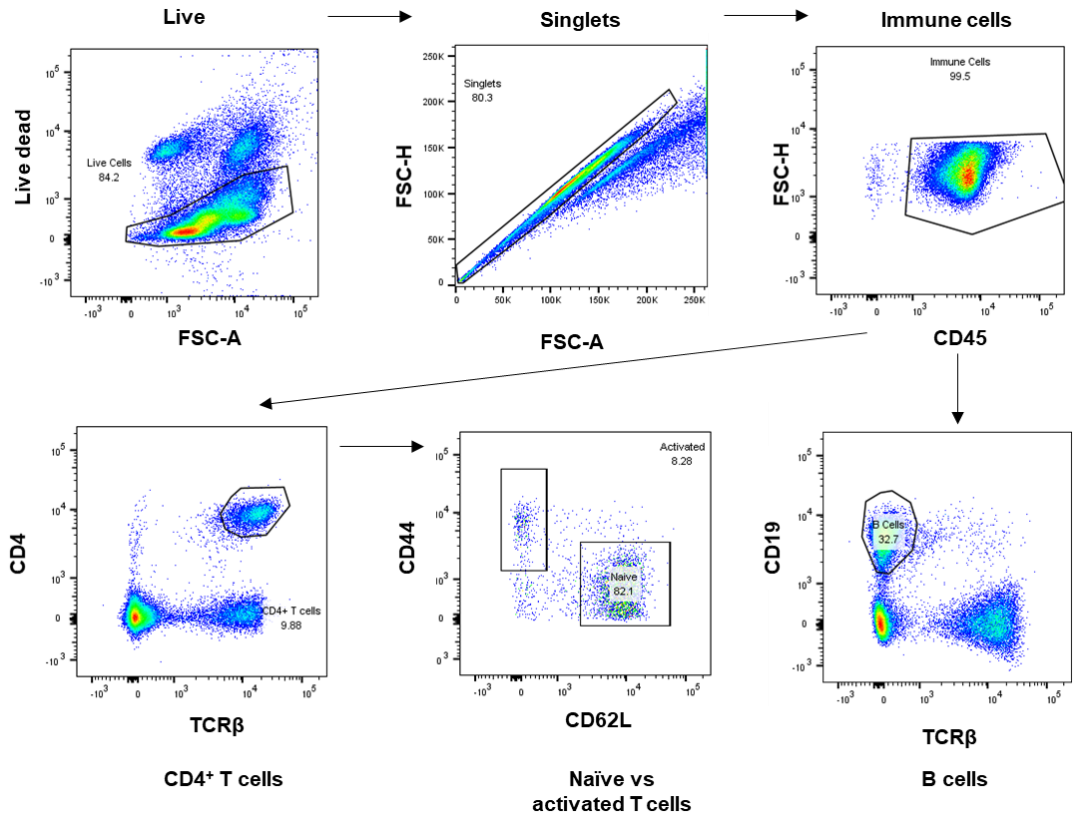


Figure 2.4: Gating strategy used to identify lymphocytes.

Example gating strategy used to identify naïve and activated $CD4^+$ T cells and B cells Example shown is from a WT lung egg injected sample of an *S.mansoni* egg model of inflammation.

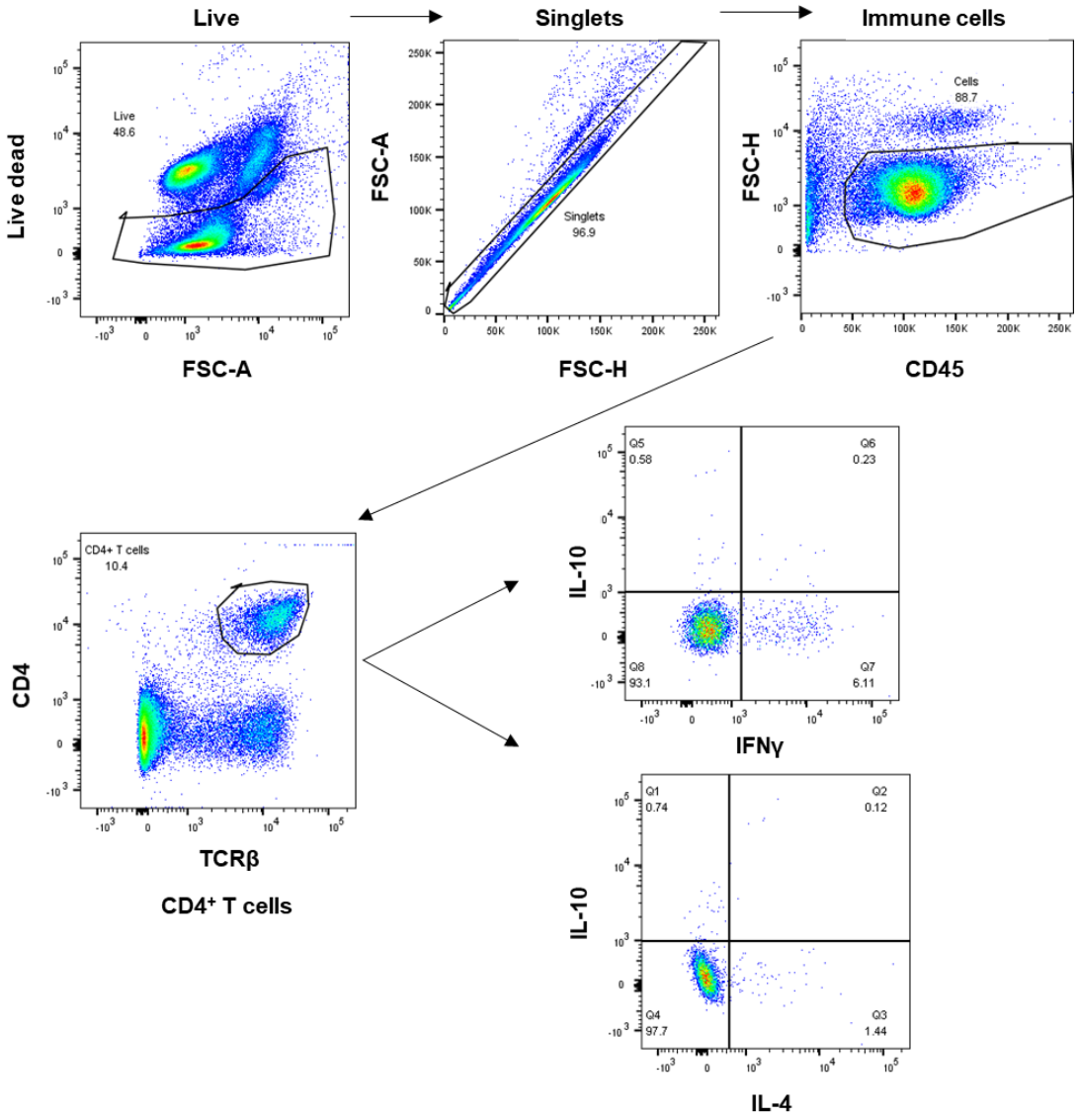


Figure 2.5: Intracellular cytokine staining gating strategy.

Example gating strategy used to identify IL-10, IL-4 and IFNγ expression derived from $CD4^+$ T cells. Example shown is from a WT lung egg injected sample of an *S.mansoni* egg model of inflammation.

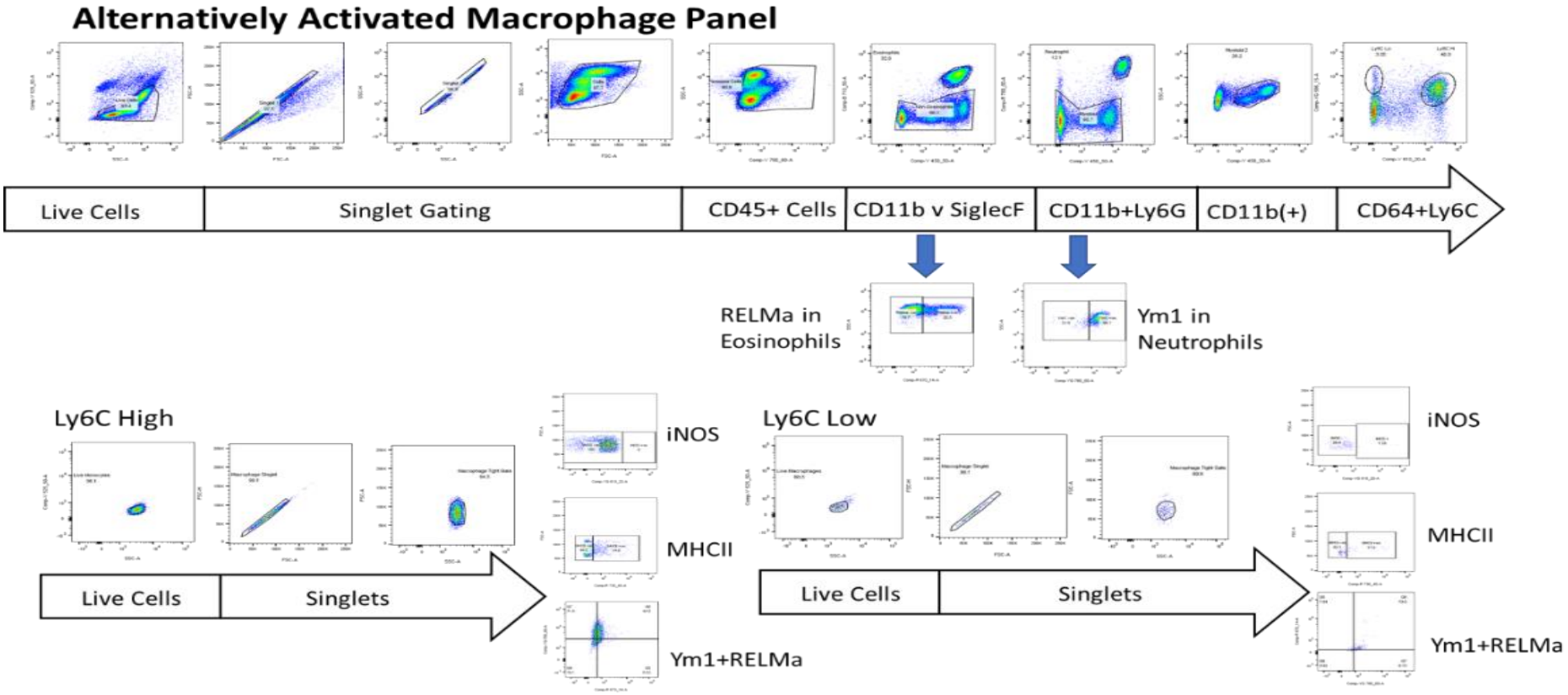


Figure 2.6: Myeloid gating strategy. Gating strategy used to identify different myeloid sub populations. Example shown is from a WT lung injected sample of an *S.mansoni* egg model of inflammation. N.B this figure was provided by Magnus Gwynne who performed the gating for this experiment.

2.8 Western blotting

2.8.1 Western blotting buffers

1 x Transfer buffer: 50 ml 10x Tris/Glycine stock (0.25 M Tris, 1.92 M Glycine) (Geneflow), 50 ml methanol, 400 ml ddH_20

1x Running buffer: 100 ml 10x Tris/Glycine/SDS stock (0.25 M Tris, 1.92 M Glycine, 1% SDS) (Geneflow) 900 ml ddH_20

Blocking buffer: 2% BSA TBST (0.1%) pH 8.0

1 x TBST: 100 ml 10x TBS, 900 ml H_2O and 1 ml Tween

2.8.2 Generation of cell lysates

Cells were isolated and washed once in PBS by centrifugation at 250 g. The supernatant was removed, and cells were lysed in 100 µl of radio immunoprecipitation assay buffer (RIPA) buffer containing 1 µl of protease inhibitors at a density of ~1 x 10^6 cells per 30 µl of lysis buffer. Once collected samples were centrifuged at 10,000 g, 4 ºC for 5 minutes. The soluble fraction of the lysates was isolated leaving the pellet undisturbed.

2.8.3 Assessment of protein concentration

When appropriate protein concentration was calculated using a bicinchoninic acid assay (BCA) (ThermoFisher). Samples were diluted 1 in 6 with PBS and 5 µl added to a flat bottomed 96 well plate in duplicate. 95 µl of assay reagents (A 93.1 µl + B 1.9 µl) was added to the samples and left to incubate at 37 ºC for 30 minutes. Calibration curves were generated from recombinant protein standards using step wise dilutions. Plates were read at a 562 nm wavelength.

2.8.4 Separation of proteins

Samples were diluted in 4x loading buffer and H_2O. The proteins were denatured by heating at 95^{oC}, for 10 minutes with regular pulsing and vortexing. The proteins were separated by size sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) using an 8% polyacrylamide gel, 1.5mm, 12 well (in house) at 120 volts (v), 300 watts (W) for 1 hour 30 minutes.

2.8.5 Transfer to membranes

The proteins were transferred to a PVDF membrane using the Trans-blot SD semi dry transfer cell. Initially the membrane was immersed in methanol for 1 minute, then immersed in H_2O for 1 minute and subsequently soaked in 1x transfer buffer until required. Whattman paper and the SDS-PAGE gel were also soaked in 1x transfer buffer before use. The sandwich was assembled and any air bubbles removed by rolling with a cylinder. The transfer took place at 25 (v), 0.2 amps, for 1 hour 30 minutes. Once completed success of transfer was confirmed by checking no ladder was left on the gel. The membrane was rinsed briefly in TBST then blocked for 1 hour at room temperature with gentle rocking.

2.8.6 Western blotting antibodies

Blocking solution was removed and the membrane was transferred to a 50 ml falcon tube containing a 5 ml of the desired antibody in TBST 5% BSA (***Table 2.9***).

Table 2.9. Primary antibodies used for western blotting.

Antibody	Company	Clone	Host species	Dilution used
αMAF (mouse)	Proteintech	Polyclonal	Rabbit	1 in 1000
αSTAT4 (mouse)	Proteintech	Polyclonal	Rabbit	1 in 1000
αPhosphoSTAT4 (mouse)	Thermofisher	Polyclonal	Rabbit	1 in 1000
α βActin (mouse)	Abcam	AC-15	Mouse	1 in 5000
αSRSF1	Proteintech	12929-2-AP (Polyclonal)	Rabbit	1 in 1000
αhnRNPA1	Proteintech	11176-1-AP (Polyclonal)	Rabbit	1 in 1000

Membranes were incubated with primary antibody overnight at 4^{o}C with gentle rocking. After incubation the antibody solution was frozen and re-used approximately between 5 and 10 times. Membranes were rinsed in 3 x 10-minute washes in 1x TBST. After rinsing membranes were transferred to a new falcon tube and incubated in 5 ml of secondary antibody solution (***Table 2.10)*** for 1 hour at room temperature with gentle rocking.

Table 2.10. Secondary antibodies used for western blotting.

Antibody	Company	Clone	Dilution used
HRP Goat anti Rabbit Ig	Dako	Polyclonal	1 in 5000
HRP Goat anti Mouse Ig	Dako	Polyclonal	1 in 5000

2.8.7 Imaging of the membrane

After the 1 hour incubation, the membrane was washed 3 x 10 minutes in 1 x TBST followed by 2 x quick washes in ddH_2O. As the secondary antibodies are horse radish peroxidase (HRP) conjugated to enable detection ECL (Amersham) was added to the membrane. ECL reagents A + B were first mixed together at a 1 to 1 ratio and an appropriate volume added to cover the size of the blot. The blot was placed on to a glass plate, covered in cling film and imaged using the Biorad ChemiDoc MP imaging system (Biorad).

2.8.8 Re-probing the membrane

When needed, the membrane was stripped and re-probed. The membrane was incubated with 5 ml of blot stripping buffer (Thermo Scientific) for 10 minutes in a 50 ml falcon tube. To remove residual stripping buffer the membrane was rinsed 4 x 10 minute washes in 1 x TBST. The membrane was re-blocked and probed as before.

2.9. Microscopy

2.9.1 Cytospinning

To adhere suspension cells to a microscope slide cells were first pelleted by centrifugation at 450 g supernatant was removed and cells re-suspended in 1 x PBS at

a density of 1 x 10^5 cells per 250 µl. 250 µl of cell suspension was then loaded in to a cytofunnel EZ double with white filter paper (Thermo Scientific) containing a super frost plus microscope slide (Thermo scientific). Cells were spun on the slide using the Cytospin 4 (Thermo Shandon) at low acceleration, 500 rpm for 5 minutes. After spinning the slide was removed from the cytofunnel and inspected to ensure cells had successfully transferred to the slide. A hydrophobic barrier was created by circling the cells with a PAP pen.

2.9.2 Seeding adherent cells onto coverslips

NIH3T3 cells were passaged 24 hours prior to immuno-fluorescence staining and seeded into a 9 cm culture dish containing an appropriate number of circular sterile cover slips. The coverslips were gently pressed to the base of the dish using the end of a sterile pipette tip. Cells were returned to a 37 ^oC, 5% CO_2 incubator to adhere to the coverslips. To stain the cells, coverslips were carefully moved using tweezers between wells of a 24-well plate containing the appropriate buffers.

2.9.3 Staining of the nucleus

Buffers:

1 x $PBS^{-Mg2+, -Ca2+}$ (Invitrogen)

CSK:

- 10 mM PIPES pH 6.8
- 100 mM NaCl
- 300 mM sucrose
- 1 mM $MgCl_2$
- 1 mM EGTA

To assess the integrity of the nucleus cells were initially fixed and permeabilised. Once adhered to a cover slip or microscope slide celled were fixed in 8% paraformaldehyde (PFA) (Sigma) for 15 minutes at room temperature. To remove residual PFA the cells were washed twice in 1 x PBS-$^{Mg2+, -Ca2+}$. Membranes were permeabilised with CSK buffer supplemented with 0.1% Triton (ACROS) and 1 mM DTT (Invitrogen) for 10 minutes at room temperature. Next the nucleus was stained with a 1 in 5000 dilution of DAPI (Thermofisher) in 1 x PBS$^{-Mg2+, -Ca2}$ for 1 hour at room temperature. Cells were washed once in PBS the mounted in Vectasheild with DAPI (Vectorlabs).

2.9.4 SRSF1 and hnRNPA1 imaging

One million EL4 cells or primary $CD4^+$ T cells at different stages of differentiation were harvested and washed in 10 ml of PBS by centrifugation. Supernatant was removed and the cell pellet was re-suspended in 2 ml 4% paraformaldehyde (PFA). Cells were pipetted up and down 10 times to break up the cell clump and incubated at 37^{o}C for 1 hour in a water bath to fix the cells. Cells were pelleted and re-suspended in 70% ethanol at 1 x10^6 cells/ ml with thorough pipetting to create a single cell suspension. Cells were spun onto a microscope slide as previously described then the samples were dehydrated by first immersing the slides in 50% ethanol for 5 minutes are room temperature. The slides were moved to a 70% ethanol solution for 5 minutes, until being replaced with 100% ethanol and finally stored at -20^{o}C until required.

Cells were removed from -20^{o}C and air dried for 30 minutes at room temperature. Then the cells were encircled with a hydrophobic pen. Cells were permeabilised by incubating samples in 0.4% triton in PBS for 15 minutes at room temperature. Samples were washed 3 times in PBS for 5 minutes. Samples were

blocked with 1% BSA in PBS for 10 minutes at room temperature. Cells were incubated with primary antibodies for 2 hours at 37^{oc}, samples were washed 3 times with PBS (5 minute incubations). Subsequently, samples were incubated in secondary antibodies for 2 hours at 37ºC (Alexa Fluor 488 green goat a-rabbit /Alexa Fluor 568 red goat a-rabbit (Thermofisher)). Primary antibody details are listed in (***Table 2.11***). Cells were washed three times in PBS and mounted in Vectasheild medium containing DAPI (Vector Labs).

Table 2.11: Antibodies used for immune-fluorescence.

Antibody target	Clone	Species	Brand
SRSF1	12929-2-AP (Polyclonal)	Rabbit	Proteintech
hnRNPA1	11176-1-AP (Polyclonal)	Rabbit	Proteintech

2.9.5 RNA-Scope

One million cells were harvested and washed in 10 ml of PBS by centrifugation. Supernatant was removed and the cell pellet was re-suspended in 2 ml PFA. Cells were pipetted up and down 10 times to break up the cell clump and incubated at 37ºC for 1 hour in a water bath to fix the cells. Cells were pelleted once more and re-suspended in 5 ml of CSK with 0.1% Triton x100 and incubated for 5 minutes at room temperature. Cells were pelleted again and re-suspended in 70% ethanol at 1 $x10^6$ cells/ ml with thorough pipetting to create a single cell suspension. Cells were spun onto a microscope slide as previously described then the samples were dehydrated by first immersing the slides in 50% ethanol for 5 minutes are room

temperature. The slides were moved to a 70% ethanol solution for 5 minutes, until being replaced with 100% ethanol and finally stored at -20^{o}C until required.

Cells were removed from -20^{o}C and air dried for 30 minutes at room temperature. Then the cells were encircled with a hydrophobic pen. ~1-2 drops of RNAscope® Hydrogen Peroxide was added to entirely cover the cells. RNAscope® Hydrogen Peroxide was removed from the slide by gentle tapping onto absorbent paper. The slides were then immersed in 1 x PBS in a coplin jar, and washed 5 times by changing the 1 xPBS solution and moving the slides up and down.

RNAscope® 1X Target Retrieval Reagents were heated to 95^{0c} and the slides submerged into the solution and incubated for 5 minutes. The slides were then transferred to a coplin jar with 1 x PBS and rinsed for 25 seconds. The PBS was then replaced with 100% ethanol for 3 minutes. Slides were baked at 60^{o}C for 45 minutes to completely dry the slides.

Slides were treated with ~1 drop of RNAscope® Protease 3 (enough to completely cover the cell pellet). Slides were then incubated at 40^{o}C for 30 minutes then washed as before five times in 1 x PBS.

Malat1 C1 probe was warmed for 10 minutes at 40^{o}C in a water bath, then left to cool to room temperature for ~10 minutes. *Malat1* probe was hybridised to the sample by adding 45 µl of probe mix to the slide then placed in a 40^{o}C oven for 2 hours. Slides were washed as before in 1 x RNAscope® wash buffer. Slides were then submerged and stored in 5 x SSC overnight at room temperature.

The following day excess liquid was removed from the slides by flicking. To amplify the *Malat1* C1 probe 1-2 drops of RNAscope® Multiplex FL v2 Amp 1 was added to each slide. Slides were then incubated at 40^{o}C for 30 minutes. Cells were

then washed twice in 1 x RNAscope® wash buffer with 2 minute incubations in each wash. This same amplification steps were repeated using RNAscope® Multiplex FL v2 Amp 2v then RNAscope® Multiplex FL v2 Amp 3. After washing excess liquid was removed from the slide and 1-2 drops of RNAscope® Multiplex FL v2 HRP-C1 added to each slide. Slides were heated to 40^{o}C for 15 minutes then washed 2 times as previously mentioned in 1 x wash buffer and excess liquid removed from the slide.

TSA® Plus Cyanine 5 stocks were diluted 1 in 1500 with the TSA buffer provided in the RNAscope® Multiplex Fluorescent Kit v2. 45 µl of diluted TSA Cy5 was added to the slide, which was incubated for 30 minutes at 40^{oc}. Cells were washed again in 1 x wash buffer, then the nucleus counterstained with DAPI for 30 seconds. DAPI was removed and the slides were mounted using Fluroshield and sealed with nail varnish.

Slides were stored in the fridge overnight and images captured.

2.9.6 Image acquisition and analysis

Fluorescent images were captured using the Axioinvert 200 M microscope (Zeiss) or LSM 880 with Airyscan processing (Zeiss), using the x63 objective and Axiovision image acquisition software (SE64 release 4.9.1) or Zen blue (Zeiss). Images were processed using Fiji/ImageJ v1.50e (NIH Bethesda, Maryland). For co-localisation analysis images were processed using the JaCoP plugin which determined the percentage of channel A in channel B and vice versa. Thresholds were kept consistent across images.

2.10 Animal experiments

2.10.1 Mixed bone marrow chimera

All mixed bone marrow chimera mice used in experiments for this thesis were kindly generated by Dr James Hewitson York Biomedical and Research Institute (YBRI), University of York, York, UK

Female CD45.1 x CD45.2 F1 C57BL/6 mice (12 weeks old) were whole body lethally irradiated (two doses of 550 rad, 24 hours apart) using a X-Rad iR225 (PXI). Immediately after the second dose of irradiation, mice were reconstituted with 5 x 10^6 whole bone marrow cells in 200 µl of PBS through intravenous injection. The bone marrow mixture consisted of a 50:50 ratio of cells derived from WT CD45.1 C57BL/6 female mice and *Malat1*$^{-/-}$ CD45.2 C57BL/6 female mice. After 15 weeks mice were infected with either *L.donovani* amastigotes or *S.mansoni* cercariae. For visceral leishmaniasis experiments, the Ethiopian strain of *L.donovani* (*LV9)* was used – kindly provided by Prof. Paul Kaye (YBRI, University of York, York, UK). Mice were intravenously injected with 3 x 10^7 amastigotes. All *L.donovani* infected mice were housed under containment level 3 (CL3) conditions. Mice were sacrificed after 6 weeks of infection and tissues were harvested for analysis. All tissue processing of *L.donovani* infected mice took place in CL3 laboratories. For schistosomiasis experiments, mice were infected with 35 *S. mansoni* cercariae by percutaneous penetration across the shaved abdomen under anaesthesia (cercariae provided by Dr James Hewitson). Mice were sacrificed after 7 weeks of infection and tissues were harvested for analysis.

2.10.2 *S.mansoni* egg induced models of inflammation

All mouse procedures as part of the *S.mansoni* egg injection model where kindly performed by Dr James Hewitson.

Female and male WT and *Malat1*$^{-/-}$ mice (6-12 weeks old), are first primed with 8,000 *S.mansoni* eggs in 200µl of PBS through intraperitoneal injection. After two weeks mice are challenged with 6,000 *S.mansoni* eggs in 200 µl of PBS through intravenous injection. Following challenge, *S.mansoni* eggs are transported to the lung via the pulmonary artery and become trapped in the lung parenchyma. After one week, mice were sacrificed and tissues harvested for analysis.

S. mansoni eggs were kindly harvested and prepared by Dr James Hewitson.

2.10.3 Preparation of tissues derived from infected mice

Single cell suspensions were generated from spleen and mesenteric lymph nodes by passing them through a 70µm cell strainer in RPMI before being centrifuged at 450 g, 4^{o}C for 5min.

In some cases, livers were kindly prepared by Magnus Gwynne and Joanna Greenman (YBRI, University of York, UK). Livers were finely chopped in Hank's balanced salt solution (HBBS) (HyClone) and incubated at 37^{o}C for 45 minutes with 160 U/ ml DNase and 0.8 U/ ml Liberase TL (Roche Diagnostics). Digestion was stopped by adding final concentration 10mM EDTA and diluted with DMEM (Gibco) supplemented with 10% FCS, 2mM L-glutamine, 100 U/ ml penicillin G, and 100 µg/ ml streptomycin. Samples were passed through a 100 µm cell strainer, and centrifuged twice at 450 g, 4^{o}C for 5mins. The pellet was re-suspended in 8 ml of 33% isotonic percoll (GE Healthcare) / PBS and centrifuged at 700 g, room temperature for 12min

without brake or acceleration. Cells were stained for flow cytometry as previously described.

2.11 RNAseq

2.11.1 Sample preparation and analysis of RNA integrity

RNA was extracted and prepared as described in section 2.4.1. 2µl of each sample was analysed using Aligent 2100 bioanalyzer by Lesley Gilbert (Technology Facility, University of York, York, UK). RNA concentration and RIN was calculated.

2.11.2 Library preparation and sequencing

10 µl of each sample (min 100 ng) was shipped to the Wellcome Sanger Institute for library preparation and sequencing. Samples were processed as follows:

Samples were quantified with QuantiFluor RNA System, 1 ml from Promega UK Ltd using Mosquito. LV liquid platform, Bravo WS and BMG FLUOstar Omega plate reader and cherrypicked to100ng / 50ul using Tecan liquid handling platform. Library construction (poly(A) pulldown, fragmentation, 1 st and 2 nd strand synthesis, end prep and ligation) using 'NEB Ultra II RNA custom kit' on an Agilent Bravo WS automation system. PCR set-up using KapaHiFi Hot start mix and Eurofins dual indexed tag barcodes on Agilent Bravo WS automation system.PCR cycles, 14 standard cycles (1 Incubate 98C 45 secs, 2 Incubate 98C 15 secs, 3 Incubate 65C 30 secs, 4 Incubate 72C 30 secs, 5 Cycle from 2, 13 more times, 6 Incubate 72C 1 mins). Post PCR plate purified using Agencourt AMPure XP SPRI beads on Caliper Zephyr liquidhandling platform. Libraries were quantified with Biotium Accuclear Ultra high sensitivity dsDNA Quantitative kit using Mosquito LV liquid handling platform, Bravo WS and BMG FLUOstar Omega plate reader. Libraries pooled in equimolar amounts on a Beckman BioMek NX-8 liquid handling platform. Pooled libraries

quantified on an agilent bioanalyser Libraries normalised to 2.8nM ready for sequencing.

2.11.3 Read mapping and quantification

Mapping and differential gene expression analyses were performed by Kylie James (The Wellcome Sanger Institute, Cambridge and currently The Garvan Institute, Darlinghurst, Australia).

To determine differential transcript usage and differential gene expression files were kindly further processed by Joshua Lee (University of York, York, UK), using the Salmon swimming downstream packages. https://www.bioconductor.org/packages/release/workflows/vignettes/rnaseqDTU/inst/doc/rnaseqDTU.ht ml

2.11.4 Analysis of RNAseq data

CSV files were merged using RStudio (R 4.2.1) using the following script

```
A <- read.csv("filename1.csv")

B <- read.csv("filename2.csv)

C <- merge.data.frame(A, B, by="Gene")
write.table(C, "intersect.csv", row.names=FALSE)
```

To identify pathways genes of interest were involved in lists were processed using the gsea-msigdb.org/gsea/index.jsp website. Hallmark and gene ontology sets were identified.

To determine networks of genes of interest, lists were processed using the following website https://string-db.org/. Under basic settings text mining was excluded and medium to high confidence thresholds were applied.

Numbers of genes differentially expressed were calculated and filtered in Excel (Microsoft Office 2016). Volcano plots and heat maps were produced in GrapPad Prism 8 (GraphPad Software). q thresholds of <0.1 and <0.05 were used for data analysis.

Venn diagrams were generated using the following website https://bioinfogp.cnb.csic.es/tools/venny/, propotional venn diagrams were generated using Meta Chart https://www.meta-chart.com/venn

The following code was used to identify chromosome number using the Biomart database.

```
#load libraries
library(biomaRt)
library(dplyr)
library(tidyverse)
#read table, where header for GeneIDs is GeneID
file = read.csv(file = ".csv", header = TRUE, sep = ",")
head(file)
#remove all zero rows
file = file[rowSums(file[])>0,]
#add gene names & biotype
mart = useMart
ensembl = useEnsembl("ensembl", dataset = "mmusculus_gene_ensembl")
gene <- getBM(attributes =
c("ensembl_gene_id","external_gene_name","gene_biotype","chromosome_name"),
values = file$GeneID,mart = ensembl)
```

```
id <- match(file$GeneID , gene$ensembl_gene_id)
file$gene_name <- gene$external_gene_name[id]
file$biotype <- gene$gene_biotype[id]
file$chromosome_name<- gene$chromosome_name[id]
write.csv(file, file = ".csv")
```

2.12 RAP-MS

The protocol used to identify *Malat1* binding partners by RAP-MS has been minorly adapted from the work which identified *Xist* binding partners (McHugh et al., 2015)

2.12.1 RAP-MS Buffers

1 x PBS for harvesting and crosslinking

Total cell lysis buffer

- 10 mM Tris HCl pH 7.5 (Fisher)
- 500 mM LiCl (ThermoScientific)
- 0.5% Dodecyl maltoside (DDM) (Sigma)
- 0.2% SDS (Sigma)
- 0.1% sodium deoxycholate (Sigma)

200X DNase Salt Solution

- 500 mM MgCl2 (Life Technologies)
- 100 mM CaCl2 (Sigma)

1.5X Hybridization Buffer

- 15 mM Tris-HCl pH 7.5

- 7.5 mM EDTA (Life Technologies)
- 750 mM LiCl
- 0.75% DDM
- 0.3% SDS
- 0.15% sodium deoxycholate
- 6 M urea (Sigma)
- 3.75 mM Tris(2-carboxyethyl)phosphine (TCEP, Sigma)

1X Hybridization Buffer

- 10 mM Tris-HCl pH 7.5
- 5 mM EDTA
- 500 mM LiCl
- 0.5% DDM
- 0.2% SDS
- 0.1% sodium deoxycholate
- 4 M urea
- 2.5 mM TCEP

Benzonase Elution Buffer

- 20 mM Tris-HCl pH 8.0 (Life Technologies)
- 0.05% NLS
- 2 mM MgCl2
- 0.5 mM TCEP

NLS Elution Buffer

- 20 mM Tris-HCl pH 8.0

- 10 mM EDTA
- 2% NLS
- 2.5 mM TCEP

2.12.2 Biotinylated probes

Antisense probes that span the length of *Malat1* have previously been used to identify RNA-RNA interactions (Engreitz et al., 2014). These 142nt, 5' biotinylated probes were synthesised at 0.05 µmol, HPLC purification (Sigma). See supplementary tables 1 and supplementary figures 1-2 for sequences and binding locations on *Malat1*.Probes were reconstituted at 100 µM in ultra-pure water.

Each probe was diluted 1 in 10 to create a 10 µM stock. Subsequently, the individual probes were mixed to create a total probe stock used for RAP captures.

2.12.3 Ultraviolet (UV) crosslinking

To generate an *in vivo* covalent linkage between *Malat1* and its interacting proteins, the cells of interest were washed in PBS and re-suspended at 2 x 10^6 cells per 10 ml of sterile $PBS^{-Mg2+-Ca2+.}$ To ensure that a single cell suspension was generated the cells were passed through a sterile 70µm cell strainer. Subsequently, the cell suspension was then transferred to a sterile 10cm dish, and placed onto a larger dish of ice. Cells were cross-linked in the Stratalinker 2000 using two energy doses set at 4000, for a total of 8000 (x 100 µJ/cm^2) total dose. The cell suspension was shaken briefly in between the two doses to prevent clumping and ensure even crosslinking. The cell suspension was then transferred to a 50 ml falcon tube and pelleted by centrifugation at 1000 g for 10 minutes at 4^{oc}. The supernatant was removed and the cell pellets were re-suspended in 1 ml of cold PBS. The cell suspension was transferred

to an Eppendorf tube and centrifuged as above. The supernatant was removed and cell pellets were flash-frozen in liquid nitrogen and stored at -80^{oc}.

As an alternative crosslinking strategy the Minitron cross-linker (in house) was used according to (De Pablos et al., 2019). Briefly the lab area was thoroughly cleaned with 70% ethanol. The pipe was first rinsed with 1 x PBS to remove any residual dust. The cell suspension was poured into the pipe (~120 ml of cells in RPMI) and exposed to UV light for either 60 or 120 seconds. Cells were collected and the pipe rinsed with PBS. The wash was combined with the cell suspension and pelleted by centrifugation at 1000 g for 10 minutes at 4^{oc}. The supernatant was removed and the cell pellets were resuspended in 1 ml of cold PBS. The cell suspension was transferred to an Eppendorf tube and centrifuged as above. The supernatant was removed and cell pellets were flash-frozen in liquid nitrogen and stored at -80^{oc}.

To confirm that crosslinking had taken place a small percentage of the cell population was seeded into a 6 well plate and counted. Cell counts were measured at various time points post crosslinking and compared to a non-crosslinked control. Viability was also assessed by flow cytometry and propidium iodide (PI) staining.

2.12.4 Generation of cell lysates

All buffers were chilled on ice before use, in addition the cell lysate was kept on ice unless otherwise specified. The frozen 20 million cell pellet was resuspended in 900 µl of cold total cell lysis buffer. 4.6 µl of 1x protease inhibitor cocktail and 920 U of murine RNase inhibitor was added to the sample. The sample was then incubated on ice for 10 minutes. During this incubation the cell suspension was passed 5 times through a 26-gauge needle attached to a 1 ml syringe. The samples were sonicated

using a microtip for 30 seconds using intermittent pulses ensuring the samples were kept on ice.

4.8 µl of 1x DNase salt stock and 20 U of Turbo DNase was added to the sample and incubated for 10 minutes at 37^{o}C to digest the DNA. Samples were then returned to ice and 10 mM of EDTA, 5 mM EGTA and 2.5 mM TCEP immediately added to the suspension to stop the reaction. The lysate was then mixed with 2x the sample volume of 1.5x hybridisation buffer. Samples were incubated for 10 minutes on ice, then centrifuged at 16,000 g for 10 minutes at 4^{o}C to pellet the insoluble material. The supernatant was removed and moved to a fresh tube which was flash-frozen in liquid nitrogen. The cell lysate was then stored at -80^{o}C until required.

2.12.5 Pre-clearing lysate (volumes for 200 million cells)

Frozen lysates were warmed at 37^{o}C using an Eppendorf Thermomixer C (Eppendorf) with 1.5 ml of 15 ml smart blocks (Eppendorf). Where appropriate lysates were pooled into a single tube. 1.2 ml of streptavidin-coated magnetic beads per 200 million cells were transferred to an Eppendorf. The storage buffer was removed by magnetic separation. Beads were re-suspended in 1 ml of 10 mM Tris-HCL pH 7.5 and washed 4 times by magnetic separation. The beads were then washed 2x in 1x hybridisation buffer. After the final wash the beads were re-suspended in the cell lysate and incubated for 30 minutes at 37^{o}C mixing at 1100 rpm. The beads were then magnetically separated from the lysate and supernatant transferred to a fresh tube. This was repeated twice. At this stage 100,000 cells worth was lysate was removed and stored in a PCR strip. This is subsequently referred to as the RNA input sample.

2.12.6 *Malat1* captures

20 µg of mixed biotinylated probe stock was denatured per 200 million cells. Denaturing took place by heating the probe to 85^{o}C for 3 minutes. The lysate and probe were mixed and incubated for 2 hours at 67^{o}C at 1100 rpm in a thermo mixer. After the incubation 100,000 cells worth of lysate was transferred to a PCR tube. This is subsequently referred to as the RNA input plus probe sample. Streptavidin beads were prepared as above and re-suspended in the cell lysate. Lysate plus beads were incubated for 30 minutes at 67^{o}C at 1100 rpm in a thermomixer. Beads were magnetically separated and 100,000 cells worth of supernatant was taken and transferred to a PCR tube. This sample is subsequently referred to as the RNA flow-through sample. The beads were washed 6 times in one bead volume of 1 x hybridisation buffer with each washing being incubated at 67^{o}C for 5 minutes. 100,000 cells worth of sample was transferred to a PCR tube at this stage. This sample is subsequently referred to as the RNA elution sample. Beads were finally re-suspended in 1 ml of benzonase elution buffer, and 125U of benzonase (a nuclease) added to the lysate. To enable digestion of RNA the sample was incubated for 2 hours, at 37^{o}C 1100 rpm using a thermomixer. After this period had elapsed the beads were magnetically separated and transferred to a new eppendorf tube 6 times to ensure the streptavidin beads had effectively been removed.

2.12.7 Protein precipitation

To precipitate the proteins in the samples 10% final concentration of trichloroacetic acid (TCA) was added to the protein elution sample. This was incubated at 4^{o}C overnight. Samples were then centrifuged at 16,000 g for 30 minutes to pellet any proteins in the sample. The supernatant was removed and replaced with 1 ml of cold acetone. The sample was centrifuged at 16,000 g for 15 minutes,

supernatant removed and the pellet left to air dry in the fume hood. Pellets were stored at -20^{o}C until required.

In some cases, the protein pellet was re-suspended in SDS loading buffer and separated by size using a 4-15% Mini-PROTEAN TGX gradient gel (Biorad). Gels were either stained with coomassie (Thermofisher) to confirm presence of protein the bands cut out and sent for mass spectrometry analysis. Alternatively, gels were processed for western blotting as described previously.

2.12.8 RNA elution

RNA samples which had been placed in PCR tubes were diluted to 20 µl final volume with NLS elution buffer. The RNA elution sample was magnetically separated and beads re-suspended in 20 µl of NLS buffer. The RNA elution sample was heated for 2 minutes at 95^{oc}, then magnetically separated and supernatant which will contain the eluted RNA transferred to a new PCR tube. 1 mg/ ml Proteinase K was added to each sample and incubated for 1 hour at 55^{o}C to digest any protein bound to the RNA samples. Samples were then stored at -20^{o}C until required.

SILANE beads were used to clean up the RNA samples. 20 µl of beads per sample was magnetically separated and storage buffer removed. The beads were re-suspended in 60 µl of RLT buffer. Beads were added to 20 µl of RNA sample and mixed well. 120 µl of 100% ethanol was added to each sample and left for two minutes for RNA to bind to the beads. The beads were washed twice with 150 µl of 70% ethanol. The supernatant was removed and beads left to air dry for 5 minutes. The RNA sample was eluted by adding 26 µl of 1x TuboDNAse buffer. The beads were left in the same tube and 1 µl of murine RNase inhibitor was added to each sample in addition to 3 µl of TurboDNAse. The samples were incubated for 20 minutes at 37$^{oc.}$

A second round of clean-up was carried out using SILANE beads in the tube. 90 µl of RLT buffer and 180 µl of 100% ethanol was added to each sample and mixed well. Samples were left for two minutes for the RNA to bind to the beads. Beads were washed twice using 70% ethanol. The supernatant was removed and the beads allowed to air dry for 5 minutes. The RNA samples were eluted by adding 10 µl of UltraPure water, mixing well magnetically separating beads from the RNA and moving the supernatant to a new tube.

RNA samples were converted to cDNA and analysed by qRT-PCR as previously described to confirm that *Malat1* had successfully been pulled down.

2.12.9 Mass spectrometry analysis

All mass spectrometry analysis was carried out by the York Biology Technology facility by Adam Dowel. Methods were provided and are as follows: Protein was denatured by dissolving in aqueous 9 M urea buffer containing, 20 mM HEPES pH 8.0, 1 mM sodium orthovanadate, 2.5 mM sodium pyrophosphate, 1 mM β-glycerophosphate. Disulfides were reduced with 0.7 mg/ ml dithiothreitol and heating at 55^{oC}, then alkylated with 1.9 mg/ ml iodoacetamide before proteolytic digestion with 0.2 mg Promega sequencing grade trypsin/Lys-C mix, and incubation at 37^{o}C for 16 h.

Resulting peptides were desalted with Millipore C18 ZipTip before being re-suspended in aqueous 0.1% trifluoroacetic acid (v/v) then loaded onto an mClass nanoflow UPLC system (Waters) equipped with a nanoEaze M/Z Symmetry 100 Å C18, 5 µm trap column (180 µm x 20 mm, Waters) and a PepMap, 2 µm, 100 Å, C18 EasyNano nanocapillary column (75 µm x 500 mm, Thermofisher). The trap wash solvent was aqueous 0.05% (v:v) trifluoroacetic acid and the trapping flow rate was

15 µl/min. The trap was washed for 5 min before switching flow to the capillary column. Separation used gradient elution of two solvents: solvent A, aqueous 0.1% (v:v) formic acid; solvent B, acetonitrile containing 0.1% (v:v) formic acid. The flow rate for the capillary column was 300 nL/min and the column temperature was 40°C. The linear multi-step gradient profile was: 3-10% B over 7 mins, 10-35% B over 30 mins, 35-99% B over 5 mins and then proceeded to wash with 99% solvent B for 4 min. The column was returned to initial conditions and re-equilibrated for 15 min before subsequent injections.

The nanoLC system was interfaced with an Orbitrap Fusion Tribrid mass spectrometer (Thermofisher) with an EasyNano ionisation source (Thermofisher). Positive ESI-MS and MS2 spectra were acquired using Xcalibur software (version 4.0, Thermo). Instrument source settings were: ion spray voltage, 1,900 V; sweep gas, 0 Arb; ion transfer tube temperature; 275°C. MS1 spectra were acquired in the Orbitrap with: 120,000 resolution, scan range: m/z 375-1,500; AGC target, 4e5; max fill time, 100 ms. Data dependant acquisition was performed in top speed mode using a 1 s cycle, selecting the most intense precursors with charge states >1. Easy-IC was used for internal calibration. Dynamic exclusion was performed for 50 s post precursor selection and a minimum threshold for fragmentation was set at 5e3. MS2 spectra were acquired in the linear ion trap with: scan rate, turbo; quadrupole isolation, 1.6 m/z; activation type, HCD; activation energy: 32%; AGC target, 5e3; first mass, 110 m/z; max fill time, 100 ms. Acquisitions were arranged by Xcalibur to inject ions for all available parallelizable time.

Peak lists in. raw format were imported into Progenesis QI (Version 2.2., Waters) and LC-MS runs aligned to the common sample pool. Precursor ion intensities were normalised against total intensity for each acquisition. A combined peak list was

exported in .mgf format for database searching against the mouse subset of the UniProt database (16,923 sequences; 9,545,405 residues), appended with common proteomic contaminants (116 sequences; 38,371 residues). Mascot Daemon (version 2.6.0, Matrix Science) was used to submit the search to a locally-running copy of the Mascot program (Matrix Science Ltd., version 2.7.0). Search criteria specified: Enzyme, trypsin; Max missed cleavages, 1; Fixed modifications, Carbamidomethyl (C); Variable modifications, Oxidation (M), Deamidats (NQ), Acetyl (N-term); Peptide tolerance, 3 ppm; MS/MS tolerance, 0.5 Da; Instrument, ESI-TRAP. Peptide identifications were passed through the percolator algorithm to achieve a 1% false discovery rate assessed against a reverse database and individual matches filtered to require minimum expect score of 0.05. The Mascot .X ML result file was imported into Progenesis QI and peptide identifications associated with precursor peak areas and matched between runs. Relative protein abundance was calculated using precursor ion areas from non-conflicting unique peptides. Accepted protein quantifications were set to require a minimum of two unique peptide sequences. Statistical testing was performed in Progenesis QI from ArcSinh normalised peptide abundances and ANOVA-derived p-values were converted to multiple test-corrected q-values using the Hochberg and Benjamini approach.

2.13 iCLIP

2.13.1 Sample preparation and crosslinking

In vitro polarised Th2 cells were prepared as described in section 2.3. Cells were pelleted and re-suspended in batches of 6-7 million cells in 2 mL of ice-cold PBS. To ensure that a single cell suspension was generated the cells were passed through a sterile 70µm cell strainer and transferred to a sterile 6 well plate on ice. Cells

were cross-linked in the Statalinker 2000 using four energy doses set at 1000, for a total of 4000 (x 100 μJ/cm^2) total dose. The cell suspension was shaken briefly in between the two doses to prevent clumping and ensure even crosslinking. The cell suspension was then transferred to a 50 ml falcon tube and pelleted by centrifugation at 1000 g for 10 minutes at 4^{oc}. The supernatant was removed and the cell pellets were resuspended in 1 ml of cold PBS. The cell suspension was transferred to an Eppendorf tube and centrifuged as above. The supernatant was removed and cell pellets were flash frozen on dry ice. Samples were stored at -80°C and shipped to our collaborators in the Ule laboratory at the Francis Crick Institute.

2.13.2 Immunoprecipitation

Urška Janjoš (Francis Crick Institute, London) processed samples for iCLIP from this point forwards and provided the following.

Cell pellets were lysed in 1 ml of lysis buffer and subsequently sonicated using a Bioruptor. Next, the lysate was treated with 2 μL Turbo DNase (Thermo Fisher Scientific), 2.5 U /ml RNase I (Thermo Fisher Scientific) with shaking at 1,100 rpm and 37 °C for 3 min. The lysates were then incubated for 1 h with 2 μg of antibodies (SRSF1, Thermo Fisher Scientific, 103; hnRNPA1, Novus Biologicals,4B10) prebound to 100 μl of protein G Dynabeads. After a series of washes, an RNA adapter was ligated to the 3′ end of cross-linked RNAs. Subsequently, cross-linked complexes were separated by size using SDS–PAGE (NuPAGE 4–12% Bis-Tris Protein Gel, Thermo Fisher Scientific) and then transferred to a nitrocellulose membrane. The membrane portion which contained radioactively labelled protein–RNA complexes was excised. The RNA was isolated and then reverse transcribed to cDNA. To isolate the cDNA, this was selected based on size by urea denaturing gel electrophoresis and the fraction running between 145–400 nucleotides was isolated. This fraction was then

circularized by employing single-stranded DNA ligase. After final PCR amplification, multiplexed libraries were sequenced using high throughout sequencing methods.

2.14 Sequence alignments

To enable easy visual comparison of sequence conservation between species, sequences were aligned and figures generated using DNAstar Lasergene v17. Initially FASTA files were obtained from the NCBI webpages (https://www.ncbi.nlm.nih.gov/) and subsequently aligned using the MegAlign Pro 17 tool.

2.15 Schematic diagrams

Schematic depictions of pathways and models were generated using either Powerpoint Microsoft Office 2016, or using Biorender software https://biorender.com/.

2.16 Data availability

When appropriate the raw RNA-seq data files, RAP-MS and iCLIP will be deposited into appropriate online repositories. All raw data files are available upon request.

2.17 Statistical analysis

All results were analysed using GrapPad Prism 8 (GraphPad Software). Before applying statistical tests, normality of data distribution was examined by Kolmogorov-Smirnov test. Each data set was analysed using appropriate statistical testing depending on the distribution of the data. Statistical test is indicated next to each data set. Commonly used statistical tests were unpaired t-tests, paired t-tests and one-way analysis of variance (ANOVA) with post hoc tests. For example, normally distributed

data, with an n ≥3, was analysed by an unpaired t-test or paired t-test. Data was considered unpaired if there were two different experimental conditions for example treated vs untreated. Data was considered paired when measurements were taken from the same mouse for example WT vs *Malat1*$^{-/-}$ T cells in a mixed bone marrow chimera. Mann-Whitney U test was used if the data was not normally distributed for two experimental conditions. Data was analysed by ANOVA to determine if there was a statistically significant difference between more than two results of an independent variable for example uninfected, WT infected and *Malat1*$^{-/-}$ infected. Post hoc tests were used to determine which groups have statistically significant differences in means. P values of <0.05 have been considered significant and are displayed where appropriate. Of note, this study complied with the ARRIVE guidelines however, experiments were not performed blind.

3 *Malat1* suppresses immunity by promoting the expression of MAF and IL-10 in $CD4^+$ T cells

3.1 Introduction

3.1.1 Statement

Some of the work presented in this chapter is also featured in the publications below:

West and Lagos, 2019. Long Non-Coding RNA Function in $CD4^+$ T Cells: What We Know and What Next? Non-Coding RNA, 5(3), p.43.

Hewitson, J., West, K., James, K., Rani, G., Dey, N., Romano, A., Brown, N., Teichmann, S., Kaye, P. and Lagos, D., 2020. *Malat1* Suppresses Immunity to Infection through Promoting Expression of MAF and IL-10 in Th Cells. *The Journal of Immunology*, 204(11), pp.2949-2960.

3.1.2 LncRNAs in $CD4^+$ T cells

T helper cells are an ideal platform for the study of lncRNAs given they function as part of a complex cellular network and undergo remarkable and finely regulated gene expression changes upon antigenic stimulation. Using various knockdown and RNA interaction studies lncRNAs have been shown to be crucial for T helper cell function and differentiation through both *in-cis* (local gene regulation) and *in-trans* (genes distal from the site of transcription) modes of action (Gil & Ulitsky, 2019; West & Lagos, 2019). Given that RNA targeting therapeutics are rapidly gaining attention, further understanding the mechanistic role of lncRNAs in a T helper cell context is an exciting area of research as it may provide a wide range of new candidate targets for the treatment of $CD4^+$ mediated pathologies.

Indeed, cytokines and transcription factors are essential for effector $CD4^+$ T cell differentiation, yet emerging evidence suggests lncRNAs also play a vital role in

the differentiation process (***Figure 3.1***). Interestingly, RNA-seq data gathered from 42 subsets of murine thymocytes and $CD4^+$ T cells at numerous developmental time points identified 1,524 lncRNA expressing genomic regions (clusters), some of which encode more than one lncRNA. 48-57% of the identified lncRNAs were indicated to be cell-type-specific contrasting to 6-8% of mRNAs (Hu et al., 2013). The number of lncRNA clusters in each T cell subset was between 154 and 353. Many of the detected lncRNAs are genomic neighbours of important immune-regulatory genes and contain binding sites for $CD4^+$ T cell transcription factors such as STAT4 and T-bet. The unique expression profiles of lncRNAs in different T cell subsets suggests that lncRNAs have critical roles in Th cell function during homeostasis and disease.

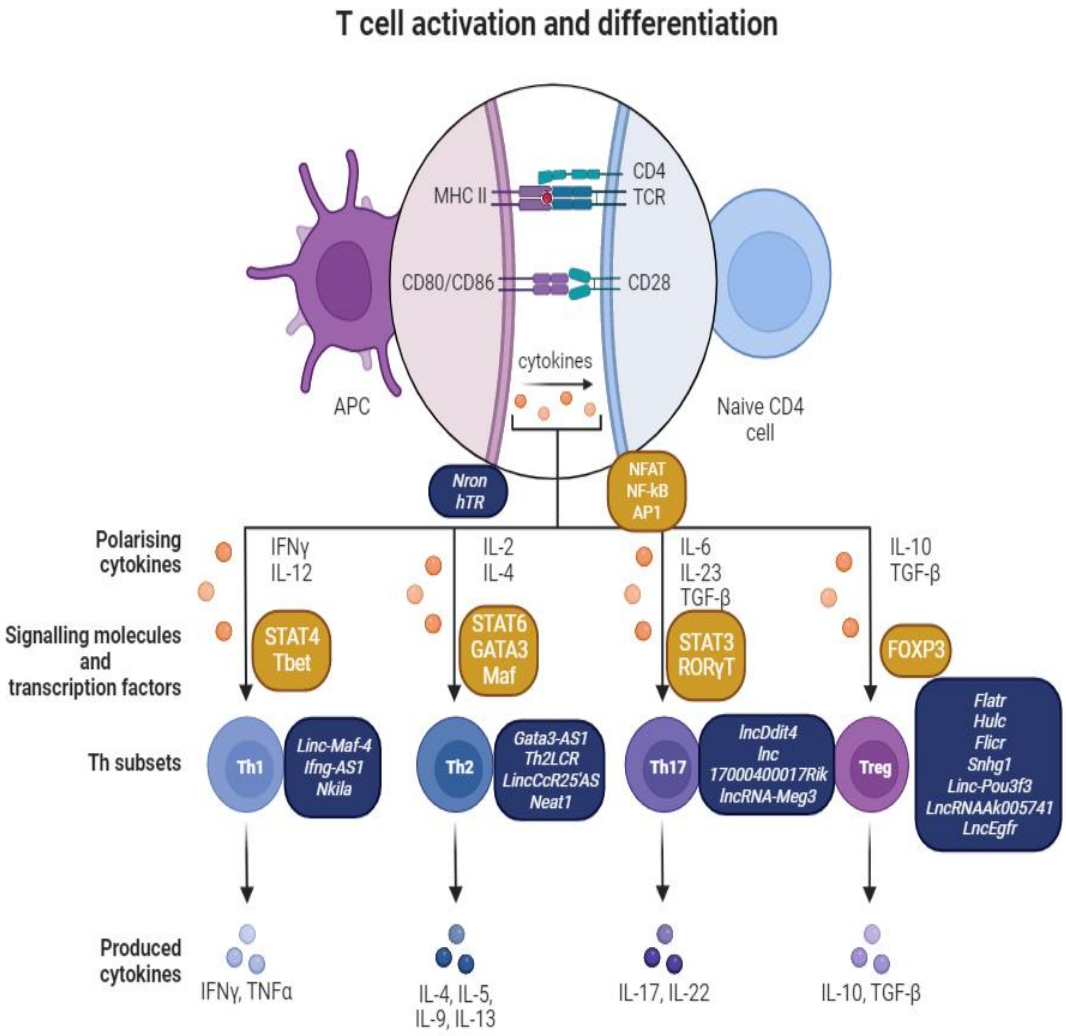


Figure 3.1: Schematic representation of T helper cell differentiation and associated lncRNAs.

Schematic representation of $CD4^+$ T cell differentiation process highlighting key secreted cytokines transcribed transcription factors (yellow boxes) and implicated lncRNAs (blue boxes) are indicated at the appropriate transitional stages.

3.1.3 *Cis-acting* lncRNAs

One example of a *cis-acting* lncRNA which has been shown to be involved in $CD4^+$ T cell differentiation is the lncRNA *Linc-MAF4*. *Linc-MAF4* was first identified by RNA-seq analysis of human $CD4^+$ T cell subsets. It was found to be a Th1-specific lncRNA which repressed avian musculoaponeurotic fibrosarcoma oncogene homolog (MAF, also known as c-MAF) expression (predominantly a Th2 transcription factor). Ranzani and colleagues showed that MAF is repressed through *Linc-MAF4* interactions with the chromatin modifiers enhancer of zeste homolog 2 (EZH2) and lysine-specific histone demethylase 1A (LSD1) which deposited H3K27me3 marks in the promotor region of *Maf,* which silenced MAF expression. Subsequent *Linc-MAF4* knockdown in human peripheral blood mononuclear cells (PBMCs) skewed $CD4^+$ T cell differentiation towards a Th2 phenotype (Ranzani et al., 2015). Supporting the data set, when naïve human $CD4^+$ T cells were transfected with synthetic *Linc-MAF4* this promoted Th1 differentiation and repressed Th2 differentiation (F. Zhang et al., 2017). Collectively, this data demonstrated the key role of *Linc-MAF4* in permitting a Th1 phenotype.

Another lncRNA which has been shown to be important in Th1 cells is *ne*ttoie *S*almonella pas *T*heiler's (cleanup Salmonella not Theiler's) (*NeST*), also known as *Ifng-AS1* and originally *TMEVPG*1. *NeST* was initially identified as a candidate genetic factor for controlling Theiler's virus persistence in the central nervous system (Vigneau et al., 2003a). This study noted that *NeST* and *Ifng* are neighbouring genes which are transcribed on opposite DNA strands in both mice and humans. Subsequent human and mouse $CD4^+$ T cell *in vitro* polarisation experiments found that IFNγ (protein level) and *NeST* (RNA level) were the most highly expressed in Th1 cells. Interestingly, *in vitro* activated Th1 cells isolated from D011.10. $STAT4^{-/-}$ and

DO11.10. *T-bet*$^{-/-}$ mice showed diminished levels of both IFNγ and *NeST*. Moreover, subsequent siRNA knockdown of *NeST* decreased the expression of IFNγ (Collier et al., 2012). This indicated that *NeST* had an activating effect on IFNγ expression. To further characterise this finding, immunoprecipitation and qPCR experiments revealed that *NeST* interacts with WD repeat-containing protein 5 (WDR5), which is part of a complex of proteins that facilitates H3K4 methylation marks (Gomez et al., 2013a). This study went on to generate *in vivo* data which revealed that in a mouse model of sepsis, mice that overexpressed *NeST* exhibited greater expression of IFNγ and displayed increased H3K4me3 marks at the *Ifng* locus (Gomez et al., 2013a). Additional analysis has begun to characterise the mechanism by which *NeST* is regulated. Interestingly, T-bet was shown to bind to the *NeST* promoter and distal enhancers in Th1 cells, with NF-kB and Ets1 also shown to be important in this process (Collier et al., 2014). Consistent with previous studies of *NeST* function, analysis of Hashimoto's thyroiditis (HT) patients revealed that HT patients have an increased number of circulating Th1 cells with *NEST*, *TBX21* and *IFNG* mRNA levels also upregulated in these patients. Most notably, *NeST* expression was positively correlated with the number of circulating Th1 cells, T-bet and IFNγ protein expression, providing an additional link between *NeST* and IFNγ (Peng et al., 2016a). Together, these studies highlight the important role of *NeST* in Th1 cells and IFNγ expression.

Genome-wide analysis of primary human CD4^{+} subsets identified a Th2-specific lncRNA *GATA3-AS1* (Zhang et al., 2013). *GATA3-AS1* is a divergent lncRNA as it is located in the same promoter as *GATA3* and they are transcribed in opposite directions. To understand the role of *GATA3-AS1* in Th2 cells, the lncRNA was knocked-down using siRNAs in total human PBMCs stimulated under Th2 polarising conditions. Knockdown reduced the expression of signature Th2 polarised genes

GATA3, IL-13 and IL-5, at both the mRNA and protein levels (Gibbons et al., 2018). Additionally, chromatin immunoprecipitation (ChIP) and PCR analysis in *GATA3-AS1* knockdown cells revealed significantly reduced levels of the activating epigenetic marks H3K4me2/3 and H3K27ac across both the *GATA3* and *GATA3-AS1* genomic loci. To uncover the mechanism by which *GATA3-AS1* alters chromatin marks further ChIP and DNA-RNA immunoprecipitation (DRIP) analysis was performed. This demonstrated that the *GATA3-AS1* RNA forms an R loop with the DNA in a central intron of *GATA3-AS1* and that *GATA3-AS1* binds to an essential component of the chromatin-modifying MLL H3k4 methyltransferase complex. Although *GATA3-AS1* could not induce a Th2 phenotype in the absence of factors such as MAF these findings provide strong evidence that the lncRNA *GATA3-AS1* plays an essential role in modifying the chromatin landscape of *GATA3* and *GATA3-AS1*, in addition to regulating the expression of GATA3 and the Th2 effector cytokines IL-5 and IL-13 (Gibbons et al., 2018).

Not only do lncRNAs play a role in Th1 and Th2 differentiation, but there is also evidence to support lncRNA function in Th17 differentiation. Microarray data from PBMCs of multiple sclerosis (MS) patients revealed that DNA-damage inducible transcript 4 (DDIT4) and lncRNA *DDIT4* (*lncDdDIT4*) were upregulated in Th17 cells compared to other T cell subsets (Zhang et al., 2018). DDIT4 is a cytoplasmic protein which is upregulated upon DNA damage and is known to inhibit mechanistic target of rapamycin complex 1 (mTORC1) activity (Tirado-Hurtado et al., 2018). The mTOR pathway is crucial for a plethora of functions such as proliferation and has been implicated in the positive regulation of Th17 differentiation (Nagai et al., 2013). *lncDDIT4* is located downstream of DDIT4 at the genome level, which suggested a potential *cis*-regulatory role of *lncDDIT4*. This *cis*-regulatory effect was observed

when *lncDDIT4* was knocked-down in naive human $CD4^+$ T cells which decreased DDIT4 expression, and when the knocked-down cells were stimulated under polarising conditions this increased DDIT4/mTOR signalling and bolstered Th17 cell numbers. Conversely, when *lncDDIT4* was overexpressed DDIT4 expression was also enhanced, this reduced the activation of the DDIT4/mTOR pathway meaning Th17 cell numbers were reduced. Importantly, the alteration of *lncDDIT4* levels did not affect Th1 or Th2 differentiation demonstrating a cell-type-specific effect (Zhang et al., 2018). Future work to confirm and characterise the mechanistic interaction between *lncDDIT4* and DDIT4 would be beneficial. Notwithstanding, it is clear that *lncDDIT4* plays a key role in Th17 differentiation.

Xist is a well-studied lncRNA, predominantly known for its function in X chromosome inactivation. In-depth analysis describes the *in-cis* mode of silencing the majority of genes along the inactive X chromosome through *Xist*-mediated recruitment of epigenetic modifiers that deposit repressive chromatin marks (Brockdorff, 2018; Brockdorff et al., 1992; Penny et al., 1996) Fascinatingly, RNA fluorescence in situ hybridisation (FISH) analysis of both human and murine naïve female $CD4^+$ T cells shows dispersed *Xist* localisation. Upon *in vitro* T cell activation *Xist* returns to the Xi in female cells, however, levels of repressive chromatin marks such as H3K27me3 are notably reduced. Knockdown experiments demonstrated that the return of *Xist* to the Xi upon T cell activation is YY1 and hnRNPU dependent. As a result of lower repressive chromatin marks, the inactive X chromosome becomes partially reactivated thus, biallelic expression of the X-linked immune genes CD40LG and CXCR3 occurs in a small percentage of cells (Wang et al., 2016). As CD40L and CXCR3 are strongly associated with the autoimmune disorder systemic lupus erythematosus (SLE), this inspired further work to characterise potential links between *Xist* localisation and the

autoimmune disease. Interestingly, splenic T cells isolated from SLE patients and late-stage NZB/W F1 female mice show dispersed *Xist* localisation. Consequently, gene expression analysis of SLE patients' splenic T cells showed an abnormal upregulation of transcription along the X chromosome and altered expression of *Xist* RNA binding proteins (Syrett et al., 2019). In further support of the importance of correct *Xist* localisation in T cells, when Cip1-interacting zinc finger protein (CIZ1) (a protein recruited to the Xi by *Xist*) is knocked-out CIZ1 null mice develop fully penetrant female-specific lymphoproliferative disorder, and *Xist* localisation is disrupted in splenic T cells in the absence of CIZ1 (Ridings-Figueroa et al., 2017). Collectively, these studies suggest that *Xist* localisation at the Xi could be a crucial factor in maintaining dosage compensation of X-linked genes in T cells. Moreover, this provides a potential argument that *Xist* may be a contributing factor in mammalian females' ability to better combat pathogens compared to males and their increased predisposition to autoimmune disorders.

3.1.4 *Trans-acting* lncRNAs

One of the first lncRNAs shown to be functional in T helper cells is the non-coding RNA nuclear repressor of NFAT (*Nron*). To uncover if lncRNAs affected the function of the Ca^{2+} regulated transcription factor NFAT, an NFAT-responsive luciferase assay was used to array a library of short hairpin RNAs (shRNAs) targeted towards 512 ncRNAs in Jurkat and HEK293 cells. Interestingly, shRNA targeting of *Nron* elevated the activation of NFAT (Willingham et al., 2005). Subsequent *in vitro* binding experiments demonstrated *Nron* interactions with three members of the importin-beta superfamily, which are involved in-transporting NFAT from the nucleus to the cytoplasm. However, later work in Jurkat and $CD8^+$ T cells instead suggest that *Nron* bound to NFAT as part of an RNA-protein scaffold complex which comprised

three NFAT kinases and the scaffold protein IQ motif containing GTPase activating protein (IQGAP). Yet, in further support of previous findings, the knockdown of *Nron* resulted in increased phosphorylation of NFAT, translocation to the nucleus and higher levels of NFAT-regulated cytokines (Sharma et al., 2011). Additional investigations which studied the scaffold indicated that NFAT became uncoupled from the *Nron* scaffold by exploiting the transcription factor E26 transformation-specific sequence 1 (Ets-1). Coimmunoprecipitation and ChIP analysis of WT and Ets-1 KO Th cells showed that Ets-1 interacted with NFAT and synergistically activated the IL-2 promoter in the nucleus. In addition, upon Ca^{2+}signalling Ets-1 moves to the cytoplasm and competes with NFAT for the binding of the *Nron* complex, this releases NFAT from the complex enabling nuclear entry. (Tsao et al., 2013). Despite the debate on the exact mechanism of *Nron* function, these studies provide compelling evidence for the *in-trans* mode of action for the lncRNA *Nron* in T helper activation.

Another example of a *trans-acting* lncRNA is the lnc-epidermal growth factor receptor (EGFR). The role of *lnc-EGFR* in Tregs was discovered in a study of hepatocellular carcinoma (HCC) in which high throughput screening examined links between lncRNAs and mRNAs in HCC patient samples (Jiang et al., 2017). *lnc-Egfr* was discovered to be highly expressed in Tregs and correlated positively with FOXP3 expression but negatively with IFNγ. Moreover, RNA immunoprecipitation (RIP) analysis revealed that *lnc-EGFR* interacts with the protein EGFR. Interestingly, overexpression and knockdown of *lnc-EGFR* suggested that *lnc-EGFR* prevents ubiquitination of EGFR by blocking the effects of c-CBL. This study went on to show binding sites for NFAT and AP1 are in the promoters of *lnc-EGFR*, EGFR and FOXP3, and expression of all three increases upon NFAT and AP1 signalling. The proposed model suggests that *lnc-EGFR* acts in a forward feedback loop *lnc-EGFR* -

EGFR-NFAT/AP1-*lnc-EGFR* which promotes Treg differentiation and clearly showcases the *in-trans* mechanism of *lnc-EGFR.*

3.1.5 Other lncRNAs

The lncRNA *Neat1* has been implicated in both Th2 and Th17 differentiation. A study by Shui and colleagues examined PBMCs from rheumatoid arthritis patients, the authors indicated that *Neat1* was elevated in Th17 cells and demonstrated that knockdown of *Neat1* inhibited Th17 differentiation. It is thought that *Neat1* regulated the differentiation process through interactions with STAT3 (Shui et al., 2019). However, another study suggested that *Neat1* promotes Th2 differentiation (Huang et al., 2019). RIP and ChIP analysis revealed that *Neat1* binds to EZH2, which is recruited to the ITCH promoter repressing ITCH, in addition, expression of STAT6 was increased along with IL-4, IL-5 and IL-13. Together these studies suggest that *Neat1* may have multiple roles in T helper cell differentiation.

Other lncRNAs have been implicated in $CD4^+$ T cell functions. Examples of these include a Th2-specific lncRNA cluster *Th2LCR* which regulates differentiation by affecting IL-4, IL-5 and IL-13 expression (Spurlock et al., 2015). A novel lncRNA *Flatr* is thought to act in the Treg differentiation pathway in mice (Brajic et al., 2018). Additionally, the lncRNA highly upregulated in liver cancer (*HULC*) can reduce P18 expression which impairs Treg differentiation (Zhao et al., 2015). Together these studies provide compelling support for the essential role of lncRNA in $CD4^+$ T cell differentiation; thus, enabling the immune system to elicit targeted responses to pathogens. A summary of the role of lncRNAs in $CD4^+$ T cells is shown in (***Table 3.12***)

Table 3.12: LncRNA functions in CD4⁺ T cells

lncRNA	Subset	Target	Mechanism	CD4+ *In vivo*?	References
			In-cis		
Linc-MAF4.	Th1	MAF	Interacts with EZH2 and LSD1 to epigenetically repress *MAF*.	No	(Ranzani et al., 2015; F. Zhang et al., 2017).
NeST	Th1	IFNγ	H3K4-methylation at *Ifng* promoter aids transcription	Yes	(Gomez et al., 2013b; Peng et al., 2016b; Vigneau et al., 2003b)
GATA3-AS1	Th2	GATA3, IL-13, IL-5	Epigenetic modification of *GATA3* chromatin landscape	No	(Gibbons et al., 2018; Spurlock et al., 2015; H. Zhang et al., 2013)
Th2 LCR (Cluster of lncRNA)	Th2	IL-4, IL-13, IL-15	Histone acetylation, methylation and demethylation in the Th2 cytokine locus.	Yes	(Hwang et al., 2012; Koh et al., 2010; Spurlock et al., 2015)
lncRNA-1700040D17Rik	Th17	RORγt.	Potential *in-cis*-regulation	No	(W. Guo et al., 2017)
lncDDIT4	Th17	DDIT4/mTOR	Potential *in-cis* regulation	No	(F. Zhang et al., 2018)
Flicr	Tregs	Foxp3	Negative regulator modifies *Foxp3* chromatin access	Yes	(Zemmour et al., 2017)

In-trans					
Lnc-EGFR	Treg	EGFR receptor	Binds EGFR directly, stabilising it and augmenting activation enabling EGFR expression	Yes	(Jiang et al., 2017)
Nron	Activated CD4+	NFAT	Sequesters NFAT in the cytoplasm	No	(Sharma et al., 2011; Willingham et al., 2005)
HULC	Treg	P18	*Hulc* binds P18 potential direct inhibition	No	(Zhao et al., 2015).
Linc-Pou3f3	Treg	TGF-β	Binds TGF-β increasing phosphorylation of SMAD 2/3	No	(Xiong et al., 2015)
Snhg1	Treg	miR-488	Facilitates Treg differentiation through potential interactions with miR-488 and IDO	No	(Pei et al., 2018)
Nkila	Th1	ACID	Binds and inhibits NFkB	No	(Huang et al., 2018)
LncRNA-Meg3	Th17	RORyt	Acts as a miRNA sponge	No	(Li et al., 2016; Qiu et al., 2019)
Yet to be characterised					
Flatr	Treg	Foxp3	Unknown	Yes	(Brajic et al., 2018)
Linc-Ccr2-5'AS	Th2	CCR1, CCR3, CCR2, CCR5	Unknown	Yes	(Hu et al., 2013)
lncRNA-AK005641	Treg	Differentiation	Unknown	No	(Cheukfai et al., 2017)
hTR	$CD4^+$	Apoptosis	Unknown	No	(Gazzaniga & Blackburn, 2014)

3.1.6 Chapter hypothesis and aims

A growing body of evidence has characterised the essential role of lncRNAs in $CD4^+$ T cells. Given that Th cells are an ideal platform for the study of lncRNA function we wanted to use this model to deepen our understanding of the molecular function of the lncRNA *Malat1*. We hypothesised that *Malat1* has a non-redundant role in $CD4^+$ T cells during immune response to infections.

The specific aims of this chapter are to:

- Determine *Malat1* expression patterns in WT $CD4^+$ T cells
- Compare $CD4^+$ T cell differentiation in WT and *Malat1*$^{-/-}$ cells *in vitro*
- Compare $CD4^+$ T cell responses in WT and *Malat1*$^{-/-}$ mice *in vivo*

3.2 Results

3.2.1 *Malat1* expression is diminished in gene targeted mice

To study the role of *Malat1* in $CD4^+$ T cells, we obtained *Malat1*$^{-/-}$ mice which had been generated by the Prasanth laboratory (Nakagawa et al., 2012b). Nakagawa and colleagues used a gene targeting strategy to insert the lacZ gene followed by a polyadenylation signal downstream of the transcriptional start site (TSS) of *Malat1* (***Figure 3.2A***). Mice used for experiments were genotyped through either genomic DNA PCR screening of ear clips or qRT-PCR analysis of cells and tissues derived from mice (***Figure 3.2B/D***). This confirmed that *Malat1* expression was diminished in knockout mice. Of note, we observed >99% reduction in *Malat1* expression but not complete deletion in $CD4^+$ T cells derived from *Malat1*$^{-/-}$ mice. Cycle threshold (Ct) values indicated some residual *Malat1* expression remained. In WT Th2 cells an average Ct value for *Malat1* was 18.8 compared to *Malat1*$^{-/-}$ Th2 cells where an average Ct value of 27.6 was obtained, this was above the level of background in no template control samples which had a Ct value of 32.4. Nevertheless, we continued with this model as it provides a reliable platform for understanding the role of *Malat1* in $CD4^+$ T cells.

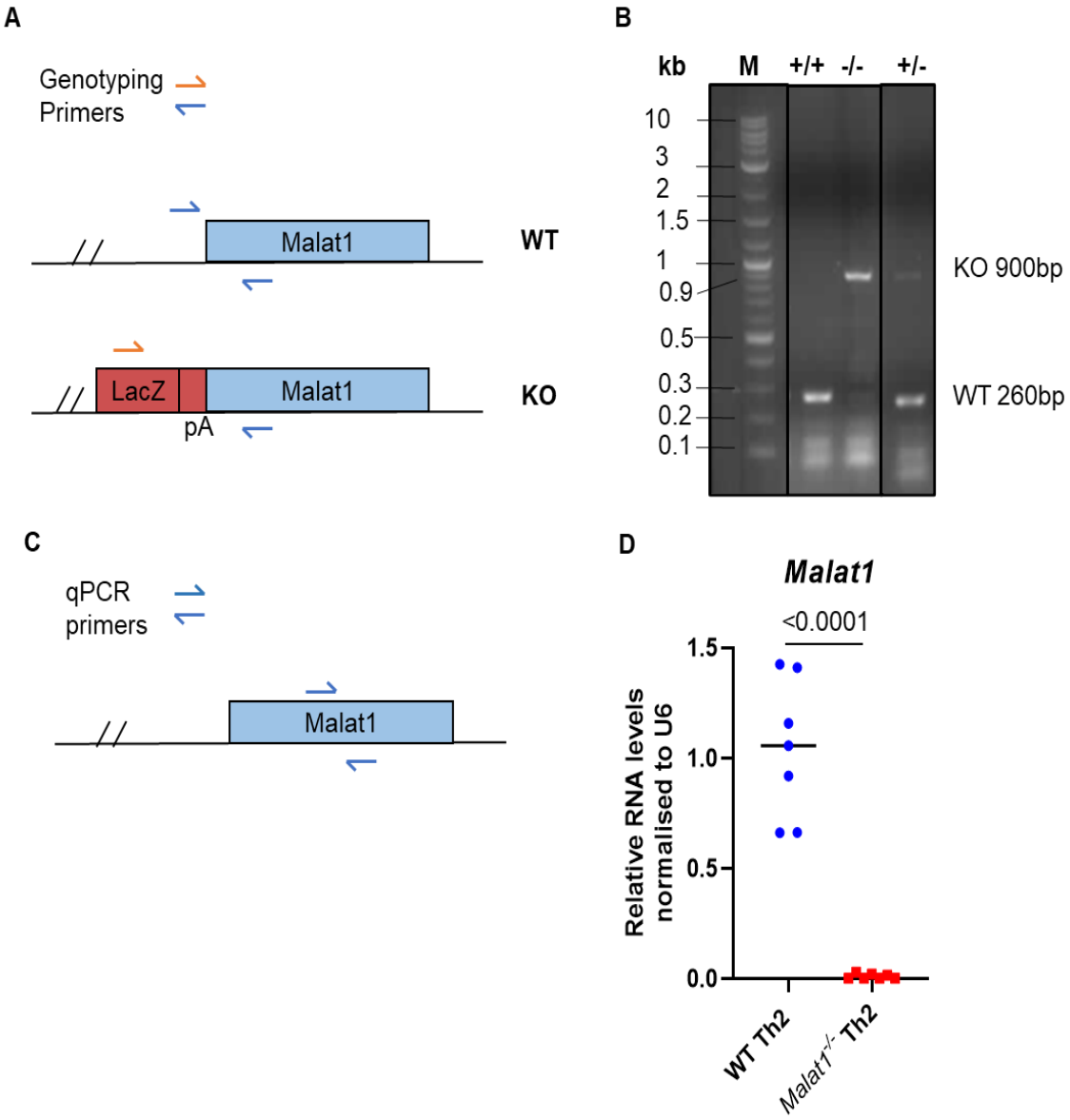


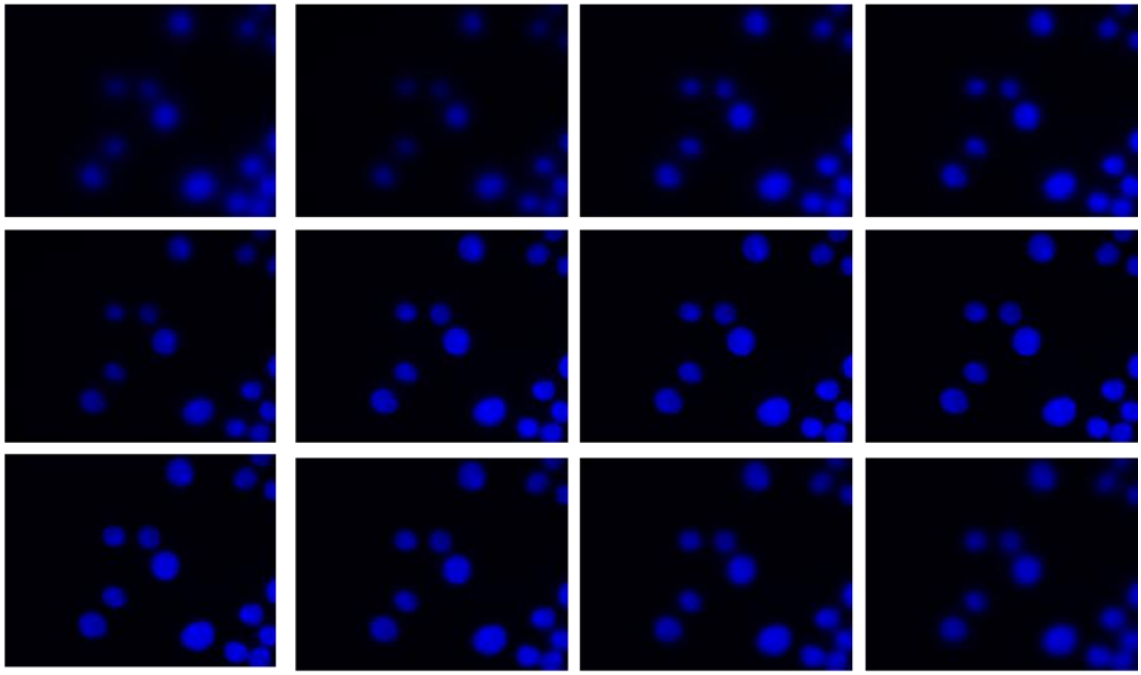
Figure 3.2: *Malat1* expression is diminished in knockout mice.

A) Schematic representation of primers and their respective binding sites used for genotyping (N.B not to scale) **B)** Genotyping of mice - the *Malat1* locus was PCR amplified from genomic DNA isolated from ear clips. PCR products were separated by size using agarose gel electrophoresis. PCR products are labelled. Examples of WT ($^{+}/^{+}$), heterozygote ($^{+}/^{-}$) and *Malat1* knockout ($^{-}/^{-}$) mice are depicted. **C)** Schematic representation of qRT-PCR primer binding sites (N.B not to scale). **D)** Relative levels of *Malat1* RNA expression detected by qRT-PCR in WT and *Malat1* KO *in vitro* polarised Th2 cells (n=7) biological replicates representative of numerous experiments. Data were analysed using an unpaired t-test. p value is displayed.

3.2.2 *Malat1* is located in the nucleus

To gain insight into the function of *Malat1* in $CD4^+$ T cells, its cellular location was examined. As $CD4^+$ T cells grow in suspension, cells were cytospun onto a microscope slide to enable visualisation. Z stack analysis confirmed that EL4 cell nuclear integrity was maintained upon cytospinning, as a similar nuclear morphology was observed in 3T3 cells (adherent cell line) (***Figure 3.3***). *Malat1* localisation was determined using RNAScope which demonstrated that *Malat1* was located in the nucleus of EL4 cells (***Figure 3.4***).

A



B

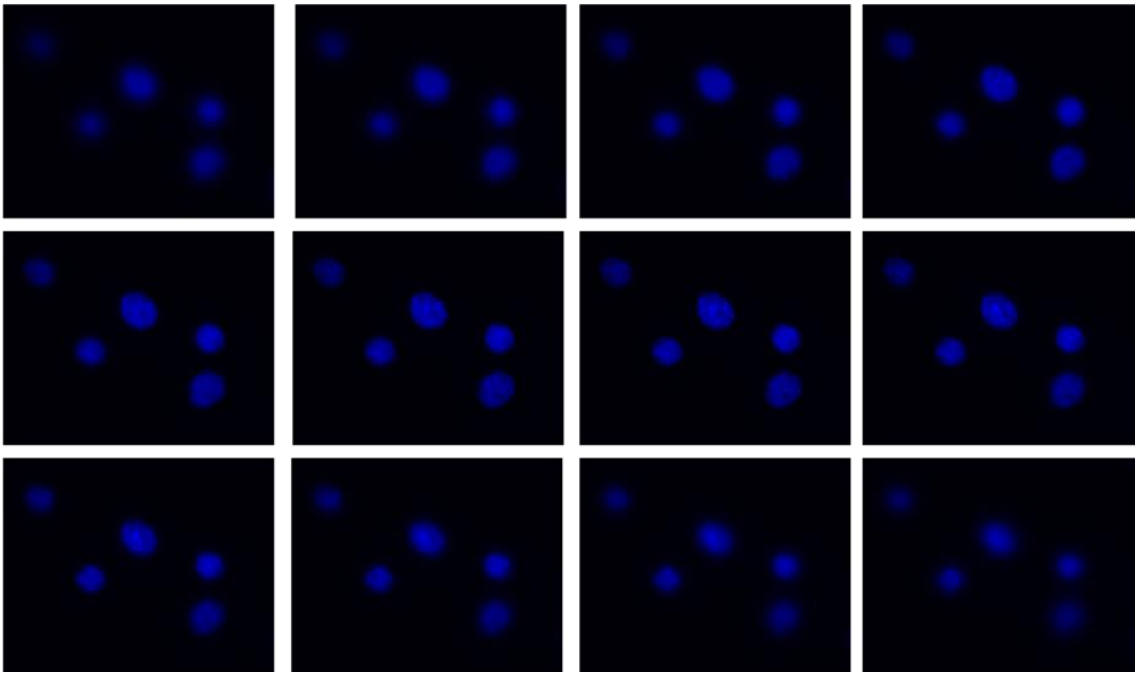


Figure 3.3: T cell nuclear integrity is maintained after cytospinning.

Example Z stack images of nuclei in **A)** EL4 (suspension) cells **B)** 3T3 (adherent) cells. Nuclei are stained with DAPI and are shown in blue. Representative image from n=1 from one experiment.

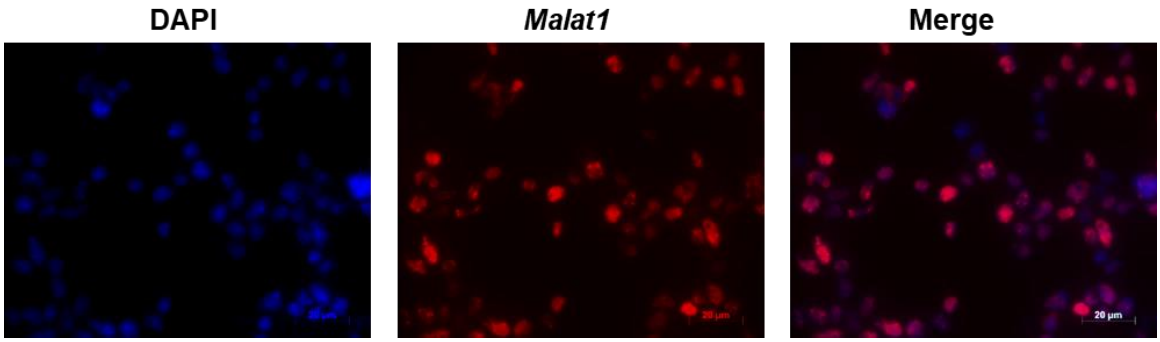


Figure 3.4: *Malat1* is located in the nucleus of $CD4^+$ T cells.

Representative images of EL4 cells showing endogenous *Malat1* (red) detected by RNAscope. Nuclei are stained with DAPI and are shown in blue. The scale bar shows 20µM. n=1 from one independent experiment.

3.2.3 *Malat1* is downregulated upon $CD4^+$ T cell activation

To gain insight into the kinetics of *Malat1* expression during $CD4^+$ T cell activation and differentiation qRT-PCR was used to assess *Malat1* levels at multiple time points during *in vitro* WT Th1 polarisations (***Figure 3.5****A*). We found that *Malat1* is quickly suppressed within 24 hours of $CD4^+$ T cell activation, and that this was sustained for up to 8 days of *in vitro* culture (***Figure 3.5****B*). *Neat1* is a lincRNA that is adjacent to *Malat1* in both human and mouse genomes. We determined that *Neat1* is also suppressed during *in vitro* Th1 polarisation, although this suppression was not as sustained as that of *Malat1* (***Figure 3.5****C*). In addition, *Cd69* was found to be upregulated within 24 hrs of $CD4^+$ T cell activation (***Figure 3.5D***).

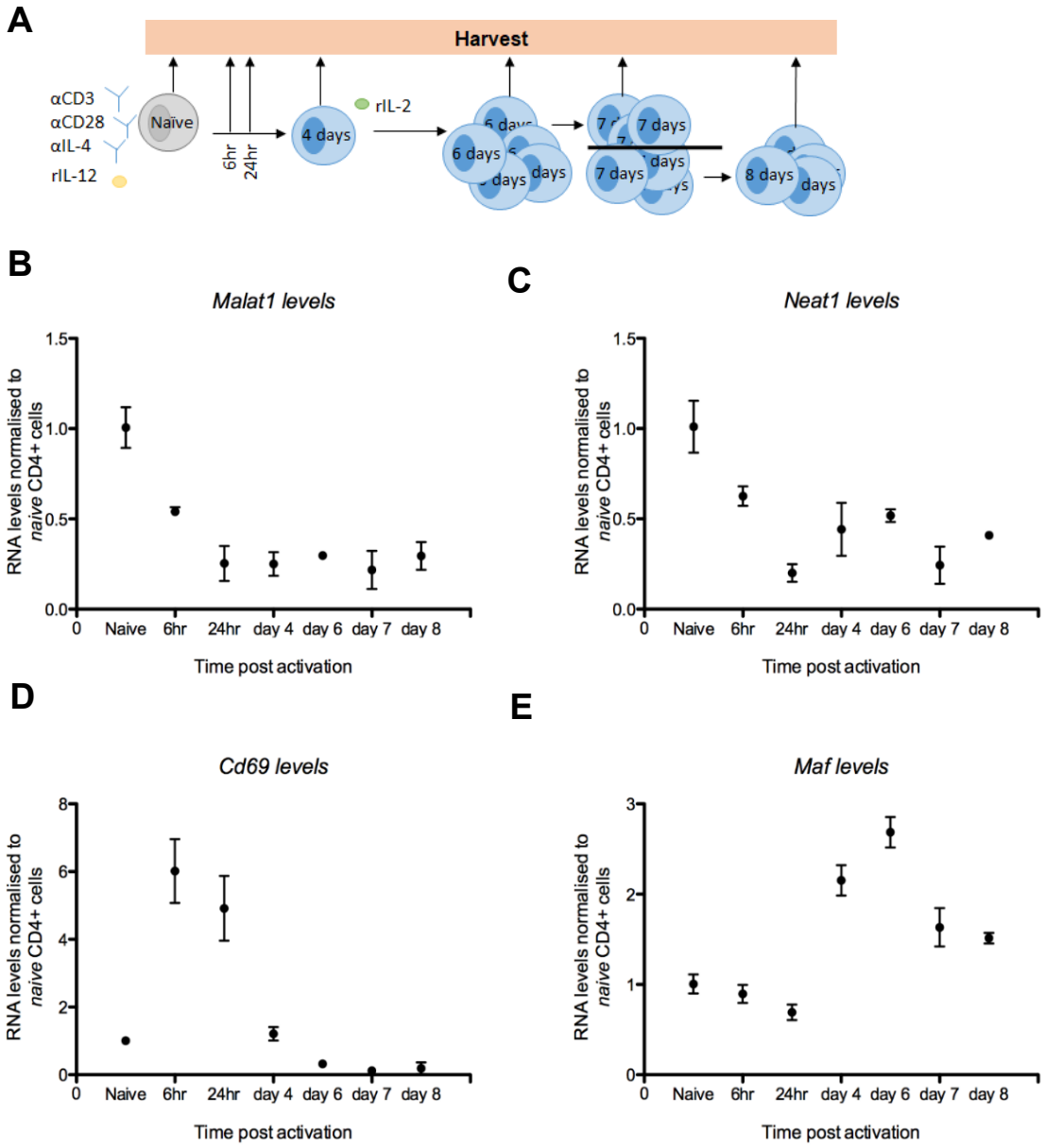


Figure 3.5: *Malat1* is downregulated upon CD4⁺T cell activation.

A) Schematic representation of *in vitro* Th1 polarisation time course. The time points at which samples were collected are depicted. Expression of **B)** *Malat1* RNA levels **C)** *Neat1* RNA levels **D)** *Cd69* RNA levels **E)** *Maf* RNA levels during *in vitro* Th1 polarisation. Data was normalised to naïve CD4⁺ T cells (relative to U6 housekeeper). RNA levels are determined by qRT-PCR. Each point represents the average of n=3 biological replicates representative of two independent experiments. Each data point represents the mean +/- the standard deviation.

3.2.4 *Malat1* does not consistently regulate the expression of chromosomally adjacent genes

Many lncRNAs have been shown to regulate the expression of chromosomally adjacent genes in $CD4^+$ T cells (Gil & Ulitsky, 2019; West & Lagos, 2019). Contrastingly, many lncRNAs also function distally to the site of transcription. Thus, to determine if *Malat1* prominently regulated local gene expression, we began our investigation by examining the expression of *Neat1* and SCY1-like pseudokinase 1 (SCYL1) which are chromosomally adjacent genes to *Malat1*. We examined the expression of these genes in different $CD4^+$ T cell states including Th0, Th1 and Th2 *in vitro* polarised cells derived from WT and *Malat1*$^{-/-}$ mice ***(Figure 3.6A)***.

The expression of *Neat1* was not significantly affected by the loss of *Malat1* in in *vitro* polarised Th0, Th1 or Th2 cells (***Figure 3.6C***). Contrastingly, loss of *Malat1* increased the expression of *Scyl1* specifically in *in vitro* polarised Th2 cells. This suggests that *Malat1* has a limited *in-cis-regulatory* effect, which differs between cell types (***Figure 3.6D***) Given the seemingly limited adjacent regulatory functions of *Malat1,* it appeared logical to next investigate *Malat1* functions in a broader Th cell context and develop tools in order to further probe the mechanistic relevance of *Malat1*.

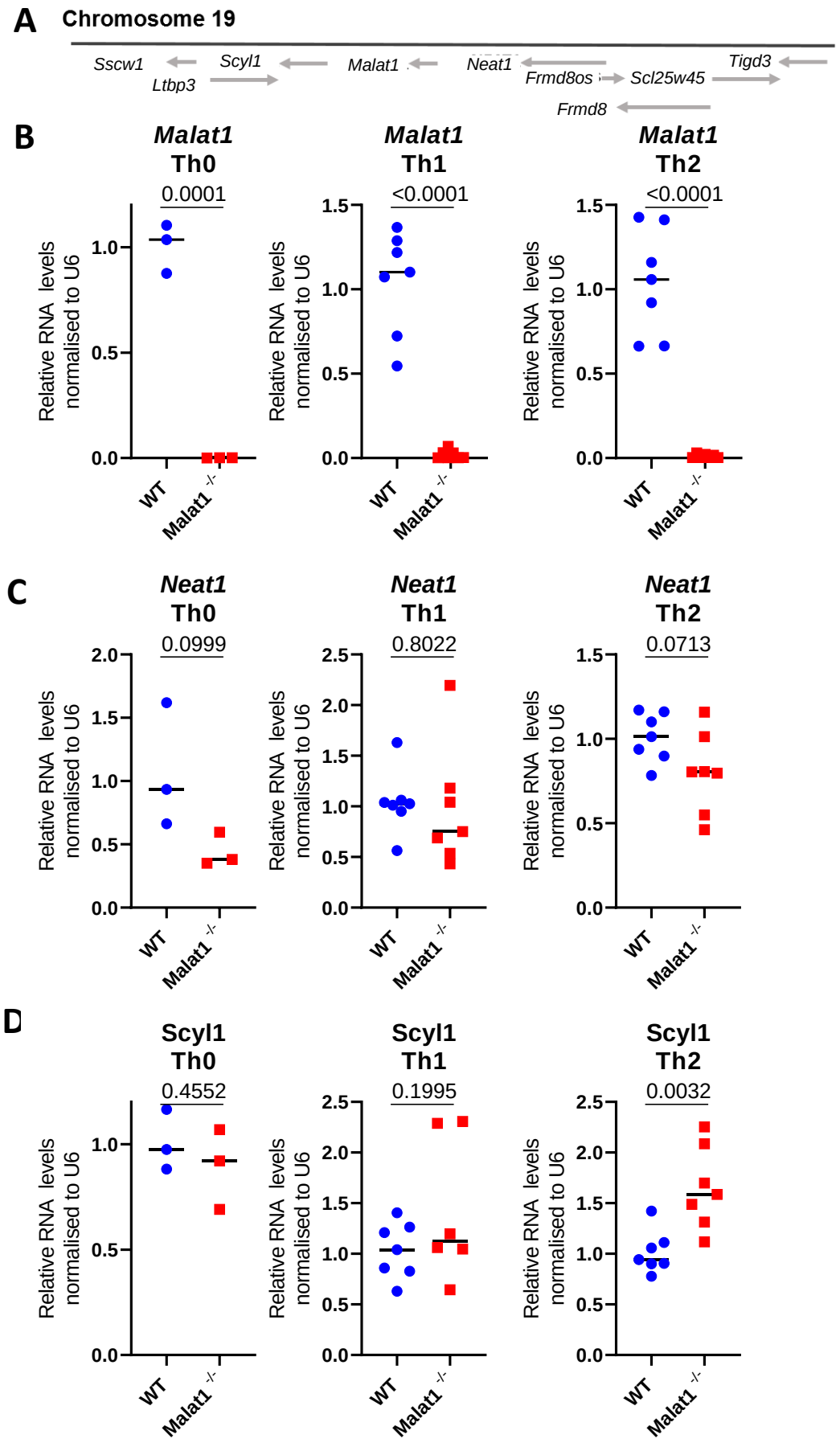

A
Chromosome 19
Sscw1
Ltbp3
Scyl1
Malat1
Neat1
Frmd8os
Scl25w45
Tigd3
Frmd8
B
Malat1 Th0
0.0001
Malat1 Th1
<0.0001
Malat1 Th2
<0.0001
C
Neat1 Th0
0.0999
Neat1 Th1
0.8022
Neat1 Th2
0.0713
D
Scyl1 Th0
0.4552
Scyl1 Th1
0.1995
Scyl1 Th2
0.0032
Relative RNA levels normalised to U6
WT
Malat1 -/-

Figure 3.6: *Malat1* does not consistently regulate the expression of its chromosomally adjacent genes in CD4^{+} T cells.

A) Schematic representation of *Malat1* DNA locus and neighbouring genes. Not to scale. Expression of **B)** *Malat1*, **C)** *Neat1* and **D)** *Scyl1* RNA levels in *in vitro* polarised WT Th1(n=7) (blue), *Malat1*$^{-/-}$ Th1 (n=7) (red), WT Th2 (n=7) (blue), *Malat1*$^{-/-}$ Th2 (n=7) (red), WT Th0 (n=3) (blue), and *Malat1*$^{-/-}$ Th0 (n=3) (red). RNA levels are determined by qRT-PCR. All samples are normalised to housekeeper U6 and the WT of each group. Data were analysed using an unpaired t-test. p value is displayed. Each point represents a biological replicate and is pooled data from two independent experiments (Th1 and Th2 cells). Or one independent experiment (Th0 cells). The mean value is depicted in each condition.

3.2.5 *Malat1* promotes IL-10 expression in Th cells *in vitro*

Next, in order to have tools available to validate our findings from the knockout model efforts were made to knockdown *Malat1* in $CD4^+$ T cells using a multi-pronged approach. It is important to be able to knock down a lncRNA to determine if the RNA product itself is responsible for any potential observed changes. Initial attempts using siRNAs successfully reduced *Malat1* expression in 3T3 cells in a time-dependent manner (***Figure 3.7A***). *Malat1* levels were reduced ~40% after 72 hours of siRNA treatment. However, it was not possible to knockdown *Malat1* using siRNAs in primary $CD4^+$ T cells (***Figure 3.7C***). In contrast, it was possible to knockdown *P300* in primary $CD4^+$ T cells using siRNAs (***Figure 3.7D***). Approximately 80% of *P300* was knocked down using siRNAs after 72 hours of siRNA treatment.

Notably, *Malat1* was significantly knocked down in both EL4 cells and *in vitro* polarised Th1 and Th2 cells when using 100nM of *Malat1* targeting GapmeR (***Figure 3.7E/F***). GapmeR-mediated knockdown of *Malat1* provided a useful platform for examining the relevance of the *Malat1* transcript in $CD4^+$ T cells.

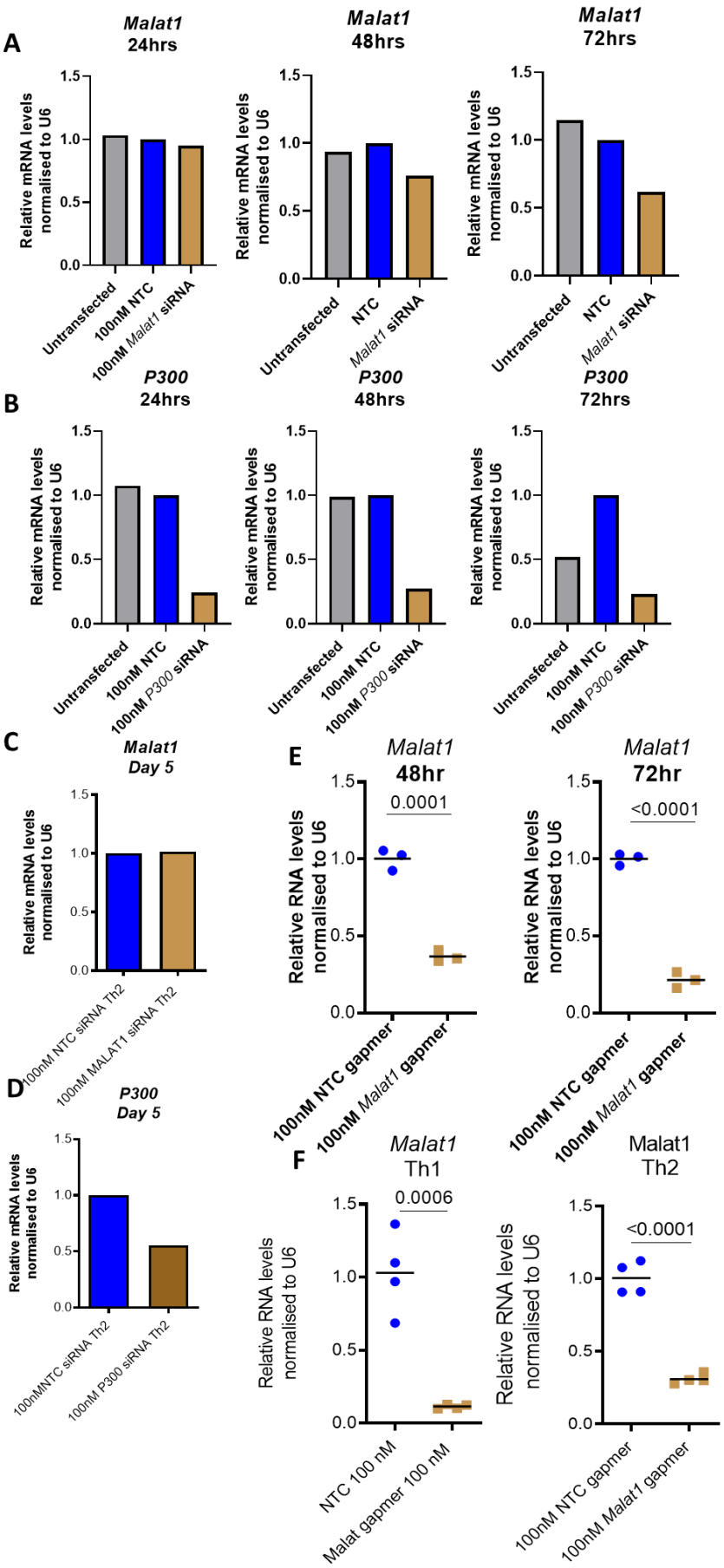

A
Malat1
24hrs
Malat1
48hrs
Malat1
72hrs
Relative mRNA levels normalised to U6
1.5
1.0
0.5
0.0
Untransfected
100nM NTC
100nM Malat1 siRNA
NTC
Malat1 siRNA
B
P300
24hrs
P300
48hrs
P300
72hrs
100nM P300 siRNA
C
Malat1
Day 5
100nM NTC siRNA Th2
100nM MALAT1 siRNA Th2
D
P300
Day 5
100nMNTC siRNA Th2
100nM P300 siRNA Th2
E
Malat1
48hr
0.0001
Relative RNA levels normalised to U6
100nM NTC gapmer
100nM Malat1 gapmer
Malat1
72hr
<0.0001
F
Malat1
Th1
0.0006
NTC 100 nM
Malat gapmer 100 nM
Malat1
Th2
<0.0001

Figure 3.7: *Malat1* is knocked-down using GapmeRs in primary CD4^{+} T cells

A) Relative *Malat1* RNA levels in 3T3 cells when transfected with 100 nM of NTC siRNA (blue) or *Malat1* targeting siRNA (brown) at 24 hours, 48 hours and 72 hours. n=1. **B)** Relative P300 RNA levels in 3T3 cells when transfected with 100 nM of NTC siRNA (blue) or P300 targeting siRNA (brown) at 24 hours, 48 hours and 72 hours. n=1 **C)** Relative *Malat1* RNA levels in primary *in vitro* polarised Th1 cells when transfected with 100nM of NTC siRNA (blue) or *Malat1* targeting siRNA (blue) 5 days post-transfection. n=1 **D)** Relative *P300* RNA levels in primary *in vitro* polarised Th1 cells when transfected with 100 nM of NTC siRNA (blue) or *Malat1* targeting siRNA (brown) 5 days post-transfection. n=1 **E)** Relative *Malat1* RNA levels in EL4 cells when treated with 100 nM of NTC GapmeR (blue) or *Malat1* targeting GapmeR (brown) at 48 and 72 hours post treatment. n=3 biological replicates from 3 independent experiments **F)** Relative *Malat1* RNA levels in *in vitro* polarised Th1 or Th2 cells when treated with 100nM of NTC GapmeR (blue) or *Malat1* targeting GapmeR (brown) at day 6 post-treatment. n=4 biological replicates representative of two independent experiments. RNA levels were determined by qRT-PCR and normalised to U6 of NTC controls. Where appropriate data were analysed using an unpaired t-test. p values are displayed. Where n>1 the mean value is depicted in each condition.

Having found that *Malat1* expression is rapidly reduced upon Th activation and that *Malat1* had limited immediate *in-cis*-regulatory effects, we next examined the role of *Malat1* on Th cell cytokine production in *in vitro* $CD4^+$ T cell polarised Th1 and Th2 cells. We found that following *in vitro* Th1 polarisation loss of *Malat1* reduced the protein expression of IFNγ (although not significantly), but significantly reduced the protein expression of IL-10 (***Figure 3.8A/B***). Upon *in vitro* differentiation of Th2 cells, loss of *Malat1* also resulted in a reduction in the percentage of IL-10 expression, expression of IL-4 remained unchanged in the absence of *Malat1* (***Figure 3.8C/D***). The reduction in IL-10 expression was more prominent in Th2 cells than in Th1 cells – a likely reflication of the differences in baseline IL-10 expression. IL-10 was also found to be downregulated at the protein level in *Malat1*$^{-/-}$ *in vitro* polarised Th17 cells (***Figure 3.9***) A reduction in IL-17 expression was also observed in *Malat1*$^{-/-}$ Th17 cells. (***Figure 3.9).***

In addition to experiments where *Malat1* expression was diminished through genetic deletion, we complemented these studies by using antisense oligonucleotides to target the RNA product of *Malat1*. We found that knockdown of *Malat1* also significantly diminished IL-10 protein expression in *in vitro* polarised Th2 cells (***Figure 3.8E***) Knockdown of *Malat1* did not significantly reduce IL-10 expression at the protein level in Th1 cells but decreased IFNγ expression (***Figure 3.8E/F***).

A similar effect on IL-10 expression is also seen at the mRNA level. The loss of *Malat1* reduced mRNA levels of IL-10 in both Th1 and Th2 cells (***Figure 3.10A***). However, this was more pronounced in Th2 cells. Additionally, knockdown of *Malat1* also significantly reduced *Il10* mRNA levels in Th1 and Th2 cells (***Figure 3.10B***). Collectively this data provides compelling evidence that the *Malat1* transcript regulates IL-10 expression in *in vitro* polarised $CD4^+$ T cells.

We repeated these experiments under reduced polarising conditions where we found a statistically significant reduction in IFNγ expression in Th1 cells. However, no change in IL-10 or IL-4 was observed upon loss of *Malat1* under weakly polarising conditions in Th2 cells (***Figure 3.11***).

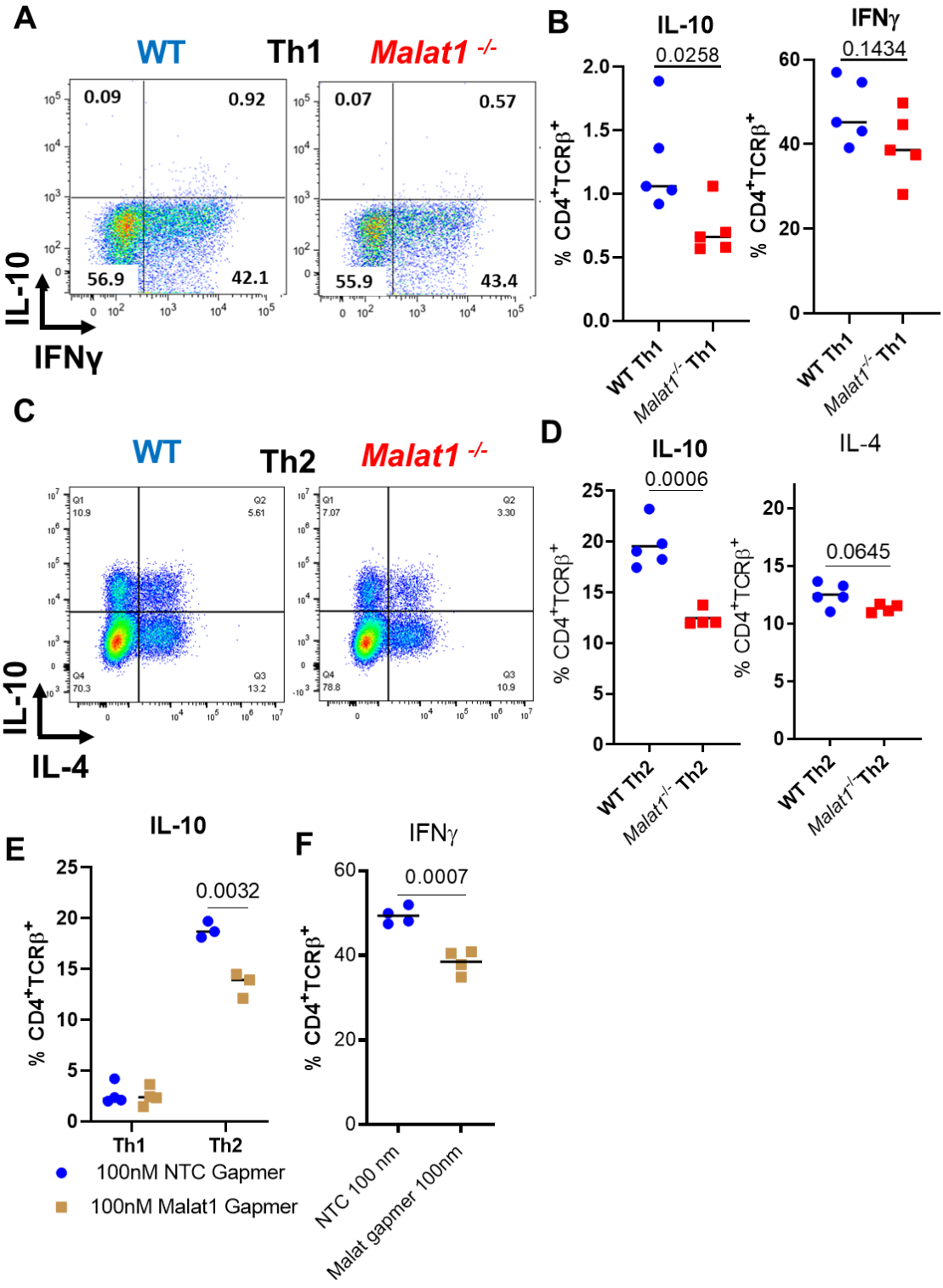

A
WT
Th1
Malat1 -/-
0.09
0.92
56.9
42.1
0.07
0.57
55.9
43.4
IL-10
IFNγ
B
IL-10
0.0258
IFNγ
0.1434
% CD4+TCRβ+
WT Th1
Malat1-/- Th1
C
WT
Th2
Malat1 -/-
Q1 10.9
Q2 5.61
Q4 70.3
Q3 13.2
Q1 7.07
Q2 3.30
Q4 78.8
Q3 10.9
IL-10
IL-4
D
IL-10
0.0006
IL-4
0.0645
WT Th2
Malat1-/- Th2
E
IL-10
0.0032
Th1
Th2
100nM NTC Gapmer
100nM Malat1 Gapmer
F
IFNγ
0.0007
NTC 100 nm
Malat gapmer 100nm

Figure 3.8: **Loss of *Malat1* reduces the expression of IL-10 in *in vitro* polarised Th1 and Th2 cells.**

A) Representative FACs plots of IL-10 and IFNγ expression in *in vitro* polarised Th1 (day 6) in WT or *Malat1*$^{-/-}$ cells determined by intracellular cytokine staining **B)** Quantification of the percentage of live, CD4^{+} TCRβ^{+} cells that express IL-10, IFNγ in WT (blue) or *Malat1*$^{-/-}$ (red) *in vitro* polarised Th1 cells (day 6). n=5 biological replicates representative of at least 3 independent experiments **C)** Representative FACs plots of IL-10 and IL-4 expression in *in vitro* polarised Th2 (day 6) in WT or *Malat1*$^{-/-}$ cells determined by intracellular cytokine staining **D)** Quantification of the percentage of live, CD4^{+} TCRβ^{+} cells that express IL-10, or IL-4 in WT (blue) or *Malat1*$^{-/-}$ (red) *in vitro* polarised Th1 cells (day 6) n=5.biological replicates representative of at least 3 independent experimentst **E)** Quantification of the percentage of live, CD4^{+} TCRβ^{+} cells that express IL-10, in *in vitro* differentiated Th1 cells (n=4 biological replicates) or Th2 cells (n=3) biological replicates representative of two independent experiments (day 6) transfected with control (blue) or *Malat1*-targetting GapmeR (brown) **F)** Quantification of the percentage of live, CD4^{+} TCRβ+ cells that express IFNγ, in in vitro differentiated Th1 (day 6) transfected with control (blue) or *Malat1*-targetting GapmeR (brown) n=4 biological replicates, representative of two independent experiments. Data were analysed using an unpaired t-test. p values are displayed**.**

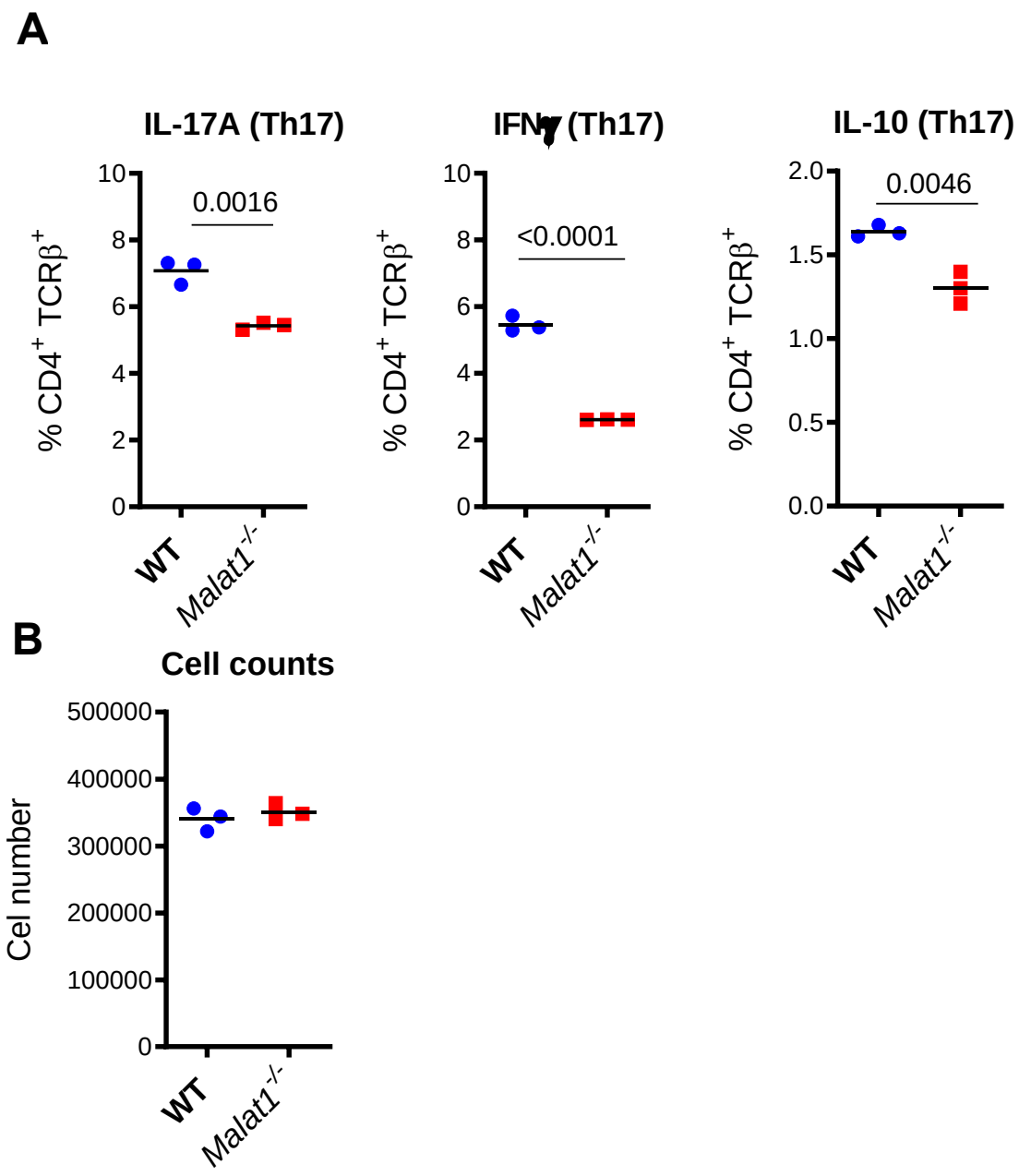


Figure 3.9: Loss of *Malat1* impairs *in vitro* Th17 differentiation.

A) Percentage of IL-17$^+$, IFNY$^+$, IL-10$^+$ live TCRβ$^+$ CD4$^+$ WT (blue) or *Malat1*$^{-/-}$ (red) in vitro differentiated Th17 cells. Levels determined by intracellular cytokine staining. **B)** Cell counts of WT (blue) or *Malat1*$^{-/-}$ (red) in vitro polarised Th17 cells. n=3 biological replicates representative of 1 independent experiment. Data were analysed using an unpaired t-test. p value is displayed.

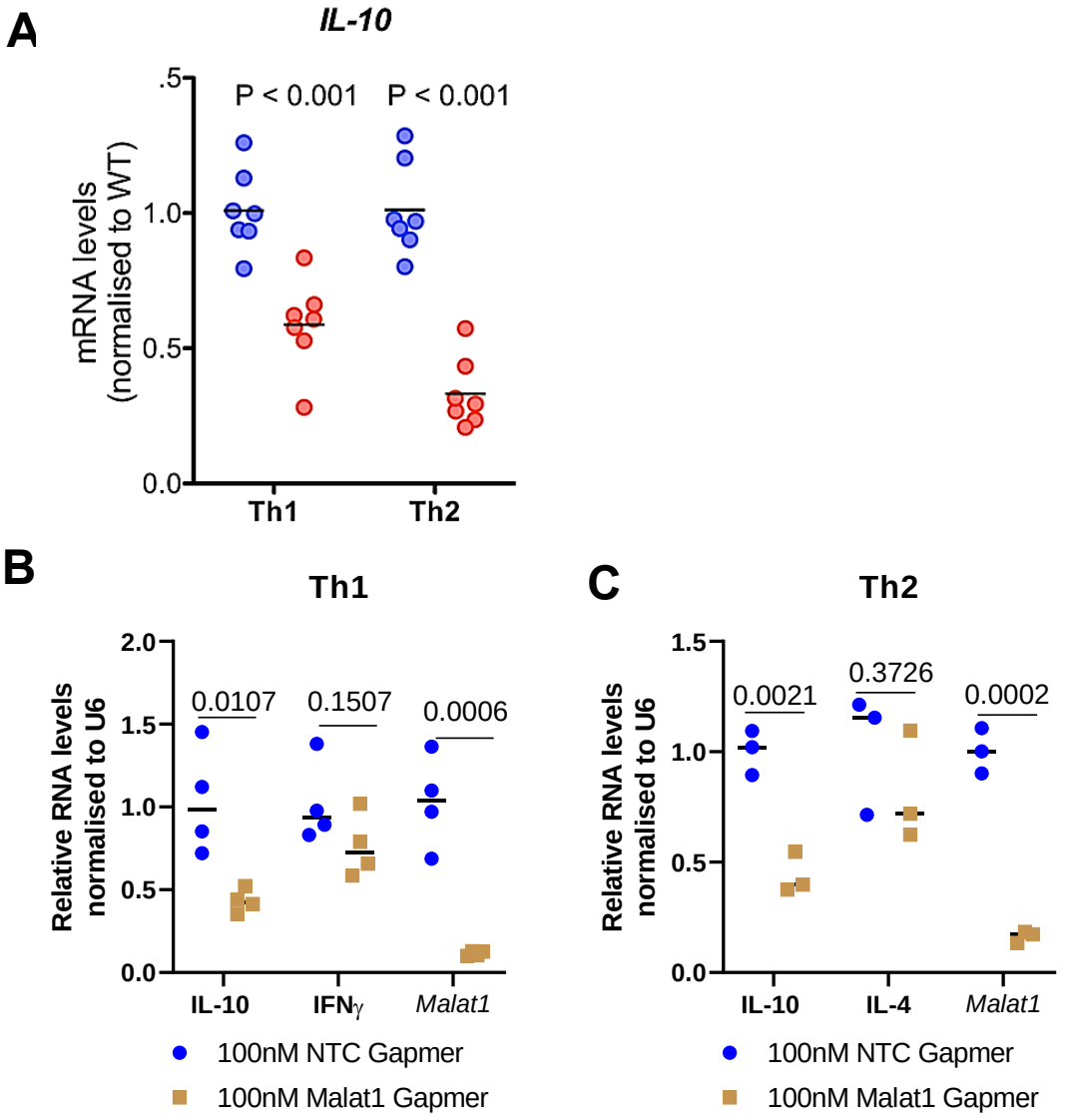


Figure 3.10: Loss of *Malat1* reduces mRNA levels of IL-10 in *in vitro* polarised CD4⁺ T cells.

A) Relative IL-10 mRNA levels from in vitro polarised WT (blue) or *Malat1*$^{-/-}$ (red) Th1 and Th2 cells, n=7 biological replicates representative of at least 3 independent experiments. **B)** IL-10, IFN-γ, and *Malat1* RNA levels in in vitro polarised Th1 cells treated with 100nM control (blue) or 100nM *Malat1*-targeting (brown) GapmeR, n=4. RNA levels were normalized to U6 and average levels in cells treated with control GapmeR. **C)** IL-10, IL-4, and *Malat1* RNA levels in in vitro polarised Th2 cells, transfected with 100nM control (blue) or 100 nM *Malat1*-targeting (brown) GapmeR n=3. All RNA levels normalized to U6 and average levels in either WT cells or cells treated with control GapmeR. Data were analysed using an unpaired t-test. p values and the mean are displayed

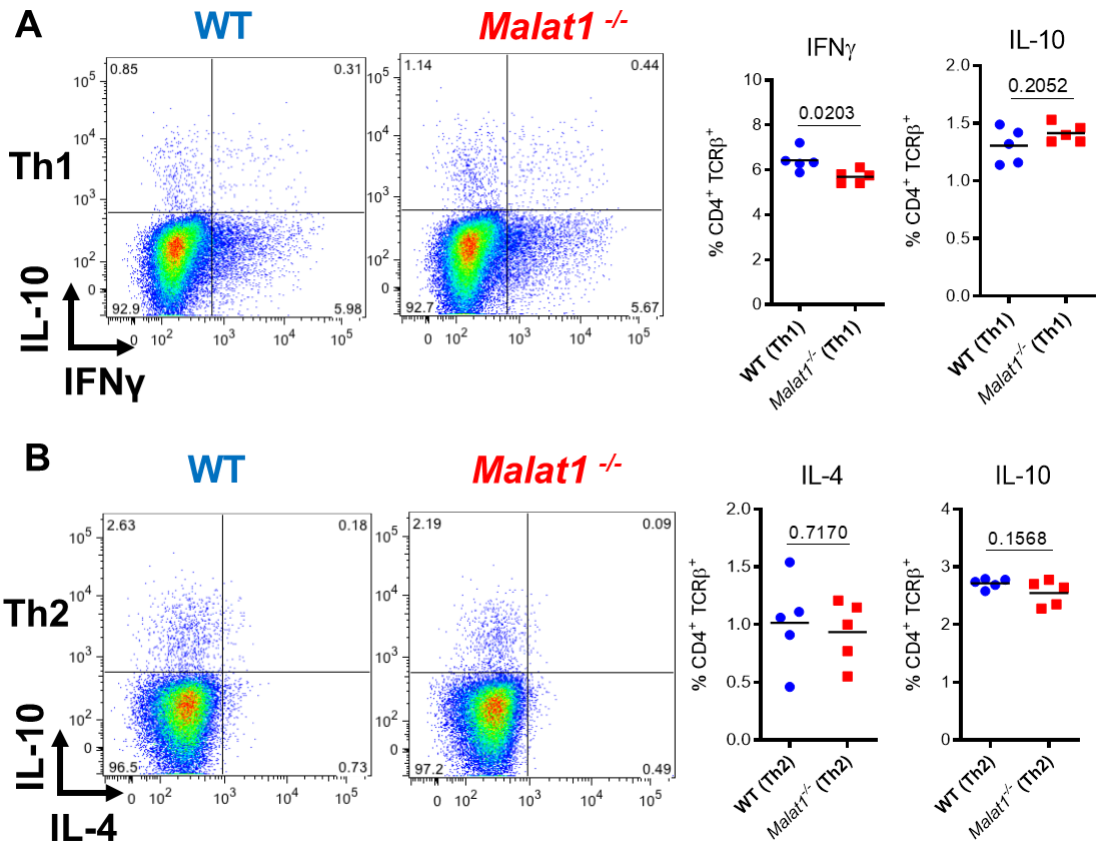


Figure 3.11: Differentiation of Th1 and Th2 cells is not affected under reduced polarising conditions in *Malat1*$^{-/-}$ cells.

Representative flow cytometry plots of WT (blue) or *Malat1*$^{-/-}$ *in vitro* polarised Th cells (red) (left) Quantification of the percentage of IL-10^{+}, IFNγ^{+} or IL-4^{+} live TCRβ^{+} CD4^{+} WT (blue) or *Malat1*$^{-/-}$ (red) under conditions inducing suboptimal **A)** Th1 or **B)** Th2 differentiation (conditions are defined in materials and methods chapter). Levels determined by intracellular cytokine staining. n=5. Biological replicates. Representative of two independent experiments. Data were analysed using an unpaired t-test. p values and the mean are displayed.

3.2.6 The transcription factor MAF is downregulated in *Malat1*$^{-/-}$ *in vitro* polarised Th cells

To explore the potential mechanism of *Malat1*-mediated IL-10 regulation, we next investigated transcription factors associated with IL-10 regulation. Many transcription factors are known to regulate IL-10 expression in CD4^{+} T cells in both mice and humans (***Figure 3.12A*).** The transcription factor MAF (also known as c-MAF) has been deemed a master regulator of IL-10 in both myeloid and T cells (Saraiva et al., 2019). Thus, we next examined the expression of MAF in WT and *Malat1*$^{-/-}$ *in vitro* polarised Th1 and Th2 cells.

Indeed, *Malat1*$^{-/-}$ *in vitro* polarised Th1 and Th2 cells showed reduced mRNA levels of MAF which reaches statistical significance in Th2 cells. Additionally, when *Malat1* is knocked down using GapmeRs this also reduced MAF mRNA levels in *in vitro* polarised Th1 and Th2 cells (***Figure 3.12C).*** A more prominent downregulation of MAF is seen at the protein level in Th1 cells (***Figure 3.12D).*** We only observed a minor downregulation of the transcription factor STAT4 at the protein level which is also known to promote IL-10 transcription in Th1 cells (***Figure 3.12D).*** Of note, Th2 cells express much higher RNA levels of *Malat1*, *Maf* and *Il10* than Th1 or Th0 cells *in vitro* (***Figure 3.13),*** which is likely why a stronger phenotype is observed in this cell type.

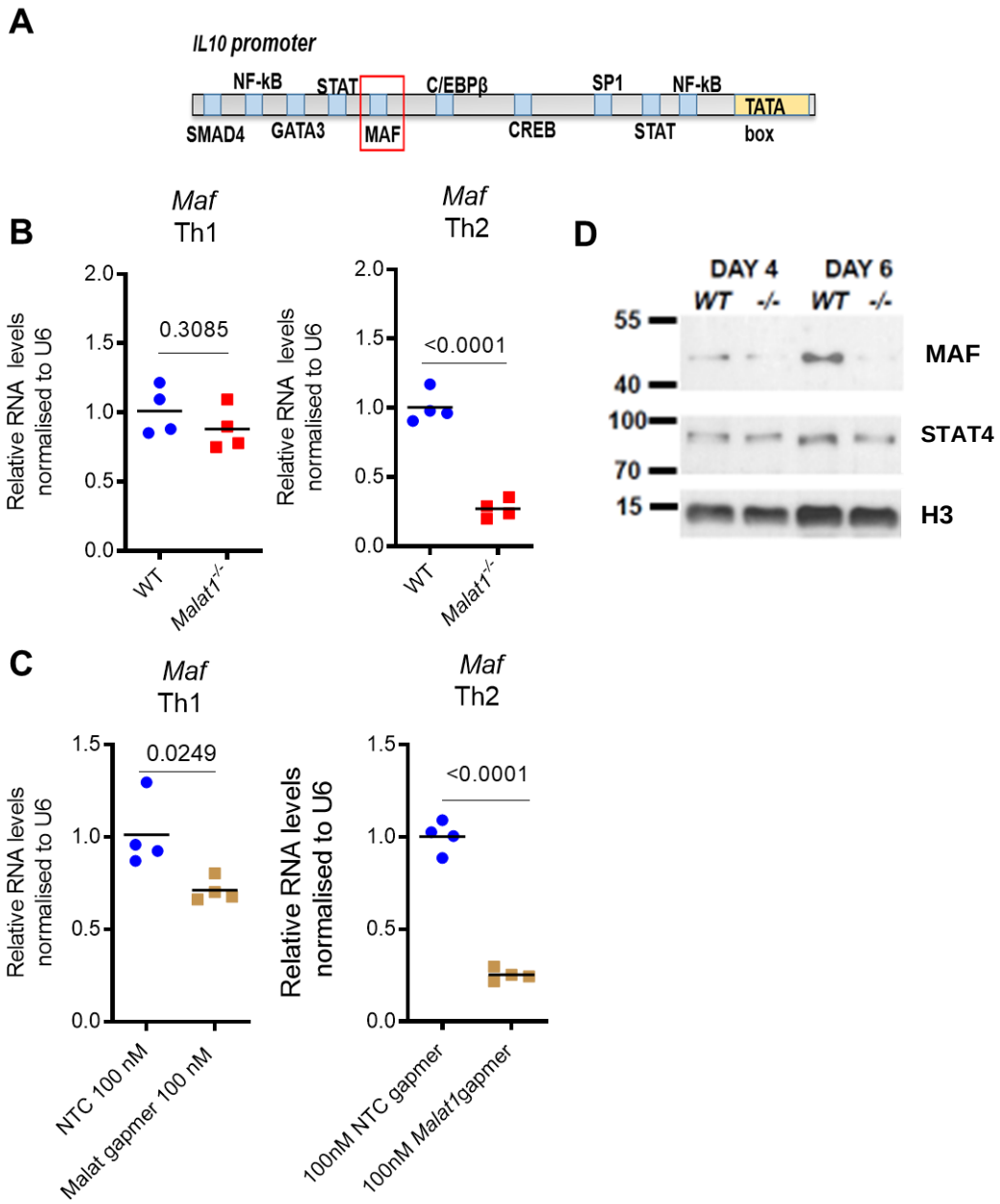


Figure 3.12: *Malat1* promotes the expression of MAF in CD4⁺ T cells.

A) Depiction of the *il-10* promoter and known transcription factor binding sites. **B)** Relative *Maf* mRNA levels from *in vitro* polarised WT (blue) or *Malat1*$^{-/-}$ (red) Th1 and Th2 cells, n=4 biological replicates representative of 2 independent experiments day 6 **C)** *Maf* RNA levels in in vitro polarised Th1 and Th2 cells treated with 100nM control (blue) or 100nM *Malat1*-targeting (brown) GapmeR, n=4 biological replicates representative of 2 independent experiments day 6. **D)** Protein levels of MAF and STAT4 in WT or *Malat1*$^{-/-}$ in *in vitro* polarised Th1 cells at days 4 or 6 *in vitro* polarisation determined by western blot. Representative blot of 2 independent experiments. Where appropriate data were analysed using an unpaired t-test. p values are displayed. The mean is also displayed.

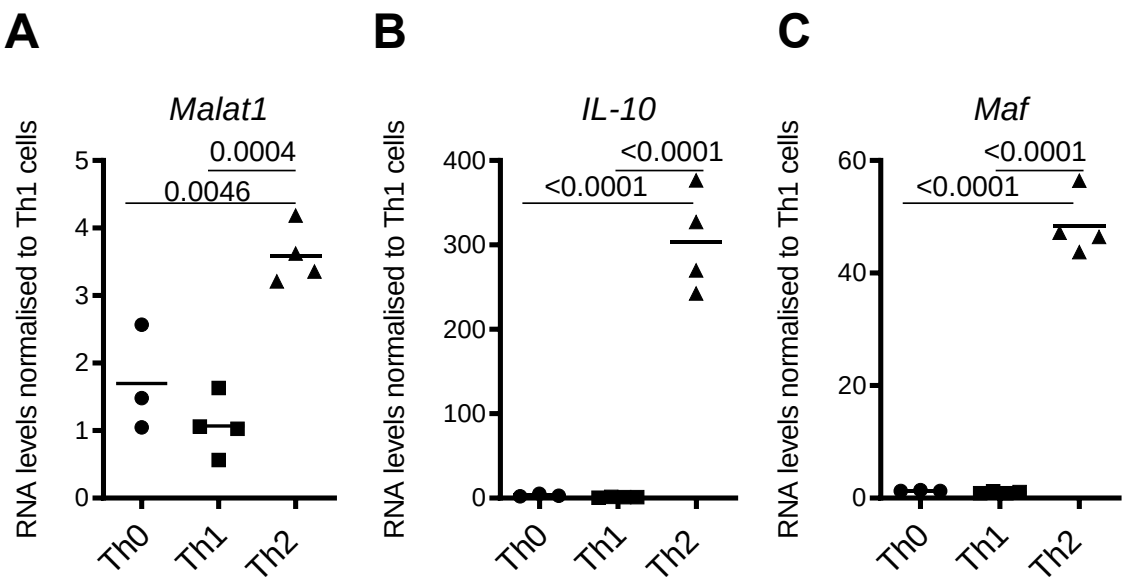


Figure 3.13: *Malat1* is most abundant in WT *in vitro* polarised Th2 cells.

Relative abundance of **A)** *Malat1* **B)** *Il10* and **C)** *Maf* RNA levels in *in vitro polarised* Th1, Th2 and Th0 cells. RNA levels were determined by qRT-PCR. n=3-4 biological replicates representative of one experiment. All samples are normalised to housekeeper U6 and Th1 cells. Data were analysed using a one-way ANOVA with the Tukey post hoc test. p value and the mean is displayed.

We next examined the RNA levels of key lineage transcription factors. No effect on the expression of *Tbet* in *Malat1*$^{-/-}$ *in vitro* polarised Th1 cells was observed. (***Figure 3.14***). Nor were the mRNA levels of *Gata3* or *Stat6* significantly altered in *Malat1*$^{-/-}$ *in vitro* polarised Th2 cells (***Figure 3.14***).

In addition, we also screened for the expression of other transcription factors that have been linked with the transcriptional regulation of IL-10. One example of this is basic helix loop helix family member e40 (BHLHE40) which is known to transcriptionally suppress IL-10 (Huynh et al., 2018). Here, upon loss of *Malat1*, there was no significant difference in the mRNA expression of *Bhlhe40* between WT and *Malat1*$^{-/-}$ *in vitro* polarised Th1 and Th2 cells which indicated that BHLHE40 is not responsible for the observed reduction in IL-10 expression (***Figure 3.14***).

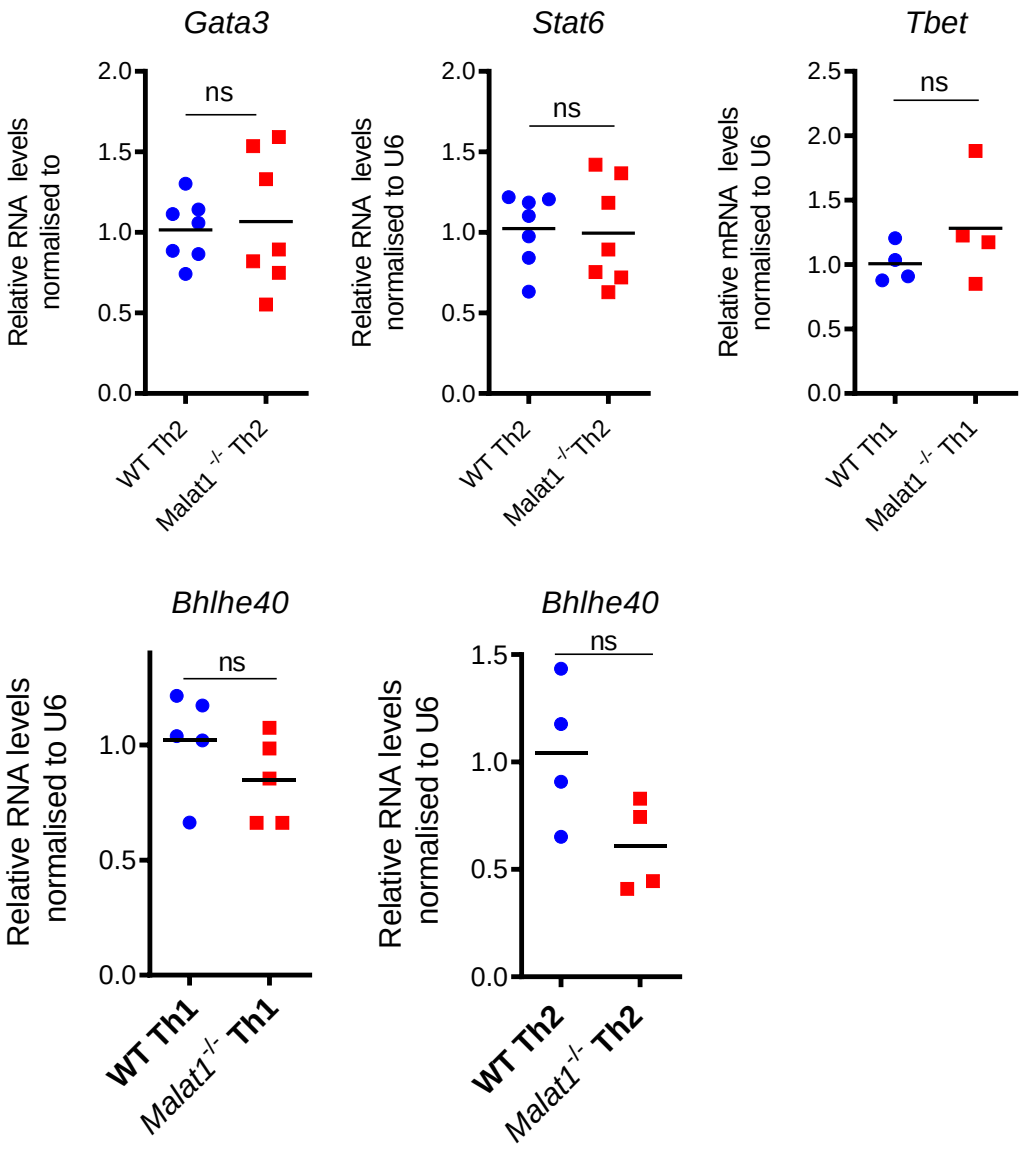


Figure 3.14: Expression of key T cell lineage transcription factors does not change on loss of *Malat1.*

Relative abundance of **A)** *Gata3* **B)** *Stat6* **C)** *Tbet* **D)** *Bhlhe40* RNA levels in *in vitro* polarised Th2 or Th1 cells derived from WT (blue) and *Malat1*$^{-/-}$*(red)* mice. RNA levels were determined by qRT-PCR. All samples were normalised to the U6 housekeeping gene. Data were analysed using an unpaired t-test. Statistical significance is noted in the graphs. n=4-8 biological replicates for representative of upto two independent experiments. ns= not significant. The mean is displayed.

To attempt to validate the link between *Malat1* and IL-10, *Malat1* was knocked down using GapmeRs in EL4 cells and transcriptional changes were assessed at multiple timepoints post-GapmeR addition by qRT-PCR. Here, *Malat1* was significantly suppressed by the GapmeRs in a time-dependent manner. However, no change in the RNA levels of *Maf*, *IL-10*, *Neat1* or *Bhlhe40* was observed (***Figure 3.15***). This is likely due to fundamental differences between primary cells and cultured cell lines.

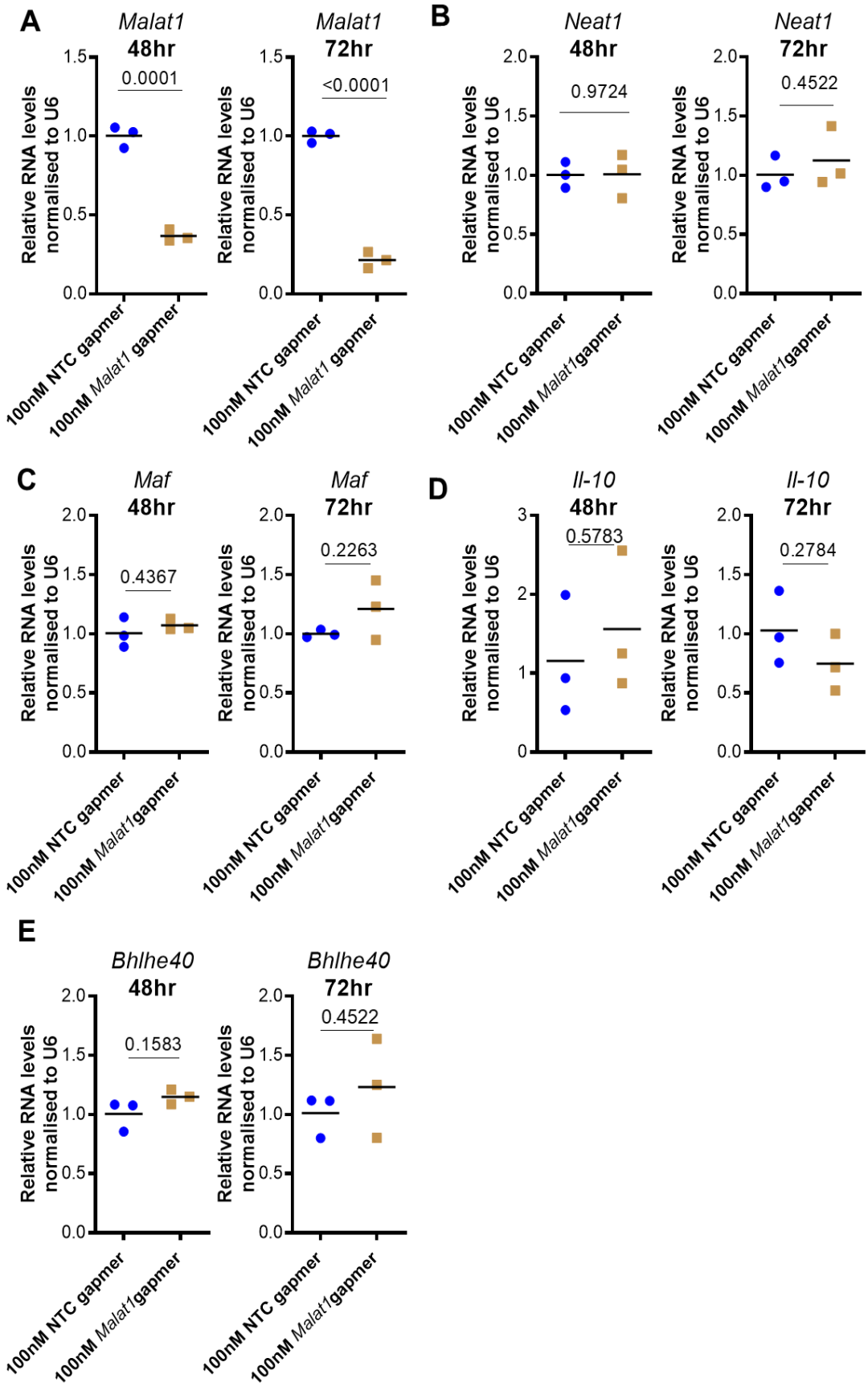

A
Malat1
48hr
0.0001
Malat1
72hr
<0.0001
B
Neat1
48hr
0.9724
Neat1
72hr
0.4522
C
Maf
48hr
0.4367
Maf
72hr
0.2263
D
Il-10
48hr
0.5783
Il-10
72hr
0.2784
E
Bhlhe40
48hr
0.1583
Bhlhe40
72hr
0.4522
Relative RNA levels normalised to U6
100nM NTC gapmer
100nM Malat1 gapmer

Figure 3.15: Knockdown of *Malat1* in EL4 cells does not affect *Il-10* or *Maf* mRNA levels.

Relative abundance of **A)** *Malat1* **B)** *Neat1* **C)** *Maf* **D)** *IL-10* and **E)** *Bhlhe40* RNA levels in EL4 cells (day 6) transfected with control (blue) or *Malat1*-targeting (brown) GapmeR (100 nM) for 48 or 72 hours. Levels normalized to U6 and average levels in NTC-treated cells. RNA levels determined by qRT-PCR. n=3 biological replicates. Data were analysed using an unpaired t-test. P values and the mean are displayed.

3.2.7 *Malat1* promotes IL-10 expression *in vivo*

Following our publication (Hewitson et al., 2020) and to further explore the role of *Malat1 in vivo,* we used a competitive mixed bone marrow model followed by a type 1 or type 2 infection model to determine if *Malat1* also regulated IL-10 expression under *in vivo* conditions. Dr James Hewitson performed all procedures for this experiment. (B6.CD45.1 x B6.CD45.2) F1 mice mice were irradiated and reconstituted with a 50:50 mix of C57BL/6 CD45.1 (WT) and C57BL/6 CD45.2 (*Malat1*$^{-/-}$) bone marrow derived from female mice. The mice were subsequently infected with *S. mansoni* or *L. donovani* and tissues were examined for IL-10 expression (***Figure 3.16***). We chose to study a model of visceral leishmaniasis as IL-10 produced by Th1 cells has previously been shown to be a critical determinant of disease outcome and pathology in humans (Anderson et al., 2007). We chose to study a helminth infection as chronic *S. mansoni* infection leads to increased production of IL-10 and is representative of a type 2 immune response (Araújo et al., 2004). Example gating strategies are shown in the methods section – (Figures 2.3-2.5)

S. mansoni infection revealed that *Malat1*$^{-/-}$ CD4^{+} T cells produced less IL-10 compared to their WT counterparts in the spleen (***Figure 3.17A***). Loss of *Malat1* did not affect the expression of IL-13 or IL-4 in either spleen or lymph nodes of infected mice (***Figure 3.17B/C***).

Analysis of CD4^{+} T cells derived from *L. donovani* infected mice revealed that *Malat1*$^{-/-}$ cells also produced less IL-10 in both the spleen and lymph nodes (**Figure *3.18*A**). IFNγ expression was not significantly changed in the absence of *Malat1* but showed a trend towards reduced expression in lymph nodes (**Figure *3.18*B**).

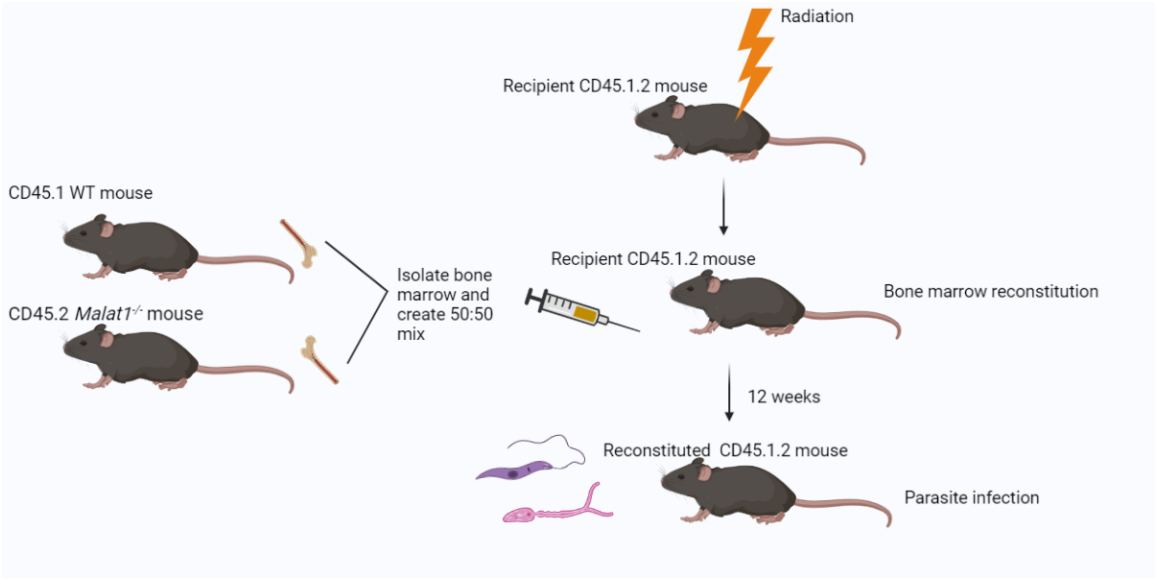


Figure 3.16: Schematic representation of the mixed bone marrow chimera infection model.

Briefly, bone marrow is isolated from CD45.1 WT mice and CD45.2 *Malat1*$^{-/-}$ mice and a 50:50 mix is generated. WT (B6.CD45.1 x B6.CD45.2) F1 mice (shown as 45.1.2 in diagram) are irradiated and their bone marrow reconstituted with the 50:50 bone marrow mix and used for infection models of leishmaniasis or schistosomiasis.

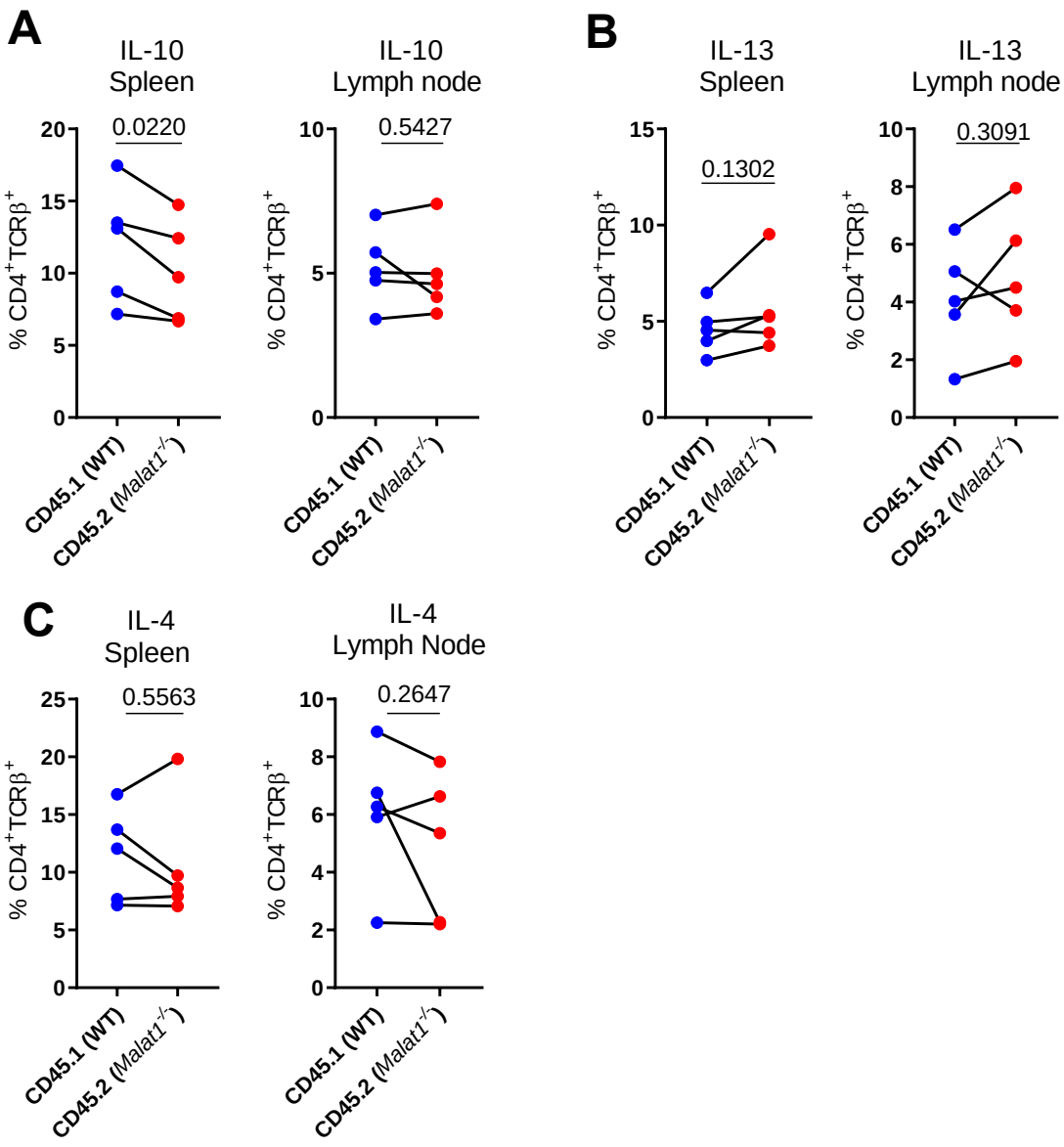


Figure 3.17: Loss of *Malat1* reduces IL-10 expression in schistosomiasis infection models in mixed bone marrow chimera mice.

A) The Percentage of IL-10$^+$, **B)** IL-13$^+$, **C)** IL-4$^+$ live TCRβ$^+$ CD4$^+$ WT CD45.1 (blue) or *Malat1*$^{-/-}$ CD45.2 (red) in *S. mansoni* infected bone marrow chimera mice. Levels determined by intracellular cytokine staining. n=5 biological replicates representative of one independent experiment. Data were analysed using a paired t-test. p value is displayed.

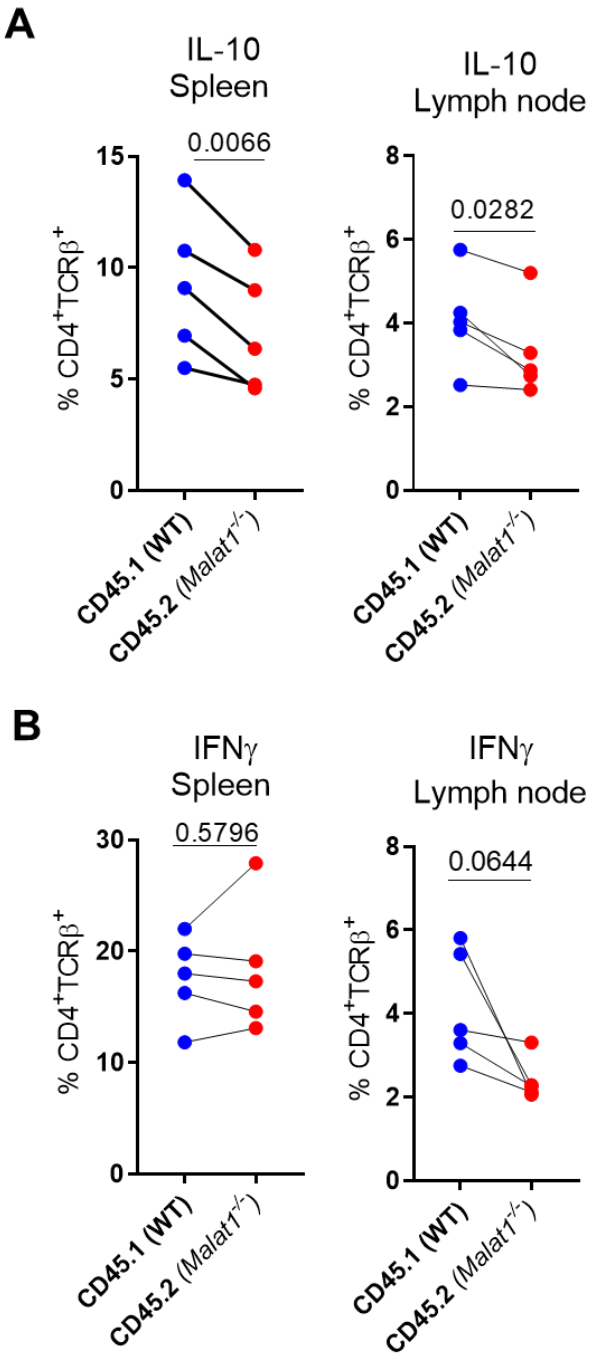


Figure 3.18 Loss of *Malat1* reduces IL-10 expression in leishmaniasis infection models in mixed bone marrow chimera mice.

A) Percentage of IL-10$^+$, **B)** IFNγ$^+$, live TCRβ$^+$ CD4$^+$ WT CD45.1 (blue) or *Malat1*$^{-/-}$ CD45.2 (red) in *L. donovani* infected bone marrow chimera mice. Levels determined by intracellular cytokine staining. n=5 biological replicates representative of one independent experiment. Data were analysed using a paired t-test. P value is displayed.

Notably, when we compared the percentage of $CD4^+$ T cells, *Malat1*$^{-/-}$ (CD45.2) cells were more abundant than their WT (CD45.1) counterparts in both *L. donovani* infection and *S. mansoni* infection (***Figure 3.19)***. Some residual host CD45.1.2 cells remained; however, these were only a small percentage of the total cells (***Figure 3.19)***. This observation was not unique to $CD4^+$ T cells and other lymphocyte populations such as *Malat1*$^{-/-}$ (CD45.2) B cells preferentially expanded in both infection models in comparison to their WT counterparts ***(Figure 3.19 C/D).***

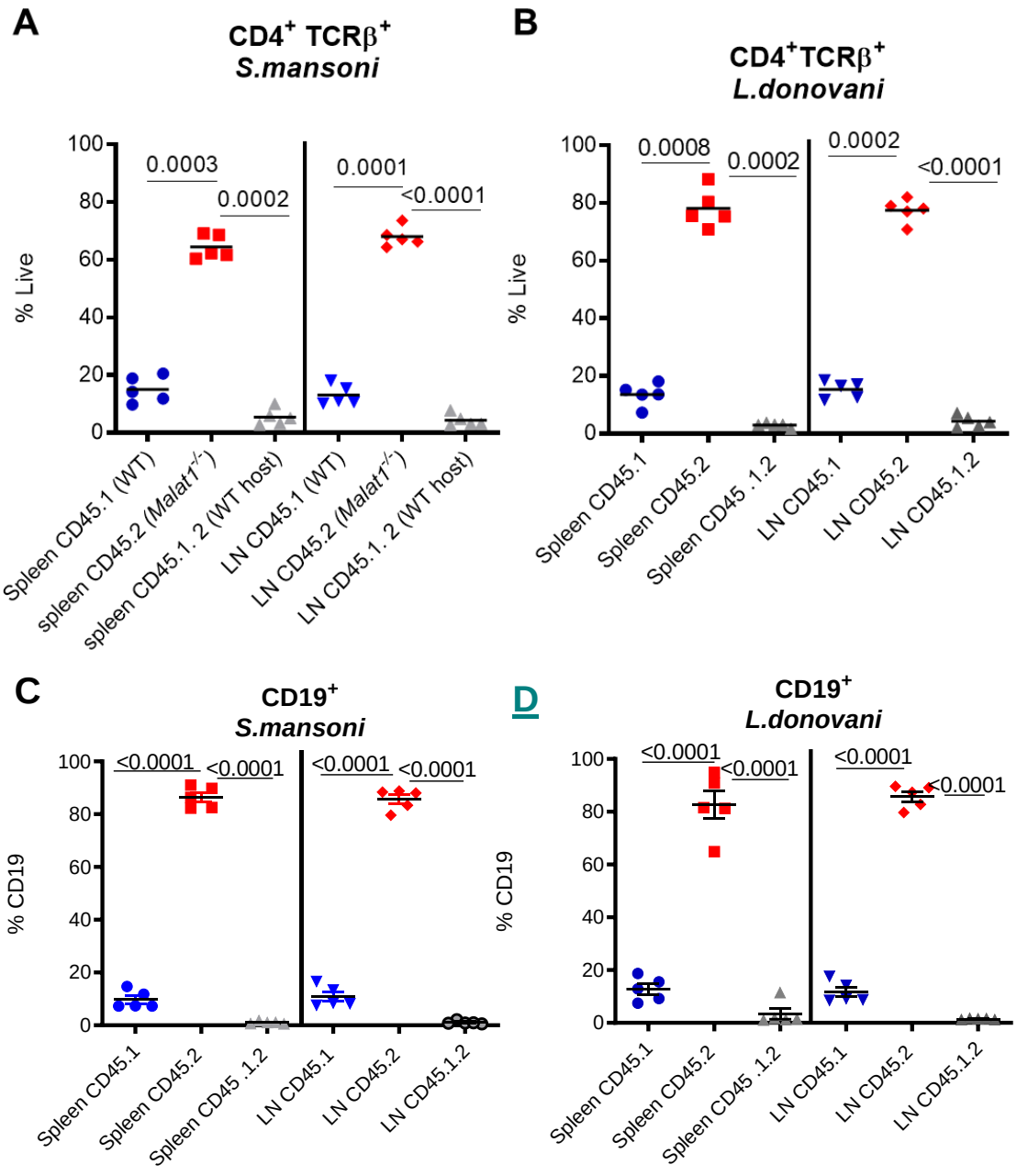


Figure 3.19 *Malat1*$^{-/-}$ CD4^{+} T cells preferentially repopulate lymph nodes and spleens in *L. donovani* and *S. mansoni* infected mixed bone marrow chimeric mice

Percentage of live CD4^{+} TCRβ^{+}cells in bone marrow chimeric mice infected with **A)** *S. mansoni*, or **B)** *L. donovani.* Cells which are WT CD45.1 (blue) or *Malat1*$^{-/-}$ CD45.2 (red) or host CD45.1.2 (grey) are shown.The proportion of CD19^{+} cells are shown in **C)** *S.mansoni* or **D)** *L.donovani* infection. Colours as above**.** The percentage of cells was determined by flow cytometry. n=5 biological replicates representative of one independent experiment. Data were analysed using a repeated measure one-way ANOVA with post hoc Tukey test. p value and the mean is displayed.

3.3 Discussion

Despite its high abundance and high degree of conservation across mammalian species, the physiological function of *Malat1* at the whole organism level has remained elusive. In this chapter and further supported by the work of Dr James Hewitson and colleagues, we demonstrated that *Malat1* is essential for the function and regulation of adaptive immunity by regulating $CD4^+$ T cell responses (Hewitson et al., 2020).

In this chapter, we began our studies by confirming that *Malat1* expression was diminished in *Malat1*$^{-/-}$ mice. We found that *Malat1* expression was reduced by >99%. However, low levels of *Malat1* were still detected by qRT-PCR. Of note, Nakagawa and colleagues profiled tissues of the homozygous *Malat1* knockout mice through northern blotting. This identified a diminished but specific *Malat1* band in numerous tissues including the brain (Nakagawa et al., 2012). The authors postulated that an internal promoter downstream of the TSS is activated in the *Malat1*$^{-/-}$ mice which drives *Malat1* expression at low levels. It is feasible that a similar mode of action has arisen in $CD4^+$ T cells which is responsible for the low levels of *Malat1* expression in the knockout mouse $CD4^+$ T cells. Yet, as *Malat1* expression is significantly diminished in this model we deemed it suitable for analysing $CD4^+$ T cell functions in the absence of *Malat1*.

We next determined that upon $CD4^+$ T cell activation *Malat1* was rapidly downregulated. This finding was further supported by a bulk RNA-seq study used to determine changes in lncRNA expression which took place after 24 hours of $CD4^+$ T cell activation (Hewitson et al., 2020). This study found that *Malat1* was within the top 2% of transcripts expressed in naïve $CD4^+$ T cells, additionally, *Malat1* was one of the most highly expressed lncRNAs that was downregulated after 24 hours of $CD4^+$

T cell activation and conserved between mammalian species. This suggested that rapid suppression of *Malat1* is a key feature and signature of early $CD4^+$ T cell activation.

It is possible that *Malat1* is not actively downregulated in the first 24 hours of $CD4^+$ T cell activation, instead, *Malat1* may be diluted due to the large number of transcriptomic changes that occur during the initial transcriptional burst in the early stages of $CD4^+$T cell activation (Pereira et al., 2017). However, this would require *Malat1* to be specifically excluded from the transcriptional burst. Many other lncRNAs are also up or down-regulated during the first 24 hours of differentiation including *Neat1* (Hewitson et al., 2020). Additionally, *Malat1* suppression is sustained throughout 6 days of *in vitro* Th1 or Th2 polarisation indicating active suppression or post-transcriptional degradation of *Malat1*, rather than dilution of the transcript. This suggested that suppression of *Malat1* in the early stages of $CD4^+$T cell activation might be required to enable Th cell differentiation.

We next made efforts to knockdown *Malat1* in $CD4^+$ T cells. Attempts to knockdown *Malat1* using siRNAs were unsuccessful. However, the use of GapmeRs significantly reduced *Malat1* expression in $CD4^+$ T cells. It is likely that as *Malat1* is a nuclear RNA it is more easily knocked-down in $CD4^+$ T cells when GapmeRs are used as this uses the RNAseH mechanism of degradation. RNAseH is detected in both the nucleus and the cytoplasm, whereas the RNA induced silencing complex (RISC) complex involved in siRNA degradation is cytoplasmic (Fluiter et al., 2009).

3.3.1 *Malat1* promotes the expression of the anti-inflammatory cytokine IL-10

One of the most striking observations to emerge from this data set is that loss of *Malat1* reduced expression of the anti-inflammatory cytokine IL-10 both *in vitro*

and *in vivo*. Work which occurred in parallel to this study by Dr James Hewitson demonstrated that *Malat1* loss resulted in reduced IL-10 expression in $CD4^+$ T cells in two different *in vivo* models of infection (Hewitson et al., 2020). Hewitson and colleagues found that loss of *Malat1* resulted in enhanced immunity and pathogen clearance upon infection with *L. donovani* and that loss of *Malat1* resulted in more severe immunopathology when knockout mice were infected with *Plasmodium chabaudi chabaudi* AS.

The differences in pathogen clearance and immunopathology between the two infections can be explained by considering the need for a balanced immune response during infection. An activated immune system must maintain a fine balance to ensure that a sufficient inflammatory response is generated to eradicate invading pathogens, but not so strong that it causes unnecessary tissue damage. The immune system prevents damage by producing anti-inflammatory cytokines. However, anti-inflammatory cytokines may not sufficiently control the pro-inflammatory environment of disease or overcompensate and hinder pathogen clearance. As such a dynamic balance of pro-inflammatory and anti-inflammatory molecules exists in the immune response. Consequently, the effect of a cytokine depends on the timing of expression and the local environment in which it acts with receptor density, competing cytokines, and tissue responsiveness influencing effectiveness.

IL-10 is an essential anti-inflammatory cytokine (Opal & DePalo, 2000). During infection, IL-10 has been shown to inhibit the activity of Th cells, $CD8^+$ T cells, NK cells and macrophages meaning that detrimental tissue damage as a result of pathogen destruction is prevented. However, given that successful implementation of the immune system is a fine balancing act between pathogen destruction and prevention of autoimmunity, the anti-inflammatory properties of IL-10 in cases can

hinder the effectiveness of pathogen clearance (Ouyang et al., 2011). Therefore, appropriate IL-10 production is key in ensuring a variety of pathogens are effectively destroyed by the immune system.

A robust immune response is essential for eliminating the parasites which cause malaria, as such, the immune response must be tightly regulated to prevent unnecessary tissue damage. Consequently, IL-10 plays an important role in determining the outcome of malaria infection by suppressing inflammation. IL-10 knockout mice which are infected with *Plasmodium chabaudi chabaudi AS* display enhanced disease pathology including loss of body weight and increased pro-inflammatory cytokine production as a result of excessive pro-inflammatory conditions (Butler et al., 2019). A reduction in IL-10 expression due to the loss of *Malat1* resulted in a similar response causing increased immunopathology. Crucially, the *Malat1*$^{-/-}$ mice showed more pronounced weight loss compared to WT controls and the experiment had to be terminated early. (Hewitson et al., 2020) .

In contrast, in visceral leishmaniasis, IL-10 production correlates with ongoing disease. This is because high IL-10 expression can reduce macrophage activation which is essential for parasite elimination. Of note, mice which lack IL-10 expression are highly resistant to *L. donovani* infection and display 10-fold lower parasite burdens and an increase in macrophage activation (inducible NO synthase expression (iNOS) expression) (Murphy et al., 2001). This is consistent with findings from Dr James Hewiston who found that upon *L. donovani* infection the *Malat1*$^{-/-}$ mice showed reduced IL-10 expression, which was accompanied by increased expression of iNOS expression in myeloid cells. This was particularly prominent in inflammatory monocytes ($CD11b^{+}$ /$CCR2^{+}$ /Ly-$6C^{high}$). This was complemented by enhanced

pathogen clearance and reduced splenomegaly in the *Malat1*$^{-/-}$ mice (Hewitson et al., 2020).

Initial experiments identified IL-10 as a cytokine synthesis inhibitory factor (CSIF) (Fiorentino et al., 1989). When Th2 clones were stimulated with the lectin concanavalin A (ConA) an inducer of TCR crosslinking and T cell activation, CSIF was discovered in the supernatant. Cell stimulations including CSIF showed its inhibitory effects on the production of Th1 cytokines including IFNγ (Fiorentino et al., 1989). Later cloning work resulted in the renaming of CSIF as IL-10 (Moore et al., 1990; Vieira et al., 1991). It is now known that IL-10 is produced by multiple populations of cells including macrophages, dendritic cells (DC), natural killer cells (NK), eosinophils, neutrophils, B cells, $CD8^+$ T cells, epithelial cells in addition to $CD4^+$ subsets Th1, Th2, Th17 and Treg in response to a single infection (Couper et al., 2008). Of note, Hewitson and colleagues found that dysregulation of IL-10 in *Malat1*$^{-/-}$appeared to $CD4^+$ T cell-specific, no differences were observed in IL-10 levels in B cells or monocytes during infection (Hewitson et al., 2020).

Additionally, knockdown of *Malat1* confirmed that the RNA transcript is responsible for the regulation of IL-10 expression. Some minor differences are observed between the knockout and knockdown models for example IFNγ is significantly reduced in the knockdown models while only non-significantly trends downwards in the knockout models, except under reduced polarising conditions. This could be because *Malat1* loss occurs before naïve $CD4^+$ T cell activation in the knockout models, whereas the knockdown of *Malat1* takes place during the differentiation process. Yet, both models provide compelling evidence for the role of *Malat1* in regulating IL-10.

3.3.2 Transcriptional regulation of IL-10

Several transcription factors are known to regulate IL-10 expression in $CD4^+$ T cells. As the *Il10* gene is well conserved between humans and mice, and there is a high degree of homology in transcription factor binding sites, IL-10 is regulated similarly in mice and humans (Iyer & Cheng, 2012). The *Il10* promoter region is characterised by a TATA box and CCAAT box in both mice and humans. Binding sites in the *Il10* promoter exist for many transcription factors, including as STATs, BHLHE40, NF-kβ, BLIMP-1 and MAF. MAF is the master transcriptional regulator of IL-10 in $CD4^+$ T cells and has been shown to be involved in a wide range of biological processes including, bone and lens development (Gabryšová et al., 2018). *In vivo* experiments of malaria, allergy and autoimmunity demonstrated that MAF was essential for IL-10 expression in Th1, Th2 and Th17 cells. The expression of MAF correlates with IL-10 expression in all subsets of $CD4^+$ T cells (Gabryšová et al., 2018).

In this chapter, we found that loss of *Malat1* reduced the expression of MAF in both Th1 and Th2 cells. This was further supported by the work of Dr James Hewitson who found that 25 transcription factors showed a significant correlation with *Malat1* in single-cell RNA-seq data derived from $CD4^+$ T cells of *PcAS-infected* mice (Hewitson et al., 2020). *Maf* was only one of three transcription factors to show a positive correlation with *Malat1* expression. In addition, *Maf* was one of only five genes which positively correlated with *Il-10* expression at the single-cell level (Hewitson et al., 2020). Of note, STAT4 is also known to regulate IL-10 expression. Dr James Hewitson found that *Stat4* levels were reduced in the *Malat1*$^{-/-}$ Th1 cells (Hewitson et al., 2020). In this chapter, we determined that STAT4 expression is only modestly reduced at the protein level in Th1 cells. In contrast, the expression of MAF

was more prominently reduced at the protein level in Th1 cells. Other transcription factors such as *Bhlhe40* were shown to anti-correlate with *Malat1* expression at the single-cell level. In addition, the RNA levels of additional transcription factors have been tested by a previous member of the laboratory group. Jonny Cope (Master's student University of York, York, UK), who found no change in B lymphocyte-induced maturation protein 1 (BLIMP1) expression in the absence of *Malat1 (*data not shown). Taken together, the data demonstrates that the loss of *Malat1* suppresses MAF which is a central transcriptional regulator of IL-10 in Th cells.

3.3.3 The role of *Malat1* in $CD4^+$ T cells

A growing number of studies have begun to examine the role of *Malat1* in $CD4^+$ T cells, with conflicting reports. Our work somewhat aligns with observations from Liang and Tang who demonstrated that overexpression of *Malat1* enhanced IL-10 expression (Liang & Tang, 2020). However, they found that overexpression of *Malat1* reduced IFNγ and T-bet levels (Th1 cells) but enhanced IL-4 and GATA3 (Th2 cells). The discrepancies in findings could be because this study focuses on the role of *Malat1* in human $CD4^+$ T cells rather than mice.

Yao and colleagues suggested that *Malat1* was not required for thymic T cell development nor T cell antiviral responses to acute LCMV infection (8 days) (Yao et al., 2018). Of note, IL-10 only alters susceptibility to chronic LCMV infection (Richter et al., 2013). Additionally, this paper prominently focused on the role of *Malat1* in $CD8^+$ rather than $CD4^+$ T cells as such perhaps the full extent of the role of *Malat1* in $CD4^+$ T cells was not explored in this context (Yao et al., 2018). In contrast, more recent work by Kanbar and colleagues found that $CD8^+$ T cells where *Malat1* was knocked-down using shRNA (CD45.1) or non-targeting shRNAs (CD45.1.2) and adoptively co-transferred in to CD45.2 recipient mice, after 7 days of LCMV infection,

the number of *Malat1* knockdown $CD8^+$ T cells was significantly reduced. This study also found that *Malat1* is essential for the differential of T cell effector memory cells during primary LCMV infection and repressed secondary T cell effector memory generation during infection rechallenge (Kanbar et al., 2022b). This study did not report on any potential changes in IL-10 expression in $CD8^+$ T cells in the absence of *Malat1*.

Masoumi and colleagues found that siRNA depletion of *Malat1* enhanced Th1 and Th17 polarisation (Masoumi et al., 2019). Yet, other studies which used *shMalat1* reduced the percentage of Th17 cells in *in vivo* models of acute viral myocarditis (AVMC) (Xue et al., 2022). This aligns with our findings which indicatied loss of *Malat1* impaired Th17 differentiation *in vi*tro. Despite reported variation in the role of *Malat1* in Th cells, we note that our data as our findings demonstrated consistent results in both *in vivo* and *in vitro* models.

Several studies have linked *Malat1*-dependent gene regulation with *Malat1*-miRNA interactions (Arun et al., 2020). This extends to IL-10, where *Malat1* has been implicated in IL-10 regulation in a miRNA dependent manner. This includes work by Liang and Tang who suggested that *Malat1* sponged miR-155, which altered the Th/Th2 balance of cells in asthma patients (Liang & Tang, 2020). Other work has suggested that *Malat1* induced tolerogenic dendritic cells through promoting dendritic cell-specific intercellular adhesion molecule-3 grabbing nonintegrein (DC-SIGN) by sponging miR-155, which in turne increased IL-10 expression (Wu et al., 2018). These studies are particularly interesting, as they allude to a role of *Malat1* in sponging miRNAs. It is unclear in these studies if this is attributed to the role of nuclear *Malat1* as miRNAs are enriched in the cytoplasm. It is feasible that *Malat1* sequesters these miRNAs in the nucleus during the early stages of pre-miRNA processing. A previous

member of the lab Edward Muscutt (University of York, York, UK), profiled a panel of ~20 miRNAs in *Malat1*$^{-/-}$ cells and found no difference in expression levels (data not shown). This indicates that the functional role of *Malat1* in CD4^{+} T cells may be independent of miRNAs.

3.3.4 Mixed bone marrow chimera

The mixed bone marrow chimera experiments revealed that *Malat1*$^{-/-}$ CD4^{+} T cells produce lower levels of IL-10 *in vivo* when infected with either *L. donovani* or *S. mansoni* than their WT counterparts (**Figure *3.17***). This suggested that *Malat1* plays a cell-intrinsic role in regulating IL-10. This finding is consistent with other *in vivo* models where the loss of *Malat1* reduced the expression of IL-10 in *in vivo* models of malaria and leishmaniasis (Hewitson et al., 2020).

Notbaly, we found that *Malat1*$^{-/-}$ CD4^{+} T cells appeared at a higher frequency in the mixed bone marrow chimera models (***Figure 3.19)***. This could be because these cells express lower levels of IL-10 which would confer an advantageous nature to the host when infected with *L. donovani* or *S. mansoni*. However, a caveat to using CD45.1 and CD45.2 as lineage markers is that CD45.1 cell have a slower rate of fully reconstituting the immune compartment compared to CD45.2 cells, due to their lower homing capabilities (Ferreira et al., 2019). Therefore, it is unlikely that the expansion of CD45.2 *Malat1*$^{-/-}$ cells are due to their advantageous role during infection. To gain confidence in this finding, the experiment could be repeated using WT mice from a CD45.2 background and *Malat1*$^{-/-}$ mice from a CD45.1 background.

3.3.5 Future work

In this chapter, we found that *Malat1* appeared to have limited *in-cis*-regulatory capabilities, which differed between Th cell subpopulations. This

difference could be explained by potential alterations in the local chromatin landscape between cell types However, we only examined its effects on its immediately neighbouring genes. Of note, *cis-acting* lncRNAs are capable of acting over long genomic distances (Gil & Ulitsky, 2019). *Xist* for example can exert its repressive effects across the whole X chromosome. *Cis-acting* lncRNAs that have a more contained effect can also function over large genomic distances, for example, the lncRNA *Peril* which is transcribed from a *Sox2* super-enhancer positively regulates the expression of two genes found in separate locations over ~1.5 Mb away (Groff et al., 2018). It is possible that the previously demonstrated *in-cis* capabilities of *Malat1* extend beyond that of regulation of *Neat1* and could regulate other genes along chromosome 19. We explore this possibility further in Chapter 4 where we utilise RNA-seq.

It is thought that *Malat1* regulates gene expression through interactions with RBPs. Engreitz and colleagues used RAP-RNA, which involves *in vivo* crosslinking followed by RNA capture and high throughput RNA sequencing to identify RNA-RNA interactions (Engreitz et al., 2014b). Different crosslinking strategies were utilised to identify different interactions - formaldehyde was to create RNA-protein and protein-protein interactions, 4’ aminomethyltrioxalen which creates crosslinks between uridine bases and does not interact with proteins, and finally formaldehyde plus disuccinimidyl glutarate a strong protein-protein crosslinker to identify RNAs linked indirectly through protein intermediates. They found that *Malat1* interacted directly with U1, but mainly interacted with nascent pre-mRNAs, snRNAs and mRNAs indirectly through protein intermediates. They expanded their analysis to RAP-DNA and found that *Malat1* interacted with thousands of sites of active transcription across the genome indirectly through protein and pre-mRNA

intermediates. RBPs have cell-type specific RNA targets and functions. This is further supported by data from West and colleagues who found that *Malat1* bound to active gene regions using CHART-seq. They also use CHART-MS to identify *Malat1* protein interaction partners and postulate that interactions with these proteins mediate *Malat1* interactions with chromatin (West et al., 2014a). Other lncRNAs such as *Xist* have been shown to function through a protein intermediate, for example, *Xist* mediates X chromosome inactivation through interactions with the polycomb repressive complex. We postulate that *Malat1* regulates $CD4^+$ T cell functions through a network of interactions between *Malat1* and RBPs. It is plausible that these *Malat1* RBP interactions are responsible for the regulation of Maf and IL-10 and other potential regulatory roles of *Malat1* in $CD4^+$ T cells. We explore the *Malat1* protein interactome in chapter 5.

3.3.6 Conclusions

Collectively, the findings in this chapter demonstrate that *Malat1* is a negative regulator of the immune system through promoting the expression of IL-10 which has functional implications in infectious diseases. It can be speculated that through *Malat1-dependent* regulation of Th cell function *Malat1* likely modulates immune function in numerous infections and inflammatory conditions.

4. *Malat1* dependent gene regulation in $CD4^+$ T cells

4.1 Introduction:

4.1.1 *Malat1* dependent regulation of the transcriptome:

Several studies have performed RNA-seq on cells which lack *Malat1* in different cell contexts. Notably, Zhang and colleagues generated a *Malat1*$^{-/-}$ mouse model and performed RNA-seq on WT and *Malat1*$^{-/-}$ brains and livers. Despite its high abundance loss of *Malat1* only resulted in differential gene expression of 10 genes in the liver and 5 genes in the brain. (Zhang et al., 2012). It is possible that as whole tissues were examined for gene expression changes, the heterogeneity of different cell populations within tissues may mask any effects of *Malat1* in specific areas of the brain. Nevertheless, this study indicated that *Malat1* has limited effects on the transcriptome in brains and livers.

Similarly, a recent study by Shaath and colleagues in MDA-MB-231 triple-negative breast cancer negative cells found that upon CRISPR/Cas9 deletion of *Malat1*, 32 genes were up-regulated and 129 genes were down-regulated in the absence of *Malat1* (Shaath et al., 2021). Pathway analysis revealed that these genes were involved in networks which involved IFNG, NFKB and TNF pathways Additionally, a separate study knocked-down *Malat1* using ASO oligos in A431 cells (epithelial cells from an epidermoid carcinoma patient) (Y. Zhang, Gao, et al., 2019). They found that 147 genes were differentially up-regulated and 195 genes were down-regulated upon loss of *Malat1*. These genes were involved in regulation of the cell cycle, and the endomembrane system. These studies collectively indicated that *Malat1* has a limited regulatory role at the transcriptome level.

However, in contrast to the studies which found that *Malat1* had limited transcriptome wide effects, others have found that loss of *Malat1* alters the expression

of thousands of genes. For example, one study knocked-down *Malat1* using ASOs in C3H10 cells and performed RNA-Seq. 1,120 genes were differentially expressed ($p<0.05$) with 403 up-regulated and 717 genes down-regulated (J. Han et al., 2021). Gene set enrichment analysis revealed that *Malat1* knockdown enriched several biological processes including lipid and fatty acid metabolism, fat cell differentiation and the PPAR signalling pathway. This suggested that *Malat1* played a role in adipogenesis.

Similarly, Arun and colleagues aimed to determine gene expression changes upon *Malat1* knockdown in tumour organoids and primary tumours (Arun et al., 2016). They found that ASO treatment of *Malat1* changed the expression of 478 genes. Many of these genes were involved in focal adhesion, and cell adhesion pathways. This highlighted the important role of *Malat1* in tumour progression and the potential role of *Malat1* in regulating gene expression changes.

Of note, another study investigated the role of *Malat1* in $CD8^+$ T cells (Kanbar et al., 2022). $CD8^+$ T cells were transduced with *Malat1* targeting (CD45.1) or non-targeting (CD45.1.2) shRNA. Cells were co-transferred into a CD45.2 mouse and infected with LCMV. After 7 days of infection, the cells were isolated using FACs and analysed by scRNA-seq. 1,486 genes were found to be differentially expressed in the absence of *Malat1*, these genes were distributed into 3 scRNA-seq clusters. The results suggested that *Malat1* typically reduced the expression of genes which promoted $CD8^+$ T cell memory.

A complex picture has emerged where cell type and potentially knockout or knockdown method showed profound differences in the effect of *Malat1* on gene expression changes. Therefore, it is challenging to predict if *Malat1* will have limited

or wide spread effects on transcription in $CD4^+$ T cells as such further investigation is required to determine if loss of *Malat1* would have a significant impact on gene expression levels.

In mice, the loss of *Malat1* has no overt phenotype (Eißmann et al., 2012; Nakagawa et al., 2012b; Zhang et al., 2012). However, we found that *Malat1* null mice display a more potent immune response in the context of infection (Hewitson et al., 2020). This phenomenon is a result of the lower expression of the anti-inflammatory cytokine IL-10 in $CD4^+$ T cells derived from *Malat1* null mice. Crucially, these experiments demonstrated that *Malat1* played a non-redundant role in mammalian immunity.

As *Malat1* is important for $CD4^+$ T cell function, understanding changes in the transcriptional landscape in the absence of *Malat1* could help explain the regulation of IL-10 and MAF, in addition to gaining insight into previously unknown roles of *Malat1*. To the best of our knowledge, at the time of writing no RNA-seq data set is available in the public domain which examines transcriptomic changes in the absence of *Malat1* in $CD4^+$ T cells.

4.1.3 Aims:

As we have previously shown that *Malat1* regulates IL-10 in $CD4^+$ T cells, we hypothesise that *Malat1* deletion results in changes across the $CD4^+$ T cell transcriptome.

By utilising NGS technology this chapter aims to

- Identify transcriptomic changes in different $CD4^+$ T cell populations in the presence of absence of *Malat1*
- Characterise potential pathways that *Malat1* regulates in $CD4^+$ T cells

4.2 Results:

4.2.1 Experimental design:

To deepen our molecular understanding of *Malat1* in Th cells, we isolated naïve $CD4^+$ T cells and generated in *vitro* polarised Th1 and Th2 cells from 4-5 independent mice of both WT and *Malat1*$^{-/}$ mice. Whole genome RNA-seq identified genes that are controlled by *Malat1* in the different cell states, as well as genes that change differently in comparison to WT during differentiation. Successful polarisation in WT cells was confirmed by flow cytometry and examination of the log count per million (LogCPM) of signature Th1 and Th2 genes in the different cell states (***Figure 4.1***). We chose these subsets of Th cells as we have previously observed changes in T cell function in both Th1 and Th2 cells *in vitro*. Our work previously has primarily focused on the role of *Malat1* in Th1 cells, but it is beneficial to include Th2 cells as a comparator to understand if *Malat1* plays a similar regulatory role across $CD4^+$ T cell subsets.

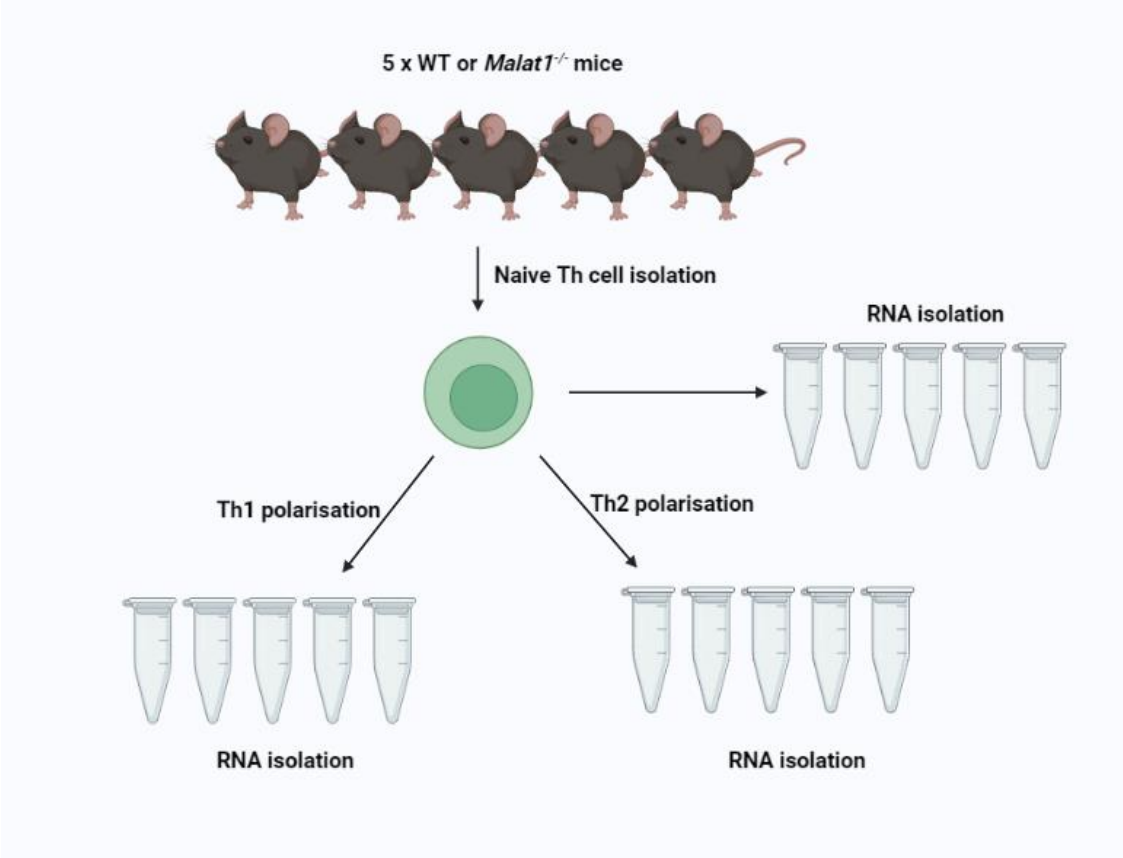


Figure 4.1: Schematic representation of RNA-seq experimental setup.

Lymph nodes and spleens were isolated from WT and *Malat1*$^{-/-}$ mice and pooled. Naïve CD4^{+} T cells were purified by FACs sorting and RNA was isolated. Naïve T cells were cultured in the presence of cytokines and antibodies to generate a Th1 or Th2 phenotype then RNA was extracted. 4-5 biological replicates were generated to create this dataset.

4.2.2 RNA integrity and quality analysis:

To ensure that undegraded and high-quality RNA was to be sent for RNA-seq analysis, RNA samples were first analysed using the Alignet BioAnalyser by the University of York Genomics Facility. Samples were mixed with a fluorescent dye and separated by size based on electrophoresis. The fluorescence was recorded and electropherograms created. An RNA integrity number (RIN) was generated by analysis of peaks across an electropherogram trace. The RIN algorithm compares 100 features of the electrophoretic trace including the area of ribosomal bands, the height of the 18s peak, and the ratio of the fast area of the electropherogram to the total area of the electropherogram (Imbeaud et al., 2005). A RIN value between 1 and 10 is generated, with 10 representing the least degraded RNA. An rRNA value was also generated which compares the ratio of 28s to 18s rRNA, with a ratio of above 1.8 considered to be an undegraded RNA sample. Here, Bioanalyzer results indicated high-quality RNA was isolated for all samples barring WT Th2 sample 2 which unfortunately did not survive culture conditions (**Figure *4.2***). All other samples were sent to the Wellcome Sanger Institute (Hinxton, England) to be processed for RNA-seq analysis. RNA-seq, read mapping and initial quantification was performed by Dr Kylie James (Garvan Institute of Medical Research, Darlinghurst, Australia).

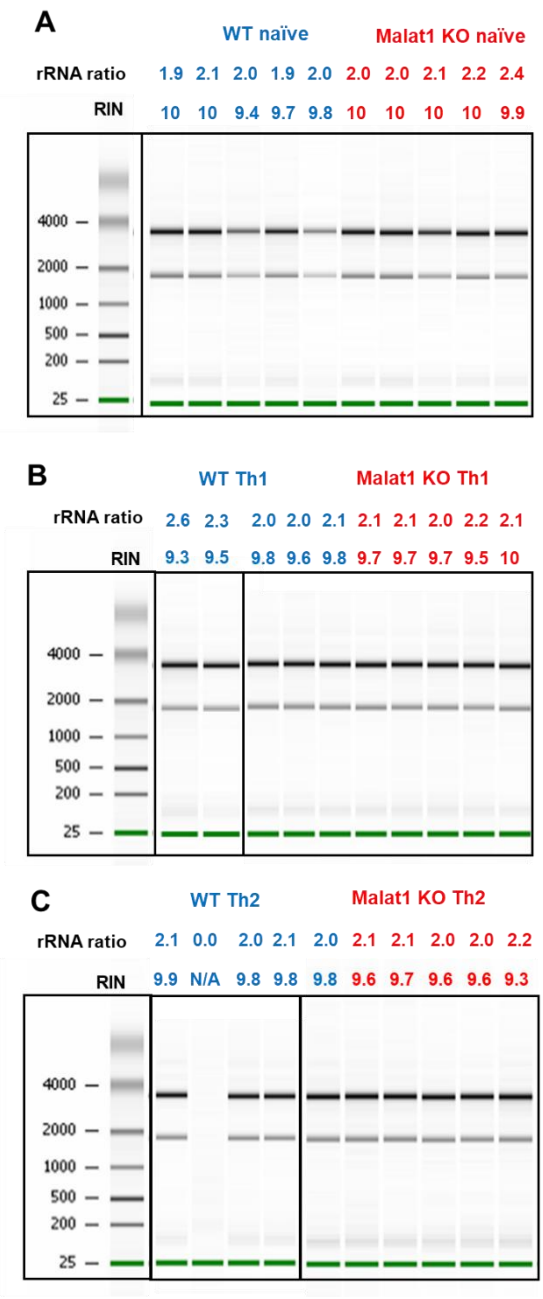


Figure 4.2: Assessment of RNA integrity and quality.

Aligent Bioanalyzer results of all RNA samples including RIN and rRNA ratio. Each lane represents an individual sample.

4.2.3 Confirmation of Th cell differentiation:

Initial analysis of RNA-seq data sets confirmed Th cell differentiation by assessing changes in signature Th1 and Th2 genes. Expression of Th1 genes such as *Tbx21* and *Ifng* are strongly up-regulated in Th1 cells, contrastingly expression of Th2 genes such as *Il-4* and *Gata3,* are more strongly up-regulated in Th2 than Th1 polarised cells (***Figure 4.3***). As the samples expressed genes in a manner that was expected, the data was then further analysed to assess the role of *Malat1* in $CD4^+$ T cell function.

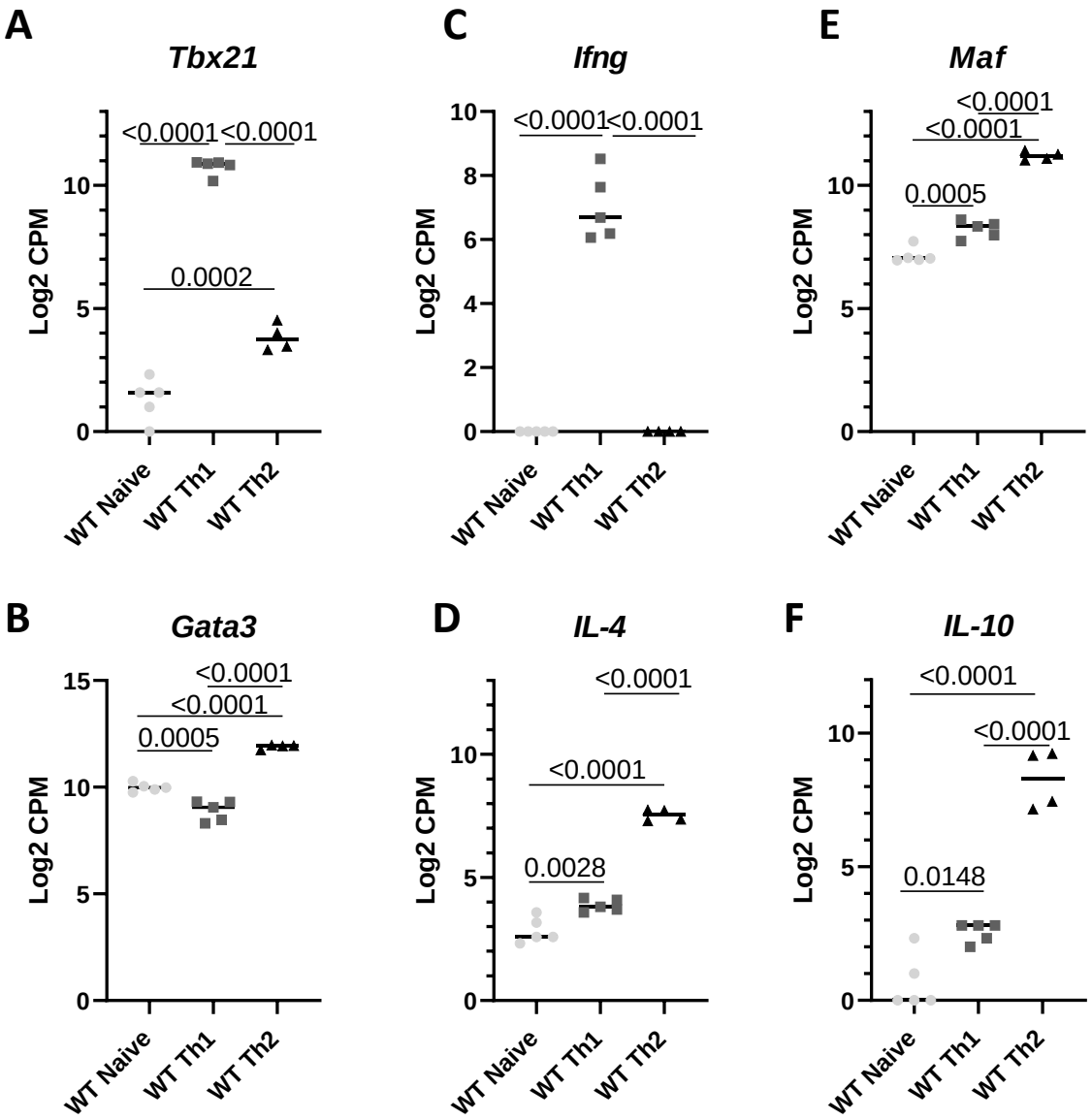


Figure 4.3: Expression of signature genes in *in vitro* polarised Th1 and Th2 cells.

Log_2CPM of signature Th1 and Th2 genes in WT naïve, Th1 or Th2 cells. **A)** *Tbx21* **B)** *Gata3* **C)** *Ifng* **D)** *Il-4* **E)** *Maf* **F)** *IL-10***.** Determined by RNA-seq. n= 5 naïve, n=5 Th1, n= 4 Th2 all samples are biological replicates. Data were analysed using a One-way ANOVA with post-hoc Tukeys test. p value is displayed where significant, the mean is depicted.

4.2.4 Loss of *Malat1* prominently alters gene expression in Th2 cells

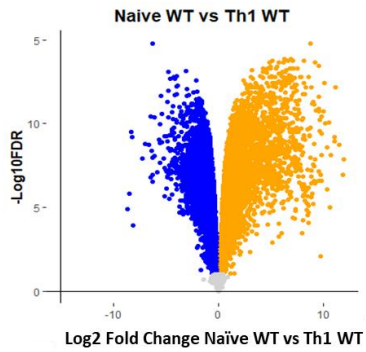
Next, differential gene expression was examined during $CD4^+$ T cell differentiation. We found that thousands of genes changed during the differentiation process. During WT Th1 differentiation at an FDR adjusted P value (q value) of <0.1, 9,365 genes were differentially expressed with a similar number of genes either up or down-regulated (***Figure 4.4A***). Similarly, in WT Th2 differentiation at a q <0.1 8,744 genes were differentially regulated with 4,225 genes up-regulated and 4,519 genes down-regulated (***Figure 4.4B***).

When comparing genes that were differentially expressed during *Malat1*$^{-/-}$ Th1 differentiation 7,970 genes were differentially regulated, and 4,081 genes were up-regulated and 3,869 genes were down-regulated (***Figure 4.4C***).

This is ~ 85% of the gene expression changes which occurred in WT Th1 differentiation. During *Malat1*$^{-/-}$ Th2 differentiation at a q <0.1 7,392 genes were differentially expressed with a similar number of genes up or down-regulated (***Figure 4.4D***). When compared to WT Th2 differentiation the number of genes which were differentially expressed in *Malat1*$^{-/-}$ Th2 differentiation is~84%.

We next analysed the chromosome location of genes which are differentially expressed during $CD4^+$ T cell differentiation (***Figure 4.5)***. A similar number of genes changed on each chromosome when normalised to chromosome length in each of the $CD4^+$T cell differentiation states (WT Th1, WT Th2, *Malat1*$^{-/-}$ Th1, *Malat1*$^{-/-}$ Th2.) When normalised to the chromosome length, the X chromosome tended to have the fewest number of gene expression changes, and chromosome 11 tended to have the highest number of gene expression changes.

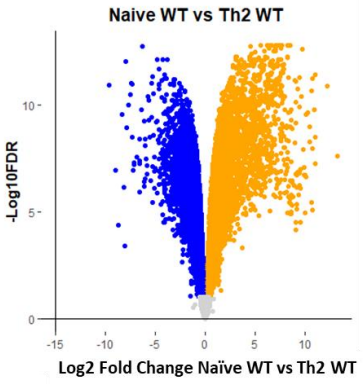
A



Sig	Total	Up	Down
FDR <0.1	10,522	5,173	5,349
FDR<0.05	10,174	5,005	5,169
FDR <0.01	9,365	4,618	4,747

Down in Th1 Up in Th1

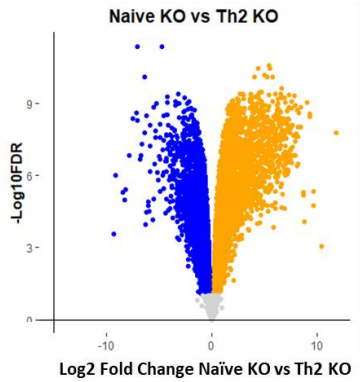
B



Sig	Total	Up	Down
FDR <0.1	10,036	4,904	5,132
FDR<0.05	9,620	4,681	4,939
FDR <0.01	8,774	4,255	4,519

Down in Th2 Up in Th2

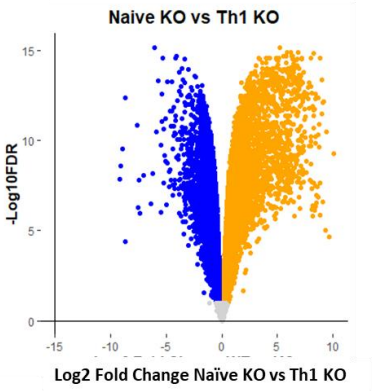
C



Sig	Total	Up	Down
FDR <0.1	9,136	4,518	4,614
FDR<0.05	8,600	4,211	4,389
FDR <0.01	7,392	3,548	3,844

Down in Th2 Up in Th2

D



Sig	Total	Up	Down
FDR <0.1	9,058	4,593	4,465
FDR<0.05	8,729	4,437	4,292
FDR <0.01	7,970	4,081	3,869

Down in Th1 Up in Th1

Figure 4.4: Similar numbers of genes significantly change upon Th cell differentiation in the presence or absence of *Malat1.*

Volcano plots comparing genome-wide gene expression changes upon Th cell differentiation, in addition to a table depicting the total number of genes which changed (ie up-regulated or down-regulated) at different false discovery rate (FDR) thresholds q<0.01, q<0.05 and q <0.1. **A)** Gene expression changes from WT Naïve to WT Th1 cells **B)** Gene expression changes from *Malat1*$^{-/-}$ Naïve to *Malat1*$^{-/-}$ Th1 cells **C)** Gene expression changes from WT Naïve to WT Th2 cells **D)** Gene expression changes from *Malat1*$^{-/-}$ Naïve to *Malat1*$^{-/-}$ Th1 cells. Blue represents down-regulated genes (q <0.05), gold represents up-regulated genes (q <0.05), grey (q >0.05).

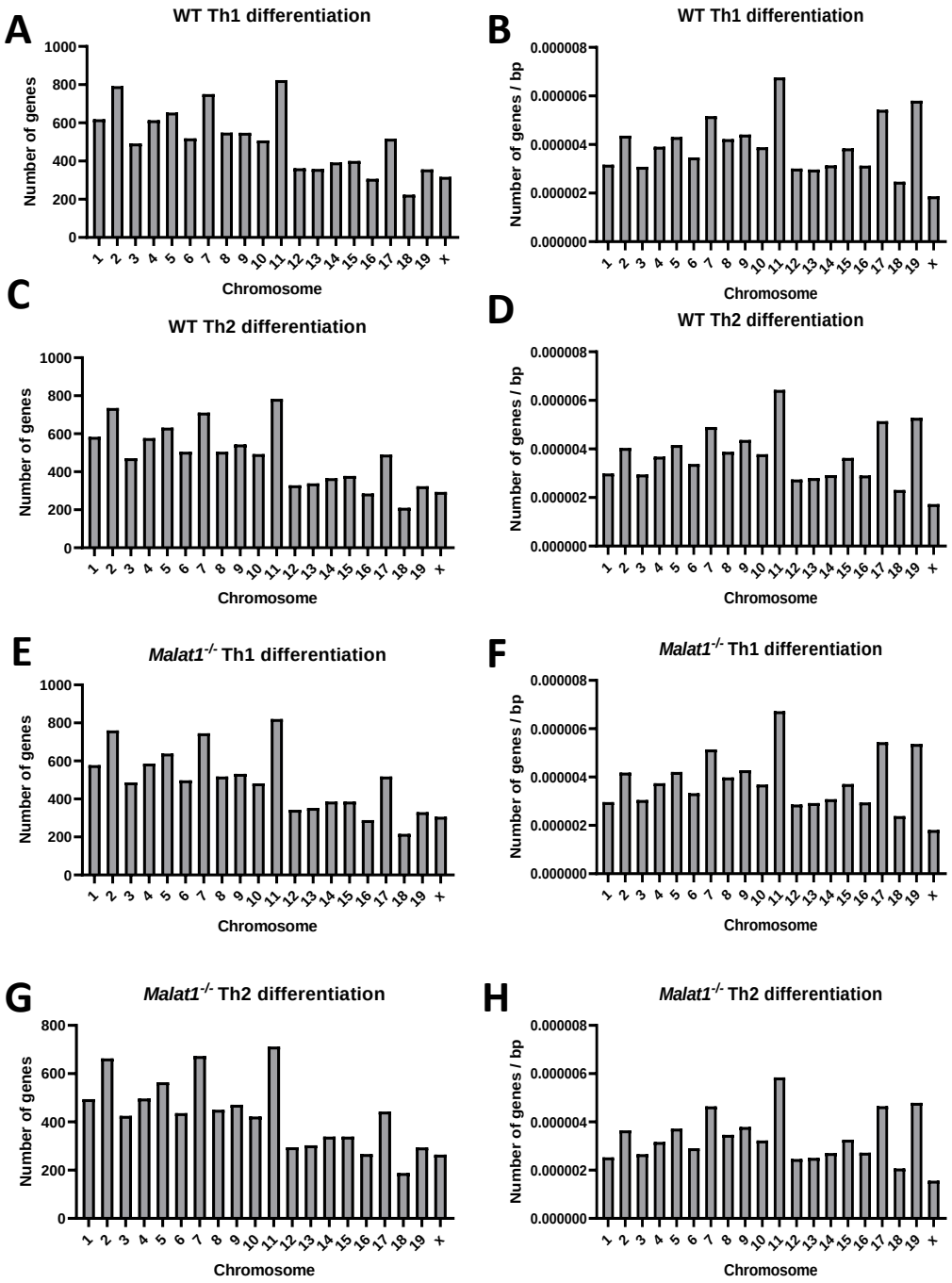
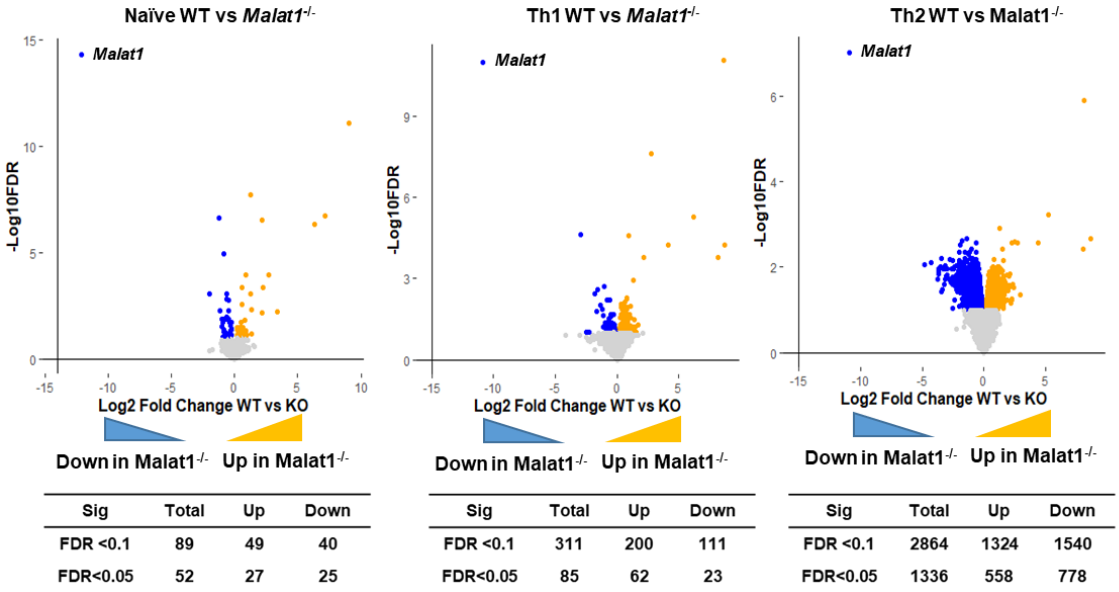


Figure 4.5: Differentially expressed genes are distributed across the genome during CD4$^+$ T cell differentiation.

Chromosome location of differentially expressed genes which change during CD4$^+$ T cell differentiation in **A)** WT Th1 differentiation (expressed as the total number of genes), **B)** WT Th1 differentiation with the number of differentially expressed genes normalised to chromosome length. **C)** As in A in WT Th2 differentiation **D)** As in B in WT Th2 cells **E)** As in A in KO Th1 differentiation **F)** As in B in KO Th1 differentiation **G)** As in A in *Malat1*$^{-/-}$ Th2 differentiation **H)** As in B in *Malat1*$^{-/-}$ Th2 differentiation.

We next compared genes which are differentially expressed between WT and *Malat1*$^{-/-}$ cells in different T cell states at a q <0.1. We identified 89 genes that were controlled by *Malat1* in naïve T cells, with 49 genes that were differentially up-regulated and 40 genes which were differentially down-regulated (***Figure 4.6***A). 311 genes were controlled by *Malat1* in and Th1 cells with 200 genes differentially up-regulated and 111 genes down-regulated (***Figure 4.6***B). 2,864 genes were found to be differentially expressed in the absence of *Malat1* in Th2 cells with 1,324 genes differentially up-regulated and 1,540 genes differentially down-regulated (***Figure 4.6***C)**.**

lncRNAs have been reported to have *in cis* effects in CD4^{+} T cells i.e., act in a manner dependent on the location of the site of their own transcription with varying distances from the site of transcription (West & Lagos, 2019). To determine if *Malat1* regulated genes that are local or distal to its site of transcription we next examined the chromosome distribution of genes which were differentially expressed in the absence of *Malat1*. Notably, we found that *Malat1* prominently affected gene expression of genes derived from chromosome 19 in naïve, Th1 and Th2 cells (***Figure 4.7***). This was particularly evident when the number of differentially expressed genes were normalised to chromosome length. This is a particularly interesting finding as *Malat1* is also located on chromosome 19 in mice and suggested that *Malat1* was acting *in cis* to regulate gene expression changes.



Sig	Total	Up	Down
FDR <0.1	89	49	40
FDR<0.05	52	27	25

Sig	Total	Up	Down
FDR <0.1	311	200	111
FDR<0.05	85	62	23

Sig	Total	Up	Down
FDR <0.1	2864	1324	1540
FDR<0.05	1336	558	778

Figure 4.6: Gene expression is strongly altered in Th2 cells in the absence of *Malat1.*

Volcano plots comparing genome-wide gene expression changes upon Th cell differentiation, in addition to a table depicting the total number of genes changed (up or down) at different FDR thresholds - q <0.1 or q <0.05 **A)** Gene expression changes from WT Naïve to *Malat1*$^{-/-}$ naïve cells **B)** Gene expression changes between WT Th1 cells and *Malat1*$^{-/-}$ Th1 cells **C)** Gene expression changes between WT Th2 cells and *Malat1*$^{-/-}$ Th2 cells. Blue represents down-regulated genes (FDR <0.05), gold represents up-regulated genes (q <0.05), grey (q >0.05).

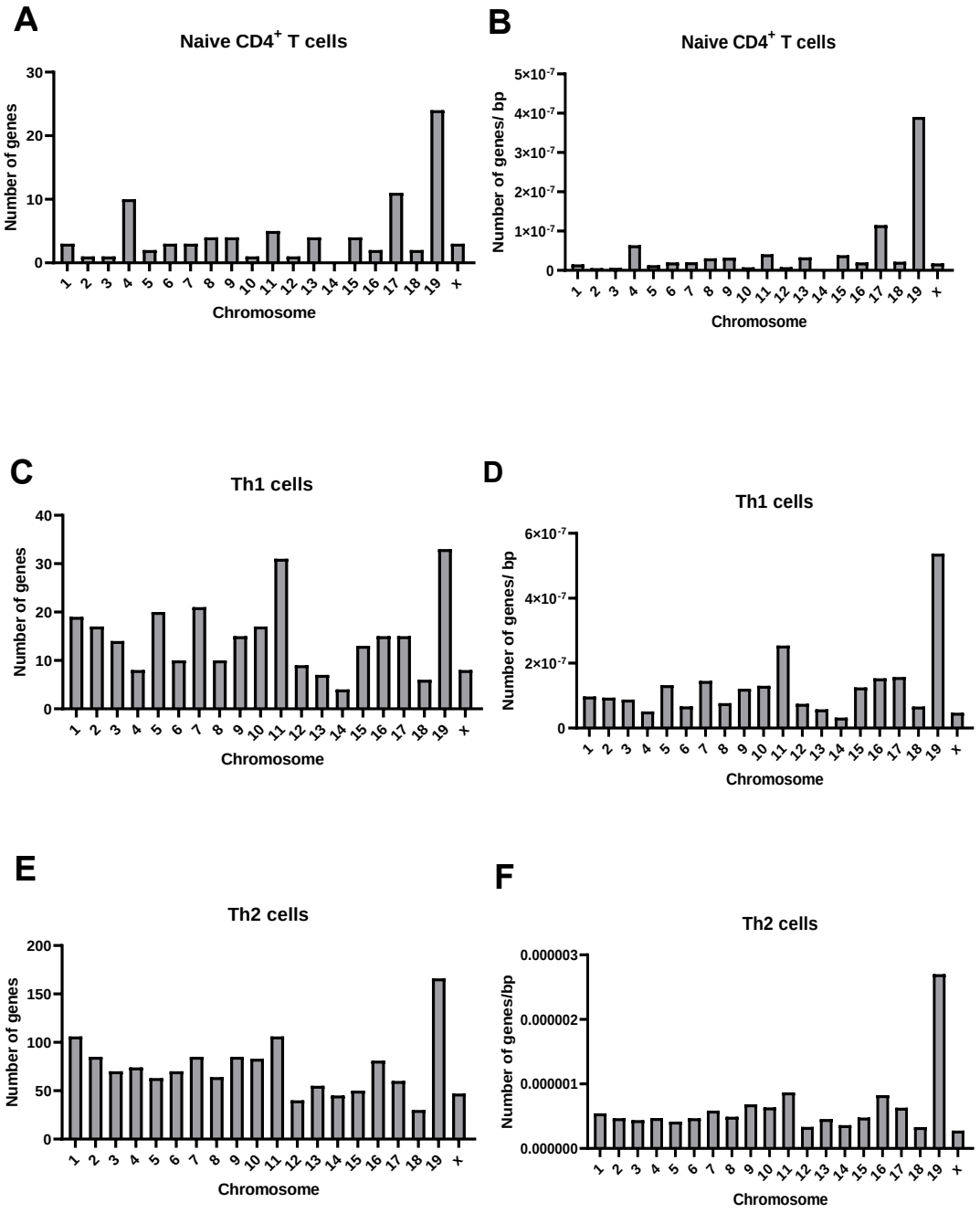


Figure 4.7: Loss of *Malat1* causes prominent differential gene expression changes on chromosome 19.

Chromosome location of differentially expressed genes which change in the absence of *Malat1* in different CD4$^+$ T cell populations **A)** WT naïve vs *Malat1*$^{-/-}$ naive cells (expressed as a total number of genes), **B)** WT naïve vs *Malat1*$^{-/-}$ naive cells the number of differentially expressed genes normalised to chromosome length. **C)** As in A comparing WT Th1 with *Malat1*$^{-/-}$ Th1 cells **D)** As in B comparing WT Th1 with *Malat1*$^{-/-}$ Th1 cells **E)** As in A comparing WT Th2 with *Malat1*$^{-/-}$ Th2 cells **F)** As in B comparing WT Th2 with *Malat1*$^{-/-}$ Th2 cells.

To identify the functional role of *Malat1* targets, we carried out gene set enrichment analysis (GSEA). Hallmark gene sets represent well-defined biological states or processes, gene ontology gene sets represent genes which are derived from specific biological processes, cellular compartments or molecular functions. The top ten gene sets with an FDR q value of <0.05 were analysed. Hallmark GSEA analysis revealed that in naïve cells *Malat1* up-regulated genes involved in the mTOR signalling pathway and those involved in hypoxia (***Figure 4.8***A). Similarly, genes involved in the hypoxia pathway are also down-regulated in the absence of *Malat1*, *i*n addition to IL-2 STAT5 signalling pathway genes and the TFNα signalling pathway in naïve cells (***Figure 4.8***A). Gene ontology GSEA analysis revealed that *Malat1* up-regulated genes in the vacuole and MHC protein complex, and down-regulated genes involved in DNA binding in naïve cells (***Figure 4.8***B). As only a handful of genes are differentially regulated by *Malat1* in naïve $CD4^+$ T cells a limited number of gene sets were identified.

Next, we explored the role of *Malat1* targets in Th1 cells. Hallmark GSEA analysis revealed that *Malat1* up-regulated genes involved in the IFNα, IFNϒ, inflammatory and apoptosis responses ***(Figure 4.9A).*** In contrast *Malat1* down-regulated hallmark genes involved in the peroxisome (enzymes involved in metabolism), P53 signalling and TNFα signalling pathway amongst others in Th1 cells ***(Figure 4.9A*)*.*** Upon examination of GO GSEA sets *Malat1* up-regulated genes with hydrolase activity, catalytic activity and immune response pathway gene sets. *Malat1* also down-regulated genes which are involved in apoptosis and regulation of signalling and metabolic processes (***Figure 4.9B).***

Examination of hallmark GSEA pathways that are up-regulated by *Malat1* in Th2 cells identified genes involved in the IFNα, UV and oestrogen responses (***Figure***

4.10A) Similar to Th1 and naïve cells genes involved in hallmark hypoxia were dysregulated. IL-2 STAT5 and glycolysis signalling were also down-regulated in the absence of *Malat1* (***Figure 4.10***A). Analysis of GO gene sets, identified genes involved in the regulation of cellular synthetic processes, DNA binding and the immune response to be up-regulated by *Malat1*. Those involved in apoptosis and regulation of cell death were down-regulated (***Figure 4.10***B).

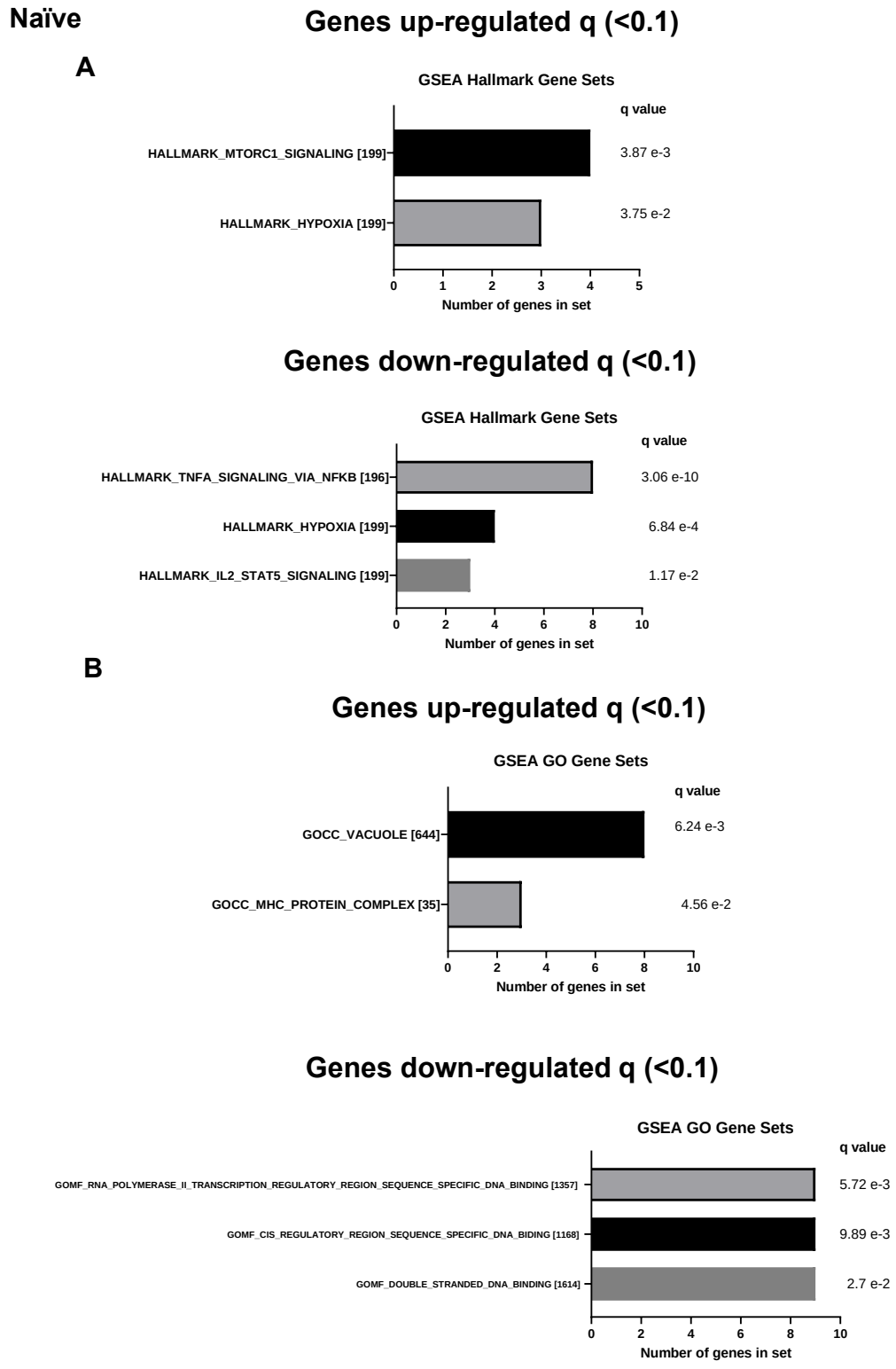


Figure 4.8: *Malat1* regulates the hypoxia pathway in naive CD4⁺ T cells

A) GSEA hallmark set analysis of genes which are either up or down-regulated by *Malat1* in naïve CD4⁺ T cells **B)** GSEA GO set analysis of genes which are either up or down-regulated by *Malat1* in naïve CD4⁺ T cells. A threshold of q <0.1 was used.

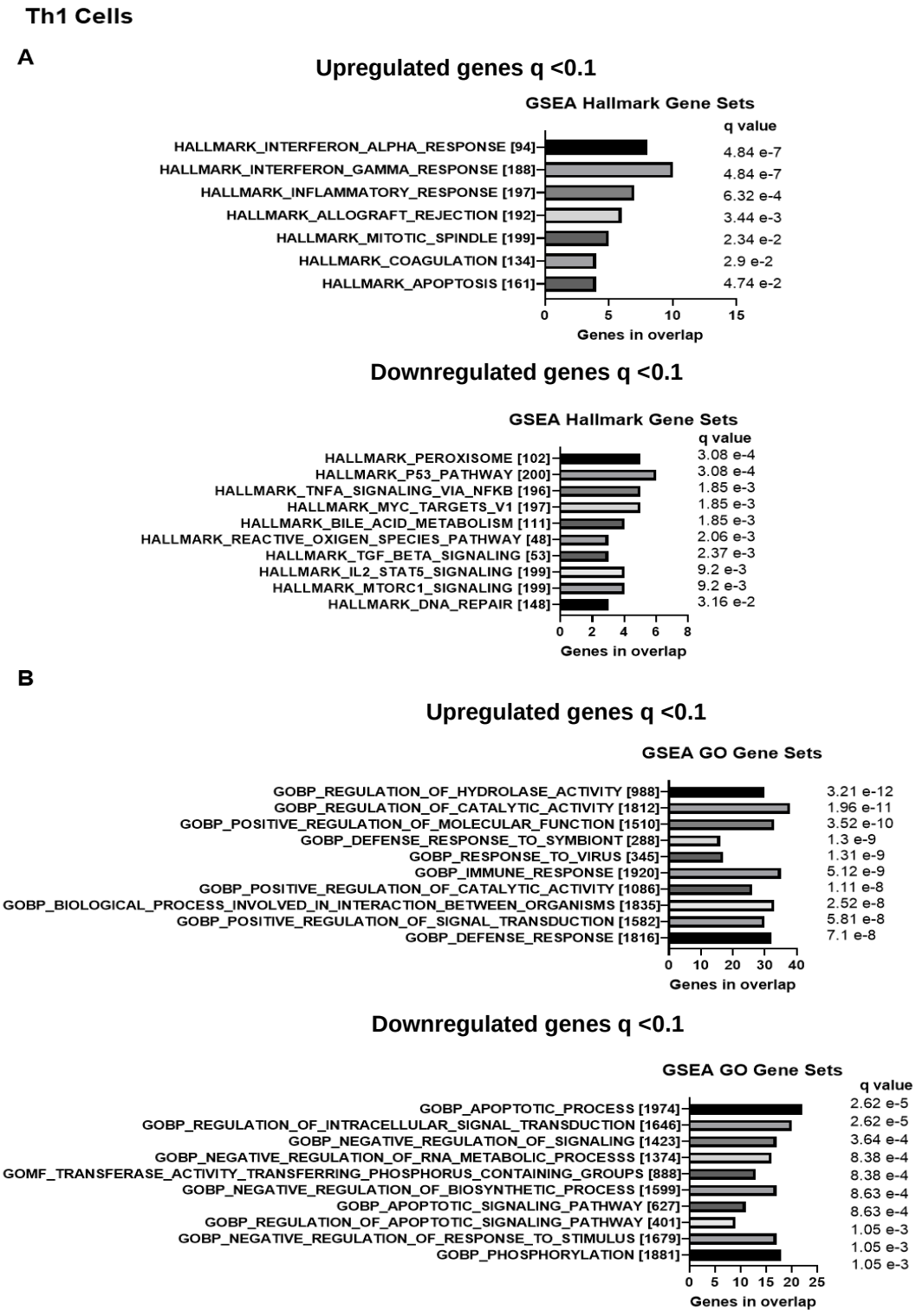
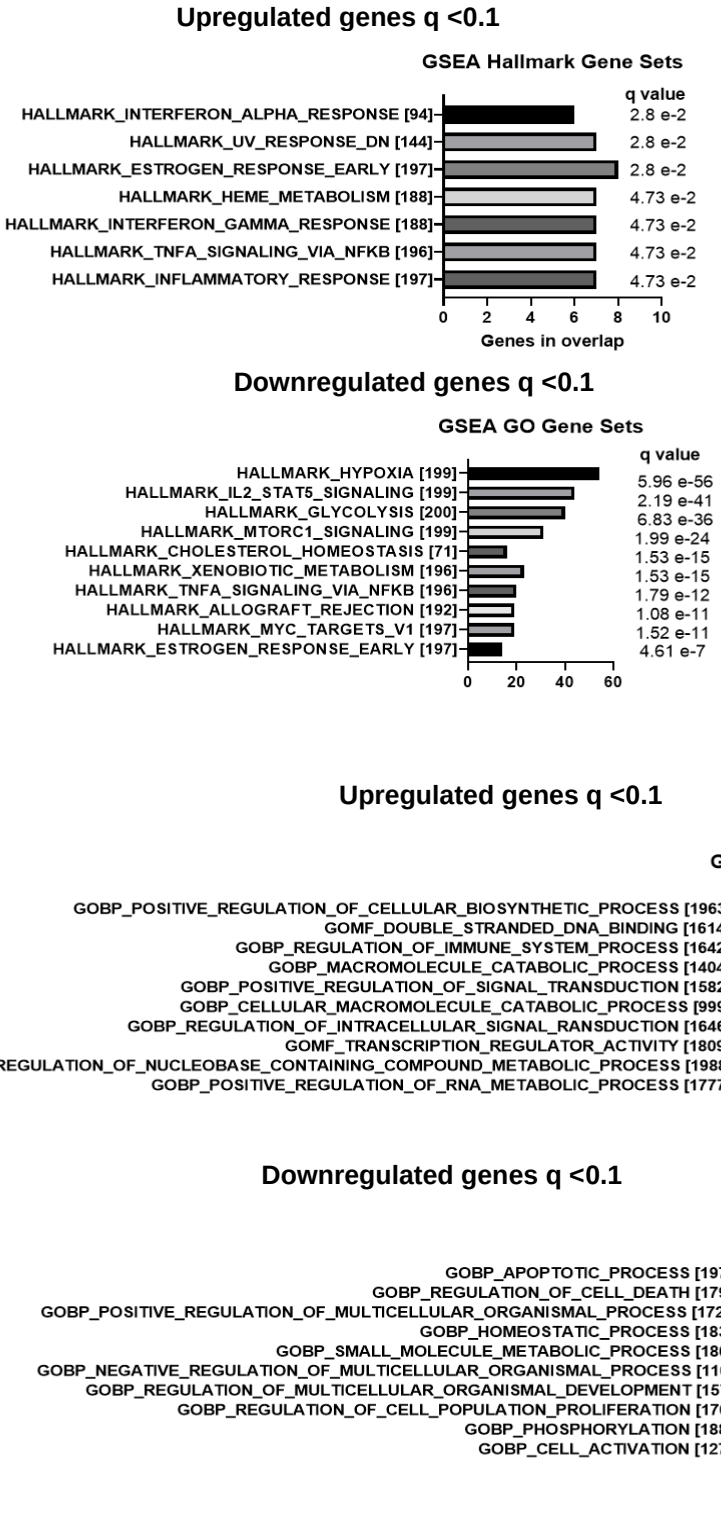


Figure 4.9: *Malat1* regulates genes involved in immune response pathways in Th1 cells

A) GSEA hallmark set analysis of genes which are either up or down-regulated by *Malat1* in Th1 cells **B)** GSEA GO set analysis of genes which are either up or down-regulated by *Malat1* in Th1 cells. A threshold of q <0.1 was used.

Th2 Cells

A



B

Figure 4.10: *Malat1* regulates immune pathways in Th2 cells.

A) GSEA hallmark set analysis of genes which are either up or down-regulated by *Malat1* in Th2 cells **B)** GSEA GO set analysis of genes which are either up or down-regulated by *Malat1* in Th2 cells. The top 100 genes were used

To gain insight into cell-specific functions of *Malat1* GSEA was used to compare pathways controlled by *Malat1* between naïve, Th1 and Th2 cells. To enable easier comparison total gene expression changes were examined this was because gene sets may not be regulated in the same manner (up-regulated or down-regulated) between cell types. Broadly speaking gene sets which were enriched in Th1, Th2 and naïve cells that are controlled by *Malat1*, surround functions such as metabolism, RNA/protein binding and regulation of cell death (***Figure 4.11***). Three gene sets were found to be enriched in both Th1 and Th2 cells, including regulation of the apoptotic process, regulation of identical protein binding and positive regulation of signalling. No gene sets were shared between naïve and Th1 cells. Only one gene set overlapped between naïve and Th2 cells this was the small molecule metabolic process gene set.

Unique gene sets were identified for each of the cell populations (***Figure 4.11***D), with the most significant gene sets being enriched in Th2 cells. This is not a particularly unexpected finding, given the greater number of genes which are altered in this cell type. However, no exact gene sets overlapped between all three cell types indicating potential differences in function between cell types.

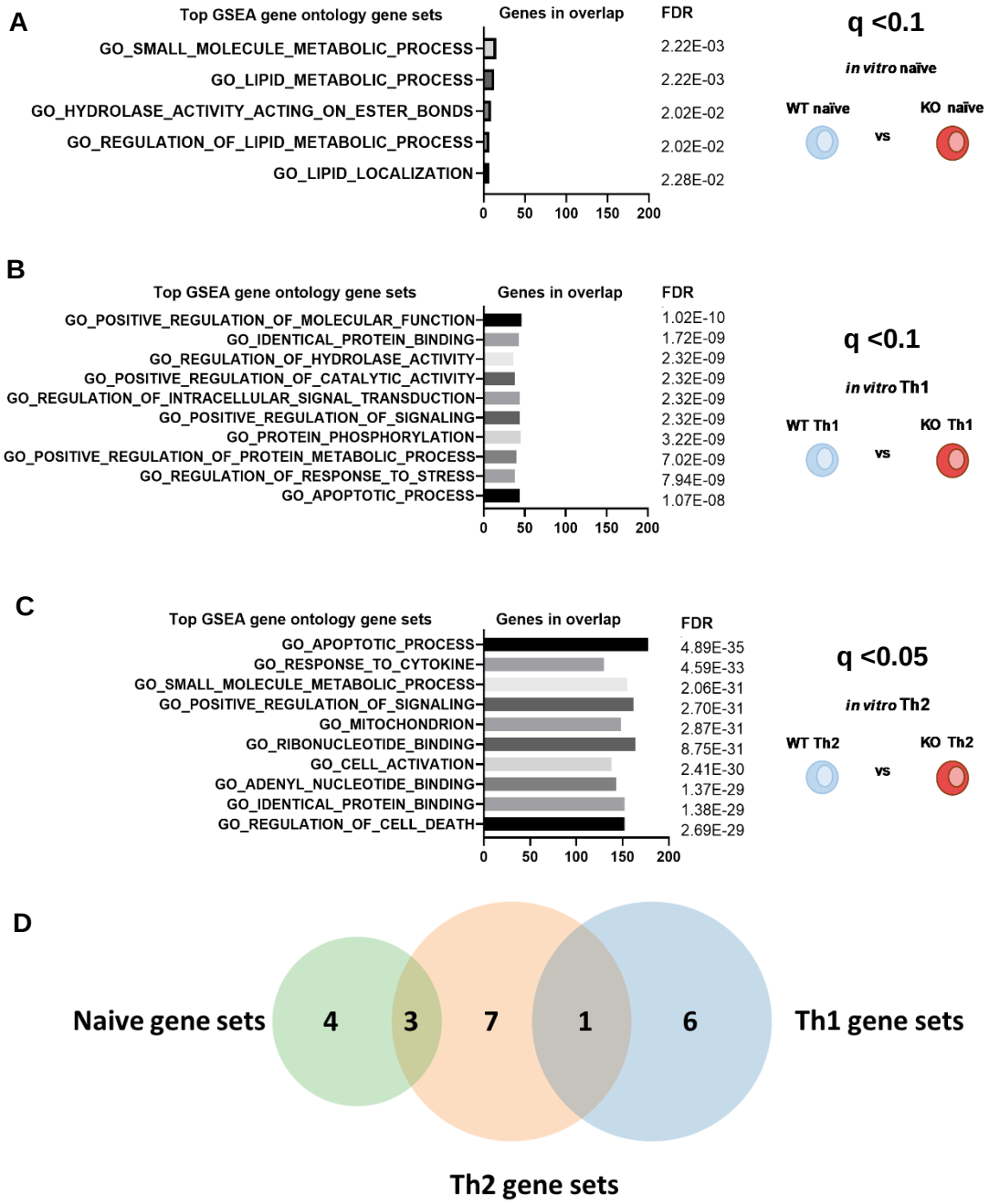


Figure 4.11: GSEA of genes controlled by *Malat1* in Th1, Th2 and naïve cells.

Genes which are significantly differentially expressed upon loss of *Malat1* in Th1, Th2 and naïve cells were analysed by GSEA. The top 10 gene ontology gene sets are displayed in addition to the significance value q value for **A)** Genes controlled by *Malat1* in naïve $CD4^+$ T cells with a q threshold of <0.1 **B)** Genes controlled by *Malat1* in Th1 cells with a q threshold of <0.1 **C)** Genes controlled by *Malat1* in Th2 cells with a q threshold of <0.05 top 200 genes **D)** Venn diagram depicting the overlap of gene ontology sets between the different cell types.

To gain further insight into core *Malat1*-dependent gene expression changes, differentially regulated genes were compared between datasets (***Figure 4.12***). Here, 13 genes are commonly controlled by *Malat1* in Th1, Th2 and naïve $CD4^+$ T cells. Of note, two of the most significant and highly up-regulated genes were transmembrane protein 181b pseudo gene (*Tmem181b-ps)* and Dynien light chain Tctex-type 1b (*Dynlt1b).* These genes are located next to one another in the genome, indicating that the strong upregulation of the pseudo gene *Tmem181b-ps* could be a possible off-target effect of the regulation of *Dynlt1b*. A likely possibility is that the creation of the *Malat1*$^{-/-}$ mice has altered this particular region of the genome.

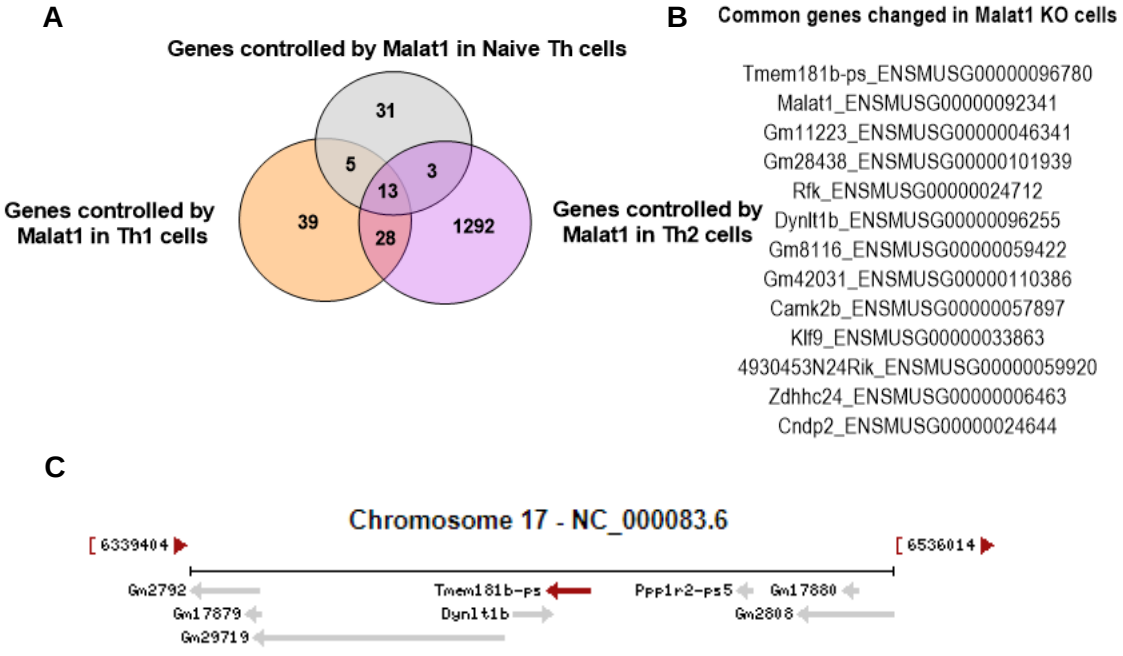


Figure 4.12: Some genes are commonly regulated by *Malat1* in CD4$^+$ T cells.

Genes which are significantly changed in *Malat1* KO cells q <0.05 and overlap are depicted in **A)** Venn diagram **B)** List format **C)** Screenshot of genomic localisation of *Tmem181b-ps* and *Dynlt1b* taken from the ncbi webpages.

4.2.4 Loss of *Malat1* impairs CD4$^+$ T cell differentiation:

Our initial analyses focused on genes which were controlled by *Malat1* in either undifferentiated or terminally differentiated states. To uncover if *Malat1*-regulated genes which changed during Th cell differentiation, genes which were controlled by *Malat1* in a terminally differentiated state (WT vs *Malat1*$^{-/-}$ Th1 or Th2) were compared to genes which are differentially expressed during WT Th1 or Th2 differentiation. Interestingly, most genes controlled by *Malat1* in either Th1 or Th2 cells overlaped with genes that were important for normal Th cell differentiation (***Figure 4.13*** A/C). This suggested that *Malat1* may help regulate CD4$^+$ T cell differentiation. When the log fold change (logFC) of genes controlled by *Malat1* in terminally differentiated cells was compared to the logFC of genes that were important in Th differentiation a striking effect was observed (***Figure 4.13*** B/D). Most genes which were up-regulated in Th1 differentiation were down-regulated upon loss on *Malat1*. Similarly, most genes which were down-regulated upon Th differentiation were up-regulated in *Malat1*$^{-/-}$ Th cells. This effect was particularly prominent in Th2 cells as more genes are controlled by *Malat1* in this cell type. Together this data suggested that *Malat1* KO cells have a differentiation defect as genes which are either up or down-regulated during WT differentiation are not up or down-regulated to the same extent in the absence of *Malat1*, as represented in a schematic (***Figure 4.13***E/F). Interestingly, genes which were differentially expressed in the absence of *Malat1* in Th1 cells, that are also differentially expressed during WT Th1 differentiation, mostly overlap with genes which are differentially expressed in the absence of *Malat1* in Th2 cells and that are differentially expressed during in WT Th2 differentiation (***Figure 4.14***).

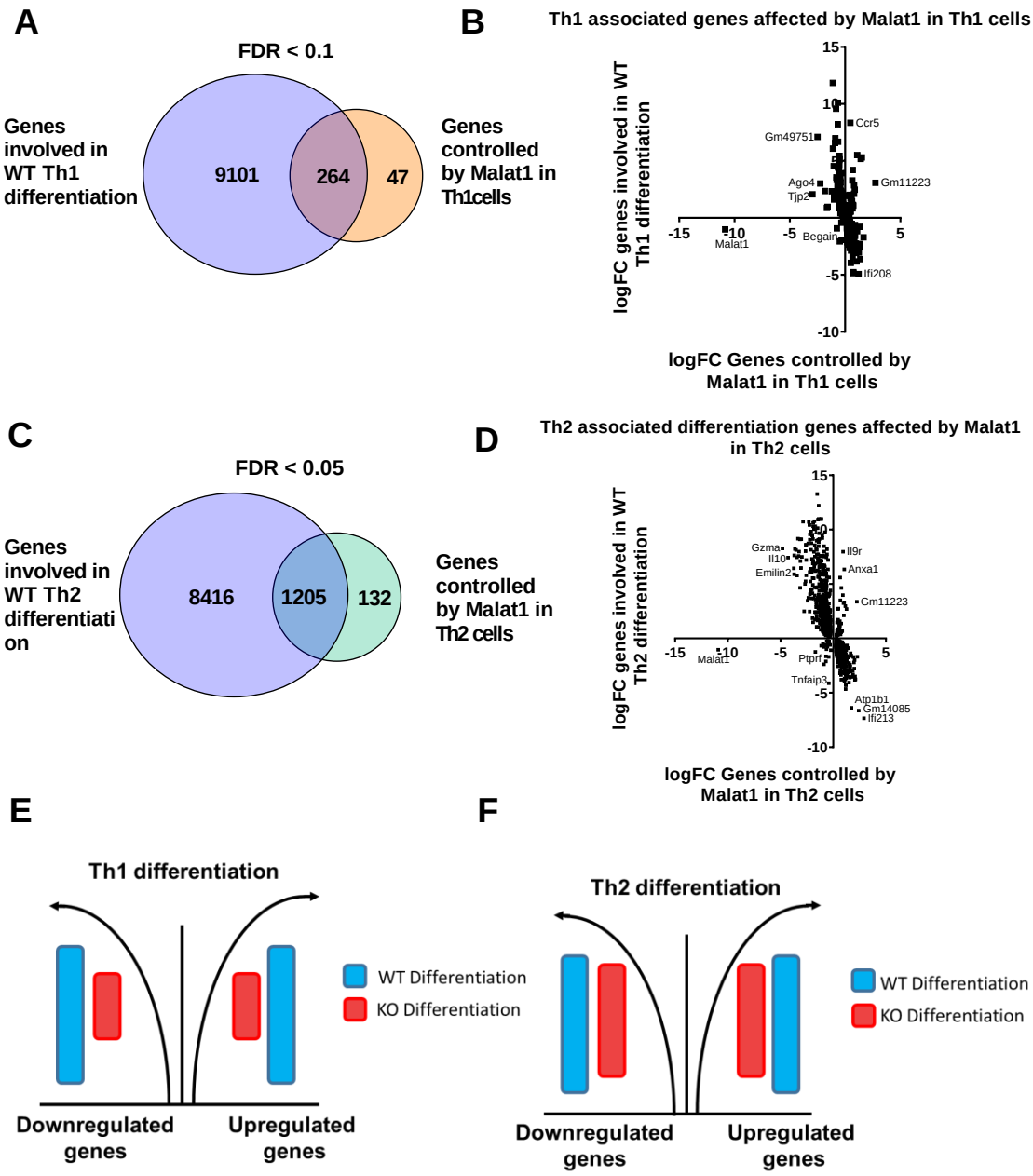


Figure 4.13 Loss of *Malat1* impairs Th cell differentiation.

A) Venn diagram showing overlap of genes controlled by *Malat1* in Th1 cells and genes which are differentially expressed during normal (WT) Th1 differentiation q<0.1 **B)** LogFC of genes controlled by *Malat1* in Th1 cells plotted against logFC of genes involved in WT Th1 differentiation q<0.05 **C)** Venn diagram showing overlap of genes controlled by *Malat1* in Th2 cells and genes which are differentially expressed during normal (WT) Th2 differentiation **D)** LogFC of genes controlled by *Malat1* in Th2 cells plotted against logFC of genes involved in WT Th2 differentiation. **E)** Schematic representation of differentiation defect in Th1 cells **F)** Schematic representation of differentiation defect in Th2 cells

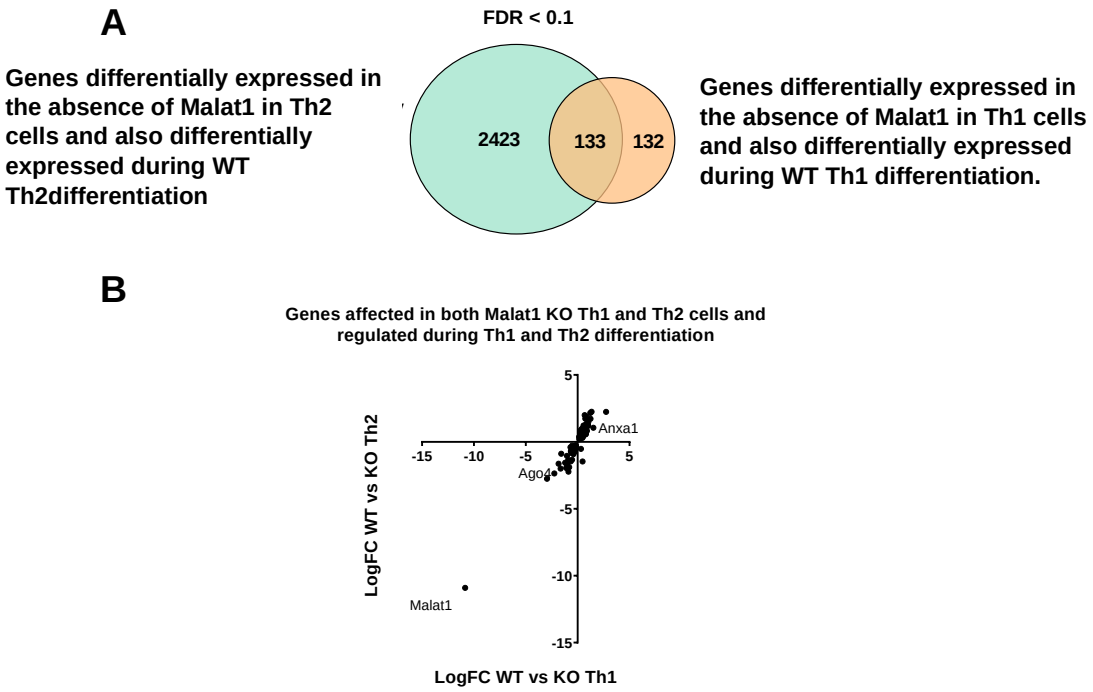


Figure 4.14: *Malat1* controls similar genes during CD4⁺ T cell differentiation in Th1 and Th2 cells.

A) Venn diagram showing overlap of genes controlled by *Malat1* in Th1 cells and genes which are differentially expressed during normal (WT) Th1 differentiation with genes controlled by *Malat1* in Th2 cells and genes which are differentially expressed during normal WT Th2 differentiation **B)** LogFC of genes controlled by *Malat1* in Th1 cells and involved in Th1 differentiation plotted against logFC of genes controlled by *Malat1* in Th2 cells and involved in Th2 differentiation.

Given the prominent effect on gene expression in Th2 cells upon loss of *Malat1*, we focused our study on the role of *Malat1* in Th2 differentiation. Next, we compared genes that were controlled by *Malat1* in Th2 cells with a previously published RNA-seq data set in which our collaborators at the Teichmann group used single-cell RNA-seq to identify genes that correlated with *Malat1* at the single-cell level in *in vitro* polarised WT Th2 cells (***Figure 4.15*****)** (Mahata et al., 2014). Here, we identified 161 genes that are both controlled by *Malat1* in *Malat1* null cells and significantly correlated with *Malat1* (either positively or negatively) at the single-cell level in WT *in vitro* polarised Th2 cells. These genes were then separated into coherent and incoherent regulatory loops, as depicted in (***Figure 4.15B*****).** Genes were allocated to coherent regulatory loops if they positively correlated with *Malat1* at the single-cell level but were down-regulated in *Malat1*$^{-/-}$ *cells* and vice versa. Genes were allocated to incoherent regulatory loops if they positively correlated with *Malat1* at the single-cell level but were up-regulated in *Malat1*$^{-/-}$ cells and vice versa (***Figure 4.15B*****).** Of those identified, the genes were similarly split between the coherent and incoherent regulatory loops (88 vs 110) (***Figure 4.15C*****)**. A list of genes which are in coherent or incoherent regulatory loops are shown in supplementary table 2 in the appendix.

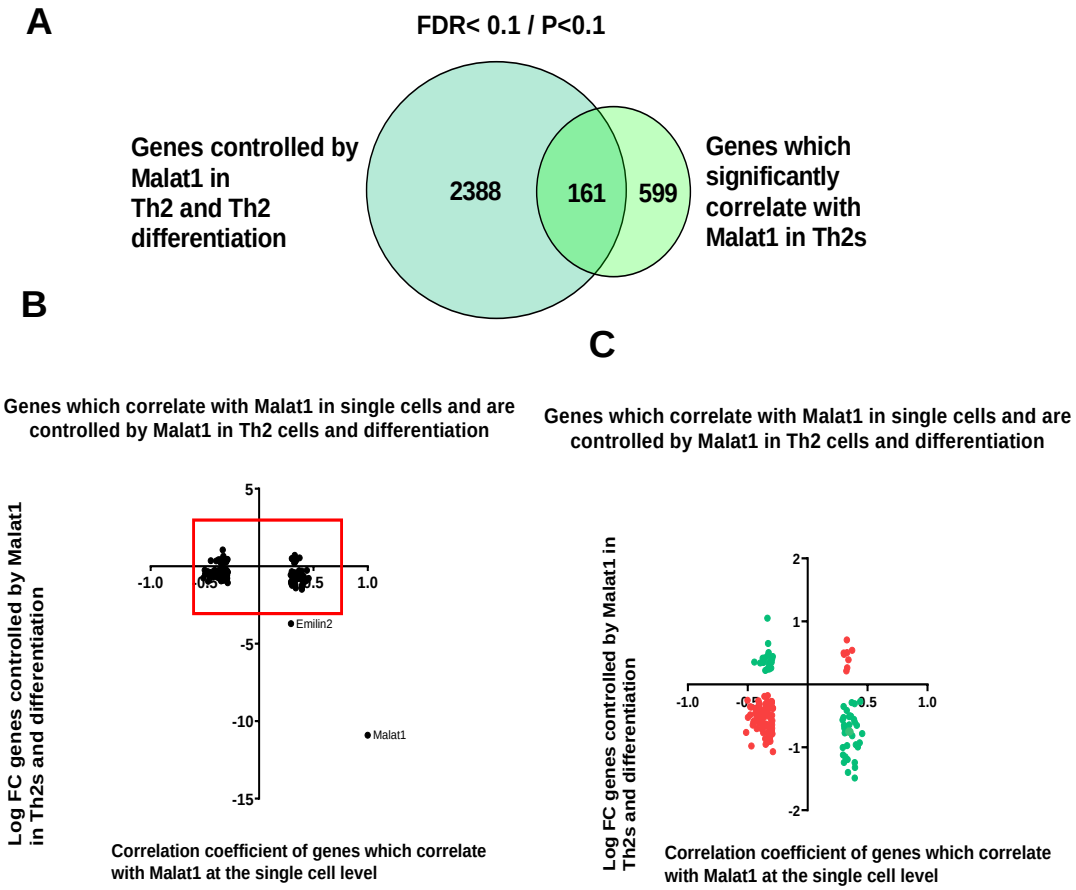


Figure 4.15: Genes controlled by *Malat1* in CD4⁺ T cell differentiation correlate with *Malat1* at the single-cell level.

A) Venn diagram showing overlap of genes controlled by *Malat1* in Th2 cells and genes which are differentially expressed during normal (WT) Th2 differentiation with genes that correlate with *Malat1* in *in vitro* polarised Th2 cells at the single-cell level **B)** LogFC of genes controlled by *Malat1* in Th2 cells and are involved in WT Th2 differentiation cells plotted against the correlation coefficient of genes that correlate with *Malat1* in *in vitro* polarised Th2 at the single-cell level **C)** Magnified view of graph depicted in B. Genes which are in coherent loops are coloured in green (genes which correlate positively with *Malat1* are down-regulated upon loss of *Malat1* and vice versa) and genes which are in incoherent loops are coloured in red (genes which correlate positively with *Malat1* are up-regulated upon loss of *Malat1* and vice versa).

4.2.5 *Malat1* regulates genes in a CD4$^+$ T cell-specific manner

To elucidate the cell type specificity of *Malat1*-dependent regulation of gene expression in CD4$^+$ T cells we compared differentially expressed genes with other published RNA-seq datasets. First, we compared our data to that of Zhang and colleagues who created a *Malat1*$^{-/-}$ mouse model and performed RNA-seq on mouse livers and brains, we used the longer list of genes when sex specific differential gene expression was used (Zhang et al., 2012) Only a handful of genes were differentially expressed in each organ with *Malat1* being the only overlapping differentially expressed gene (***Figure 4.16****A*). GSEA analysis failed to identify any enriched gene sets in the liver and identified genes which are involved in stem cell proliferation in the brain (**Figure *4.16*B**). In comparison with our data, a very limited number of genes were found to overlap with either of the datasets. *Malat1* was often the only gene found to overlap between the datasets (***Figure 4.16***C). Notably, DYNLT1B was absent from these lists in the Zhang model.

We next compared genes which were differentially regulated by *Malat1* in CD4$^+$ T cells with genes that were differentially expressed in CD8$^+$ T cells when *Malat1* is knocked-down – an experiment by Kanbar and colleagues (Kanbar et al., 2022a). 0.47% of genes which were controlled by *Malat1* in CD8$^+$ T cells were also regulated by *Malat1* in naïve CD4$^+$ T cells (***Figure 4.17***A). 2.15% of genes which were controlled by *Malat1* in CD8$^+$ T cells were also regulated by *Malat1* in Th1 cells (***Figure 4.17***A). 13.05% of genes which are differentially expressed in the absence of *Malat1* in CD8$^+$ T cells are also differentially regulated by *Malat1* in Th2 cells (***Figure 4.17***A). A small number of genes are regulated by *Malat1* in all of the different T cell subpopulations – naïve, Th1, Th2 and CD8$^+$ T cells (***Figure 4.17***C).

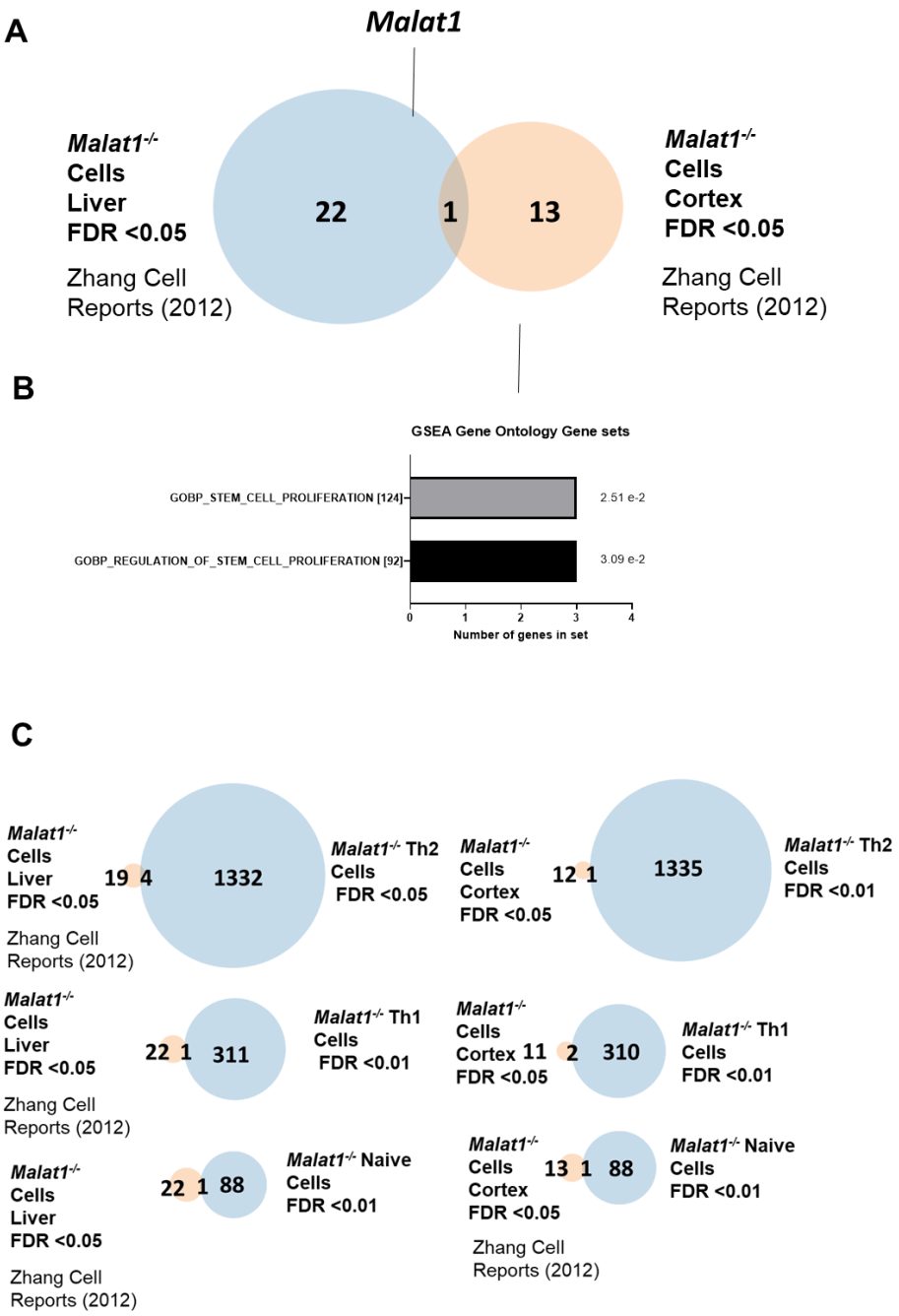


Figure 4.16: *Malat1* regulates different genes in CD4⁺ T cells than in the brain or liver.

A) Overlap of genes which are differentially regulated by *Malat1* in the liver or brain as described by Zhang and colleagues. **B)** GSEA analysis of genes which are differentially expressed in the absence of *Malat1* in the brain **C)** Venn diagrams showing the overlap of genes which are differentially regulated by *Malat1* in either the brain or liver with genes that are differentially regulated by *Malat1* in naïve, Th1 or Th2 cells.

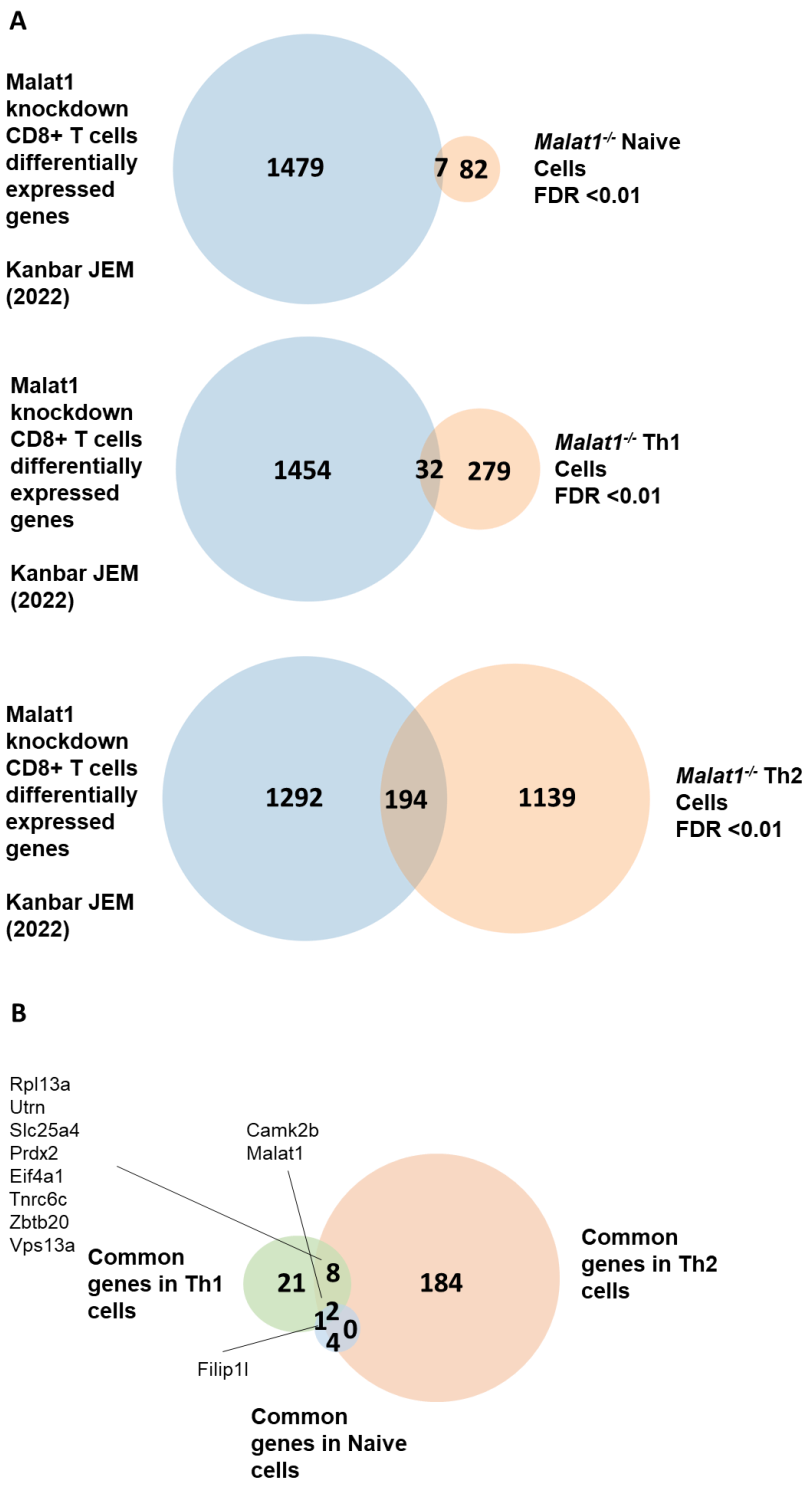


Figure 4.17: *Malat1* differentially regulates unique genes in CD4⁺ and CD8⁺ T cells.

A) Overlap of genes which are differentially regulated in the absence of *Malat1* in naïve Th1 or Th2 cells compared with genes which are differentially regulated by *Malat1* in CD8⁺ T cells as described by Kanbar and colleagues. **B)** Overlap of genes which are differentially regulated by *Malat1* in naïve, Th1 or Th2 cells and CD8⁺ T cells.

4.2.5 Validation of *Malat1*-dependent gene expression changes

To determine if the genes identified by RNA-seq that were differentially expressed in the absence of *Malat1* were truly regulated by the *Malat1* transcript we chose a selection of genes which had significant changes with high logFC and validated the gene expression changes by qRT-PCR in 4-5 biological replicates of WT and *Malat1*$^{-/-}$ Th2 cells. These were the same samples used for RNA-seq. The majority of tested genes followed the same pattern observed in the sequencing data with many differences found to be significant when comparing WT and *Malat1*$^{-/-}$ Th2 cells (***Figure 4.18***). As previously discussed in chapter 4, to confirm that changes in gene expression were due to the *Malat1* transcript and not transcription across the locus we tested the expression of these genes in cells where *Malat1* had been knocked-down using GapmeRs administered at day 0 of Th2 polarisation (***Figure 4.18***). We also compared gene expression changes in EL4 cells which had been treated with GapmeRs for 3 days. We found that knock-down of *Malat1* in primary CD4^{+} T cells reduced the RNA levels of *Emilin2, Tnfsf2, Sox5* and *Eif4a1* in a similar manner to the knock-out models (***Figure 4.18***). Reduced expression of *Malat1* in EL4 cells did not significantly alter the expression of any of these genes. Notably, we found that in some cases knock-down of *Malat1* did not change the expression of *Il21, Eef1g* or *Ago4* in the same manner as the knock-out cells (***Figure 4.19)***.

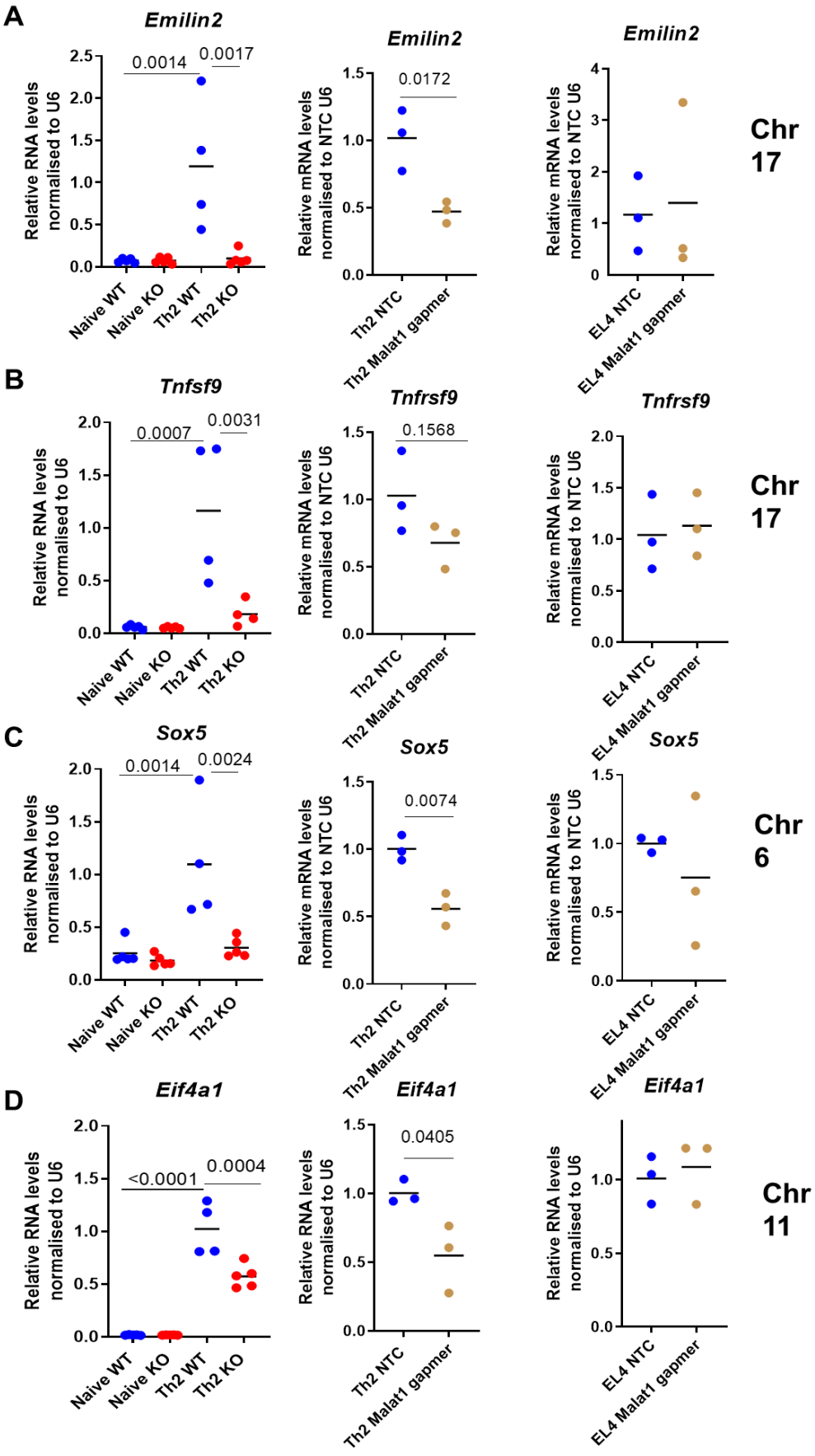

A
Emilin2
Relative RNA levels normalised to U6
0.0014
0.0017
Naive WT
Naive KO
Th2 WT
Th2 KO
Emilin2
Relative mRNA levels normalised to NTC U6
0.0172
Th2 NTC
Th2 Malat1 gapmer
Emilin2
Relative mRNA levels normalised to NTC U6
EL4 NTC
EL4 Malat1 gapmer
Chr 17
B
Tnfsf9
Relative RNA levels normalised to U6
0.0007
0.0031
Naive WT
Naive KO
Th2 WT
Th2 KO
Tnfrsf9
Relative mRNA levels normalised to NTC U6
0.1568
Th2 NTC
Th2 Malat1 gapmer
Tnfrsf9
Relative mRNA levels normalised to NTC U6
EL4 NTC
EL4 Malat1 gapmer
Chr 17
C
Sox5
Relative RNA levels normalised to U6
0.0014
0.0024
Naive WT
Naive KO
Th2 WT
Th2 KO
Sox5
Relative mRNA levels normalised to NTC U6
0.0074
Th2 NTC
Th2 Malat1 gapmer
Sox5
Relative mRNA levels normalised to NTC U6
EL4 NTC
EL4 Malat1 gapmer
Chr 6
D
Eif4a1
Relative RNA levels normalised to U6
<0.0001
0.0004
Naive WT
Naive KO
Th2 WT
Th2 KO
Eif4a1
Relative RNA levels normalised to U6
0.0405
Th2 NTC
Th2 Malat1 gapmer
Eif4a1
Relative RNA levels normalised to U6
EL4 NTC
EL4 Malat1 gapmer
Chr 11

Figure 4.18: qRT-PCR validates differential gene expression identified by RNA-seq.
A) *Emilin2* mRNA levels in naïve and *in vitro* differentiated Th2 cells from WT (blue) and *Malat1*$^{-/-}$ (red) mice (left). *Emilin2* mRNA levels in *in vitro* differentiated Th2 cells treated with either 100 nM of NTC (blue) or *Malat1* (brown) targeting GapmeR (middle). *Emilin2* mRNA levels in El4 cells treated with either 100 nM of NTC (blue) or *Malat1* (brown) targeting GapmeR right. **B)** as in A examining *Tnfsf9* mRNA levels. **C)** as in A examining *Sox5* mRNA levels. **D)** as in A examining *Eif4a1* mRNA levels. RNA levels were determined by qRT-PCR and normalised to U6 and average levels in WT or NTC treated cells. WT vs Malat1$^{-/-}$ naïve or Th2 cells n=4 biological replicates representative of 1 experiment. NTC vs GapmeR Th2 cells. n=3 biological replicates representative of 1 experiment. NTC vs Gapmer EL4 cells n=3 biological replicates representative of 3 independent experiments. Data were analysed by either a t-test or one-way ANOVA. P values are displayed where significant. The mean is displayed.

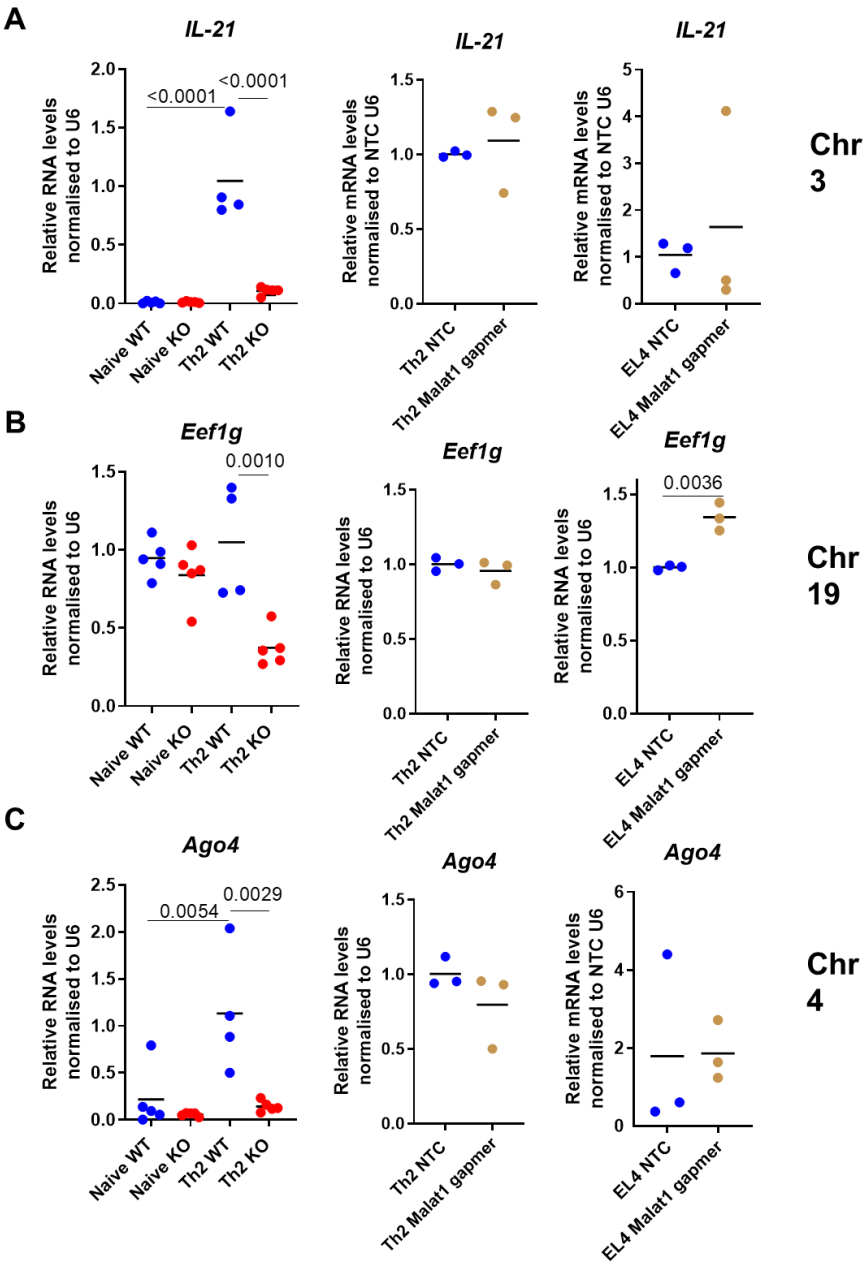

A
IL-21
Relative RNA levels normalised to U6
<0.0001
<0.0001
Naive WT
Naive KO
Th2 WT
Th2 KO
IL-21
Relative mRNA levels normalised to NTC U6
Th2 NTC
Th2 Malat1 gapmer
IL-21
Relative mRNA levels normalised to NTC U6
EL4 NTC
EL4 Malat1 gapmer
Chr 3
B
Eef1g
Relative RNA levels normalised to U6
0.0010
Naive WT
Naive KO
Th2 WT
Th2 KO
Eef1g
Relative RNA levels normalised to U6
Th2 NTC
Th2 Malat1 gapmer
Eef1g
Relative RNA levels normalised to U6
0.0036
EL4 NTC
EL4 Malat1 gapmer
Chr 19
C
Ago4
Relative RNA levels normalised to U6
0.0054
0.0029
Naive WT
Naive KO
Th2 WT
Th2 KO
Ago4
Relative RNA levels normalised to U6
Th2 NTC
Th2 Malat1 gapmer
Ago4
Relative mRNA levels normalised to NTC U6
EL4 NTC
EL4 Malat1 gapmer
Chr 4

Figure 4.19: The *Malat1* transcript is not always responsible for differential gene expression.

A) *Il21* mRNA levels in naïve and *in vitro* differentiated Th2 cells from WT (blue) and *Malat1*$^{-/-}$ (red) mice (left).*Il21* mRNA levels in *in vitro* differentiated Th2 cells treated with either 100 nM of NTC (blue) or *Malat1* (brown) targeting GapmeR (middle).*Il21* mRNA levels in El4 cells treated with either 100 nM of NTC (blue) or *Malat1* (brown) targeting GapmeR (right). **B)** as in A examining Eef1g mRNA levels. **C)** as in A examining Ago4 mRNA levels. RNA levels were determined by qRT-PCR and normalised to U6 and average levels in WT or NTC treated cells. WT vs *Malat1*$^{-/-}$ naïve or Th2 cells n=4/5 biological replicates representative of 1 experiment. NTC vs GapmeR Th2 cells. n=3 biological replicates representative of 1 experiment. NTC vs Gapmer EL4 cells n=3 biological replicates representative of 3 independent experiments. Data were analysed by either a t-test or one-way ANOVA with post-hoc Tukeys test. p values are displayed where significant. The mean is displayed

We next attempted to validate the gene expression changes of *Dynlt1b* which we found to be upregulated between 7-8 LogFC in the absence of *Malat1*. Upon analysis of *Dynlt1b* levels by qRT-PCR we observed a similar degree of *Dynlt1b* upregulation in the *Malat1*$^{-/-}$ cells **(*Figure 4.20*).** This was also recapitulated in mouse tail fibroblasts isolated from WT and *Malat1*$^{-/-}$ mice. Interestingly, knockdown of *Malat1* using GapmeRs failed to alter the expression of *Dynlt1b* in either EL4 cells or *in vitro* polarised Th2 cells **(*Figure 4.20*)**.

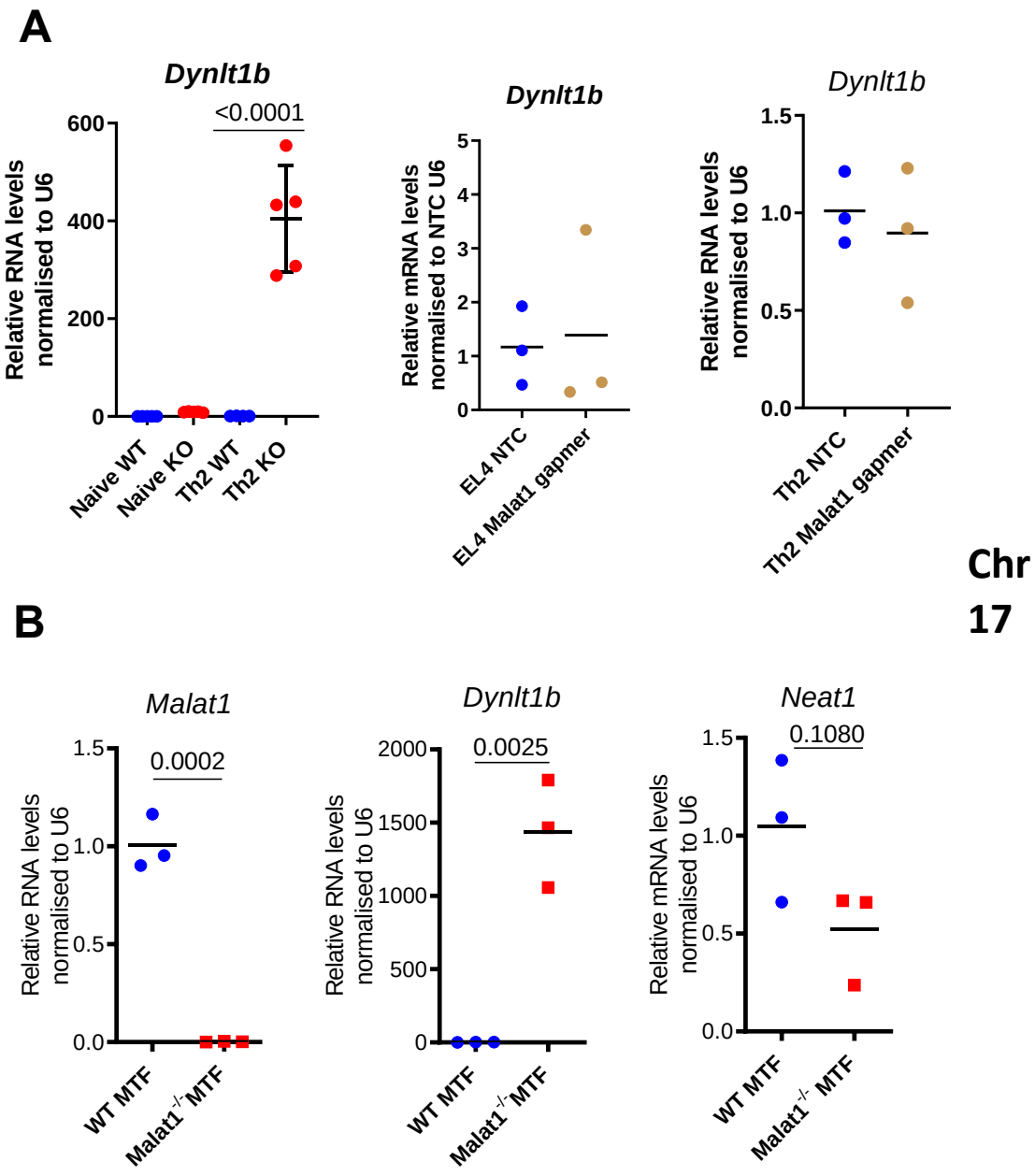

A
Dynlt1b
<0.0001
Relative RNA levels normalised to U6
600
400
200
0
Naive WT
Naive KO
Th2 WT
Th2 KO
Dynlt1b
Relative mRNA levels normalised to NTC U6
5
4
3
2
1
0
EL4 NTC
EL4 Malat1 gapmer
Dynlt1b
Relative RNA levels normalised to U6
1.5
1.0
0.5
0.0
Th2 NTC
Th2 Malat1 gapmer
Chr 17
B
Malat1
0.0002
Relative RNA levels normalised to U6
1.5
1.0
0.5
0.0
WT MTF
Malat1 -/- MTF
Dynlt1b
0.0025
Relative RNA levels normalised to U6
2000
1500
1000
500
0
WT MTF
Malat1 -/- MTF
Neat1
0.1080
Relative mRNA levels normalised to U6
1.5
1.0
0.5
0.0
WT MTF
Malat1 -/- MTF

Figure 4.20: Dynlt1b expression is elevated specifically in *Malat1*$^{-/-}$ mice.

A) Dynlt1b mRNA levels in naïve and *in vitro* differentiated Th2 cells from WT (blue) and *Malat1*$^{-/-}$ (red) mice (left). Dynlt1b mRNA levels in *in vitro* differentiated Th2 cells treated with either 100 nM of NTC (blue) or *Malat1* (brown) targeting GapmeR (middle). Dynlt1b mRNA levels in El4 cells treated with either 100 nM of NTC (blue) or *Malat1* (brown) targeting GapmeR right. **B)** *Malat1,* Dynlt1b and *Neat* mRNA levels in MTF from WT (blue) of *Malat1*$^{-/-}$ (red) mice. RNA levels were determined by qRT-PCR and normalised to U6 and average levels in WT or NTC treated cells. Naïve CD4^{+} T cells and Th2 WT vs *Malat1*$^{-/-}$ cells n=4/5 biological replicates representiative of 1 experiment. NTC vs GapmeR Th2 cells n=3 biological replicates representative of 1 experiment. n=3 biological replicates representative of 3 indpendent experiments NTC vs Gapmer EL4 cells. MTF WT and *Malat1*$^{-/-}$ cells n=3 biological replicates, data representative of 1 experiment. Data were analysed by either a t-test or one-way ANOVA. P values are displayed where significant. The mean is displayed.

4.2.6 Timing of GapmeR addition affects Th cell responses

We next asked the question if the timing of GapmeR addition affected gene expression in $CD4^+$ T cells. Under *in vitro* Th2 polarising conditions, 100 nM of NTC or *Malat1* targeting GapmeR was added to the cells at day 0 of the polarisation or at day 1 and day 4 (***Figure 4.21***). A panel of genes shown to be differentially regulated by *Malat1* as determined by RNA-seq were then examined. Surprisingly, several genes including *Emilin2, Eif4a1, Tnfsf9, Il21, Sox5 and Anxa1* were regulated differently in the absence of *Malat1* depending on the timing at which the GapmeR was added. In many cases, genes were down-regulated when the GapmeR was added on day 0 and up-regulated when the GapmeR was added on day1 and 4 (***Figure 4.21***). In some cases, no change in mRNA levels was observed in either the day 0 or day 1 and day 4 conditions. This disparity in response could be explained by differences in *Malat1* expression kinetics under different conditions (***Figure 4.22***). For example, knockout cells lack the majority of *Malat1* expression throughout development of the mouse and from the beginning of the *in vitro* experiments. WT cells express high levels of *Malat1,* and this is rapidly downregulated upon $CD4^+$ T cell activation. Addition of the GapmeR at day 0 will further reduce the levels of *Malat1* in relation of WT controls but remains higher expression levels than knockout cells. In contrast, addition of the GapmeR at day 1 and 4 will also reduce the expression of *Malat1* further in relation to WT cells upon activation but will occur later than when the GapmeR is added at day 0.

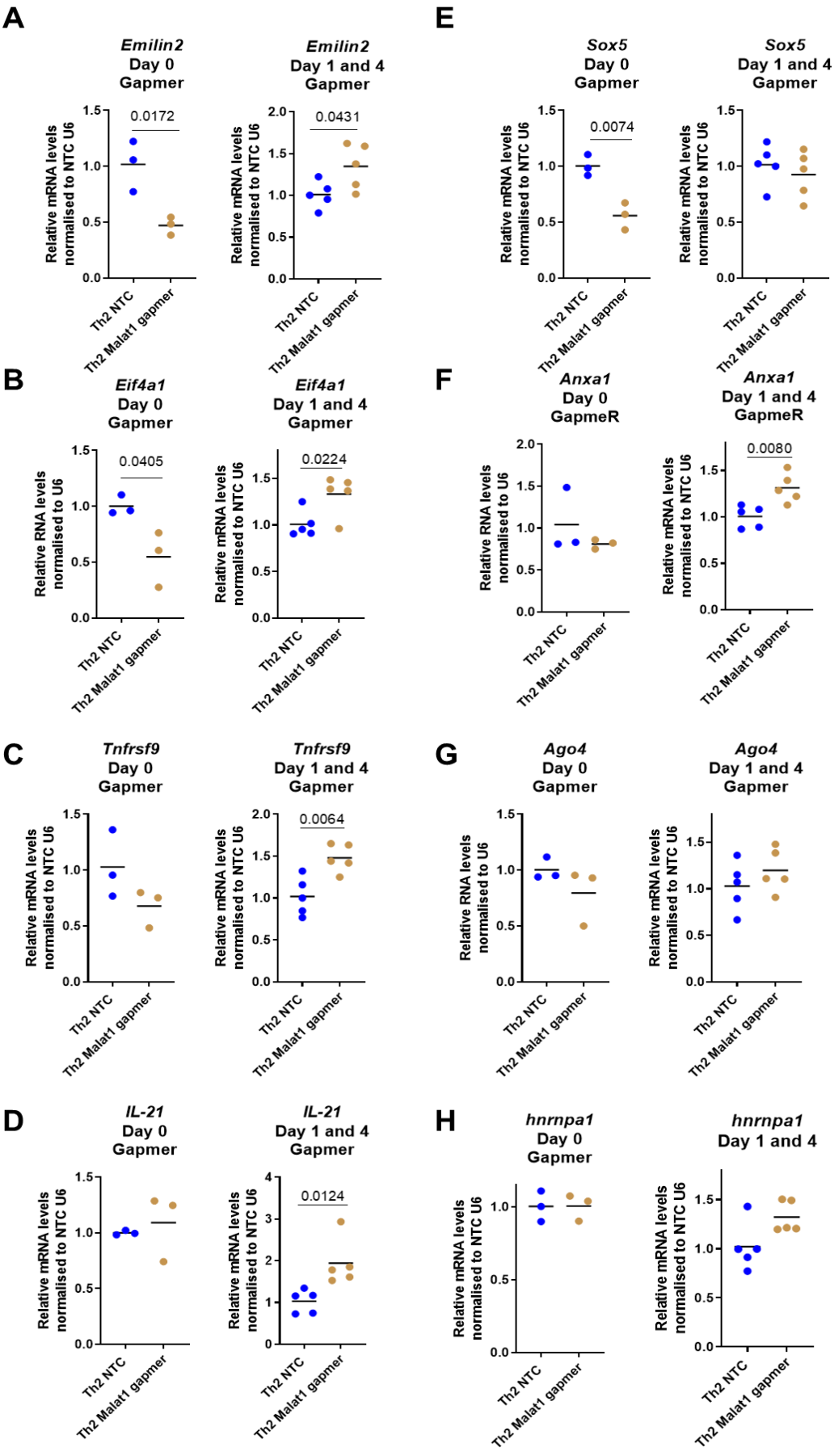

A
Emilin2
Day 0
Gapmer
Relative mRNA levels normalised to NTC U6
0.0172
Th2 NTC
Th2 Malat1 gapmer
Emilin2
Day 1 and 4
Gapmer
0.0431
B
Eif4a1
Day 0
Gapmer
Relative RNA levels normalised to U6
0.0405
Eif4a1
Day 1 and 4
Gapmer
0.0224
C
Tnfrsf9
Day 0
Gapmer
Tnfrsf9
Day 1 and 4
Gapmer
0.0064
D
IL-21
Day 0
Gapmer
IL-21
Day 1 and 4
Gapmer
0.0124
E
Sox5
Day 0
Gapmer
0.0074
Sox5
Day 1 and 4
Gapmer
F
Anxa1
Day 0
GapmeR
Anxa1
Day 1 and 4
GapmeR
0.0080
G
Ago4
Day 0
Gapmer
Ago4
Day 1 and 4
Gapmer
H
hnrnpa1
Day 0
Gapmer
hnrnpa1
Day 1 and 4

Figure 4.21: Timing of GapmeR addition affects Th cell responses.

mRNA levels in *in vitro* polarised Th2 cells treated with either 100nM of NTC (brown) or 100 nM of *Malat1* targeting GapmeR. GapmeR was introduced either a day 0 or day 1 and day 4 and is indicated in the figure. Cells were harvested at day 6 of polarisation and RNA levels were determined by qRT-PCR. The levels of the following RNAs were examined **A)** Emilin2 **B)** Eif4a1 **C)** Tnfsf9 **D)** IL-21 **E)** Sox5 **F)** Ago4 **G)** hnRNPA1. RNA levels were normalised to U6 and average NTC treated cells. NTC vs GapmeR day 0 Th2 cells n=3 biological replicates representative of 1 independent experiment. NTC vs Gapmer day 1 and day 4 n=5 technical replicates representative of 1 independent experiment. Data were analysed using an unpaired t-test. p values are displayed where significant. The mean is depicted.

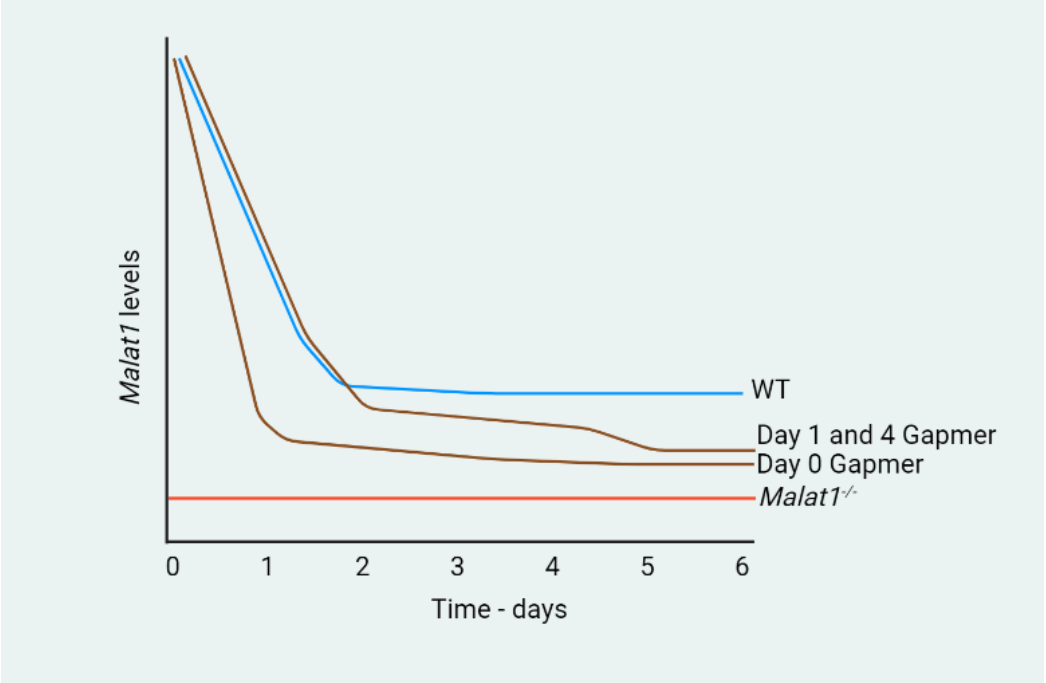


Figure 4.22: Schematic representation of *Malat1* levels upon GapmeR addition.

Schematic depicting suggested *Malat1* RNA levels at different time points during *in vitro* $CD4^+$ T cell polarisation under different conditions.

4.3 Discussion:

In this chapter, we have provided unique insight into the role of *Malat1* in differential gene expression in $CD4^+$ T cells. We demonstrated that loss of *Malat1* has different effects on gene expression depending on the $CD4^+$ T cell subset in which *Malat1* expression is deleted. We identified pathways of genes which are differentially expressed in the absence of *Malat1*. Finally, we demonstrated that loss of *Malat1* reduced gene expression changes upon $CD4^+$ T cell differentiation.

Previous studies have reported a range of differential gene expression changes in the absence of *Malat1* spanning from very limited gene expression changes to thousands of genes identified as differentially expressed in the absence of *Malat1*. In our data, we found that the effect of loss of *Malat1* varied depending on the subset of $CD4^+$ T cells which were examined. One of the first observations from this study is the prominent effect on gene expression that occurs in Th2 cells upon loss of *Malat1***.** This is an intriguing finding, as we have previously characterised the non-redundant role of *Malat1* in type 1 models both *in vitro* and *in vivo* which included visceral leishmaniasis *(Leishmania donovani)* and malaria (*Plasmodium Chabaudi)* which induce a strong but not unique Th1 response. Yet only a fraction of genes were affected upon loss of *Malat1* in Th1 cells. Further, investigating the role of *Malat1* in type 2 *(*Th2) *in vivo* models may reveal an even more striking phenotype and deepen our understanding of the role of *Malat1* in adaptive immunity. Why this type 2 bias occurs is unclear. *Malat1* is known to interact with several RBPs and miRNAs. Perhaps the local changes in RBP binding or miRNA binding could explain the distinct role in gene expression changes in Th2 cells (Arun et al., 2020) .

Upon further investigation of *Malat1*-dependent gene expression changes we uncovered a bias towards altered gene expression along chromosome 19 in naïve, *in vitro* polarised Th1 and Th2 cells. *Malat1* is also located on chromosome 19 in mice which suggested a potential *in cis* regulatory role. Several lncRNAs have been shown to have an *in cis* or local regulatory in $CD4^+$ T cells (West & Lagos, 2019). A classic example of a lncRNA which regulates genes locally is the lncRNA *Ifng-as1* which neighbours the *Ifng* gene. *Ifng-AS1* and *Ifng* are transcribed on opposite DNA strands. In mouse models of sepsis mice which overexpress *Ifng-as1* show greater expression of IFNγ (Gomez et al., 2013a) Similarly, *Malat1* has been shown to have *in cis* regulatory effects and has been shown to alter the expression of the neighbouring gene *Neat1* (Zhang et al., 2012). We have also previously shown that *Malat1* appears to have limited *in cis* regulatory effects and can regulate gene expression of its neighbouring genes in a context-dependent manner (Hewitson et al., 2020). Previous work in our lab has also shown a significant correlation between lncRNA expression and their chromosomally adjacent genes in $CD4^+$ T cells (Hewitson et al., 2020). As such it is not entirely surprising that *Malat1* may be showing *in cis* regulatory effects across chromosome 19. An alternative possibility is that these gene expression changes are a result of an off-target effect which took place in the generation of the knockout mouse models. As expression of *Malat1* is much higher in Th2 than Th1 cells this could explain the greater effect on chromosome 19 genes in this cell type – however, *Malat1* is most highly expressed in naïve cells. As the extent of chromosome 19 gene alteration is limited in naïve cells and most pronounced in Th2 cells it may be less likely that an off-target effect is responsible for these changes.

GSEA pathway analysis of genes differentially regulated by *Malat1* identified hallmark hypoxia pathways to be some of the most enriched pathways in naïve and

Th2 cells. Examination of the data sets found that genes such as the hypoxia inducible factors (HIF) including HIF-1α are dysregulated in the absence of *Malat1*. Interestingly, during cancer development blood vessels which form around tumours have poor blood flow. This creates a hypoxic microenvironment and is a hallmark of highly proliferative cells (Harris, 2002). *Malat1* has been linked to metastasis severity in a plethora of cancers including lung cancer (Ji et al., 2003). Several studies have shown that hypoxia up-regulates *Malat1* levels in a HIF-1α and HIF-2α dependent manner which gives potential insight into the mechanism driving *Malat1* dysregulation in cancer (Lelli et al., 2015; Shih et al., 2021). *Malat1* has also previously been shown to impair HIF-1α hydroxylation and promote glycolysis. Interestingly, TCR stimulation results in a robust HIF-1α protein stabilisation, which can be enhanced by exposure to hypoxic conditions. Additionally, if $CD4^+$ T cells are stimulated *in vitro* under hypoxic conditions (2-5% O_2) they show reduced proliferation and diminished cytokine expression (Mcnamee et al., 2013) . HIF-1α has also been linked to promoting IL-10 expression in B cells (Meng et al., 2018). It could be interesting to repeat the *in vitro* polarisation experiment under hypoxic conditions to see if loss of *Malat1* has an even more pronounced effect on IL-10 expression.

GSEA pathway analysis identified genes involved in the IL-2 STAT5 signalling pathway to be dysregulated in naïve, Th1 and Th2 cells in the absence of *Malat1*. During *in vitro* $CD4^+$ T cell polarisation IL-2 is added to the cell culture medium to enhance $CD4^+$ T cell proliferation at day 4 this indicates that changes which occur at the later stages of the differentiation process may be responsible for some of the phenotypes identified.

GSEA pathway analysis also identified apoptosis as a commonly enriched pathway altered in the absence of *Malat1*. During *in vitro* $CD4^+$ T cell polarisation no

obvious differences in cell viability were observed in this case. However, *Malat1* has previously been linked to the promotion of apoptosis in $CD8^+$ T cells during infection (Kanbar et al., 2022). Further experiments which included apoptosis markers such as Annexin V could determine if *Malat1*$^{-/-}$ $CD4^+$ T cells are more or less apoptotic.

Upon examination of gene expression changes in the knockout mice during $CD4^+$ T cell differentiation, we found that genes are not up or down-regulated to the same extent as WT cells. This indicates that *Malat1*$^{-/-}$ $CD4^+$ T cells do not differentiate to the same extent as their WT counterparts. This effect was particularly prominent in *in vitro* polarised Th2 cells. In chapter 3, we showed that *Malat1* is quickly suppressed upon T cell activation, this begs the question of why *Malat1* is being suppressed if loss of *Malat1* impairs differentiation. There are several possibilities, perhaps quick suppression of *Malat1* is essential for the early stages of differentiation. An alternative suggestion is that *Malat1* is suppressed to enable the correct up or downregulation of a small number of genes shown in the top right and bottom left quadrants of (***Figure 4.13***B/D). However, given that most genes which are affected by *Malat1* are not up or down regulated to the same extent as WT cells it is likely that the foremost reason *Malat1* is suppressed is mainly to affect these specific genes which are involved in Th cell differentiation. This indicates that the quick suppression of *Malat1* during Th cell activation is to act as a break for Th cell differentiation, meaning that *Malat1* may be important for controlled licensing of naïve $CD4^+$ T cells to differentiation ensuring that differentiation occurs at appropriate times, such that aberrant differentiation does not occur.

Of the genes identified by RNA-seq, qRT-PCR showed that the majority followed the same expression pattern as the RNA-seq data in WT and *Malat1*$^{-/-}$ Th2 cells**.** Experiments with day 0 *Malat1* inhibitors demonstrated that the *Malat1*

transcript is responsible for the regulation of the genes *Sox5, Emilin2 and eIF4a1* in $CD4^+$ T cells**.** However, this was the case for approximately 50% of the genes which were validated by qRT-PCR. This suggests that in many cases transcription across the *Malat1* locus is responsible for regulating gene expression changes or alternatively, complete loss of *Malat1* is required to see these changes in gene expression. As only a handful of genes were validated by qRT-PCR this may skew our perception of what percentage of genes are regulated by the *Malat1* transcript and which genes are regulated by transcription or which are a result of complete *Malat1* deletion. To gain a better understanding of this *Malat1* could be knocked-down using GapmeRs under different polarising conditions and gene expression changes compared between knockout and knockdown cells using RNA-seq.

One gene in particular stood out in this data set; this gene was *Dynlt1b* which was found to be significantly up-regulated in the absence of *Malat1*. DYNLT1b is a member of a family which is associated with the t-complex of mice. Defects in DYNLT1b expression have been linked to male sterility in mice and flies. Its expression is 200-fold higher in mouse testes than in any other tissues (Indu et al., 2015). We examined the expression of DYNLT1b in MTF derived from WT and *Malat1*$^{-/-}$ mice and also observed a significant increase in expression. As *Malat1* targeted GapmeR treatment of either naïve $CD4^+$ T cells or EL4 cells did not alter *Dynlt1b* expression, this suggested that the altered expression of *Dynlt1b* was an off-target effect of the knockout mouse generation. We have not tested the mRNA levels of *Dynlt1b* in the testes of WT and *Malat1* null mice. However, no significant effects are observed on mouse fertility in any of the *Malat1* knockout mouse models (Eißmann et al., 2012; Nakagawa et al., 2012b; Zhang et al., 2012). *Dynlt1b* was not found to be up-regulated in the Zhang *Malat1* knockout mouse mode suggesting that

this is a specific effect of the Nakagawa model (Zhang et al., 2012). *Dynlt1b* could be knocked-down in $CD4^+$ T cells to see if it has an impact on Th cell function. It is beneficial to note that we know that at least some of the changes observed in our study are certainly due to the *Malat1* transcript as it the case with IL-10.

Interestingly, in this study, we found that timing of GapmeR addition (either at day 0, or day 1 and day 4) had strikingly different impacts on *Malat1*-dependent gene regulation. Typically, when GapmeRs were added at the beginning of the $CD4^+$ T cell differentiation process this resulted in gene expression being altered in the opposite direction to if GapmeRs were added at day 1 and day 4. There are several possible explanations for this observation. Perhaps these gene expression changes are only observed with early knockdown of *Malat1* as it more closely resembles a knockout cell, as after just 24 hours of T cell stimulation the expression of thousands of genes change including RBPs. An alternative possibility is that these *Malat1*-dependent gene expression changes are part of a very early event in $CD4^+$ T cell differentiation we see some changes after 24 hours of stimulation which is explored in chapter 7. Another explanation is that as GapmeRs have been added twice to the day 1 and day 4 condition in comparison to the day 0 condition, this means they have been treated with double the concentration of GapmeR. It is a possibility that an off-target effect of increased GapmeR in either the NTC or *Malat1* targeting condition could be altering the cells in some way. Similar knockdown efficiencies were observed between each condition so this cannot explain the differences observed. However, as technical replicates were used for the day 1 and day 4 conditions, and these experiments took place on different days this experiment would need to be repeated before any significant conclusions could be drawn. Interestingly, some of the genes were not significantly altered in any of the GapmeR conditions, this could be due to the low n

number used for this preliminary experiment. An alternative possibility is that it is transcription across the *Malat1* locus which affects these genes, not the RNA product itself or complete deletion of *Malat1* is required. It would be beneficial to reproduce this experiment and increase the number of time points used for GapmeR addition to determine how the kinetics of reducing *Malat1* expression can impact $CD4^+$ T cell function.

4.3.2: Future work

One interesting avenue for future exploration could be to examine the role of *Malat1* in a sex-specific manner in $CD4^+$ T cells. Zhang and colleagues found limited gene expression changes in the brains on livers of *Malat1* null mice (Zhang et al., 2012). Interestingly, the authors note greater differences in differential gene expression when the mice are categorised by sex. The author's state they found *Neat1* to be more highly expressed in males. When the male and female data was pooled only one gene was found to be significantly altered in the absence of *Malat1* – Serum amyloid A3. However, when grouped by sex this list expanded to 22 genes which are differentially expressed in the absence of *Malat1*. This suggests that *Malat1* may have some sex-specific functions and could be an interesting area for further research – we explore this concept in chapter 7.

One disadvantage of RNA-seq technologies is that although it gives comprehensive information on RNA abundance at a quantitative level mRNA and protein levels are only modestly correlated (Wang et al., 2014). The use of proteomic approaches would provide essential confirmation for the functional relevance of findings from these RNA-seq data sets.

In summary, for the first time, we identified *Malat1*-dependent differentially expressed genes in $CD4^+$ T cells. This revealed a strong Th2 bias in gene regulation, in addition to identifying a potential role of *Malat1* in licensing Th cell differentiation. Further understanding how *Malat1* regulates these genes in specific Th cell contexts warrants further investigation.

5. RAP-MS identifies *Malat1*-protein interactions in $CD4^+$ T cells

5.1 Introduction

5.1.1 RNA binding proteins

An RBP is a protein that interacts with RNA via one or more RNA binding domains (RBDs) to form an RNA-protein complex (Glisovic et al., 2008). These interactions play a key role in many stages of the RNA life cycle including transcription, splicing, RNA localisation, post-transcriptional regulation, translation, and RNA turnover rate. Regulation of these processes ultimately impacts numerous cellular functions (Gehring et al., 2017). 1,914 murine and 4,257 human RBPs have been experimentally validated and are described in RBPbase (Gebauer et al., 2020). This growing number of RBPs may exceed previous estimates of ~7.5% representation of the human proteome (Hoefig et al., 2021). RBP structure and binding capabilities are incredibly diverse, which enables the formation of numerous unique RNA-protein complexes with specific functions (Glisovic et al., 2008; Turner & DÍaz-Muñoz, 2018). Thus, identifying the RNAs and proteins that comprise these complexes gives essential insight into the function of the complex and its components.

RNA-protein complex composition is determined by numerous factors including RNA structural elements, epigenetic RNA modifications, specific sequence motifs, and protein modifications (Rissland, 2017). Consequently, RNA transcripts can form widely different complexes. Additionally, RNA-protein complex composition is dynamic and can change dramatically throughout an RNAs lifecycle. One suggestion for the dynamic nature of the RNA-protein complex is that RBPs act as “clothes” for the RNA transcript, ensuring that specific regions or structures of the RNA are covered or uncovered enabling the RNA to function appropriately during the different stages of its life cycle (Hentze et al., 2018; Singh et al., 2015). It is well-

established that lncRNAs interact with DNA, RNA, and proteins (Ma et al., 2013). Yet, lncRNAs are involved in the recruitment of transcription factors, chromosome remodelling, and protein sequestering. This indicates that in some cases the reverse situation is taking place, in which RNA within a complex is acting to affect protein function rather than be affected by RBPs (Hentze et al., 2018). This lends further weight to the importance of studying RNA-protein interactions to understand function.

RNA binding domains (RBDs) are the functional regions of RBPs that enable RNA-protein interactions. RBPs must contain at least one RBD, however, it is common for RBPs to comprise multiple RBDs which can enhance their RNA binding capabilities (Glisovic et al., 2008). Common RBDs include RNA recognition motif (RRM), K-homology domain (KH), and zinc finger (Znf) domains (Hentze et al., 2018). Yet over 600 structurally different RBDs have been defined in canonical RBPs which also often contain intrinsically disordered regions (IDRs). Several of these domains are known to bind specific types of RNA with RRM, PUF repeats, ZnF and KH domains known to bind specifically to ssRNA (Auweter et al., 2006). Other domains such as the double-stranded RNA binding domain commonly recognises dsRNA (Masliah et al., 2013). However, RBDs can be difficult to classify, with many experimentally identified RBPs lacking known/canonical RBDs. Surprisingly, 71 metabolic enzymes with well-defined metabolic functions have been identified as moonlighting as RBPs (Hentze et al., 2018). Consequently, the number of RBDs continues to grow as new RBPs are discovered. The emergence of newly defined RBPs is partially due to the expansion of RIC experiments.

There are several methods of identifying RBDs, for example, mutagenesis followed by electrophoretic mobility shift assay (EMSAs). Alternative global approaches such as RBDmap can identify regions involved in binding. This is an

expansion of the RNA interactome method which uses UV crosslinking, oligo(dT) capture, proteolysis, further oligo (dT) capture, and mass spectrometry to distinguish binding sites (Castello et al., 2017).

RBP functional diversity is achieved through the structural modularity of both RNA binding domains and auxiliary regions. For example, many RBPs have multiple copies of the RRM RBD which is capable of binding up to 6 nucleotides. Increasing the number of copies of this domain enables the recognition of specific and more complex RNA targets (Lunde et al., 2007). In parallel, RBPs have modular auxiliary domains such as kinase domains or deaminase domains which contribute to RBP functionality (Glisovic et al., 2008). Collectively, the modular structures of both the RNA binding domains and auxiliary domains facilitate the unique specificity and functions of RBPs within a cell.

Notably, RBP complexes are often found in granules lacking any membrane and form due to liquid-liquid phase separation (LLPS). Examples include nucleoli, stress granules, paraspeckles, and nuclear speckles. The multivalent properties (multiple binding sites) of RBPs are essential for driving phase separation. These can arise through interactions in ordered or IDR of the proteins (Gomes & Shorter, 2019). Notably, lncRNAs can act as a scaffold for the RBP complex due to their secondary structures and ability to reduce protein viscosity this drives phase separation formation, for example, *Neat1* is required for paraspeckle formation (Naganuma & Hirose, 2013). Environmental factors such as temperature, salt concentration and pH can have a large influence on the LLPS formation. Additionally, LLPS has been implicated in several neurological disorders such as frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS) (Luo et al., 2021).

5.1.2 RNA binding proteins and CD4$^+$ T cell function

To combat invading pathogens, cells of the immune system undergo rapid and complex gene expression changes – as is the case with CD4$^+$ T cell activation and differentiation. Emerging evidence suggests RBPs are important in controlling gene expression changes required for immune cell differentiation and function (Díaz-Muñoz & Turner, 2018; Turner & DÍaz-Muñoz, 2018).

Unsurprisingly, RBPs have been linked to CD4$^+$ T cell differentiation and function. When CD4$^+$ T cells are stimulated, the expression of numerous RBPs is altered. This is essential for enabling the initial stages of CD4$^+$ T cell activation as reviewed by (Turner & DÍaz-Muñoz, 2018). One example is the RBP ROQUIN1 which has been shown to play an essential role in the post-transcriptional regulation of key Th17 genes (Jeltsch et al., 2014). ROQUIN1 acts as part of an RNA-protein complex that comprises key Th17-promoting mRNAs and the nuclease regulatory RNase-1 (REGNASE-1). This complex degrades key Th17 mRNAs, which in turn impairs Th17 differentiation. However, when T cells are activated a paracaspase known as mucosa-associated lymphoid tissue lymphoma translocation protein 1 (MALT1) cleaves REGNASE-1 and ROQUIN-1 thus de-repressing the expression of their mRNA targets and enabling Th17 differentiation (Jeltsch et al., 2014)

Another example is poly (C) binding protein 1 (PCBP1) which is upregulated upon Th cell activation. Interestingly, T cell-specific deletion of PCBP1 revealed it is upregulated to prevent the conversion of T effector cells into T regulatory cells. In addition, RNA-seq identified that PCBP globally represses genes that have immune-

suppressive functions (Ansa-Addo et al., 2020). This is important for maintaining optimal Th effector cell function and aiding anti-tumour immunity.

Interestingly, many genes which encode cytokines expressed by $CD4^+$ T cells such as *Il2, Il3, Il4,* and *Il13* have adenylate uridylate-rich elements (AREs) which are regulatory elements involved in mRNA turnover and translation. Many RBPs bind to these ARE elements which can promote or hinder mRNA translation and stability. The ubiquitous RBP ELAV like RNA binding protein 1 (ELAVL1/HuR) aids mRNA stability and translation through AIRE binding sites. ELAVL1 has been shown to positively regulate *Gata3* and increased both its mRNA and protein expression upon Th2 stimulation in both human and mouse Th cells (Stellato et al., 2011). ELAVL1 has also been shown to post-transcriptionally regulate *IL4, IL13* and maintain IL-2 homeostasis (Stellato et al., 2011).

Another study investigated the role of the novel lncRNA lincRNA00892 in SLE patients. Through microarray screening lincRNA00892 was found to be aberrantly upregulated in patient samples. mRNA-lncRNA co-expression analysis indicated that lincRNA00892 targeted CD40L. As such overexpression of lincRNA00892 enhanced the expression of CD40L at the protein level. Pull down of lincRNA00892 followed by mass spec identified hnRNPK as its main interaction partner. The authors postulated that lincRNA00892 interacted with hnRNPK to drive CD40L expression post-transcriptionally (X. Liu, Lin, et al., 2021).

Recent work has defined the core T cell RBPome using orthogonal organic phase separation (OOPs) and RNA interactome capture (RIC), which identified 798 mice and 801 human RBPs (Hoefig et al., 2021). STAT1 and STAT4 were identified as RBPs and confirmed by protein-immunoprecipitation binding assays. This provides

exciting insight into the potential moonlighting functions of STAT proteins. Further work would be required to establish binding sites and understand which of these RBPs which regulate mRNAs are important for T helper cell function. Nevertheless, this growing body of evidence provides strong support for the relevance of RBPs in T helper cells.

5.1.3 Methods to detect RNA protein interactions

One of the major questions surrounding lncRNAs is, how they exert their function. Considerable literature has grown around the theme of lncRNA-protein interactions in gene regulation and chromatin structure modifications. Therefore, identifying lncRNA-protein interaction partners would provide deeper insight into how these lncRNAs are functional in the context of a cell (McHugh & Guttman, 2018). *Malat1* has previously been shown to interact with several RBPs such as EZH2, methyltransferase like 16 (METTL16) and TDP-43 (Brown et al., 2016; Kanbar et al., 2022; Liu et al., 2020). However, the *Malat1*-protein interactome remains largely unexplored in immune cells. Thus, understanding *Malat1*-protein interactions in $CD4^+$ T cells would deepen our understanding of how *Malat1* functions in the immune system and determine if this is through protein-dependent or independent mechanisms.

Typically, an RBP forms an RNA-protein complex via well-characterised binding between RNA and a RBPs RBD such as the RRM domain. However, more recently interaction studies are starting to uncover non-traditional RBDs as reviewed by (Hentze et al., 2018). Thus, it is important to take unbiased approaches to identify RNA–protein interactions.

Several methods to identify RNA-protein interaction partners have been developed as reviewed by (McHugh et al., 2014). These approaches can be broadly

classified as either protein or RNA-centric. Protein-centric methods (native or denaturing) such as cross-linking and immunoprecipitation (CLIP), typically involve UV crosslinking to establish a covalent link between RNA and its interacting proteins. Subsequently, specific antibodies are used to pull down a protein of interest and the RNA to which it is bound. The associated RNA is identified by utilising sequencing approaches. However, the ability to identify new RBPs is often limited with this method. Therefore, using RNA-centric approaches enables the identification of new RNA-RBP interactions (McHugh & Guttman, 2018).

The main classes of RNA-centric approaches involve either *in vivo* or *in vitro* purification typically followed by mass spectrometry (MS). *In vitro* purification uses cellular extracts and synthetic RNA bait to identify interaction partners. However, these *in vitro* methods often give rise to false positive results and identify binding in the context of cellular contents in solution; and may not truly represent the interactions that take place inside a cell (McHugh & Guttman, 2018). To overcome these barriers *in vivo* crosslinking can be employed which preserves the true cellular RNA-protein interactions (Barra & Leucci, 2017). However, crosslinking methods such as the use of formaldehyde can prove technically difficult for peptide identification by MS. To overcome some of the challenges posed by chemical crosslinking numerous techniques have adopted UV crosslinking to determine *in vivo* binding partners.

One example of an RNA-centric approach is RNA interactome capture (RIC). UV crosslinking is used to create a covalent crosslink between RNA and protein. Subsequently, RNA protein complexes are purified using oligo(dT) beads and the bound proteins are identified (Castello et al., 2012). However, this is limited to RNAs that have a polyA tail, meaning this approach is not suitable for bacterial systems or non-polyadenylated RNAs such as *Malat1*. Other approaches have been developed to

mitigate these limitations such as orthogonal organic phase separation (OOPS). This approach is based on phenol-chloroform separation where RNA migrates to the top aqueous phase of the solution and proteins to the bottom. When RNA-protein complexes are stabilised by exposure to UV, these will migrate to the interface of the phenol-chloroform complex. The RNA and proteins can be digested and analysed by various techniques including RNA-seq and mass spectrometry (Queiroz et al., 2019). This is an efficient technique that can be used to determine the total RNA interactome (which includes scRNAs).

Targeted RNA-centric approaches enable the identification of specific RNA-protein interactions. Numerous *in vivo* techniques have been identified as reviewed by (Ramanathan et al., 2019). These include, capture hybridisation analysis of RNA targets (CHART), comprehensive identification of RNA binding proteins by mass spectrometry (ChIRP), peptide nucleic acid (PNA)-assisted identification of RBPs (PAIR) tandem RNA isolation procedure (TRIP), and RNA antisense purification followed by mass spectrometry (RAP-MS).

Similar to RIC, TRIP uses UV to create covalent cross-linkages between RNA and protein, and then polyadenylated mRNAs are purified using oligo (dT) beads, however, a subsequent step is used to purify specific mRNAs of interest using 3' biotinylated antisense RNA oligonucleotides and streptavidin beads. As with RIC, this is limited to understanding polyadenylated mRNAs and would not be suitable for understanding some ncRNA-protein interactions such as *Malat1* (Matia-González et al., 2017).

A different technique is PAIR, which uses UV crosslinking to create a covalent crosslink between RNA and protein. Cells are then treated with peptide nucleic acid

(PNA) probes which can penetrate the cell membrane and hybridising the RNA of interest (Zeng et al., 2006). The resulting RNA-RNA-RBP complex is purified using antisense oligonucleotides and magnetic beads. Associated RBPs are then identified by mass spectrometry.

In contrast, CHART uses formaldehyde to reversibly crosslink RNA to DNA, RNA and other proteins. As the crosslinking is reversible this means that DNA, RNA and proteins can be examined from the same enriched samples. Antisense C-oligos which are linked to biotin are used to purify the RNA complex of interest using streptavidin beads. However, RNase H digestion steps are required to identify accessible probe sites. This technique has been used to identify protein and DNA interactions for several ncRNAs including *Malat1* (Simon et al., 2011; West et al., 2014b).

Another technique that has been used to identify specific lncRNA-protein interactions is ChIRP-MS. ChIRP-MS uses formaldehyde to create an *in vivo* cross-linkage between RNA and protein, the RNA complex is purified using short antisense biotinylated oligos and streptavidin beads. The bound proteins are then identified by mass spectrometry. This has successfully identified RNA binding proteins for numerous ncRNAs such as *Xist* (Chu et al., 2011). However, as the probes are relatively short ~20bp this increases the chance of pulling down nonspecific targets.

RAP-MS was initially developed by the Guttman lab to identify binding partners of the lncRNA *Xist* (McHugh et al., 2015). The method involves the use of UV light to covalently cross-link RNA-protein interactions *in vivo*. Subsequently, the cells are lysed and the RNA of interest is mixed with a pool of antisense biotinylated oligonucleotide probes which overlap and span the entire length of the RNA (~120nt

long). These long probes form extremely stable RNA-DNA hybrids which enable the purification of lncRNA-protein complexes under denaturing and reducing conditions (McHugh et al., 2015). Additionally, as the probes span the entire length of the RNA this ensures that all potential hybridisation locations are fully exploited, and beneficially no prior knowledge of the interaction domains is required (Barra & Leucci, 2017). However, as so many probes are required for this approach, this significantly increases the expense of the technique. The RNA-protein biotinylated probe complex is isolated from the lysate using magnetic streptavidin beads. Next, the complex is subjected to several highly denaturing and reducing washes to ensure only the covalently linked proteins are isolated. In the initial approach, stable isotope labelling of amino acids in cell culture (SILAC) was used to further enable distinction between specific protein binding and background noise. However, non-crosslinked samples have previously been used as a suitable alternative. Notwithstanding, the RAP-MS approach is a highly sophisticated and sensitive method of detecting RNA-protein interactions.

The Guttman lab has built on the RAP technique to uncover further RNA interactions. RNA antisense purification followed by DNA sequencing (RAP-DNA) is used to purify chromatin-lncRNA interactions and subsequently sequence the associated DNA to determine the genomic locations of lncRNA-DNA interactions. RAP-RNA has also been developed this technique has shown that *Malat1* indirectly interacts with pre-mRNAs through protein intermediates (Engreitz et al., 2014a). Later work has gone on to identify numerous RNA-RNA interactions including the lncRNA *Linc01285* and *lnc-C2orf63-4-1* (Gu et al., 2022; S. Zhang et al., 2021).

A Comparison of the different RNA-centric approaches is discussed in (***Table 5.13)***

Table 5.13: Comparison of RNA-centric pull-down approaches for identification of RBPs

Approach								
	RAP-MS	**CHART-MS**	**iDRiP**	**HyPR-MS**	**ChIRP-MS**	**PAIR**	**TRIP**	**IVT**
Acronym	RNA affinity purification	Capture hybridisation analysis of RNA targets - mass spectrometry	Identification of direct RNA interacting proteins	Multiplexed hybridisation purification of RNA-protein complexes for MS	Comprehensive identification of RNA binding proteins by mass spectrometry	Peptide nucleic acid (PNA)-assisted identification of RBPs	Tandem RNA isolation procedure	In vitro transcribed
Example published RNA	*Xist*	*Malat1, Neat1*	*Terra, U1*	*Malat1, Neat1, Norad*	*Xist*	*Ankylosis*	*Cep, mpk, Pgk*	*Malat1*
Cell numbers	200-800 million	100 million	150 million	100 million	100-500 million	2-10 million	Not specified	n/a
Crosslinking	UV	3% FA 30 minutes	UV	1% FA 10 minutes	3% FA 30 minutes	UV	UV	n/a
Probe size	90	25	20-25	25-30	20	PNA probes	21-24	500 bases of RNA of interest fused to biotin Aptamer
Number of probes	142	2	9	2	43	1	1	14

MS labelling approach	SILAC	Label free	Isobaric labelling	Label free	Label free	Label free	Label free	DML labelling
MS cut off	Three-fold Xist vs U1	Ten-fold over input	Two-fold over control	Five-fold over control	Ten-fold over background	Not specified	Immunoblot analysis	1.7-fold over control
Advantages	*In vivo* Fast, do not have to spend time mapping probes UV crosslinking identifies primary interactions	*In vivo* Reversible crosslinking	*In vivo* Can be used for small RNAs due to probe size UV crosslinking identifies primary interactions	*In vivo* Can analyse multiple RNA targets from the same cell culture.	*In vivo,* easy probe design, reversible crosslinking can analyse the same sample by multiple methods.	*In vivo,* PNA have high specificity and form stable complexes with RNA	Reduces background through polyA if interested in mRNA interactions DNA-RNA oligos used for capture which have high specificity	Speed and ease, can use mutagenesis studies to discover binding sites

Disadvantages	High cell number input, expensive probes due to length and number, can only look at one isoform of RNA	RNAse H assay required for probe design, FA crosslinking could identify secondary interactions	Time to design and test probes, potential false positive interactions due to probe size	FA crosslinking could identify secondary interactions.	Short probe length may increase non-specific interactions, FA use also increases chances of identifying secondary interactions	PNA probes not as widely accessible	Uses oligodT purification, this would exclude some lncRNAs. Knowledge of secondary structure required for probe design	Not *in vivo*, lacks modifications and structure, potential non-specific interactions
References	(McHugh et al., 2015; McHugh & Guttman, 2018)	(West et al., 2014a)	(H. P. Chu et al., 2021)	(Spiniello et al., 2018)	(C. Chu, Zhang, Da Rocha, et al., 2015)	(Zielinski et al., 2006)	(Matia-González et al., 2017)	(Scherer et al., 2020)

5.1.4 Aims

To gain insight into the mechanism by which *Malat1* alters gene expression in Th cells understanding the *Malat1*-protein interactome is imperative. This chapter aims to identify which proteins interact with *Malat1* by employing RAP-MS tools in primary and oncogenically transformed $CD4^+$ T cells.

5.2 Results

5.2.1 RBP expression changes during Th cell differentiation

To deepen our molecular understanding of the role of RBPs in $CD4^+$ T cells we began by analysing RNA sequencing data from *in vitro* polarised Th cells (chapter 4), and whole cell proteomics data from the Immunological Proteomic Resource (ImmPRes) database. RBPs were defined by RBPBase an online database of experimentally validated RBPs (Gebauer et al., 2020). This determined the protein and RNA levels of total proteins and RBPs in $CD4^+$ T cells **(*Figure 5.1)***. Interestingly, the RNA transcripts of RBPs are some of the most abundant transcripts in $CD4^+$ T cells **(*Figure 5.1A)***. Further, on analysis of whole cell proteomic datasets in naïve, *in vitro* polarised Th1 and Th2 cells RBPs were again among the most highly expressed proteins within Th cells **(*Figure 5.1B)*.**.** The abundance of RBPs in different $CD4^+$ T cell states highlights their potential importance in $CD4^+$ T cell functions.

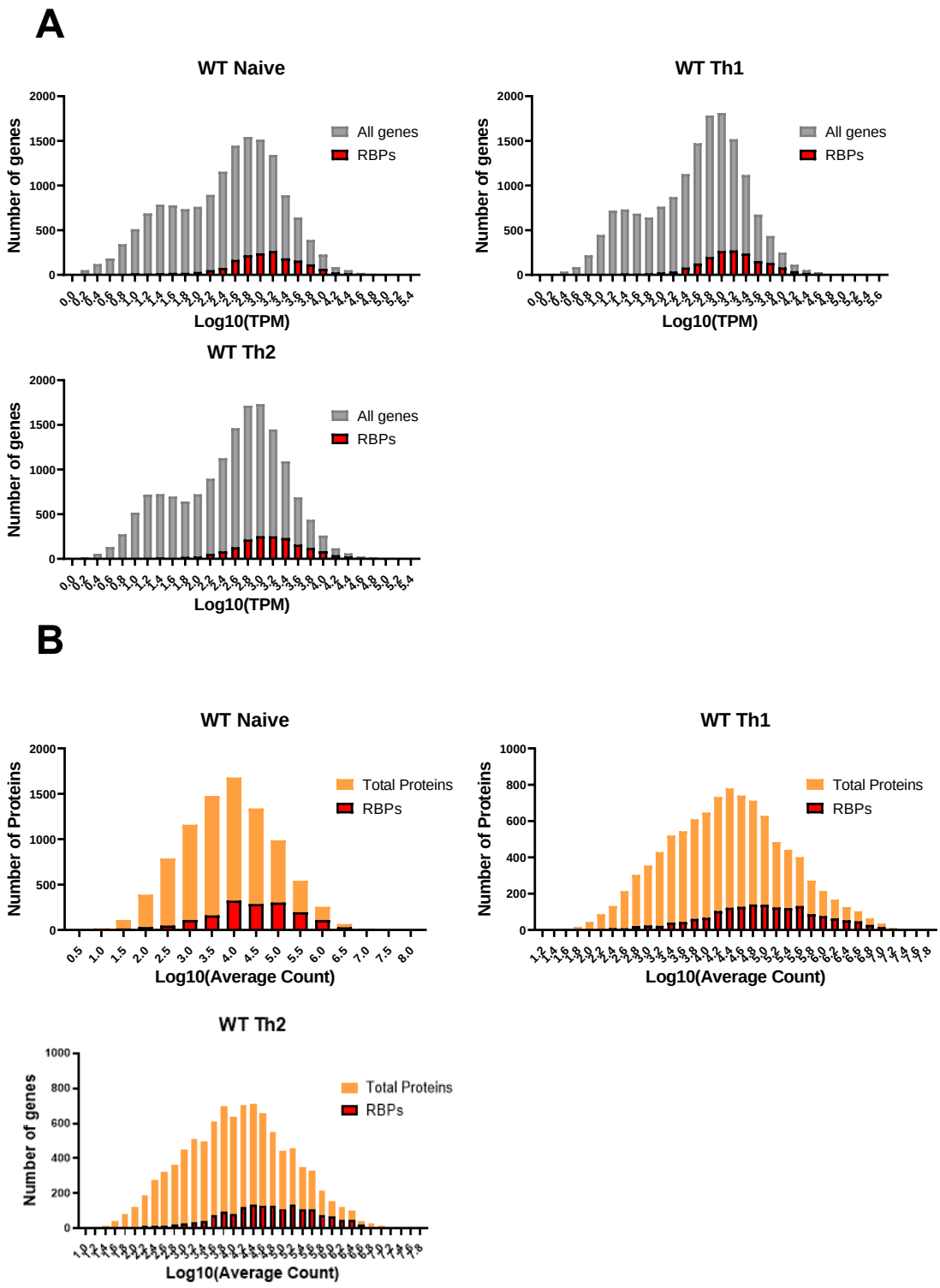

A
WT Naive
WT Th1
WT Th2
All genes
RBPs
Number of genes
Log10(TPM)
B
WT Naive
WT Th1
WT Th2
Total Proteins
RBPs
Number of Proteins
Number of genes
Log10(Average Count)

Figure 5.1: RBPs are highly expressed in CD4$^+$ T cell subsets at both the protein and RNA levels.

A) Histogram depicting log10 of transcript count per million of genes detected in WT naïve, Th1 and Th2 CD4$^+$ T cells by RNA sequencing as described in chapter 4. Total detectable genes are highlighted in grey. Murine RBPs defined by RBPbase are shown in red. An RNA detection threshold of >10 CPM was applied. **B)** Histogram depicting log10 of average protein count of proteins detected in WT naïve, Th1 and Th2 CD4$^+$ T cells by whole-cell proteomics as defined by the ImmPRes database. Total detectable proteins are highlighted in orange. Murine RBPs defined by RBPbase are shown in red.

RNA sequencing data were further analysed to determine differential expression of RBPs during Th cell differentiation. We identified 850 RBPs that were differentially expressed during Th1 differentiation and 896 RBPs that were altered during Th2 differentiation at a LogFC >0.5 and q<0.1 (***Figure 5.2****A).* A similar number of RBPs are upregulated and downregulated during Th1 and Th2 differentiation. The expression of 727 RBPs were altered during both Th1 and Th2 differentiation. However, distinct RBPs were also differentially expressed during Th1 or Th2 differentiation (***Figure 5.2****B).* It is plausible that distinct networks of RBPs are responsible for optimal Th cell differentiation. In chapter 4 we have shown that *Malat1*$^{-/-}$ T cells have a differentiation defect. As *Malat1* is thought to function through interaction with proteins and RBPs have previously been shown to have a functional role in CD4^{+} T cells we next examined the *Malat1* protein interactome in a CD4^{+} T cell context (Brown et al., 2016; Kanbar et al., 2022; Liu et al., 2020).

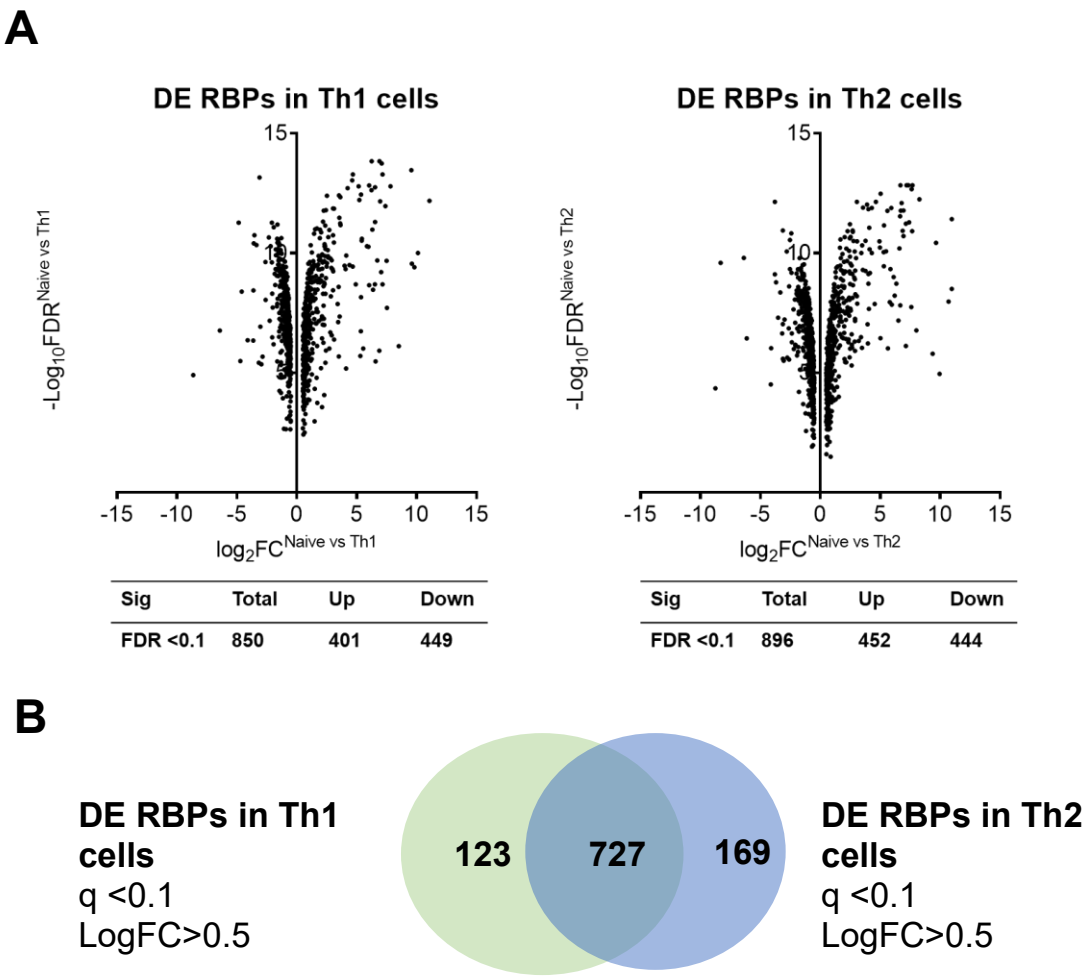


Figure 5.2: RBPs are differentially expressed during Th1 and Th2 *in vitro* polarisation.

A) Volcano plots depicting RBP expression profile of *in vitro* polarised WT Th1 and Th2 cells from RNA-sequencing data set described in chapter 4. RBPs were defined by RBPbase and a logFC >0.5 and q<0.1 was applied. Fold change is determined as log2 of mean FPKM (WT Naïve vs WT Th1/Th2). Associated tables depict the total number of RBP genes which change (up or down) at a q <0.1. **B)** Venn diagram showing overlap of RBP genes that are differentially expressed during WT Th1 or Th2 differentiation LogFC>0.5, q<0.1.

5.2.2 RAP-MS identifies *Malat1* protein interaction partners in EL4 cells

We began initial RAP-MS experiments using EL4 cells. These are a murine T cell tumour cell line that was derived from chemically induced lymphoma in C57BL/6 mice. This was for several reasons; large cell numbers can be acquired in a short time scale and with relative ease. This enabled protocol optimisation and elucidated the potential feasibility of RAP-MS in primary cells. A simplified schematic of the RAP-MS workflow is outlined below (***Figure 5.3***). Briefly, cells are exposed to UV light to create a covalent cross-linkage between RNA and protein, the cells are lysed and *Malat1* is pulled down using long antisense biotinylated probes, after denaturing purification samples are sent for mass spectrometry. This chapter will compare the *Malat1* protein interactome in EL4 cells and primary naïve $CD4^+$ T cells. Initially, we will describe the process to identify the *Malat1*-protein interactome in EL4 cells.

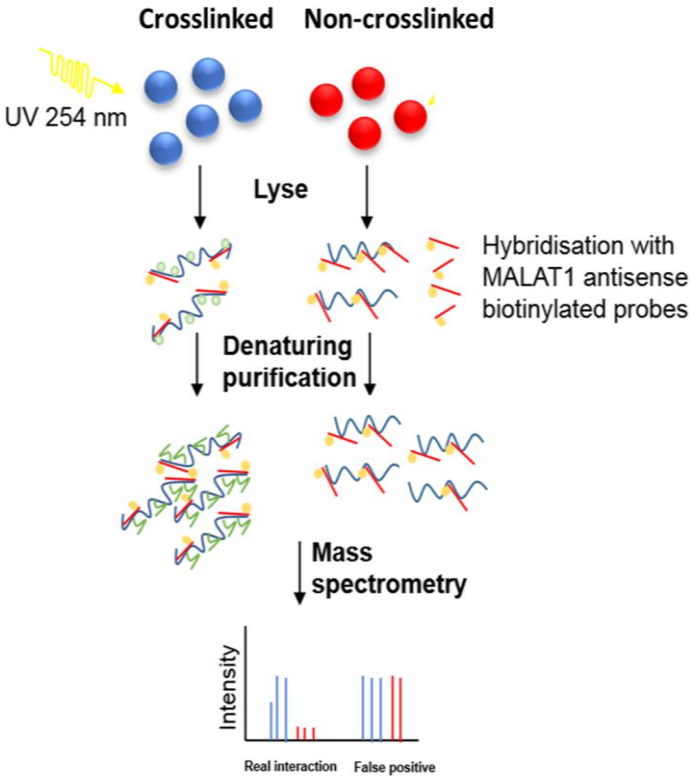


Figure 5.3: Schematic representation of the proposed RAP-MS workflow.

Briefly, cells are *in vivo* UV-irradiated to generate covalent bonds between RNA and its interacting proteins. *Malat1* is then hybridised with long antisense biotinylated probes and pulled down using streptavidin beads. After stringent denaturing purification samples are analysed by mass spectrometry. Crosslinked samples will be compared to a non-crosslinked control. Proteins which are significantly enriched in the crosslinked samples are deemed *Malat1,* interaction partners.

The initial step of the RAP-MS protocol generates a covalent cross-linkage between RNA and its interacting proteins by exposing cells to UV light. Initially, two crosslinking strategies were compared using either a Minitron or Stratalinker 2000 machine. As crosslinking halts the cellular translational machinery, it is anticipated that cell growth will be impaired (De Pablos et al., 2019). Thus, to determine if crosslinking had been successful cell growth was compared between crosslinked and non-crosslinked samples (***Figure 5.4***). In both conditions, crosslinking was successful and cell growth was impaired upon UV exposure. However, a small portion of crosslinked cells (XL) restored culture growth after either 60s or 120s UV-exposure, with cell viability largely unaffected using the Minitron. Contrastingly, growth and viability are almost completely abolished post-irradiation using the Stratalinker. This is likely due to the greater heat stress produced by the Stratalinker (~100x the heat (~150mJ/cm^2)). As the ultimate goal of this project is to identify *Malat1*-RBP interactions in primary CD4$^+$ T cells the Stratalinker was used for future crosslinking as it caters to the use of lower cell numbers.

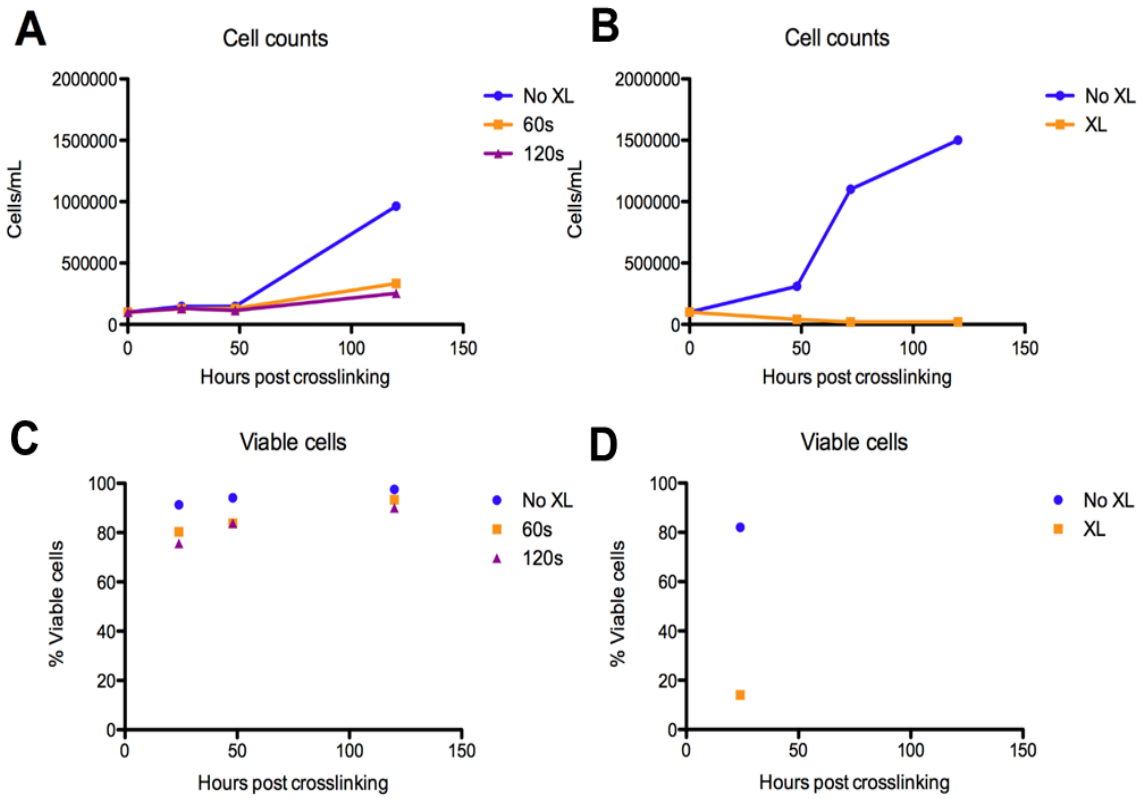


Figure 5.4: Validation of UV crosslinking through impairment of cell growth.

A) Cell counts at various time points post crosslinking using the Minitron **B)** Cell counts at various time points post crosslinking using the Stratalinker 2000. **C)** Percentage viable cells measured by PI staining using flow cytometry after crosslinking using the Minitron **D)** or Stratalkiner. n=3 technical replicates of 1 idependent experiment.

We proceeded with the RAP-MS protocol. Analysis of RNA fractions revealed that *Malat1* was enriched up to ~8,000-fold in the eluted samples (***Figure 5.5***). Moreover, *Malat1* was not enriched in the flow-through samples, adding confidence to the success of *Malat1* pulldown. *Malat1* enrichment was slightly elevated in the crosslinked samples compared to the non-crosslinked controls. Creating a strong covalent bond between *Malat1* and its protein interaction partners during crosslinking likely helps protect the *Malat1* transcript from degradation throughout the heating steps of the RAP capture. This is reflected in the observation that the *Malat1* Ct values increased after heating in the non-crosslinked samples which suggested that the RNA may have partially degraded.

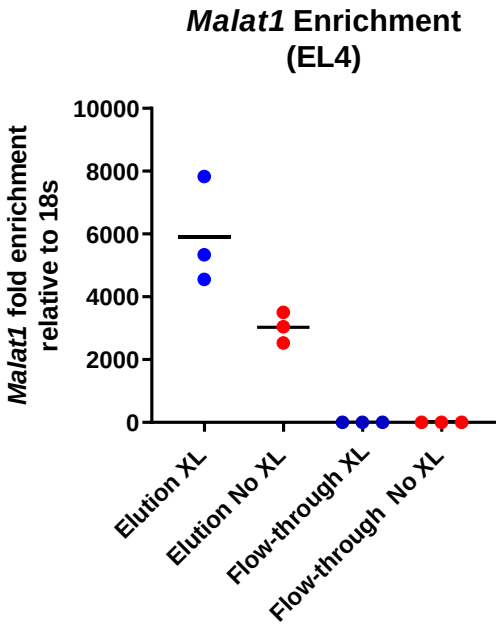


Figure 5.5: RAP enriches *Malat1* in eluted samples.

A) Fold enrichment of *Malat1* relative to 18s after RAP-MS capture in EL4 cells as determined by qRT-PCR. n=3 biological replicates of one independent experiment. The mean is displayed.

High-resolution mass spectrometry identified the captured proteins. Analysis thresholds were set at a QI Quant filter minimum of 2 unique peptides and a mascot percolator of 1% FDR. In EL4 samples 152 proteins were identified under these thresholds, many proteins appeared in both the crosslinked and non-crosslinked samples. 36 of these 150 proteins were significantly enriched in the crosslinked samples at q<0.1 (***Figure 5.6***A/B). As the fold increase of the enriched proteins was many times greater than the slight differences in *Malat1* enrichment between crosslinked and non-crosslinked samples, we are confident that the proteins identified are true *Malat1* interaction partners. In addition, it is encouraging that the enriched proteins appeared in the non-crosslinked samples at lower intensities. This is because some proteins likely form weak interactions with *Malat1* within a cell and remain intact throughout the RAP capture. Therefore, the presence of these proteins in both the crosslinked and non-crosslinked samples supports the biological relevance of the identified proteins.

To add confidence that true *Malat1* interaction partners had been identified by RAP-MS, significantly enriched proteins were compared with RBPbase (Gebauer et al., 2020). Reassuringly, of the 36 total proteins identified in EL4 cells, 35 proteins have previously been identified as RBPs. Collectively, this indicated that the RAP-MS had been successful and true interaction partners had been identified as RBPs have been enriched and no other proteins such as structural or mitochondrial proteins were identified in this dataset ***(Table 5.2).*** A literature search and analysis of online databases found that the majority of proteins had previously been experimentally shown to interact with *Malat1* in other cell types (either human or mouse). Moreover, the majority of these proteins are nuclear at some stage in their life time ***(Table 5..3)***.

STRING analysis revelaed that *Malat1* interacting RBPs are intricately linked in a network (***Figure 5.6***C). Notably, two families of RBPs appeared prominently in this dataset - the SR and hnRNP families of proteins. These are both large families of RBPs involved in nucleic acid metabolism, transcription/translational regulation and splicing (Geuens et al., 2016; Jeong, 2017). The hnRNP family represented 36.1% of enriched proteins and the SR family of proteins represented 13.9% of enriched proteins. As these proteins appeared frequently in these datasets, this indicated they could be important for *Malat1*-dependent Th cell functions.

GSEA was used to examine the molecular function of the RBPs bound to *Malat1* in EL4 cells (***Figure 5.6***D). Notably, all of the proteins in the datasets had RNA binding capabilities (***Figure 5.6***D). As expected, the proteins identified were shown to have functions requiring a wide range of RNA interactions including RNA processing, mRNA metabolism, and regulation of RNA splicing (***Figure 5.6***D). *Malat1* has previously been implicated in regulating AS (Tripathi et al., 2010). *Malat1* is located in nuclear speckles – a site enriched in splicing factors. *Malat1* has been shown to interact with SR proteins including SRSF1, SRSF2, and SRSF3 (also identified as interaction partners in both of our Th cell datasets) (Tripathi et al., 2010). Moreover, the depletion of *Malat1* impaired SR protein phosphorylation and expression (Tripathi et al., 2010). Yet, the extent to which *Malat1* plays a role in alternative splicing is unclear. Previous studies have shown alternative splicing is not impaired in embryonic fibroblasts derived from *Malat1*$^{-/-}$ mice (Zhang et al., 2012). Additionally, *Malat1*$^{-/-}$ mice have no developmental defects (Eißmann et al., 2012; Nakagawa et al., 2012; Zhang et al., 2012). In contrast, many heterogeneous nuclear ribonuclear protein splicing factor knockout mice are embryonic lethal (Komeno et

al., 2015; Liu et al., 2017). It is possible that *Malat1* interactions with these proteins affects functions outside of AS.

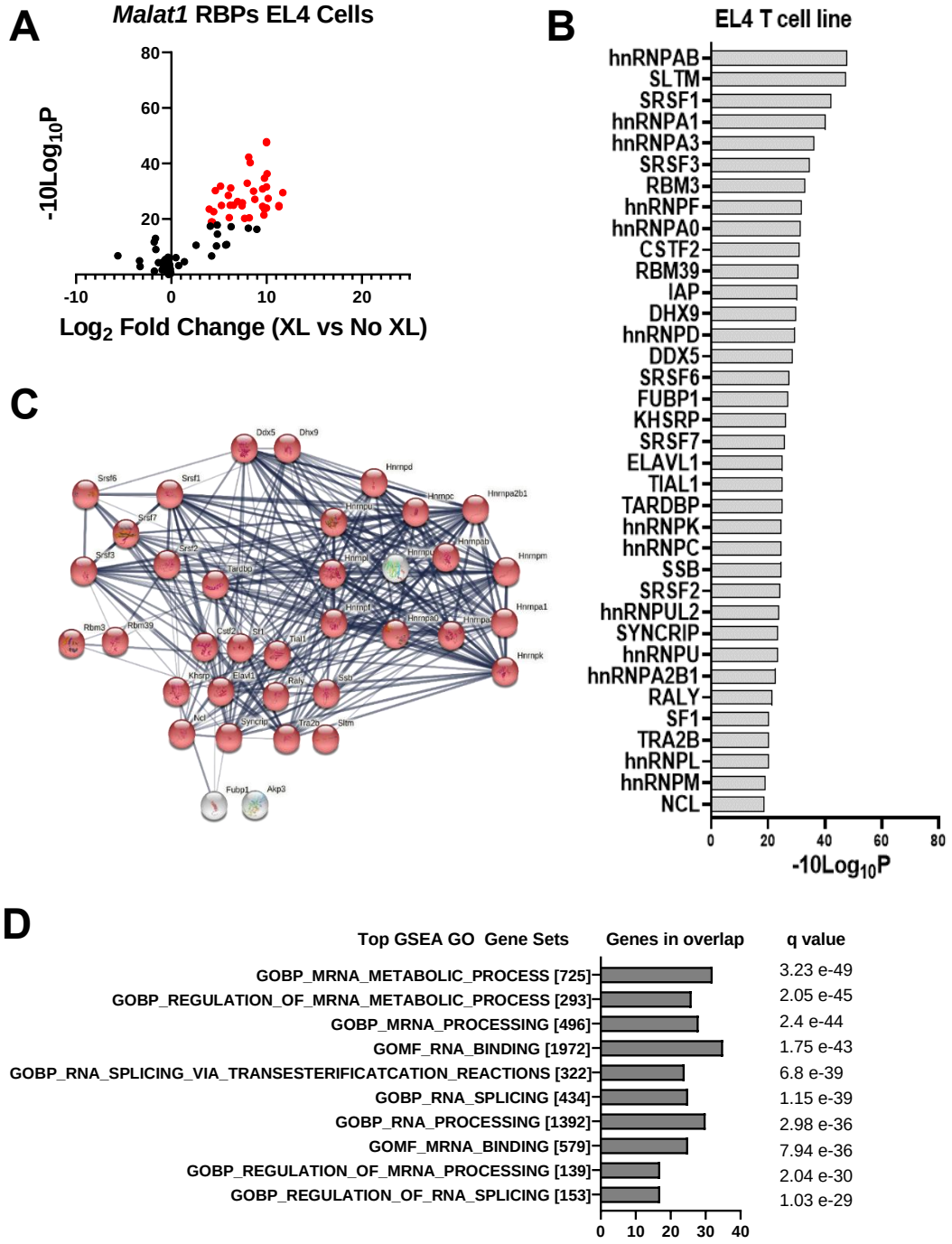


***Figure 5.6: Malat1* prominently interacts with hnRNP and SR proteins in EL4 cells.**

A) Comparison of the log2 Fold change (crosslinked samples vs non crosslinked samples) plotted against -10Log10p. Red dots indicate significantly enriched RBPs in crosslinked samples (q<0.1), in EL4 cells **B)** -10Log10p of named RBPs which were significantly enriched in crosslinked samples EL4 cells **C)** STRING interaction of RBPs interacting with *Malat1* in EL4 cells at q<0.1 are depicted. **D)** GSEA GO enrichment analysis of RBPs interacting with *Malat1* in EL4 cells at q<0.1 are depicted.

Protein motifs or RNA structures could influence the frequency of binding. To assess this possibility, RBP domains were next examined. Several common RBDs were found amongst the proteins which bound *Malat1* in EL4 cells, for example ZnF domains. Strikingly RRM domains appeared frequently in many of the RBPs shown to interact with *Malat1 (**Figure 5.7**)*.It is possible that *Malat1* preferentially binds RRM domains, however, these are the most common and well-studied of the RBDs and are thought to occur in 1% of all human proteins (Corley et al., 2020).

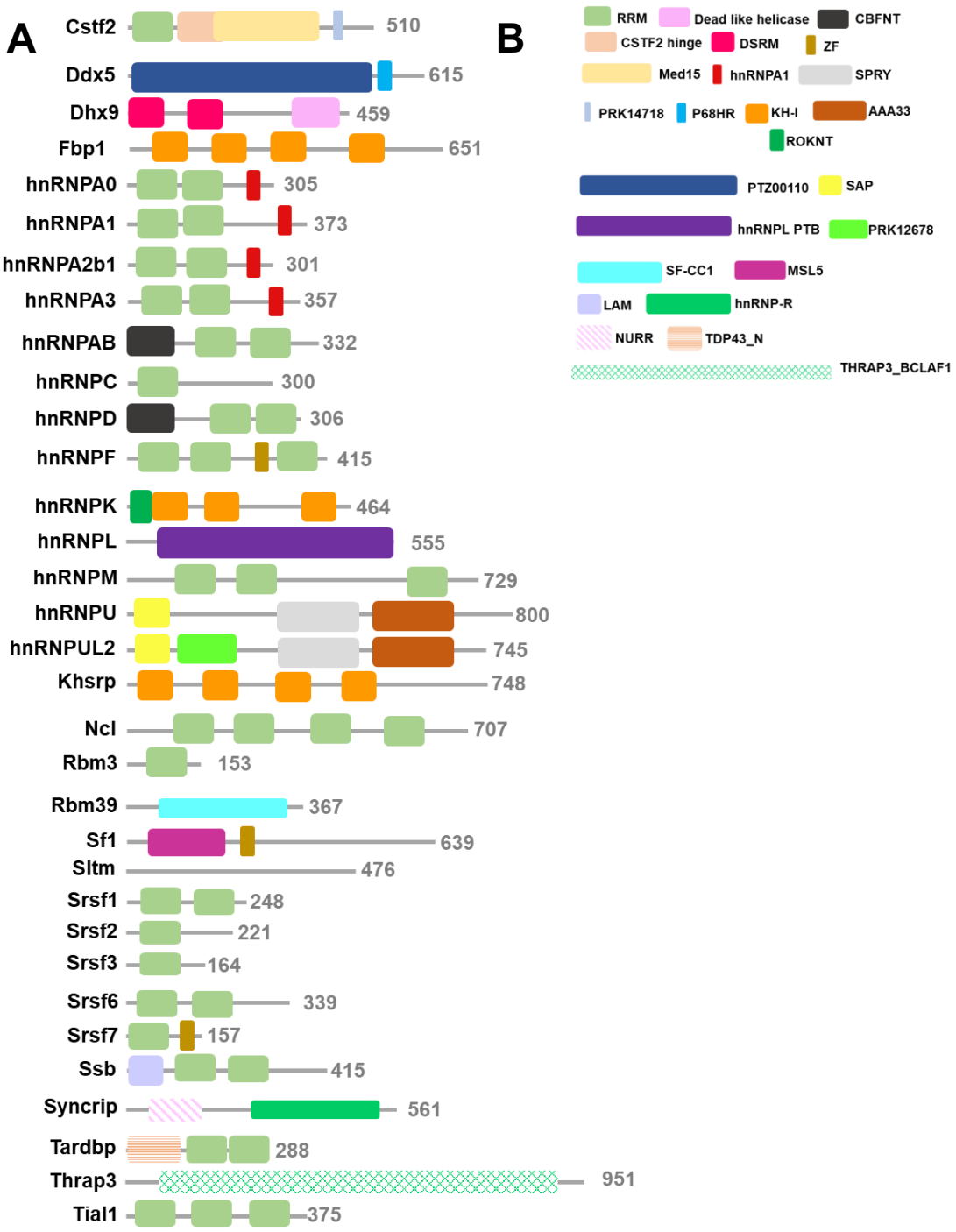


Figure 5.7: *Malat1* interacts with proteins that have RRM domains.

A) Schematic representation of *Malat1* interacting RBPs length and domains. **B)** Key displaying identified protein domains. Not to scale.

5.2.3 RAP-MS identifies *Malat1* protein interaction partners in primary naïve CD4$^+$ T cells

Given the success of RAP-MS in EL4 cells, we next wanted to explore the *Malat1* protein interactome in primary CD4$^+$ T cells. To determine feasibility, a small-scale pilot study in primary naïve CD4$^+$ T cells was performed. As primary cell numbers are more limited the aim of this pilot study was to determine if fewer cells could be used for this experiment. 10 million cells were used to create a crosslinked and non-crosslinked sample, and RAP-MS was employed. This pilot study identified many of the RBPs that *Malat1* interacted with in EL4 cells to be enriched in this sample, indicating successful pulldown ***(Figure 5.8)***. Some RBPs were not enriched in this sample – however, this could be explained by differences in binding between cell types. Of note attempts with 5 million cells did not identify an increase in RBP binding in crosslinked samples compared to non-crosslinked controls, this suggested a larger cell number would be required to be able to identify proteins. Thus, to ensure sufficient signal was generated in RAP-MS experiments in primary naïve CD4$^+$ T cells 20 million cells were used.

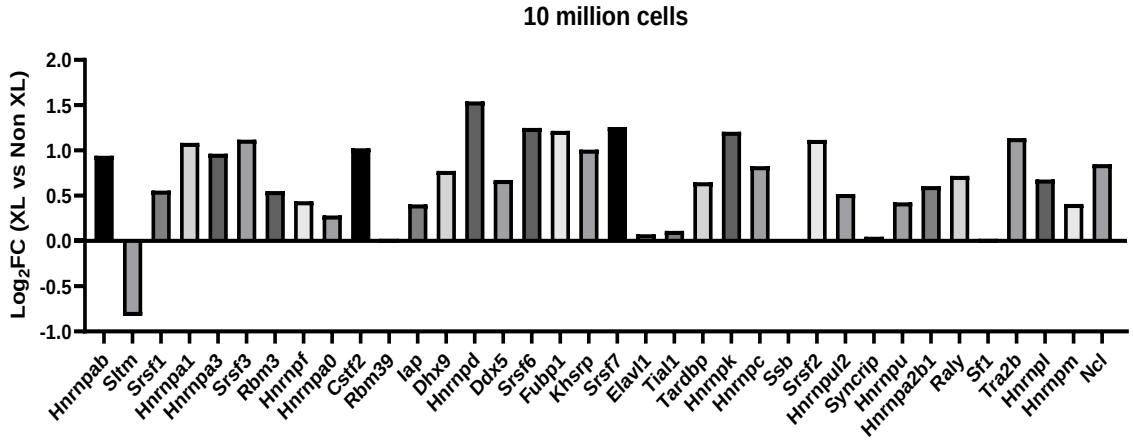


Figure 5.8: Pilot RAP-MS studies enrich RBPs in primary naive CD4$^+$ T cells.

Log_2FC (XL vs non XL samples) of RBPs which were shown to interact with *Malat1* in EL4 cells, in a pilot RAP-MS study using 10 million naïve primary CD4$^+$ T cells. n=1.

As the pilot study was successful, we proceeded with the RAP-MS protocol in primary naïve $CD4^+$ T cells which were isolated using a magnetic miltenyi $CD4^+$ naïve T cell isolation kit. This cell type was chosen as they express the highest levels of *Malat1* and they are the common origin of effector cells. Analysis of RNA fractions revealed that *Malat1* was enriched up to ~8,000-fold in the eluted samples (***Figure 5.9***). Moreover, *Malat1* was not enriched in the flow-through samples. Similarly, to the EL4 pull down, *Malat1* enrichment was slightly elevated in the crosslinked samples compared to the non-crosslinked controls, which again suggested that crosslinking prevented RNA degradation.

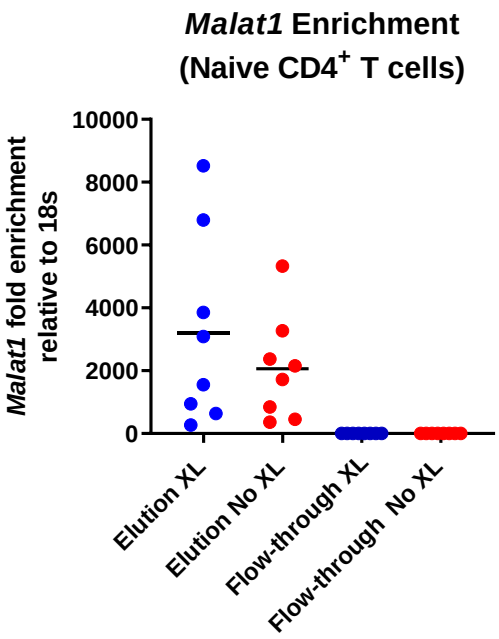


Figure 5.9: RAP enriches *Malat1* in primary naïve CD4⁺ T cells.

Fold enrichment of *Malat1* relative to *18s* after RAP-MS capture in naïve CD4⁺ T cells as determined by qRT-PCR. n=8 biological replicates representative of 1 experiment.

In primary naïve $CD4^+$ T cells, 87 proteins were identified by mass spectrometry, with 28 proteins significantly enriched in crosslinked samples at a threshold of q<0.1 ***(Figure 5.10A/C)***. Moreover, of the 28 proteins found to bind to *Malat1* in primary $CD4^+$ T cells, 26 proteins were logged in RBPbase. Additionally, a literature search and analysis of online databases found that the majority of proteins had previously been shown to interact with *Malat1* in other cell types and are commonly found in the nucleus ***(Table 5.2)*.**

STRING analysis of the identified proteins revealed that the proteins were intricately linked in a network (***Figure 5.10*C**). Similar to EL4 cells the same two families of RBPs appeared prominently in this data set the SR and hnRNP families of proteins. The hnRNP family represented 32.1% of the enriched proteins in primary Th cells The SR family of proteins represented 17.9% of enriched proteins in primary cells.

GSEA was used to examine the molecular function of the RBPs bound to *Malat1* in primary $CD4^+$ T cells (***Figure 5.10D)***. The majority of proteins in the datasets have RNA binding capabilities (***Figure 5.10D)***. The proteins identified were shown to have functions in a range of RNA centric processes such as RNA processing, mRNA metabolism, and regulation of RNA splicing (***Figure 5.10D)***.

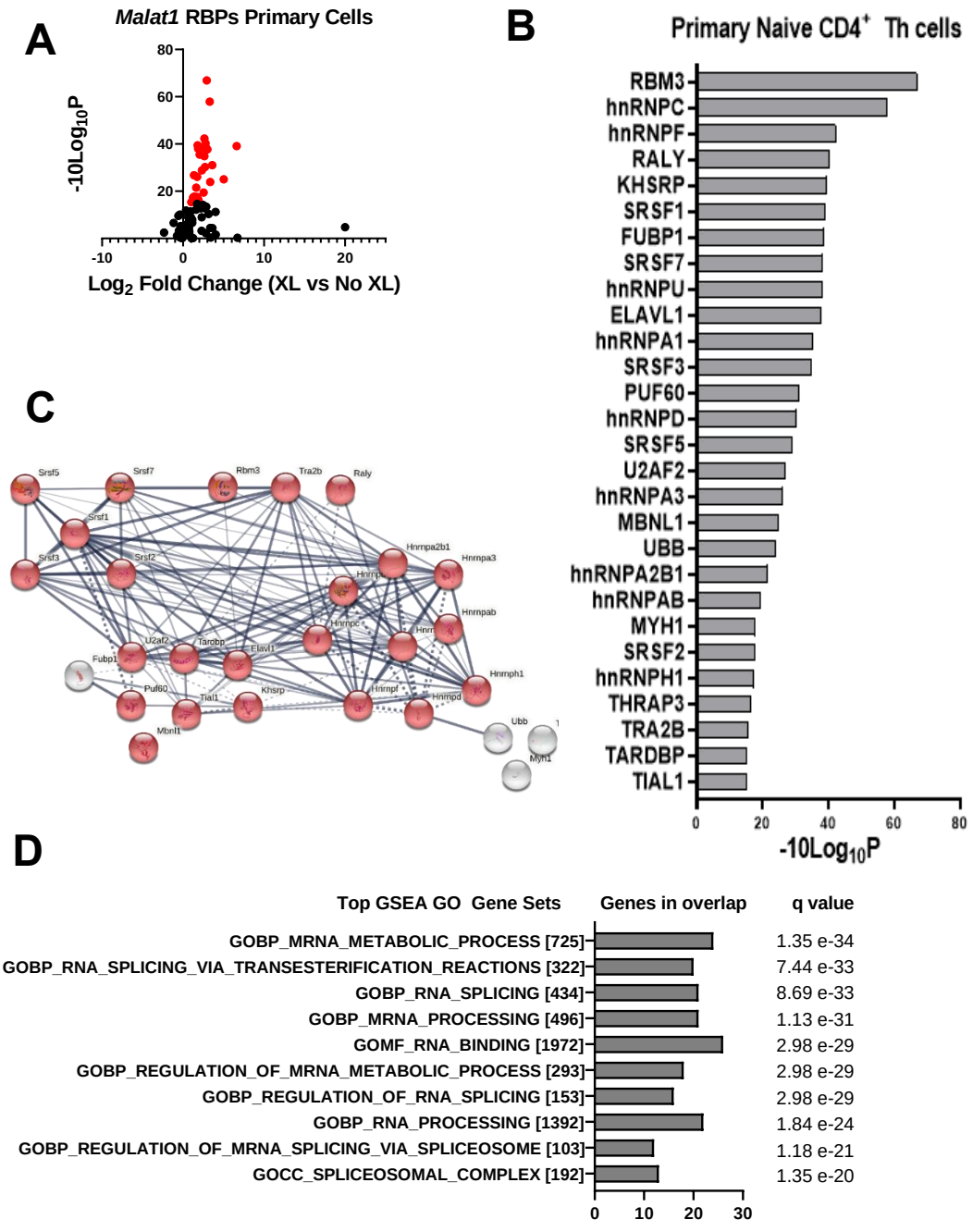


Figure 5.10: *Malat1* interacts with members of the hnRNP and SR family of proteins in primary naïve CD4⁺ T cells.

A) Comparison of log2 Fold change (crosslinked samples vs non-crosslinked samples) plotted against -10Log10P. Red dots indicate significantly enriched RBPs in crosslinked samples (q<0.1), in naïve CD4⁺ T cells **B)** -10Log10P of named RBPs which are significantly enriched in crosslinked samples naïve CD4⁺ T cells **C)** STRING interaction of RBPs interacting with *Malat1* in naïve CD4⁺ T cells at q<0.1 are depicted. **D)** GSEA GO enrichment analysis of RBPs interacting with *Malat1* in naïve CD4⁺ T cells at q<0.1 are depicted.

To determine if *Malat1* interacting RBPs in naïve $CD4^+$ T cells were differentially regulated during Th cell differentiation we further examined RNA sequencing data from *in vitro* polarised WT Th cells (***Figure 5.11****A*). Several of *Malat1* interacting RBPs were differentially regulated, and these skewed towards downregulation. As such it is possible that *Malat1* RBP interactions change during the T cell differentiation process. If the changes in RNA levels correlate with protein expression differentiated Th cells may interact with other RBPs in differentiated cells. Alternatively, the stoichiometry of interactions may be altered.

Notably, loss of *Malat1* had minimal impact on the mRNA levels of these RBPs (***Figure 5.11****B).* Only a handful of RBPs showed any significant expression changes at an $q<0.05$. mRNAs encoding RBM3, hnRNPA1, and hnRNPA3 were significantly upregulated in *Malat1*$^{-/-}$ *in vitro* polarised Th2 cells, however, this would need to be confirmed at the protein level. Yet, this does not mean RBP functionality remains unchanged, for example, SR protein phosphorylation is altered in the absence of *Malat1* (Tripathi et al., 2010c). Further investigation of RBP function in the presence and absence of *Malat1* is required.

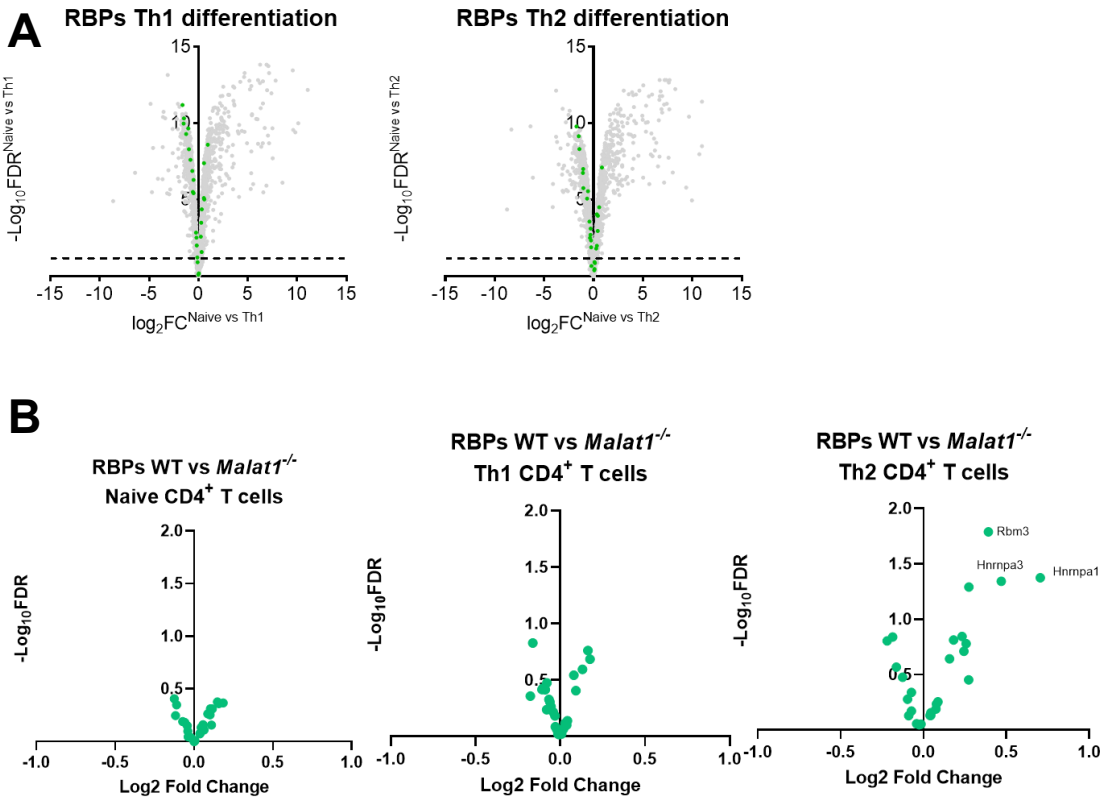


Figure 5.11: *Malat1* interacting RBPs are moderately differentially regulated during Th cell differentiation.

A) Volcano plots depicting RBP expression profile of *in vitro* polarised WT Th1 and Th2 cells from an RNA-sequencing data set. RBPs were defined by RBPbase total RBPs are shown in grey. Fold change is determined as log2 of mean FPKM (WT Naïve vs WT Th1/Th2). Coloured dots indicated RBPs which interact with *Malat1* in naïve CD4$^+$ T cells (green). **B)** Volcano plots depicting naïve CD4$^+$ *Malat1* interacting RBP expression profile of *in vitro* polarised Th1 and Th2 cells from an RNA-sequencing data set. Fold change is determined as log2 of mean FPKM (WT vs *Malat1*$^{-/-}$ Naïve, Th1/Th2). Labelled points are significantly differentially expressed at q<0.1.

To determine if *Malat1* bound to the most highly expressed proteins, the abundance at both the RNA and protein levels was compared with enrichment of proteins in the primary T cell data set. Of note, no correlation was observed between levels of binding and expression (***Figure 5. 12)***. This indicated that other factors may have influenced which proteins *Malat1* interacts with. To assess the influence of domains on binding, protein structure was next examined. Many RBDs were identified. However, RRM domains appeared the most frequently across many of the RBPs shown to interact with *Malat1* in naïve $CD4^+$ T cells ***(Figure 5.13).***

Although the expression of some proteins may be higher the nuclear to cytoplasmic ratio may differ, which would impact the intensity of *Malat1* binding. In addition, when calculating the fold change between crosslinked and non-crosslinked samples this does not take into account RBPs which may strongly interact with *Malat1* in the absence of UV light. These proteins would also be pulled down in the cross-linked samples, however, the fold change would be skewed. Therefore, the proteins with the higher fold changes may not always represent proteins that have the strongest interactions. Of note, some proteins such as UBB, PUF60, and DDX9 were not detected in the primary naïve $CD4^+$ proteomic data set.

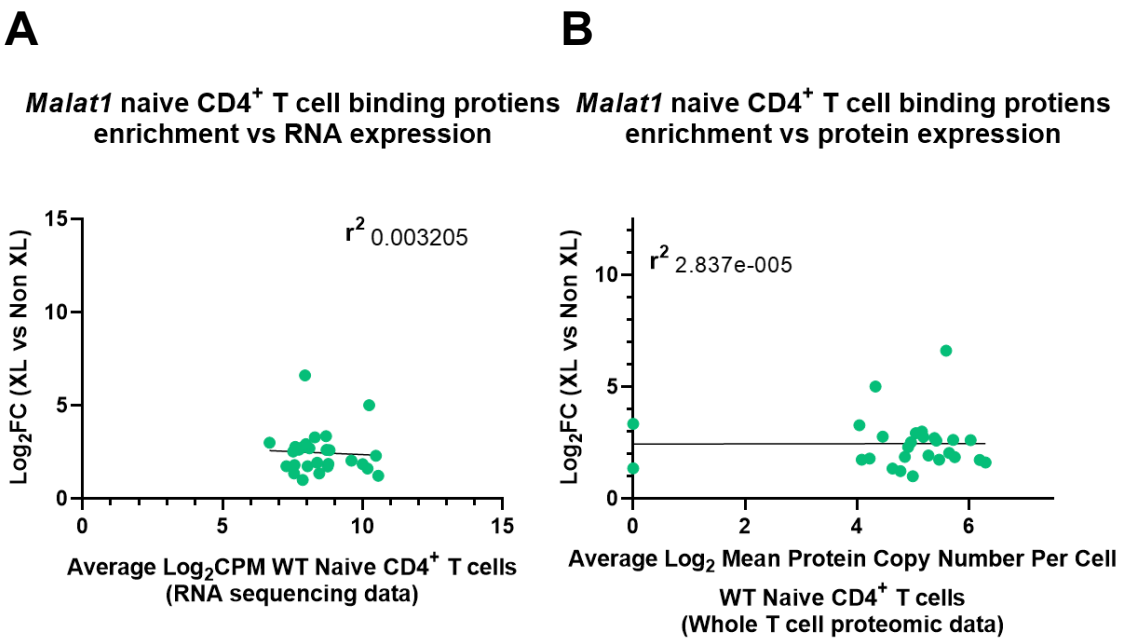


Figure 5. 12: Expression of RBPs does not correlate with enrichment in naive CD4⁺ T cells.

A) Log_2FC of RBPs which bind to *Malat1* in naïve $CD4^+$ T cells determined by RAP-MS plotted against average Log_2CPM of RNA sequencing data of WT naïve $CD4^+$ T cells. Correlation coefficient is displayed. **B)** Log_2FC of RBPs which bind to *Malat1* in naïve $CD4^+$ T cells determined by RAP-MS plotted against average Log_2 Protein copy number per cell of whole cell proteomic data of WT naïve $CD4^+$ T cells. Correlation coefficient is displayed

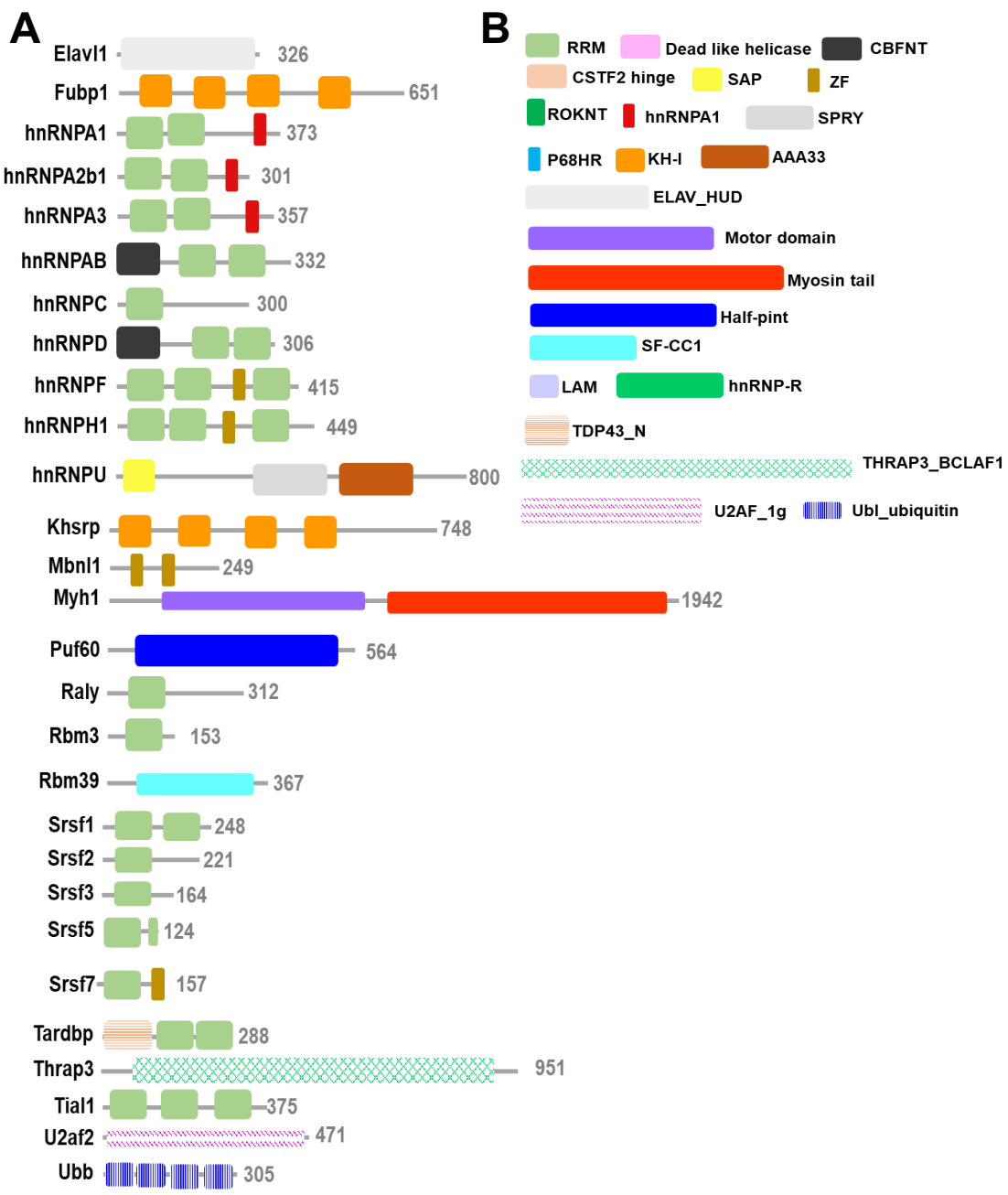


Figure 5.13: *Malat1* interacts with proteins that have RRM domains in primary naïve CD4⁺ T cells.

A) Schematic representation of *Malat1* interacting RBPs length and domains.
B) Key displaying identified protein domains. Not to scale.

5.2.4 Comparison of *Malat1* interaction partners in EL4 cells and naïve CD4$^+$ T cells

We next compared the *Malat1* interacting proteins in EL4 cells and naïve CD4$^+$ T cells to determine the degree of overlap between these cell types (***Figure 5.14***). Of note, the majority of proteins interacted with *Malat1* in both primary and EL4 cells (***Table 5.)***. This suggested that the common binders across EL4 and primary cells constitute core protein interactions of *Malat1* in CD4$^+$ T cells (***Figure 5.14)***. Yet, ~1/3 of *Malat1* binding proteins were unique to primary cells including (hnRNPH1, MBNL1, U2FA2, SRSF5, and PUF60). This indicated that the *Malat1*-protein interactome has some cell-type-specific binding. It is plausible this arose due to differences in RBP expression levels, phosphorylation status, splice variant, localisation or other interactors. Alternatively, as EL4 cells are a T cell lymphoma line significant cellular changes can arise between cells that have been grown in long-term cell culture vs primary cells, which could also impact the *Malat1* protein interactome (Stewart et al., 2019). Many of these RBPs have been shown to be functional in CD4$^+$ T cells, both proteins which bind *Malat1* in EL4 cells and naïve CD4$^+$ T cells, suggesting *Malat1* could influence CD4$^+$ T cell interactions through these RBPs (***Table 5.)***.

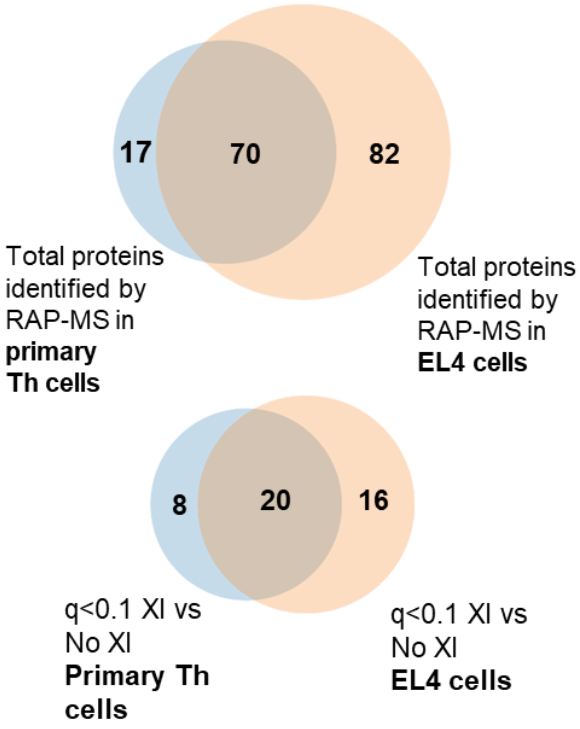


Figure 5.14: *Malat1* binds a core set of RBPs in CD4$^+$ T cells.

A) Venn diagrams depicting overlap of total proteins identified by RAP-MS in naïve CD4$^+$ T cells and EL4 cells. **B)** Venn diagrams depicting overlap of proteins which are enriched in crosslinked samples at a q value <0.1 in naïve CD4$^+$ T cells and EL4 cells identified by RAP-MS.

Table 5.2: *Malat1* binding proteins in EL4 and CD4$^+$ T cells

Table listing RBPs shown to interact with *Malat1* in EL4 cells and naïve CD4+ T cells, their localisation and basic function as defined by Uniprot database. If the protein occurs in RBPbase (RBP?) this is noted. Q Value and fold change is displayed. Previous data demonstrating binding to *Malat1* is also noted.

RBP	RBP?	NAÏVE CD4$^+$ /EL4	LOCALISATION	FUNCTION	FOLD CHANGE NAÏVE CD4$^+$	Q VALUE NAÏVE CD4$^+$	FOLD CHANGE EL4	Q VALUE EL4	BINDS *MALAT1*?	REFERENCE
CSTF2	Y	EL4	Nucleus	mRNA processing			74.96	0.02	Y	(West et al., 2014b)
ELAVL1	Y	Both	Nucleus Cytoplasmic	HuR antigen, Binds and stabalises mRNA	7.97	0.001	90.97	0.04	Y	(Chen et al., 2017a)
DDX5	Y	EL4	Nucleus, Cytoplasm Spliceosome	mRNA processing, Splicing, Transcription/translational regulation			62.91	0.02	N	-
DHX9	Y	EL4	Nucleus Cytoplasm Cytoskeleton	Immune response, Transcription regulation.			393.27	0.02	Y	(Chen et al., 2017)

FUBP1	Y	Both	Nucleus	Transcriptional regulation	6.85	0.002	431.88	0.03	Y	(West et al., 2014b)
hnRNPA0	Y	EL4	Nucleus	Splicing			1000	0.02	Y	(Chen et al., 2017)
hnRNPA1	Y	Both	Nucleus Cytoplasm Extracellular	Splicing, mRNA stability, translational and transcriptional regulation	4.15	0.002	308.54	0.01	Y	CLIPdb
hnRNPA2b1	Y	Both	Nucleus Cytoplasm, Extracellular	Splicing, mRNA stability	3.07	0.03	21.63	0.05	Y	CLIPdb (Chen et al., 2017)
hnRNPA3	Y	Both	Nuclear	Splicing, RNA trafficking	3.31	0.01	1054.25	0.01	Y	(Chen et al., 2017a; Scherer et al., 2020) CLIPdb
hnRNPAB	Y	Both	Nuclear and Cytoplasmic	RNA trafficking, translational regulation	5.78	0.05	1000	0	Y	(Scherer et al., 2020)

hnRNPC	Y	Both	Nuclear	Splicing, Transcriptional regulation	9.8	0.00007	169.63	0.04	Y	CLIPdb
hnRNPD	Y	Both	Nuclear and cytoplasmic	mRNA decay, Telomere maintenance	6.53	0.01	3304.45	0.02	Y	CLIPdb
hnRNPF	Y	Both	Nuclear	Splicing, Telomere maintenance	6.2	0.002	35.78	0.02	Y	CLIPdb
hnRNPH1	Y	Naïve $CD4^+$	Nuclear	Splicing, pre-mRNA processing	2.35	0.07			Y	CLIPdb
hnRNPK	Y	EL4	Nucleus, cytoplasm, spliceosome, cell junctions	mRNA processing, splicing, transcription and translational regulation			2478.85	0.04	N	-
hnRNPL	Y	EL4	Nuclear and cytoplasm	Splicing			205.83	0.07	Y	CLIPdb
hnRNPM	Y	EL4	Nucleus, spliceosome	mRNA processing and splicing			18.49	0.09	Y	CLIPdb (Scherer et al., 2020)

hnRNPU	Y	Both	Nuclear and cytoplasmic	Splicing, transcriptional regulation	3.62	0.002	15.64	0.04	Y	CLIPdb (Chen et al., 2017a)
hnRNPUL2	Y	EL4	Nucleus	RNA binding			1000	0.04	Y	(Chen et al., 2017a; Spiniello et al., 2018)
KHSRP	Y	Both	Nuclear and cytoplasmic	Splicing, mRNA decay, translational regulation	3.47	0.002	123.21	0.03	Y	CLIPdb
IAP	N	EL4	Nucleus and cytoplasm	Apoptosis, Wnt signalling			24.39	0.02	N	-
MBNL1	Y	Naïve $CD4^+$	Nuclear and cytoplasmic	Splicing	32.49	0.02			Y	CLIPdb
MYH1	N	Naïve $CD4^+$	Cytoplasmic	Cell movement	2.56	0.07			N	-
NCL	Y	EL4					19.67	0.09	N	-
PUF60	Y	Naïve $CD4^+$	Nuclear	Splicing and transcriptional regulation	12.1	0.01			Y	(West et al., 2014b)
RALY	Y	Both	Nuclear	Heterogeneous ribonucleoprotein	6.81	0.002	833.48	0.06	Y	CLIPdb

										(Scherer et al., 2020)
RBM3	Y	Both	Nuclear and cytoplasmic	Translational regulation	7.63	0.00001	246.86	0.02	N	-
RBM39	Y	EL4	Nucleus	mRNA processing and splicing			767.82	0.02	Y	(Chen et al., 2017)
SF1	Y	EL4	Nucleus and spliceosome	Splicing, transcription and translation regulation,			67.39	0.07	Y	CLIPdb
SLTM	Y	EL4	Nucleus	Apoptosis, transcription regulation			1000	0	Y	CLIPdb
SRSF1	Y	Both	Nuclear and cytoplasmic	Splicing, RNA trafficking	98.09	0.002	277.04	0.01	Y	CLIPdb
SRSF2	Y	Both	Nuclear	Splicing	3.33	0.07	2504.9	0.04	Y	CLIPdb
SRSF3	Y	Both	Nuclear and cytoplasmic	Splicing, RNA trafficking	6.13	0.002	869.73	0.01	Y	CLIPdb
SRSF5	Y	Naïve $CD4^+$	Nuclear	Splicing	4.94	0.01			Y	CLIPdb (Spiniello et al., 2018;

										West et al., 2014b)
SRSF6	Y	EL4					1133.74	0.03	Y	CLIPdb (Scherer et al., 2020)
SRSF7	Y	Both	Nuclear and cytoplasmic	Splicing, RNA trafficking	6.08	0.002	176.25	0.03		CLIPdb (Chen et al., 2017a; Spiniello et al., 2018)
SSb	Y	EL4	Nucleus	Transcription			762.57	0.04	Y	CLIPdb
SYNCRIP	Y	EL4					875.11	0.04	Y	(Scherer et al., 2020)
TARDBP	Y	Both	Nuclear and Cytoplasmic	Splicing, RNA stability	2.01	0.09	38.09	0.04	Y	CLIIPdb (West et al., 2014b)
THRAP3	Y	Naïve $CD4^+$	Nuclear	Splicing, RNA stability, transcriptional regulation	3.66	0.08			Y	(Spiniello et al., 2018)

TIAL1	Y	Both	Nuclear and Cytoplasmic	Splicing, translational regulation	3.33	0.09	72.21	0.04	Y	CLIPdb
TRA2b	Y	Both	Nuclear	Splicing	3.8	0.09	288.25	0.07	Y	2017; Spiniello et al., 2018)
U2AF2	Y	Naïve $CD4^+$	Nuclear	Splicing, RNA trafficking	2.55	0.01			Y	CLIPdb
UBB	N	Naïve $CD4^+$	Nuclear and Cytoplasmic	Protein degradation	10.22	0.02			N	-

Table 5.3:CD4$^+$ T cell specific functions of RBPs that interact with *Malat1*

RBP	Primary/EL4	Role in CD4+ T cells	Reference
CSTF2	EL4	No data available at the time of writing	
ELAVL1	Both	Promotes Th2 differentiation, linked to IL-4, IL-13, and Gata3 regulation. Promotes Th17 differentiation in models of asthma and EAE. Regulates c-Maf in Th17 cells. Linked to regulating CD28 pathway genes	(Chen et al., 2017, 2020; Fattahi et al., 2022; Li et al., 2020; Techasintana et al., 2015, 2017; Yu et al., 2021)
DDX5	EL4	Aids RORγT regulation of Th17 genes	(Huang & Littman, 2015) – Paper retracted
DHX9	EL4	Dhx9$^{-/-}$ mice have reduced thymic T cell output and develop spontaneous autoimmune disorders	(Dong et al., 2022)
FUBP1	Both	No data available at the time of writing	
hnRNPA0	EL4	No data available at the time of writing	
hnRNPA1	Both	Aids Treg differentiation via Foxp3, Regulates HTLV-1 replication in T cells	(Liu, et al., 2021)
hnRNPA2B1	Both	No data available at the time of writing	
hnRNPAB	Both	No data available at the time of writing	
hnRNPC	Both	Regulation of lymphocyte function-associated antigen (LFA-1) via HuR, IFNγ mRNA regulation	(Rao et al., 2018)
hnRNPD	Both	No data available at the time of writing	
hnRNPF	Both	hnrnpF impairs Treg function through FOXP3	(Du et al., 2018)
hnRNPH1	Naïve CD4$^+$	No data available at the time of writing	
hnRNPK	EL4	Regulates ERK/MAPK signalling in primary T cells. Aids IL-2 production via ERK	(Molineros et al., 2019)

hnRNPL	EL4	Regulates thymic T cell development, proliferation and migration. Bind RNAs important in T cell development TCF3, STK2B FYN.	(Gaudreau et al., 2012)(Shankarling et al., 2014)
hnRNPM	EL4	No data available at the time of writing	
hnRNPU	Both	Regulates MALT1 splicing which is important for T cell signaling and activation	(Meininger et al., 2016)
hnRNPUL2	EL4	No data available at the time of writing	
KHSRP	Both	Khsrp represses T cell proliferation and IL-5, IL-10, and IL-13 expression	(Käfer et al., 2019)
IAP	EL4	Iap inhibitors augment T cell responses. Required for survival and expansion of activated T cells	(Dougan et al., 2010; Gentle et al., 2014)
MBNL1	Naïve $CD4^+$	No data available at the time of writing	
MYH1	Naïve $CD4^+$	No data available at the time of writing	
NCL	EL4	No data available at the time of writing	
PUF60	Naïve $CD4^+$	No data available at the time of writing	
RALY	Both	Regulates CCR5 expression	(Kulkarni et al., 2019)
RBM3	Both	No data available at the time of writing	
RBM39	EL4	No data available at the time of writing	
SF1	EL4	No data available at the time of writing	
SLTM	EL4	Knockdown of Sltm reduces number of HIV-1 infected T cells	(Pedersen et al., 2022)
SRSF1	Both	Knockdown results in T cell hyperactivity. Enhanced mTOR activity, T cell activation and proinflammatory cytokine expression	(Katsuyama et al., 2019)
SRSF2	Both	No data available at the time of writing	

SRSF3	Both	No data available at the time of writing	
SRSF5	Naïve CD4+	No data available at the time of writing	
SRSF6	EL4	Linked to regulation of HIV-1 pre-mRNA splicing in CD4+ T cells	(Finley, 2015)
SRSF7	Both	No data available at the time of writing	
SSB	EL4	No data available to at the time of writing	
SYNCRIP	EL4	Knockdown of Syncrip reduced expression of IL-21	(Whisenant et al., 2015)
TARDBP	Both	Overexpression of Tardbp, increases CD4+ T cell susceptibility to HIV-1 infection	(Cabrera-Rodríguez et al., 2022)
THRAP3	Naïve CD4+	Interacts with PSF to regulate splicing of CD45 and other broad splicing events in T cells	(Yarosh et al., 2015)
TRA2B	Both	No data available at the time of writing	
U2AF2	Naïve CD4+	Knockdown of U2AF2 changes expression of activation markers and alters cytokine expression.	(Schott et al., 2021)
UBB	Naïve CD4+	No data available at the time of writing	

We next compared *Malat1* interacting proteins in EL4 cells and naïve $CD4^+$ T cells with other published RAP-MS datasets (Aznaourova et al., 2020; Gandhi et al., 2020; McHugh et al., 2015; Munschauer et al., 2018; Rea et al., 2020; Schmidt et al., 2021; Shi et al., 2021; Wang et al., 2020) (

Figure* 5.15)*. In some cases, for example on comparison with the RAP-MS dataset for lincNMR, several proteins were identified in all three datasets. Yet, in the majority of cases, for example on comparison with RAP-MS datasets for *Xist, U1, 45s, Hulc, MaiL1* no or minimal overlap was observed between the different RAP-MS data sets. This suggested that the proteins we identified by RAP-MS are not due to off-targets effects of the RAP-MS technique. This also indicated that *Malat1* binding partners are not just RBPs that bind any RNA (cytoplasmic, ribosomal, nuclear RNAs) and they have some specificity to *Malat1*.

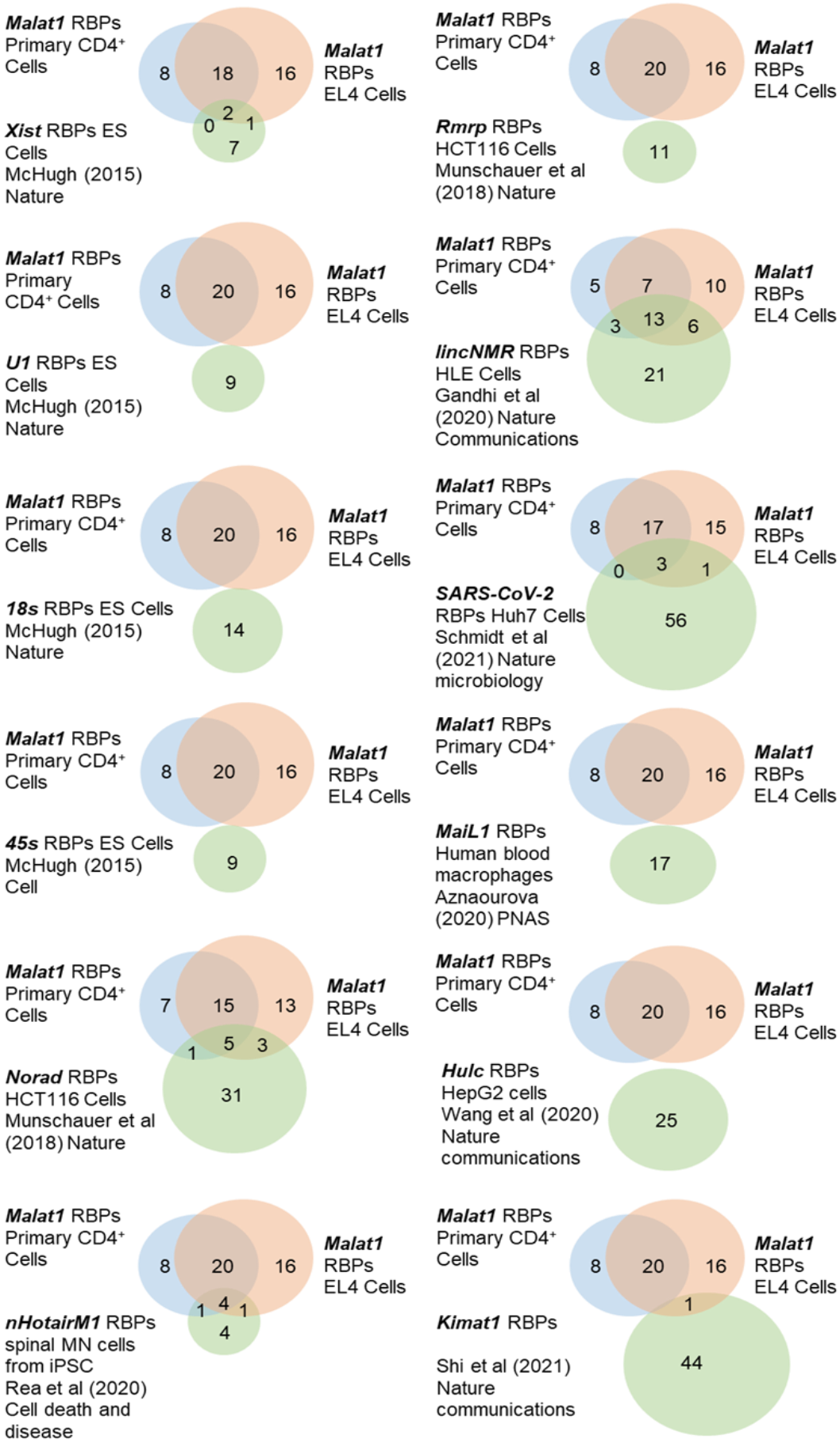

Malat1 RBPs Primary CD4+ Cells
8
18
16
0
2
1
7
Malat1 RBPs EL4 Cells
Xist RBPs ES Cells McHugh (2015) Nature
Malat1 RBPs Primary CD4+ Cells
8
20
16
11
Malat1 RBPs EL4 Cells
Rmrp RBPs HCT116 Cells Munschauer et al (2018) Nature
Malat1 RBPs Primary CD4+ Cells
8
20
16
9
Malat1 RBPs EL4 Cells
U1 RBPs ES Cells McHugh (2015) Nature
Malat1 RBPs Primary CD4+ Cells
5
7
10
3
13
6
21
Malat1 RBPs EL4 Cells
lincNMR RBPs HLE Cells Gandhi et al (2020) Nature Communications
Malat1 RBPs Primary CD4+ Cells
8
20
16
14
Malat1 RBPs EL4 Cells
18s RBPs ES Cells McHugh (2015) Nature
Malat1 RBPs Primary CD4+ Cells
8
17
15
0
3
1
56
Malat1 RBPs EL4 Cells
SARS-CoV-2 RBPs Huh7 Cells Schmidt et al (2021) Nature microbiology
Malat1 RBPs Primary CD4+ Cells
8
20
16
9
Malat1 RBPs EL4 Cells
45s RBPs ES Cells McHugh (2015) Cell
Malat1 RBPs Primary CD4+ Cells
8
20
16
17
Malat1 RBPs EL4 Cells
MaiL1 RBPs Human blood macrophages Aznaourova (2020) PNAS
Malat1 RBPs Primary CD4+ Cells
7
15
13
1
5
3
31
Malat1 RBPs EL4 Cells
Norad RBPs HCT116 Cells Munschauer et al (2018) Nature
Malat1 RBPs Primary CD4+ Cells
8
20
16
25
Malat1 RBPs EL4 Cells
Hulc RBPs HepG2 cells Wang et al (2020) Nature communications
Malat1 RBPs Primary CD4+ Cells
8
20
16
1
4
1
4
Malat1 RBPs EL4 Cells
nHotairM1 RBPs spinal MN cells from iPSC Rea et al (2020) Cell death and disease
Malat1 RBPs Primary CD4+ Cells
8
20
16
1
44
Malat1 RBPs EL4 Cells
Kimat1 RBPs Shi et al (2021) Nature communications

Figure 5.15: The *Malat1*-protein interactome demonstrates specificity when compared to other lncRNAs.

Venn diagrams depicting overlap of published RNA-protein interaction datasets identified by RAP-MS (green) with the *Malat1* protein interactome in EL4 cells (orange) and naïve $CD4^+$ T cells (blue).

To gain insight into the cell type specificity of *Malat1*-protein interactions, the EL4 and primary cell data sets were compared with published *Malat1*-protein interactomes (Chen et al., 2017; Scherer et al., 2020; Spiniello et al., 2018; West et al., 2014; Zhu et al., 2019). The majority of proteins identified by RAP-MS that interacted with *Malat1* in $CD4^+$ T cells have previously been shown to interact with *Malat1* ***(Table 5.)***. Of note, there is little overlap between the datasets (***Figure 5.16)***. Some differences in binding could be explained by experimental technique. This could be tested by directly comparing RNA centric approaches using the same cell type and target RNA to determine what if any differences in RNA-RBP interactions is identified. For example, *in vitro* transcribed "bait" RNA which is incubated with cell lysates may not identify accurate RNA-RBP interactions. This is because *in vitro* transcribed RNA may lack modifications, fold incorrectly, or come into contact with proteins that exist in different cellular compartments (Gerber, 2021). Moreover, different crosslinking methods are used across the pulldown techniques. RAP-MS requires UV crosslinking whereas HyPR-MS and CHART-MS employ formaldehyde crosslinking methods. Formaldehyde cross-linked samples can identify both direct and indirect protein interactions, which can prove challenging for functional characterisation (McHugh & Guttman, 2018). In contrast, UV based crosslinking has a relatively low efficiency (~5%) but is less likely to identify indirect interactions. However, UV crosslinking can induce cellular stress responses which could impact RNA-protein interactions (Gerber, 2021).

To determine if RAP-MS identifies unique or distinct RBP-RNA interactions, we next compared the four different *Malat1* interactome data sets. Similarly, to RAP-MS very little overlap is observed between datasets (***Figure 5.17)***. Of note, no proteins

appeared in all four data sets. Further indicating cell type and technique may influence the proteins identified by the different RNA-centric approaches.

Although use of different techniques may at least in part explain differences in *Malat1*-protein interactomes, cell type is also likely to play a role -as observed within our EL4 and naïve primary $CD4^+$ datasets, where RAP-MS has taken place under the same conditions some cell-type specificity is observed. Here, we compare *Malat1* interactions in primary immune cells, MCF-7 cells (breast cancer cell line), and PC3 cells (cell line derived from bone metastasis of prostatic adenocarcinoma). Cell type could influence protein expression, protein localisation and *Malat1* expression levels (we have already demonstrated that *Malat1* expression varies between Th cell subsets chapter 3), which in turn may impact the *Malat1* interactome. At the time of writing no published *Malat1* protein interactome datasets identified by RAP-MS, or *Malat1* protein interactome datasets in primary immune cells identified by other methods are available. As such it is difficult to conclude if cell type or technique has a greater impact on *Malat1* interactions. Further research is required to establish differences in the *Malat1* protein interactome between cell types and tissues.

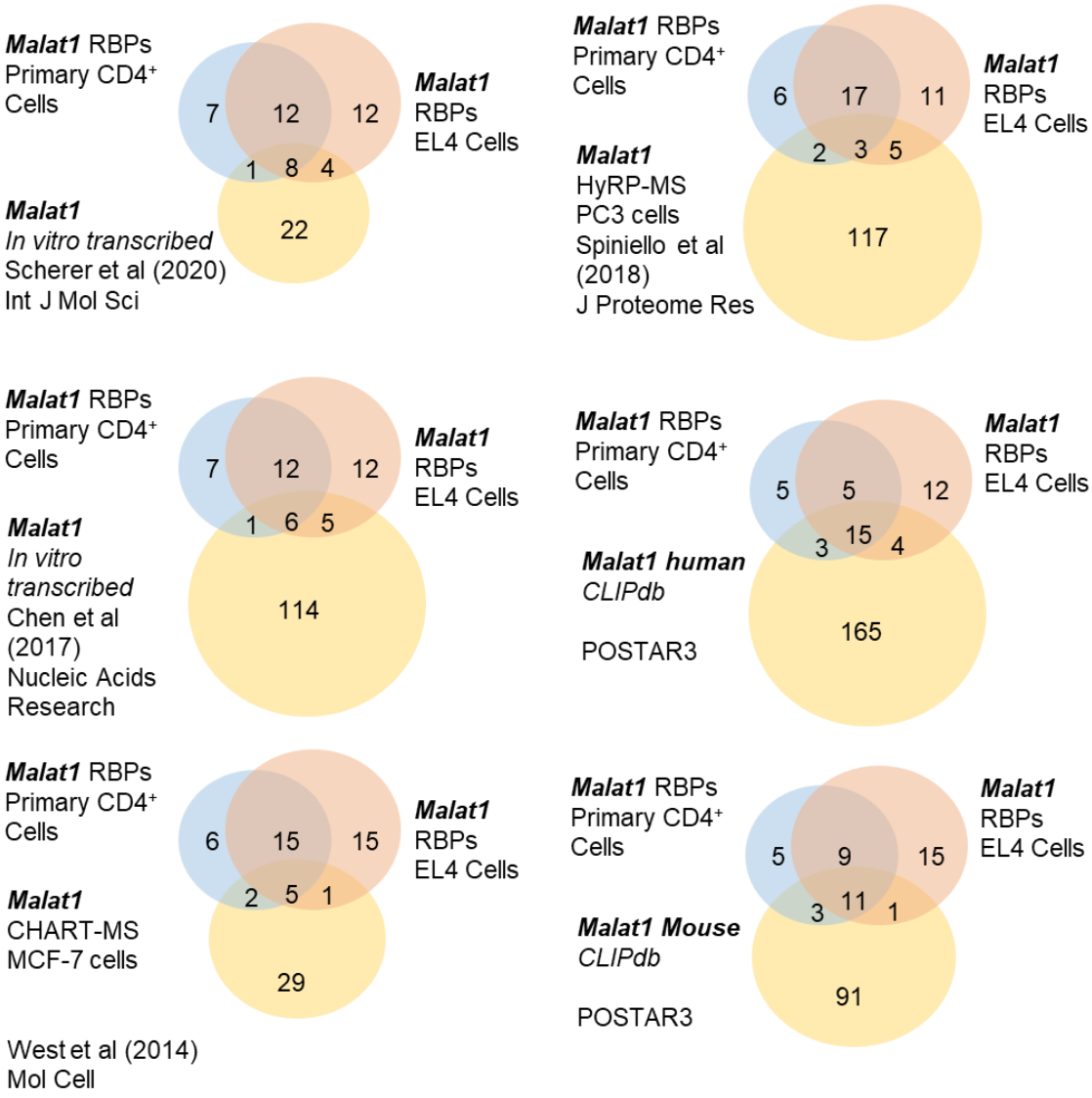


Figure 5.16: The *Malat1* interactome is cell type specific.

Venn diagrams depicting overlap of published *Malat1*-protein interaction datasets (yellow) with the *Malat1* protein interactome in EL4 cells (orange) and naïve CD4⁺ T cells (blue).

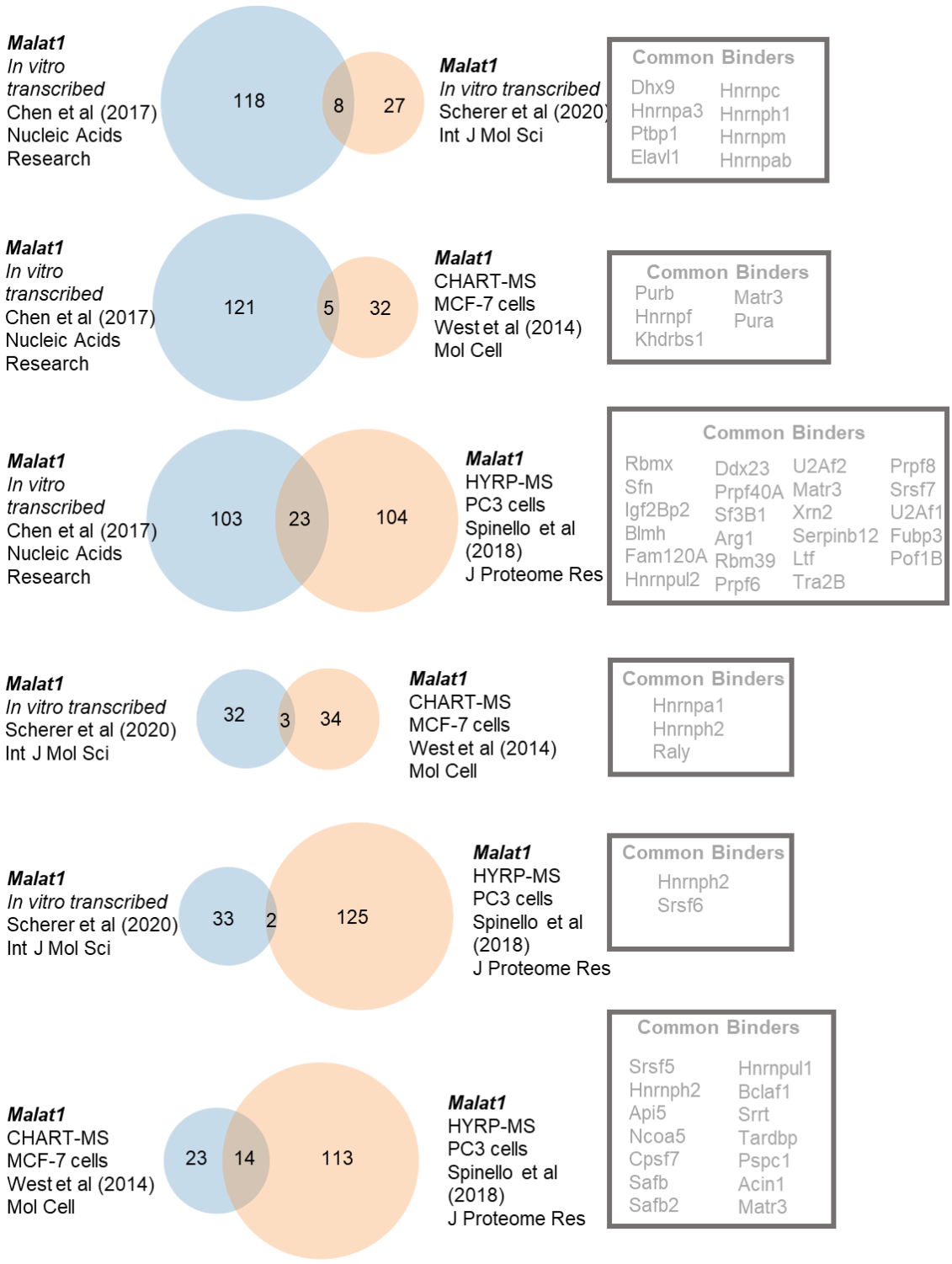


Figure 5.17: RNA-centric approaches identify unique *Malat1* RBP interactions between cell types.

Venn diagrams depicting overlap of published *Malat1*-protein interaction datasets identified using RNA-centric approaches.

Having found that *Malat1* appeared to bind RBPs in a cell specific manner we wanted to determine if this was the case for other lncRNAs. Thus, we next compared 3 independent studies using RNA-centric approaches to identify *Xist*–protein interactions (***Figure 5.18***). Chu and colleagues use ChIRP-MS to identify *Xist* interactions in embryonic stem cells (ESCs), McHugh identified the *Xist* interactome in ESCs, and Minajigi used identification of direct RNA-interacting proteins (iDRip) in mouse embryonic fibroblasts (Chu, et al., 2015; McHugh et al., 2015; Minajigi et al., 2015). Notably, RAP-MS identified the fewest RBPs of the three data sets. However, 90% of these interactions also appeared in the other datasets, with 80% of RBPs identified by RAP-MS also identified by ChIRP-MS. These experiments take place in the same cell type which could explain the similarities between the RAP-MS and ChIRP-MS datasets. Numerous factors could impact the discrepancies between published *Xist* and *Malat1* datasets. Cell type, crosslinking agent, the oligonucleotides used for pulldowns and proteomic approaches are all likely to contribute as reviewed by (Da Rocha & Heard, 2017).Of note, previously proposed *Xist* binding partners such as EZH2 and SUZ12 (members of the PRC2 complex) are not identified in any of these unbiased proteomic approaches. Instead, *Xist* could form secondary interactions with PRC2 as postulated by McHugh and colleagues (McHugh et al., 2015).

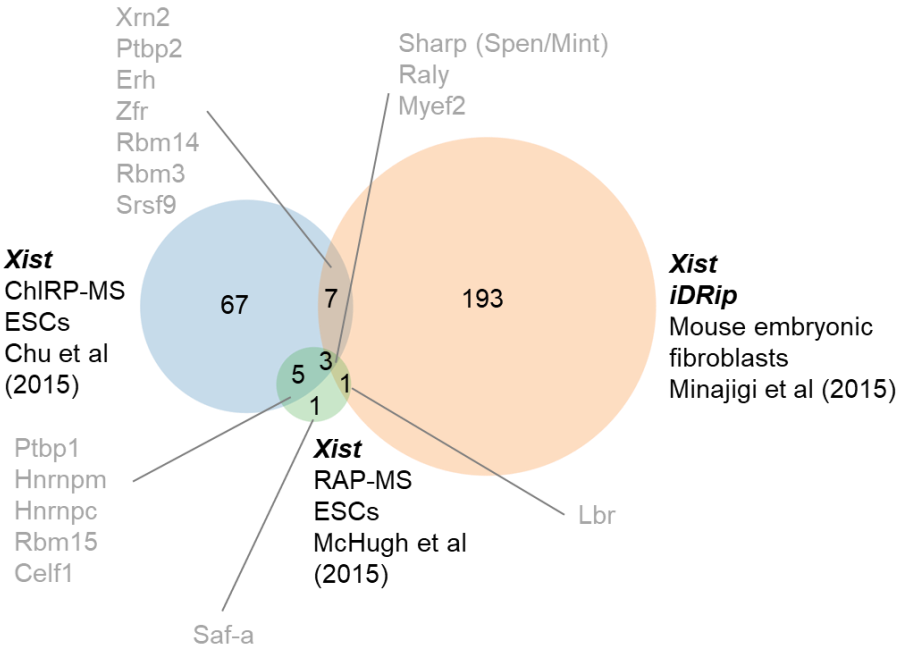


Figure 5.18: Comparison of *Xist* interacting RBPs identified through RNA-centric approaches.

Venn diagrams depicting overlap of 3 independent published Xist-protein interaction datasets identified by different RNA-centric approaches RAP-MS (green), iDRip (orange) and ChIRP-MS (blue).

5.3 Discussion

For the first time, we identified the *Malat1*-protein interactome in primary immune cells. As noted by comparing RNA-centric approaches to identify RNA-RBP interactions, there was a large degree of heterogeneity in lncRNA binding partners when comparing cell types. As such it is of great benefit to define the unique and core *Malat1* interactions in $CD4^+$ T cells.

Notably, our datasets identified that *Malat1* prominently interacted with members of the hnRNP and SR family of proteins. The hnRNPs are a large family of RBPs, comprised of 20 major hnRNPs which are functionally diverse and regulate RNA at multiple levels including splicing, stability, transcription, and translation (Geuens et al., 2016). hnRNP expression has also been linked to numerous pathologies including cancers, ALS, Alzheimer's disease, and FTLD. SR proteins are defined by their ability to interact with RNA, they are found in nuclear speckles, and are also involved in regulating pre-mRNA splicing, mRNA nuclear export, mRNA degradation, and translation. SR proteins have also been linked with numerous cancers (Shepard & Hertel, 2009).

Common structural features are found across all hnRNPs, which contain RBDs commonly in tandem (with the exception of hnRNPU). Similarly, SR proteins are structurally comparable to hnRNPs and often contain RBDs in tandem. Additionally, many hnRNPs shuttle between the nucleus and cytoplasm – and are thought to aid mRNA export from the nucleus (Han et al., 2010). This is also a common feature of SR proteins, which are well known to shuttle between the nucleus and cytoplasm. (Cáceres et al., 1998; Shepard & Hertel, 2009)

Many of these proteins have been linked to regulation of Th cell function. It is possible that *Malat1* could impact how these RBPs function which in turn could affect Th cell differentiation. To determine if *Malat1* is functioning through its interacting RBPs, and at which stage of the RNA lifecycle this takes place further investigation is required. Nevertheless, it is of great benefit to narrow down the 1000s of characterised RBPs to a short list of $CD4^+$ T cell specific *Malat1* interactions.

Several members of the hnRNP family have been linked to the regulation of $CD4^+$ T cell function. For example, hnRNPA1 has been shown to be required for Treg induction. Knockdown of hnRNPA1 in human $CD4^+$ T cells reduced the expression and ubiquitination of FOXP3 and decreased Treg differentiation (Lui, et al., 2021). hnRNPF has also been shown to regulate Treg function. Overexpression of hnRNPF reduced Treg suppressive function. hnRNPF was found to interact with FOXP3 in Tregs, which repressed its ability to bind to pre-mRNA targets thus alterative splicing was impaired (Du et al., 2018). Additional roles of the hnRNP family in Th cell function are detailed in ***(Table 5.***). Collectively, these findings demonstrated the importance of *Malat1* interaction partners in Th cells.

Members of the SR family of proteins have also been shown to be important for Th cell function – although less than the hnRNP family of proteins. For example, SRSF1 has been shown to control T cell hyperactivity. Levels of SRSF1 are reduced in T cells of patients with SLE – a disease in which hyperactive T cells play a role. Mouse models of T cell specific deletion of SRSF1, resulted in the development of systemic autoimmunity, increased the frequency of activated T cells, enhanced mTOR activity and resulted in higher levels of pro-inflammatory cytokines (Katsuyama et al., 2019).

Limited evidence has linked either the SR or hnRNP family of proteins with regulation of IL-10 – a previously established function of *Malat1* in Th cells (Hewitson et al., 2020). However, the RBP ELAVL1 (/HuR), has been shown to have a plethora of functions in $CD4^+$ T cells including the regulation of MAF a master regulator of IL-10 (Yu et al., 2021). It is possible that *Malat1* interactions with ELAVL1 influences the regulation of T cell differentiation and regulation of MAF/IL-10 however additional research is required to determine if *Malat1* affects ELAVL1 function. One possible way to determine this would be to perform iCLIP for ELAVL1 in WT and *Malat1*$^{-/-}$ cells and asses changes in RNA interactions.

ELAVL1 is a master regulator of gene expression with its dysregulation playing an important role in a myriad of diseases. ELAVL1 is able to target mRNAs to stabilise them or promote translation through interactions with 3' untranslated regions (Fattahi et al., 2022). It has been shown to regulate Th2 cells in both human and mouse models of asthma. In mice conditional $CD4^+$ T cell knockout of ELAVL1, impaired Th2 differentiation and cytokine production in a model of airway inflammation. Additionally, ELAVL1 levels were shown to be elevated in $CD4^+$ T cells of patients with asthma compared to healthy controls (Fattahi et al., 2022) Additional work has implicated ELAVL1 in the regulation of Th2 differentiation, with a conditional $CD4^+$ T cell knockout of ELAVL1 found to be required for translation of IL-2Rα, IL-2 homeostasis and Th2 differentiation (Techasintana et al., 2017). Mouse $CD4^+$ T cell specific knockouts of ELAVL1, demonstrated the RBP binds to the mRNA of CD3g and enhanced its expression. ELAVL1 was also shown to bind and stabilise IL-6Rα mRNA which in turn promoted IL-6Rα expression, and downstream phosphorylation of STAT3 which enhanced Th17 differentiation (Yu et al., 2021). Further work has supported the role of ELAVL1 in Th17 differentiation, as

when ELAVL1 was knocked-out this reduced the levels of RORγt and diminished the number of Th17 cells in models of autoimmune encephalomyelitis (J. Chen et al., 2020) Reduced CCR6 expression and Th17 cell migration was also observed in ELAVL1 $CD4^+$ T cell conditional knockouts (J. Chen et al., 2017) . Additional sequencing analysis of ELAVL1 interacting mRNAs indicated that ELAVL1 potentially regulated genes involved in T cell activation, these pathways included the CD28 pathway (Techasintana et al., 2015). Collectively, these studies provided strong support for the role of ELAVL1 in enhancing Th cell differentiation and functions. Given its importance in Th cell function, further understanding the mechanistic relevance of ELAVL1 interactions with *Malat1* may provide greater insight into the function of *Malat1* in $CD4^+$ T cells.

Notably, *Malat1* was shown to interact with TARDBP/TDP-43 in both EL4 and primary naive $CD4^+$ T cells. TDP-43 is a versatile RNA/DNA binding protein involved in RNA related metabolism, transcription alternative splicing and RNA stability (W. Liu et al., 2020). It is well known for forming aggregates in the brains and spinal cords of patients with FTLD and ALS, and known to bind to both *Malat1* and *Neat1*. *Malat1* binding to TDP-43 increased in brains of patients with FTLD which suggested a potential link between *Malat1* and neurodegenerative diseases (Amodio et al., 2018). However, the functional link between TDP-43 aggregates and disease pathology is controversial (Hergesheimer et al., 2019). The mechanistic relevance of TDP-43 in $CD4^+$ T cells is less well understood; however, it is reassuring that *Malat1* reproducibly binds TDP-43 in our datasets adding confidence to the other interaction partners observed.

Interestingly, RNA-centric datasets defining the *Malat1*-protein interactome are limited to only a handful of cell types, often using different methodologies. To

gain insight into the impact of technique and cell type on the identification of RNP complexes, multiple RNA-centric approaches should be directly compared to determine if some RNA-protein interactions are biased for ease of identification by the different techniques. This would add confidence to the cell-type-specific interactions indicated by comparisons of datasets. Previous work has compared iDRiP and ChIRP to identify the *Terra* proteome in ES cells. Surprisingly, only 27 proteins out of of 205 (ChIRP) and 134 (iDRiP) were identified in both datasets. However, of the top 30 most enriched proteins 11 appeared in both datasets. This highlighted the potential impact of technique on RNA-Protein interactome identification, and the need for validation.(Chu et al., 2021).

Yet, some cell type-specific interactions have been observed in our datasets using the same method (RAP-MS) for both cell types. Technical differences may have impacted the outcome: 50 million cells were used for each EL4 RAP capture and 20 million cells for each naïve $CD4^+$ T cell capture. Primary cell numbers are more limited as they are derived directly from mice. In addition, they are much smaller than EL4 cells ~5-7 µm vs ~12 µm. However, small pilot studies determined that RNA-protein interactions can still be identified at lower cell densities. We proceeded with these cell numbers as data sets could still be generated using fewer mice. This technical difference could explain the absence of some proteins in the naïve $CD4^+$ T cell dataset, however, some unique proteins were identified in the primary naïve $CD4^+$ T cells which indicated that *Malat1*-protein interactions can be cell-type specific. The disparity in cell numbers could also explain why *Malat1* was more enriched in pull-downs from EL4 cells rather than naïve $CD4^+$ T cells. The differences in binding between EL4 cells and primary naïve $CD4^+$ T cells implies that they could be potential

changes in nuclear speckle composition or function this concept is further explored in Chapter 6.

We chose to carry out RAP-MS in naïve $CD4^+$ T cells as they are the common origin of effector and regulatory T cells and have the highest expression of *Malat1*. In Chapter 3, we demonstrated the functional impact of *Malat1* loss in effector $CD4^+$ T cells. Given differences in *Malat1*-RBP interactions occur between EL4 cells and naïve $CD4^+$ T cells some unique interactions may arise in the different differentiated Th cell subsets eg Th1 vs Th2. As noted in (***Figure 5.2)***, unique RBP expression changes occur when comparing WT Th1 and WT Th2 *in vitro* polarised cells with some RBPs specifically upregulated during Th1 differentiation and downregulated in Th2 differentiation (and vice versa). These changes in RBP expression and potentially localisation, phosphorylation, and splice variant usage could impact *Malat1* interactions in the different T cell subsets. Although, RBP expression may have only a limited impact on *Malat1* binding, it would be beneficial to determine if these interactions hold in the context of real infection. Mice could be infected with specific pathogens to push $CD4^+$ T cell towards either a type 1 or type 2 phenotype. Limited data is available for RNA-RBP complexes in the context of infection. This would provide a unique insight into *ex vivo* infection *Malat1* -RBP interactions and determine any differences between *ex vivo* and *in vitro* polarised binders.

Other factors could influence *Malat1*-RBP interactions including RBP structure, such as RRM domain interactions as suggested by our dataset. The structure of *Malat1* likely also influences binding. Structure of lncRNAs is often better conserved than sequence (Johnsson et al., 2014). *Malat1* has a unique triple helix structure at its 3' end which has been shown to bind specific RBPs (Brown et al., 2014). RNA modifications likely impact *Malat1*-RBP interactions. The *Malat1*

transcript has m6A modifications which aid the formation of its triple helix, which makes the transcript accessible for RBPs such as hnRNPC (Liu et al., 2015).

Further experiments would be required to confirm where on the *Malat1* transcript these RBPs are binding. For example, iCLIP could be used to identify the RBP binding sites (see chapter 6). Not only would this method confirm the *Malat1*-RBP interactions, this would also determine if loss of *Malat1* impacts RBP functionality. More global RNA-centric approaches could be used to determine if loss of *Malat1* impacts formation of other RNP complexes. For example, RIC or OOPs could be used in WT and *Malat1*$^{-/-}$ cells to determine global differences in RBP binding. Although depth of these techniques could be somewhat problematic if only minor differences in binding take place in the absence of *Malat1*.

It is important to validate at least some of the *Malat1*-RBP interactions in CD4^{+} T cells. This could be done in several ways, for example RAP followed by western blotting. Alternatively, immunofluorescence could be used to confirm co-localisation of *Malat1* and its interacting proteins. This could be combined with proximity ligation assays (PLA) - a sensitive tool often used to examine protein-protein interactions (within 40nm), which can be used to confirm RNA-RBP interactions (Zhang et al., 2016).

Recent work has linked *Malat1* to promoting CD8^{+} terminal effector cell differentiation. Of note, *Malat1* was shown to directly interact with EZH2 to deposit repressive histone marks H3K27me3 at memory cell-associated genes (Kanbar et al., 2022). Numerous studies have also linked *Malat1* with EZH2 interactions, including in prostate cancer and Parkinson's disease (Cai et al., 2020; Wang et al., 2015). Interestingly, EZH2 does not appear in our datasets nor any other of the RNA-centric

approaches used to identify *Malat1* protein interactomes. This could be because these *Malat1*-EZH2 interactions are indirect and are not identified by these approaches as demonstrated with *Xist* (McHugh et al., 2015). However, targeted approaches such as RIP-seq or RIP-qRT-PCR are commonly used to identify *Malat1*-EZH2 interactions which use similar cross-linking approaches. It is well known that polycomb proteins such as EZH2 are notoriously sticky and could represent false positive interactions in these datasets. Yet, *Malat1* interactions with EZH2 have been shown to promote cellular functions. *Malat1* could be interacting with EZH2 but at low affinity intensities, that are below the level of detection in RNA-centric approaches. It is also plausible that weak interactions between *Malat1* and EZH2 may be beneficial for EZH2 function. Moreover, cell type could also play a role given the observed cell type dependence in *Malat1* binding proteins. It would be beneficial to check if *Malat1* interacts with EZH2 in $CD4^+$ T cells and determine if histone methylation marks are altered in the absence of *Malat1*.In conclusion, *Malat1* may interact with EZH2 but further investigation is required, our data suggests that EZH2 is not within the main *Malat1* binding partners.

In summary this chapter has provided a unique insight in to *Malat1* protein interactions in both primary $CD4^+$ T cells and EL4 cells. Further understanding the mechanistic relevance of these interactions would provide greater insight into the role of *Malat1* in the immune system.

6. Malat1 is a determinant of RBP function in Th2 cells

6.1 Introduction

6.1.1 Alternative Splicing (AS)

Splicing of messenger RNA precursors (pre-mRNAs) occurs in a macromolecular RBP assembly known as the spliceosome, which is responsible for the excision of >99% of introns in humans (Wilkinson et al., 2020). The spliceosome is a dynamic structure that is comprised of ~300 proteins, its exact composition changes depending on the splicing reaction, and affinity for pre-mRNAs. A subset of proteins interact with the snRNAs U1, U2, U4, U5 and U6 to form small nuclear ribonucleoprotein complexes (snRNPs) these are the best characterised components of the spliceosome and are represent its catalytic core. They are required for the selection of splice sites and to catalyse RNA cleavage. Several RBPs also comprise the spliceosome, these are capable of a sequence-specific binding. Some RBPs are essential for the splicing process, whereas others have passive roles in splicing but are required for downstream functions such as polyadenylation or nuclear export (Wilkinson et al., 2020).

The plasticity of the spliceosome enables exon skipping or inclusion. In some cases, two or more alternative 5' splice sites compete for the binding and joining of a 3' splice site, this results in AS (Eperon et al., 2000). Over 90% of human transcripts are alternatively spliced, this process is controlled by *trans*-acting splicing factors which bind to pre-mRNAs to either activate or repress the selection of splice sites (Clancy, 2008).

SR and hnRNP proteins are two major splicing regulators. The SR proteins 1-7 typically enable exon inclusion through sequence-specific binding to exon enhancers. Despite their structural similarity to SR proteins 1-7, SR proteins 9-11

typically act as repressors of AS (Akerman et al., 2015). hnRNP proteins are a diverse family of proteins and can repress or enhance AS. One study which analysed the sequence specificity of the hnRNP proteins determined that hnRNPF, H1, M and U typically activate AS and hnRNPA1 and A2B1 typically repress AS. These activators and repressors of splicing act together with the core spliceosome components and mediate AS within a cell (Akerman et al., 2015).

6.1.2 SRSF1 and hnRNPA1

SRSF1 is the archetypical member of the SR family of proteins. It was originally identified to promote spliceosome assembly and splicing in HeLa cells.(Krainer et al., 1990). Subsequently, SRSF1 has been shown to regulate mRNA transcription, stability, translation and protein sumoylation. It has also been identified as an oncogene – and has contributed to highlighting the importance of splicing in cancer progression (Das & Krainer, 2014). SRSF1 and hnRNPA1 are known to act antagonistically to one another to regulate AS. SRSF1 binds to proximal 5' splice sites (GU) regions through U1-mediated interactions, and hnRNPA1 promotes distal 5' site selection progression (Das & Krainer, 2014).

Several mechanisms have been described for hnRNPA1-dependent exon skipping. For example, hnRNPA1 competes with other RBPs including SRSF1 to bind regulatory elements that would normally bind SRSF1 (Eperon et al., 2000). Another study found that hnRNPA1 binding sites overlapped with SC-35 a nuclear speckle defining marker and an exon splicing enhancer. Upon hnRNPA1 binding to the exonic splicing silencing site, this masked the SC-35 binding site and impaired splicing of the HIV-tat exon 2 (Zahler et al., 2004). hnRNPA1 can also oligomerise and spread from an exonic splicing silencer across the region of a regulated exon which antagonises the behaviour of SRSF1. Additionally, hNRNPA1 has been shown to dimerise and

cause an exon to loop out from the pre-mRNA transcript, promoting the exon inclusion in the mature transcript (Blanchette & Chabot, 1999). It is thought that hnRNPA1 facilitates SRSF1 to reject sub-optimal splicing sites through a combination of these mechanisms.

One study which looked at protein-protein interactions of spliceosome components determined that activators from either the SR or hnRNP family of proteins are central to the spliceosome, whereas repressors are more peripheral and more loosely connected to the spliceosome (Akerman et al., 2015). This observation was confirmed by comparing hnRNPA1 (repressor) and SRSF1 (enhancer) through immunoprecipitation experiments followed by mass spectrometry. This found that SRSF1 interactions with the spliceosome are not dependent on RNA and form multiple protein-protein interactions with U2-snRNP complexes. Contrastingly, hnRNPA1 interactions with the spliceosome are RNA-dependent (Akerman et al., 2015).

SRSF1 and hnRNPA1 are also found to be dysregulated in multiple cancers. One study of a mouse model of lung cancer found that SRSF1 and hnRNPA1 expression increased in tumour tissues compared to healthy controls. Of note, hnRNPA1 expression increased substantially more than SRSF1 expression in tumour samples (Zerbe et al., 2004). Upon measurement of AS patterns of *Cd44* – which is a gene associated with many cancers – it was shown that the ratio of hnRNPA1 to SRSF1 corresponded to the presence of a metastasis-associated splice variant of *Cd44* which was not present in healthy tissues (Zerbe et al., 2004)

6.1.3 iCLIP

RBPs play an essential role in a myriad of cellular functions including regulation of pre-mRNA splicing, RNA modification, translation, stability and

localisation of RNA. To improve our understanding of RBP function it is essential to identify *in vivo* binding sites across the transcriptome. Several protein-centric approaches have been developed. Initial experiments relied on RNP/RNA immunoprecipitation (RIP). This involved formaldehyde crosslinking of protein to RNA and antibodies were used to pull down proteins from the spliceosome complex to identify snRNAs which interact with Sm proteins (Lerner & Argetsinger Steitz, 1979). After later advances in the field, microarray and transcriptomic analysis was used to determine RNA-Protein interactions and were known as RIP-chip. Later work combined RIP with high-throughput sequencing and was termed RIP-seq. The use of formaldehyde crosslinking preserves protein-protein interactions, consequently RNAs can be identified that bind other proteins using this approach. This drove the development of methods which had increased the specificity of protein-RNA interactions (Lee & Ule, 2018).

Crosslinking and immunoprecipitation (CLIP) was developed to address the issues of RIP and enable the identification of precise binding sites (Huppertz et al., 2014) (***Figure 6.1***). There have been numerous developments in the CLIP process to address different biological questions. One form of CLIP known as individual-nucleotide resolution UV crosslinking and immunoprecipitation (iCLIP) involves exposing cells to UV light to generate cross-linked protein-RNA complexes. Cells are lysed and subsequently treated with RNase. The complex is then purified for example by using antibodies and streptavidin beads. An oligonucleotide adaptor is added to the 3' end of the RNA fragments. The complex is separated by SDS-PAGE and then isolated in a size-specific manner. The RNA fragments are isolated by digesting the associated protein. This generates RNAs with a covalently bound peptide at the crosslinking site. Subsequent reverse transcription results in truncation at the cDNA

crosslinking site – consequently iCLIP cDNAs start at the nucleotide just downstream of the cross-linked peptide and terminate where RNase cleavage has occurred. Following adapter ligation, the RNA samples are analysed by high throughput RNA-seq (Huppertz et al., 2014).

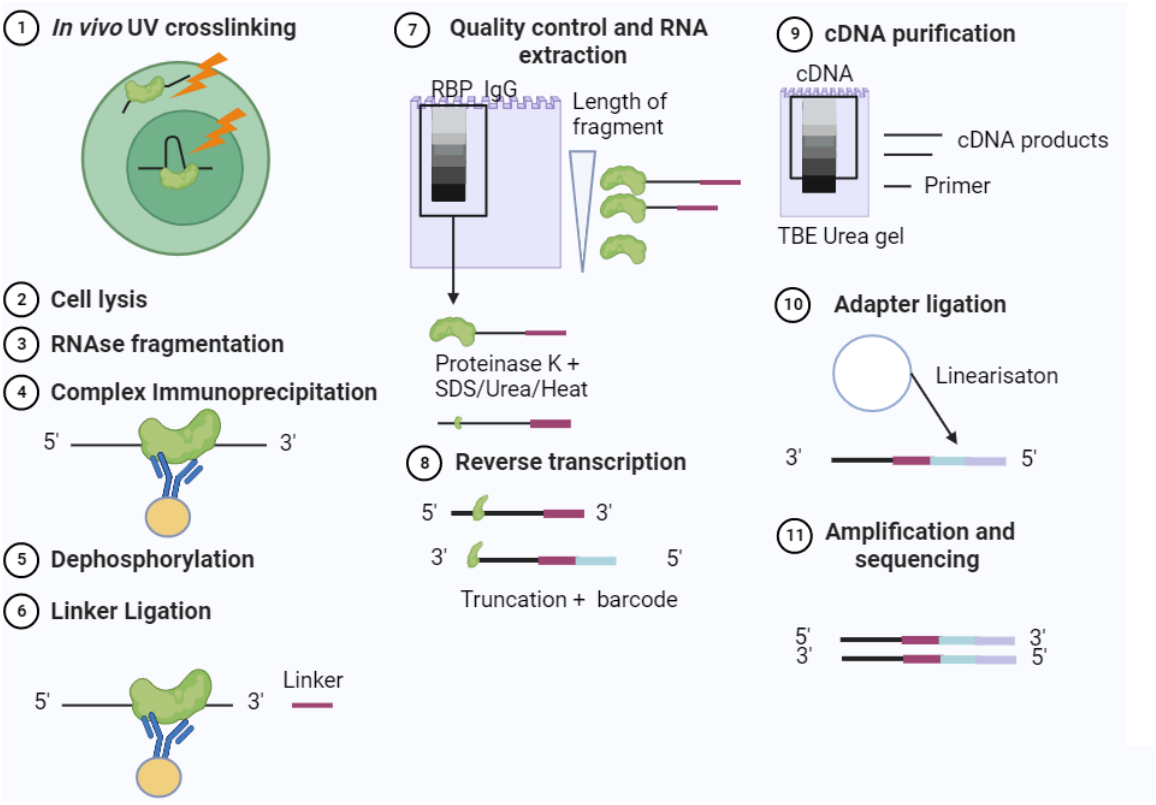


Figure 6.1: Schematic representation of the iCLIP protocol.

A simplified diagram of the iCLIP protocol. Briefly, cells of interest are exposed to UV light, creating a covalent cross-linkage between RNA and protein. Cells are then lysed and the RNA is fragmented. Subsequently, the RNA-protein complex is isolated using antibodies and separated by size on an SDS-PAGE gel. The complex is isolated from the gel, the protein is degraded using proteinase K which leaves a peptide at the crosslinking site. The sample is reverse transcribed which results in truncation at the crosslinking site. The sample is then prepared for high-throughput sequencing (König et al., 2010).

6.1.4 Aims

In chapter 5, we identified the *Malat1* protein interactome in naïve $CD4^+$ T cells. The functional relevance of *Malat1* interactions with these proteins remains unknown. We hypothesised that *Malat1* interactions with SRSF1 and hnRNPA1 can modify their interactions with target RNAs in the cell. To address this hypothesis this chapter has the following aims:

- Determine SRSF1 and hnRNPA1 interaction partners in WT and *Malat1*$^{-/-}$ Th2 cells using iCLIP
- Compare similarities and differences between interaction partners in the presence and absence of *Malat1*

6.2 Results

6.2.1 Experimental design

In Chapter 5, we identified the *Malat1* protein interactome and determined that *Malat1* prominently interacted with SR and hnRNP proteins. To determine the relevance of *Malat1* interactions with the RBPs SRSF1 and hnRNPA1 we generated 2-3 biological replicates of *in vitro* polarised Th2 cells from WT and *Malat1*$^{-/-}$ mice and performed iCLIP. Samples were isolated and cross-linked at the University of York before shipment to our collaborators Jernej Ule and Miha Modic at the Francis Crick Institute, London, UK and the Kemijski inštitut, Hajdrihova, Solvenia. Urška Janjoš and Miha Modic performed SRSF1 and hnRNPA1 iCLIP and library preparation and initial data processing using iMAPs. An overview of the experimental setup is shown in (***Figure 6.2***).

We initially aimed to determine if SRSF1 and hnRNPA1 interactions were altered during the differentiation process (naïve to Th2) and affected by the presence of *Malat1 (*WT vs *Malat1*$^{-/-}$*).* Unfortunately, pilot iCLIP experiments using naïve CD4^{+} T cells were unsuccessful. This was due to the low number of cells used for input ~1 million cells. This could be repeated at a later date with a higher cell number. Instead, here we focus on the relevance of *Malat1* interactions with SRSF1 and hnRNPA1 in Th2 cells. We chose to study these interactions in a Th2 context as we have previously found that *Malat1* had a more prominent effect on gene expression changes in a Th2 cell (chapter 4).

Of note, this work took place in parallel to that described in chapter 5 and chapter 7 where we analysed the function of *Malat1* in both male and female cells. In this chapter we only analysed *Malat1* function in female Th cells. This was because

this dataset represents a pilot experiment where we wanted to determine whether iCLIP was possible in this cell type, the antibodies were suitable for iCLIP, and sufficient library quality could be generated from these samples. In future experiments, we aim to include both male and female samples.

This results chapter will discuss preliminary data analysis of SRSF1 and hnRNPA1 interactions in WT and *Malat1*$^{-/-}$ Th2 cells.

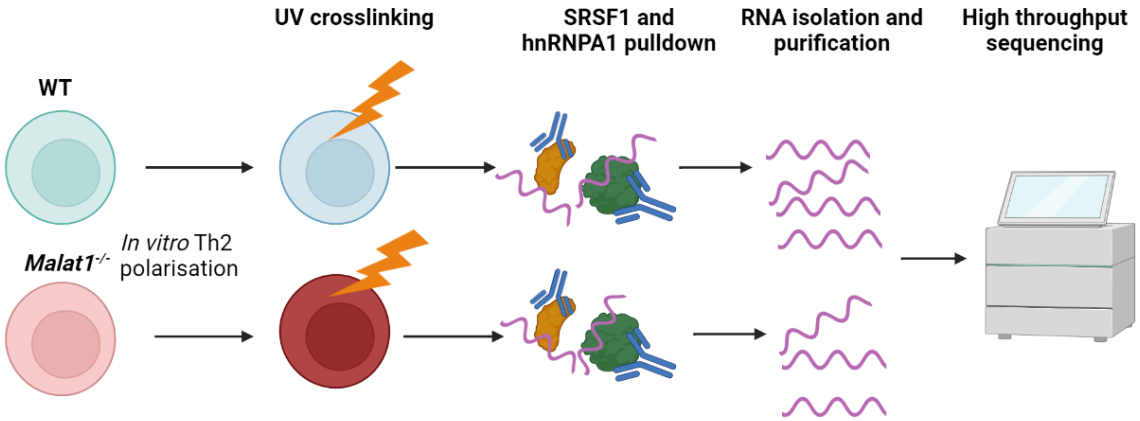


Figure 6.2: iCLIP experimental design.

Overview of experimental design. Naïve $CD4^+$ T cells were isolated from WT and *Malat1*$^{-/-}$ mice and cultured *in vitro* under Th2 polarising conditions. Samples were exposed to UV light to create a covalent cross-linkage between RNA and protein. SRSF1 and hnRNPA1 protein-RNA complexes were isolated and the associated RNA was examined using iCLIP. n=2 biological replicates SRSF1 WT and *Malat1*$^{-/-}$ representative of 1 independent experiment. n=3 biological replicates hnRNPA1 WT and *Malat1*$^{-/-}$ representative of 1 independent experiment.

6.2.2 SRSF1 interactions in Th2 cells

We began our analysis by comparing the percentage of iCLIP reads between replicates. This analysis showed that replicates derived from either the WT Th2 or *Malat1*$^{-/-}$ Th2 cells strongly positively correlated with one another and showed an r^2 value of 0.9748 and 0.9808, respectively (***Figure 6.3***). Comparison of the average percentage of reads between WT Th2 and *Malat1*$^{-/-}$ Th2 cells showed a positive correlation, however, this was weaker than correlations between sample replicates of the same condition (r^2 0.8020) this indicated that *Malat1* may have some impact on SRSF1 interaction partners (***Figure 6.4***).

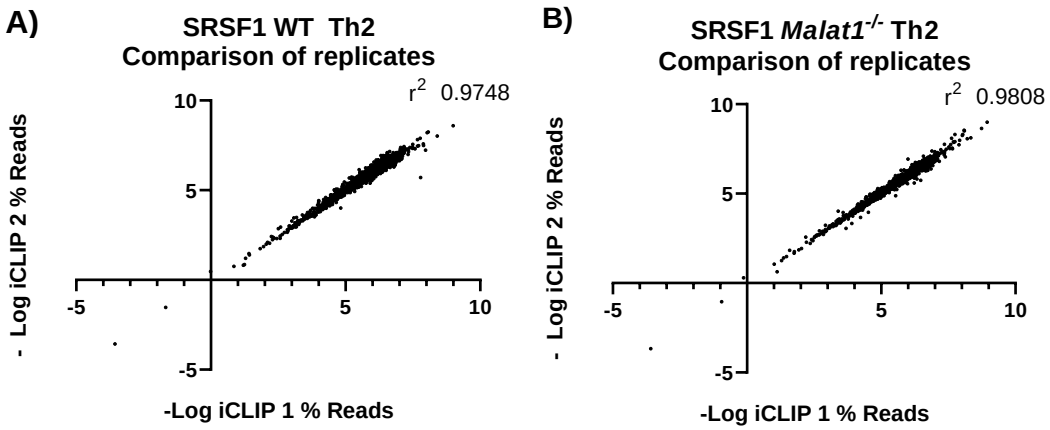


Figure 6.3: SRSF1 iCLIP replicates are positively correlated.

A) Correlation of the Log_2 percentage reads between two UV cross-linked WT Th2 SRSF1 iCLIP samples (n =2 biological replicates) **B)** Correlation of Log2 percentage of reads between two UV cross-linked *Malat1*$^{-/-}$ Th2 iCLIP samples (n =2 biological replicates). A threshold of over 0.01% reads was applied.

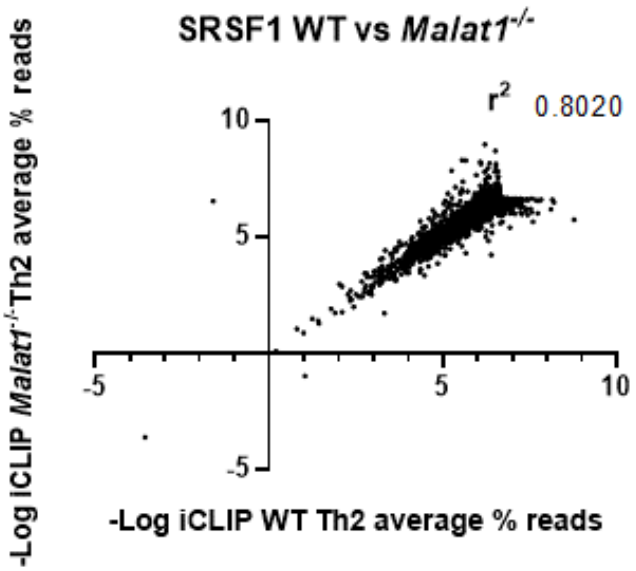


Figure 6.4: SRSF1 samples from WT and *Malat1*$^{-/-}$ iCLIP are positivley correlated.

Correlation of average Log2 percentage of reads between UV cross-linked WT Th2 SRSF1 iCLIP samples and UV cross-linked *Malat1*$^{-/-}$ Th2 iCLIP samples (n =2 biological replicates). A threshold of over 0.01% reads was applied.

Genomic annotation of cDNA sequences revealed that SRSF1 prominently bound targets in introns, this represented ~46.5% of interactions. SRSF1 was also shown to bind to coding sequence (CDS) (~22.7% of interactions). Only small changes in binding were observed in the absence of *Malat1* with a 2% drop in ncRNA binding observed in the *Malat1*$^{-/-}$Th2 cells – this drop was likely due to the loss of *Malat1*. However, further comparison of samples derived from WT and *Malat1*$^{-/-}$ Th2 cells showed that loss of *Malat1* increased SRSF1 binding to mitochondrial ribosomal (mt_r) RNA, small nuclear (sn) RNA, miRNA, ncRNA ribosomal (r) RNA, and small nucleolar (sno) RNA, perhaps these changes in type binding reflect altered SRSF1 localisation in the absence of *Malat1* (***Figure 6.5***).

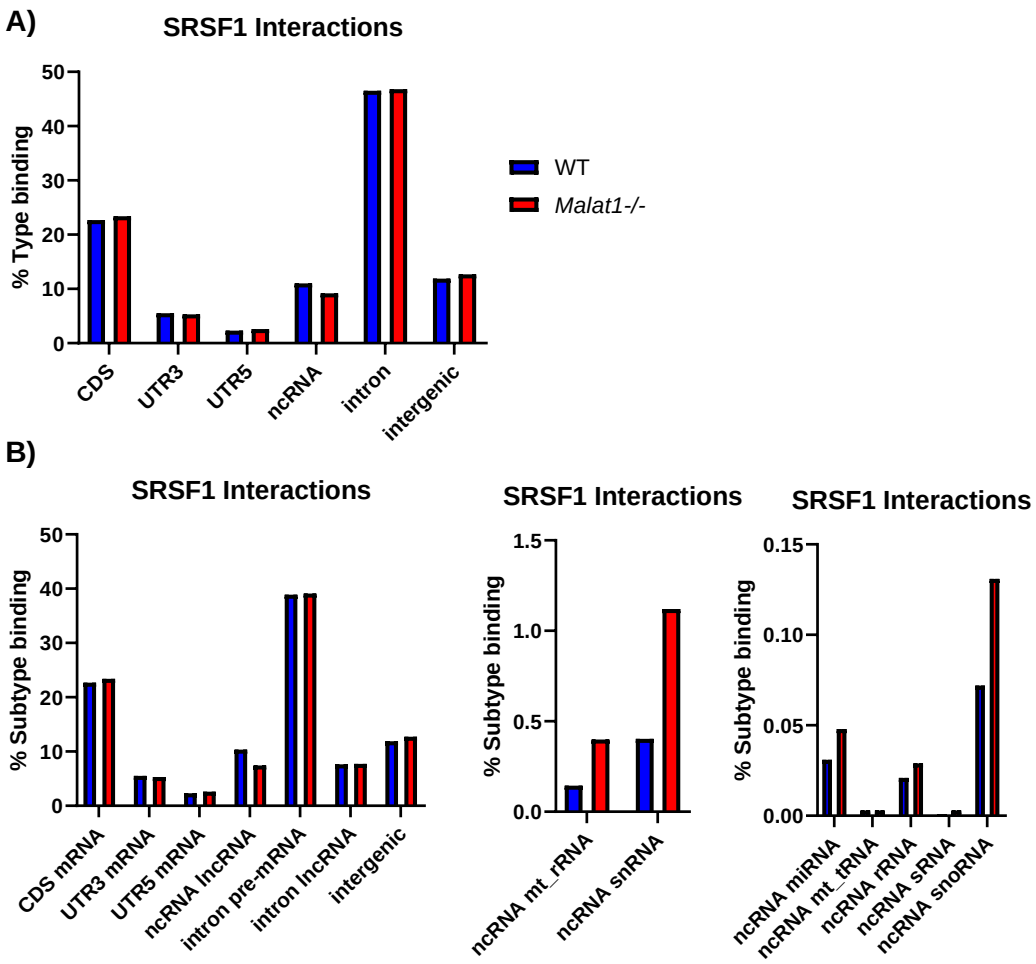


Figure 6.5: SRSF1 interacts with intronic cDNA sequences in Th2 cells.

A) The distribution of SRSF1 iCLIP crosslinking sites in different genomic regions in WT Th2 (blue) and *Malat1*$^{-/-}$ Th2 cells (red) **B)** The distribution of SRSF1 iCLIP crosslinking sites in different genomic regions categorised by subtype in WT Th2 (blue) and *Malat1*$^{-/-}$ Th2 cells (red) – split into three graphs for ease of reading due to scale.

We next examined the top interaction partners for SRSF1 derived from WT and *Malat1*$^{-/-}$ Th2 cells. *Malat1* represented 3.04% of the total reads from iCLIP and was the top interaction partner for SRSF1 (excluding intronic RNA reads) (***Figure 6.6****A*). Other top interaction partners included RNA component of 7SK nuclear ribonucleoprotein (*Rn7sk),* Rho-GTPase-activating protein 15 (*Arhgap15), Gm42418* and inositol polyphosphate 4-phosphatase type II (*Inpp4b). Rn7sk* is an ncRNA that pauses Pol II during RNA synthesis (Bandiera et al., 2021). ARHGAP15 is a GTPase which has been linked to T cell differentiation and mobility (He et al., 2021). *Gm42418* is also a lncRNA which has limited reported functions. INPP4B is involved in PI3K/Akt signalling and has been linked to tumour suppressive functions (Rodgers et al., 2021) GSEA and STRING analysis of SRSF1 top interaction partners identified several genes which are involved in T cell differentiation, immune system function, T cell activation and cell adhesion below, as these experiments took place in a $CD4^+$ T cell context, it is not unsurprising that genes linked to T cell function were identified and several of these RNAs have previously been shown to bind to SRSF1 (***Figure 6.6A/C/D).***

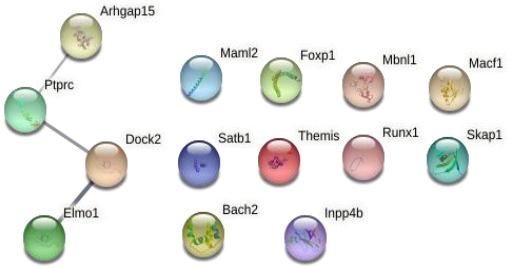
We next manually inspected the binding pattern of SRSF1 across its targets. We found that the binding pattern varied depending on the target bound. SRSF1 bound *Malat1* across the transcript with hotspots at both the 3' and 5' ends (***Figure 6.7***). In some cases, SRSF1 binding occurred at more discrete locations as with *Rn7sk*, whereas there were also transcripts for which binding occurred ubiquitously across the transcript, such as *Argap15 (**Figure 6.7**).* In all examined cases, binding patterns appeared similar between replicates.

A)

Gene ID	Average % reads	Binding in CLIPdb Mouse	Binding in CLIPdb Human
Malat1	3.039876	Yes	Yes
Rn7sk	0.865683	Yes	Yes
Arhgap15	0.573695	No	Yes
Gm42418	0.501905	No	No
Inpp4b	0.487809	No	Yes
Skap1	0.421055	No	No
Dock2	0.376265	No	No
Gm22513	0.369354	Yes	No
Satb1	0.290018	No	Yes
Ptprc	0.267192	No	No
Xist	0.245563	No	Yes
Bach2	0.232838	No	Yes
Themis	0.232076	No	No
Runx1	0.207377	No	Yes
Maml2	0.199544	No	Yes
AC140209.1	0.198338	No	No
Macf1	0.196981	Yes	Yes
Mbnl1	0.192933	Yes	Yes
Foxp1	0.184342	Yes	Yes
Elmo1	0.175943	No	Yes

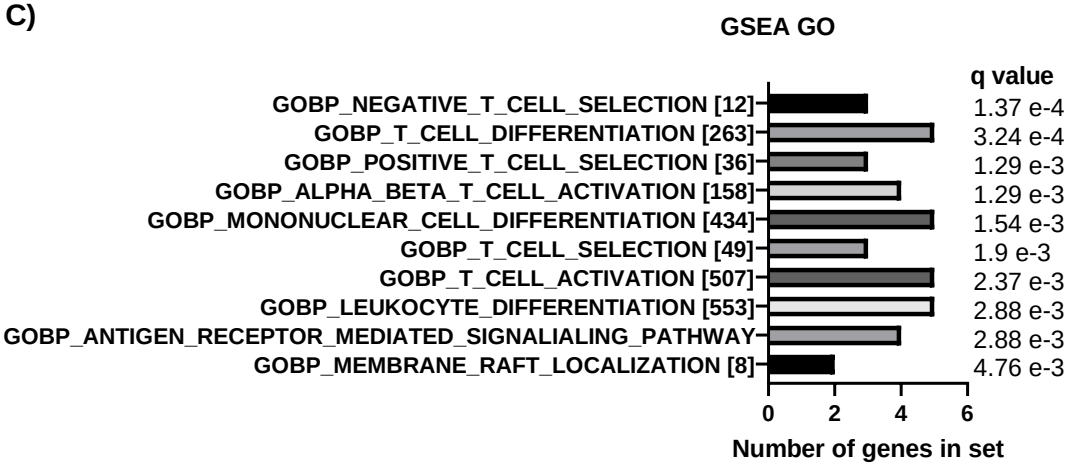
WT SRSF1

B)



Top 20 binders

C)



D)

Top 200 binders

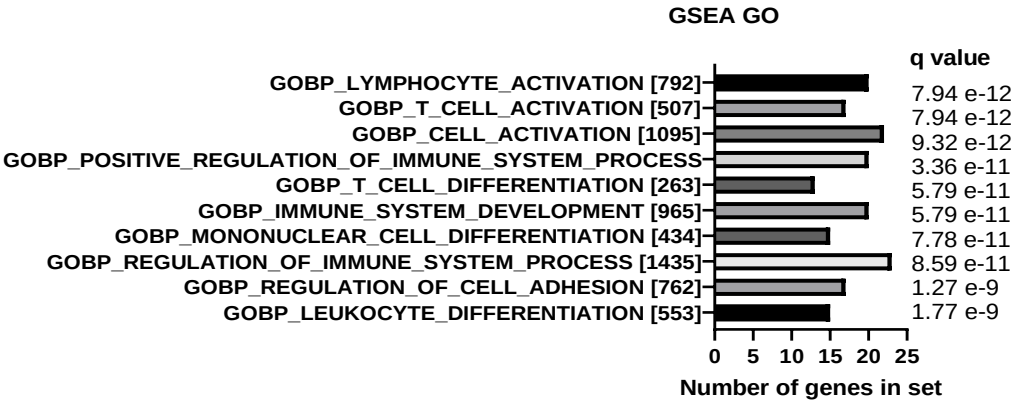


Figure 6.6: SRSF1 interactions are linked to T cell function and differentiation in WT Th2 cells.

A) Table depicting the top SRSF1 interactions in WT *in vitro* polarised Th2 cells determined by iCLIP. The average % reads are shown. Top binders were compared with CLIPdb to determine if they have previously been identified as interaction partners. **B)** STRING interaction network of the top 20 interaction partners of SRSF1 in WT Th2 cells **C)** GSEA GO enrichment analysis of the top 20 SRSF1 interaction partners in WT Th2 cells **D)** GSEA GO enrichment analysis of the top 200 SRSF1 interaction partners in WT Th2 cells.

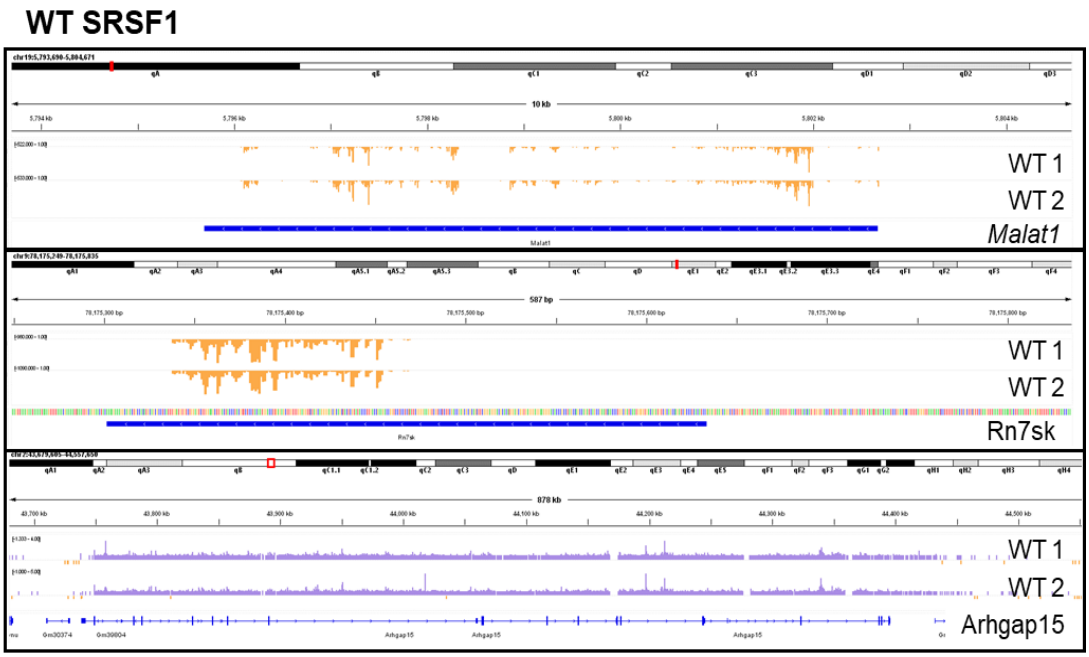


Figure 6.7: SRSF1 coverage tracks in WT Th2 cells.

Coverage plots for SRSF1 iCLIP in WT Th2 cells. Tracks for WT replicate 1 and WT replicate 2 are shown. From top to bottom coverage of the following targets is shown *Malat1, Rn7sk,* and *Arhgap15.* Tracks were generated in IGV.

We next examined the top interaction partners for SRSF1 in *in vitro* polarised Th2 cells derived from *Malat1*$^{-/-}$ mice. This identified similar top interaction targets as that of SRSF1 interaction partners derived from *in vitro* polarised WT Th2 cells with 80% of the top 20 targets appearing in both cell types (***Figure 6.8****A*). Of note, as expected *Malat1* was absent from the list of top interaction partners. Similar to WT cells many of the top interaction targets for SRSF1 in *Malat*$^{1-/-}$ Th2 cells are involved in T cell activation, differentiation and immune system function (***Figure 6.8***).

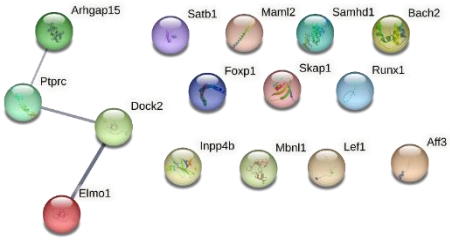
Manual inspection of the binding pattern of SRSF1 in in vitro polarised *Malat1*$^{-/-}$ Th2 cells identified similar binding between replicates (***Figure 6.9***). Binding across the transcript was observed for *Dock2* and *Arghap15*, with more discrete peaks identified for *Rn7sk*. Of the genes examined the binding pattern for SRSF1 appeared similar between WT and *Malat1*$^{-/-}$ cells.

A)

Gene ID	Average % reads	Binding in CLIPdb Mouse	Binding in CLIPdb Human
Gm22513	1.993896	No	No
Rn7sk	0.956107	Yes	Yes
Gm42418	0.553565	No	No
Arhgap15	0.488789	No	Yes
Dock2	0.413273	No	No
Skap1	0.382878	No	No
Inpp4b	0.359913	No	Yes
mt-Rnr2	0.306852	No	Yes
Ptprc	0.305034	No	No
Bach2	0.295986	No	Yes
Satb1	0.269577	No	Yes
Foxp1	0.253297	Yes	Yes
Maml2	0.211195	No	Yes
Mbnl1	0.204646	Yes	Yes
AC140209.1	0.193932	No	No
Lef1	0.187674	No	Yes
Aff3	0.187455	No	Yes
Samhd1	0.176036	No	Yes
Runx1	0.174141	No	Yes
Elmo1	0.173841	No	Yes

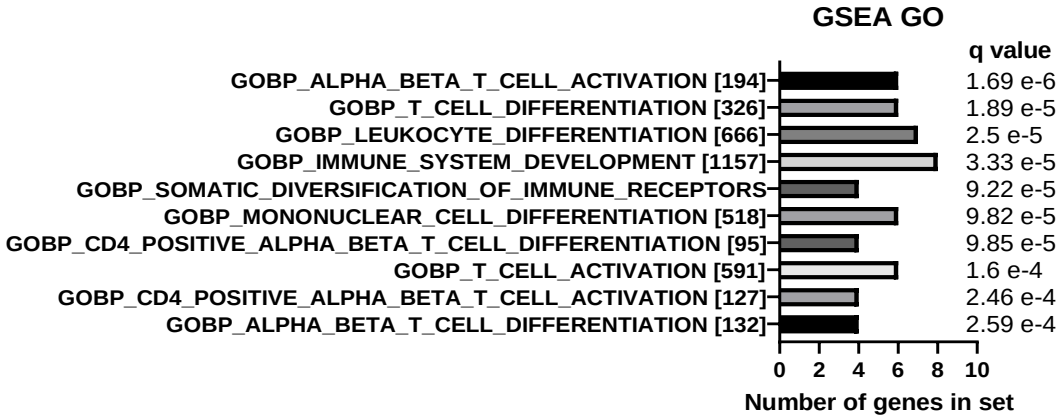
***Malat1*$^{-/-}$ SRSF1**

B)



Top 20 binders

C)



D)

Top 200 binders

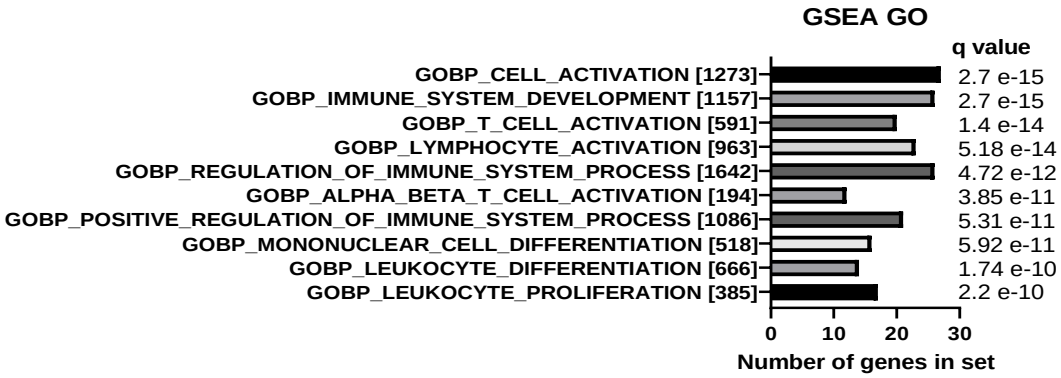


Figure 6.8: SRSF1 interactions are linked to T cell function and differentiation in *Malat1*$^{-/-}$ Th2 cells.

A) Table depicting the top SRSF1 interactions in *Malat1*$^{-/-}$ *in vitro* polarised Th2 cells determined by iCLIP. The average % reads are shown. Top binders were compared with CLIPdb to determine if they have previously been identified as interaction partners. **B)** STRING interaction network of the top 20 interaction partners of SRSF1 in *Malat1*$^{-/-}$ Th2 cells **C)** GSEA GO enrichment analysis of the top 20 SRSF1 interaction partners in *Malat1*$^{-/-}$ Th2 cells **D)** GSEA GO enrichment analysis of the top 200 SRSF1 interaction partners in *Malat1*$^{-/-}$ Th2 cells.

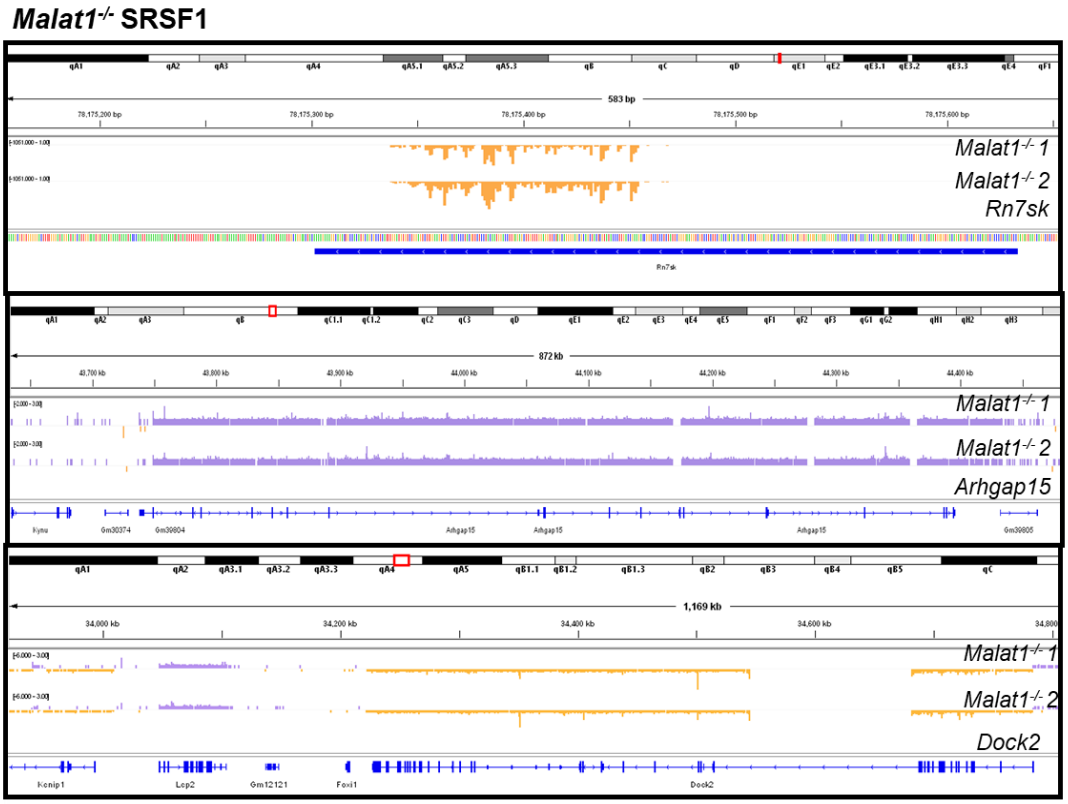


Figure 6.9: SRSF1 coverage tracks in *Malat1*$^{-/-}$ Th2 cells.

Coverage plots for SRSF1 iCLIP in *Malat1*$^{-/-}$ Th2 cells. Tracks for *Malat1*$^{-/-}$ replicate 1 and *Malat1*$^{-/-}$ replicate 2 are shown. From top to bottom coverage of the following targets is shown *Rn7sk, Arhgap1,* and *Dock2.* Tracks were generated in IGV.

In chapter 4, we determined that the loss of *Malat1* altered the transcriptomic landscape of genes derived from chromosome 19 (*Malat1* is located on chromosome 19). As *Malat1* was identified as a top interaction partner for SRSF1, we wanted to determine if loss of *Malat1* altered SRSF1 interactions with a preference for genes derived from chromosome 19. Chromosome analysis of SRSF1 interaction partners determined that SRSF1 interacted with genes on all chromosomes with no obvious bias (***Figure 6.10***). The pattern of binding for SRSF1 derived from WT and *Malat1*$^{-/-}$ Th2 cells was similar. Additionally, the loss of *Malat1* did not appear to impact the proportion of interactions with genes derived from chromosome 19 (***Figure 6.10***).

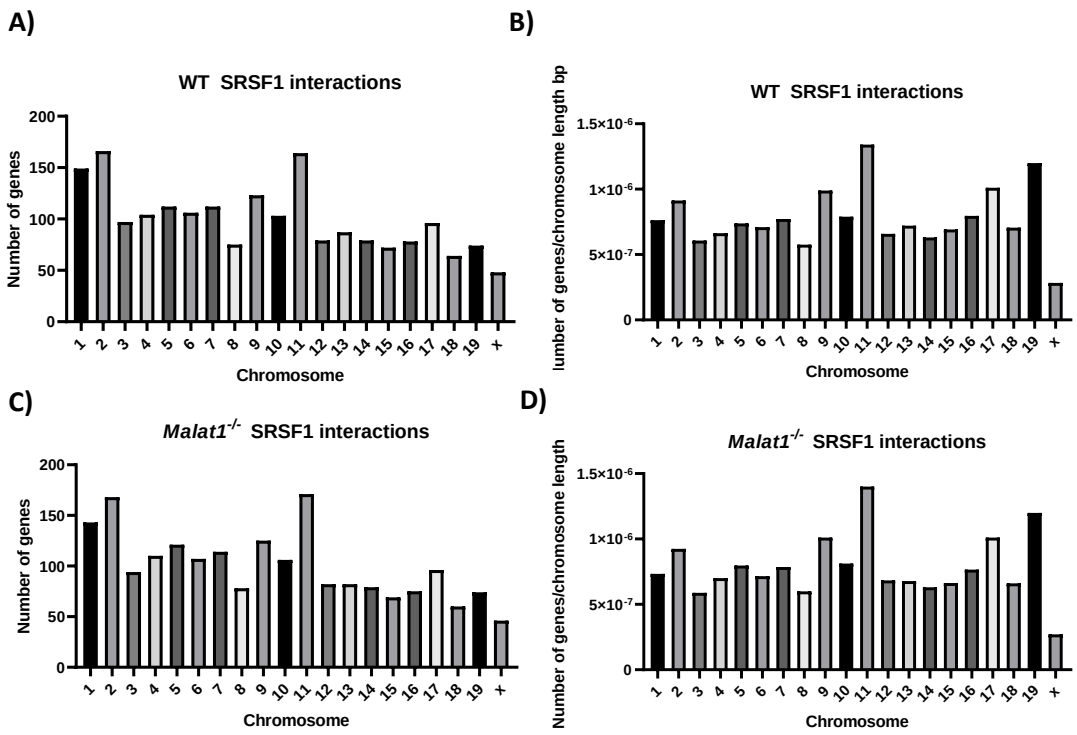


Figure 6.10: SRSF1 binds transcripts derived from chromosomes distributed across the genome.

Chromosome location of genes which bind SRSF1 in **A)** WT *in vitro* polarised Th2 cells determined by iCLIP (expressed as the total number of genes) **B)** WT *in vitro* polarised Th2 cells determined by iCLIP (number of genes normalised to chromosome length). **C)** As in A with cells derived from *Malat1*$^{-/-}$ Th2 cells **D)** As in B with cells derived from *Malat1*$^{-/-}$ Th2 cells. A threshold of >0.01% reads was applied.

We next compared the iCLIP data with our transcriptomic dataset described in chapter 4. We wanted to determine if gene expression levels correlated with the strongest binding to SRSF1. Interestingly, a comparison of the percentage reads determined by iCLIP with CPM determined by RNA-seq showed a weak correlation for both WT and *Malat1*$^{-/-}$ Th2 cells r^2 0.1353 and -0.3361, respectively (***Figure 6.11***). This indicated that factors other than expression were contributing to the binding these could include factors such as k-mer motif (nucleotide sequence), RNA modifications, and phosphorylation status of the RBP.

Thus, to determine if k-mer motif influenced binding, we next compared motif enrichment in the genes which bind SRSF1 in WT and *Malat1*$^{-/-}$ Th cells (***Figure 6.12A***). Positionally enriched k-mer analysis (PEKA), is a computational method used to examine RBPs binding preferences used to identify k-mers this analysis was performed as part of the iMAPs pipeline by Urška Janjoš and Miha Modic. We examined k-mer enrichment across the whole gene and specifically across introns as this is the major type of binding for SRSF1. The top k-mers for SRSF1 in WT Th2 cells across the whole gene were ANNNA, GAA, AGA, CAA(AG)A and AAGA ***(Figure 6.12A).*** Similar motifs were observed for SRSF1 derived from *Malat1*$^{-/-}$ Th2 cells - AAG, GGAAA, AGAC, GA, and CGA (***Figure 6.12***). Of note, the ANNNA motif which occurred most prominently in WT Th2 cells appeared to occur close to the crosslink start site. In comparison, the AAG motif which occurred most prominently in *Malat1*$^{-/-}$ Th2 cells appeared up to ~30nt downstream of the crosslink site. Additionally, the AAGA motif appeared more frequently in WT cells than *Malat1*$^{-/-}$ cells this represented ~2.5% of k-mers across the whole gene, in contrast, this motif was not reported in the top 5 k-mer motifs in *Malat1*$^{-/-}$ Th2 cells ***(Figure 6.12A).***

We next identified k-mer motifs across introns, as this was the most prominent interaction type for SRSF1 (***Figure 6.12B***). The most prominent k-mer for SRSF1 derived from WT Th2 cells was GAA[AG] which represented ~6% of k-mers across introns (***Figure 6.12B***). A similar motif was observed for SRSF1 derived from *Malat1*$^{-/-}$ Th2 cells GAAG which also represented ~6% of the k-mers. Of note the motif GAAA occurred at a greater frequency in *Malat1*$^{-/-}$ Th2 than in WT Th2 cells. A similar motif was observed in WT Th2 cells GGAA however this represented ~3% of the k-mers compared to ~6% of k-mers in *Malat1*$^{-/-}$ Th2 cells (***Figure 6.12B***). Together this data set suggested that loss of *Malat1* may have some impact on k-mer binding motifs however G/A rich motifs occurred frequently in both cell types.

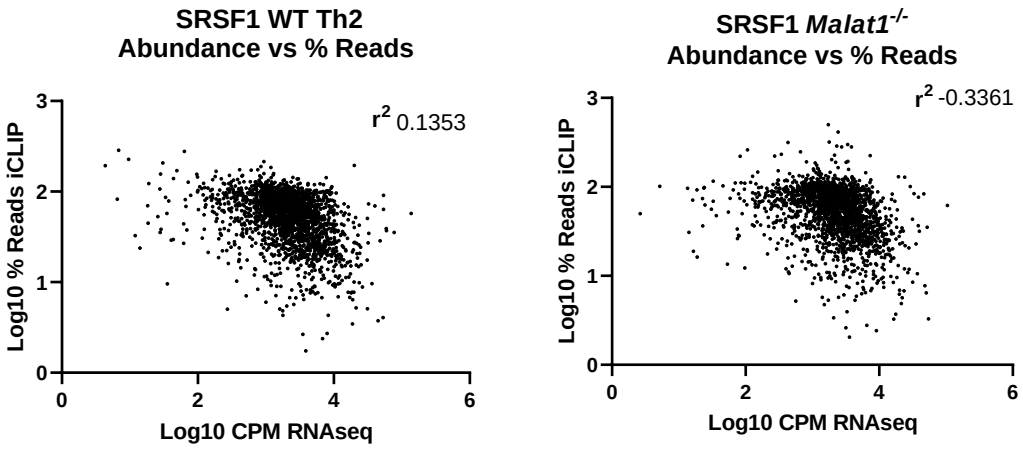


Figure 6.11: SRSF1 interactions do not correlate with RNA abundance.

Comparison of Log_{10} reads derived from SRSF1 iCLIP or Log_{10}CPM of reads from RNAseq data from Th2 cells (chapter 4). Comparison is shown for WT *in vitro* polarised Th2 cells (left) or *Malat1*$^{-/-}$ Th2 cells (right)

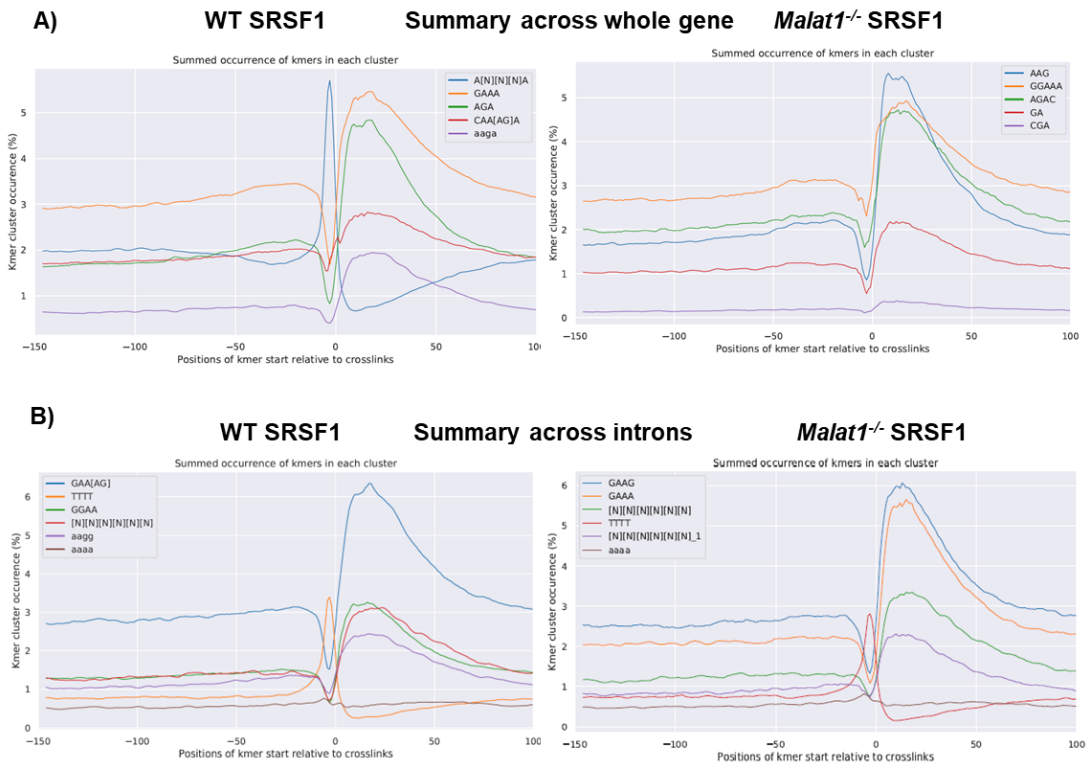


Figure 6.12: SRSF1 interacts with GA-rich motifs.

k-mer analysis across all cross-link sites across the whole gene for SRSF1 iCLIP in **A)** WT Th2 cells or *Malat1*$^{-/-}$ Th2 cells. k-mer analysis in the vicinity of all cross-link sites across genes for SRSF1 iCLIP **B)** WT Th2 cells or *Malat1*$^{-/-}$ Th2 cells are shown. k-mer analysis in the vicinity of all cross-link sites across introns for SRSF1 iCLIP The most highly enriched RNA k-mers are depicted as a percentage of occurrence. The position of k-mers is relative to the crosslink site 0. PEKA was used to determine the k-mer frequency.

We next aimed to determine if the loss of *Malat1* affected the total number or types of genes which interacted with SRSF1 in Th2 cells (***Figure 6.13***). We initially compared total interactions with SRSF1 in WT and *Malat1*$^{-/-}$ Th2 cells at a threshold of >100 reads and >0.01% reads. These thresholds represented over 72% of the total reads in WT Th2 cells. This identified 1,809 common interaction partners in WT and *Malat1*$^{-/-}$ Th2 cells. 193 genes were unique to WT Th2 cells at these thresholds (***Figure 6.13A***). GSEA GO analysis revealed these genes were involved in response to stress, protein localisation, and regulation of transport (***Figure 6.13B***). 211 genes were unique to *Malat1*$^{-/-}$ Th2 cells (***Figure 6.13C).*** These genes were involved in intracellular transport, Golgi function and regulation of enzyme activity ***(Figure 6.13C***). Although these genes were unique to each cell type at these thresholds, many of them do appear but at very low frequencies in both cell types.

Comparison of the log fold change of percentage reads when comparing WT and *Malat1* $^{-/-}$ Th2 cells for SRSF1 iCLIP, showed that loss of *Malat1* caused general decrease in RNAs that bound SRSF1 **(*Figure 6.13D).*** 194 genes reduced binding to SRSF1 at a LogFC > 0.5, and 126 genes which increased binding at a LogFC of >0.5. GSEA GO analysis of these genes determined those which increased in the percentage of reads in *Malat1*$^{-/-}$ Th2 cells were involved in Golgi function, T cell activation and cytokine production (***Figure 6.13E***). In contrast, genes which showed a reduction in the percentage of reads were involved in apoptosis, proliferation and immune system development (***Figure 6.13F***).

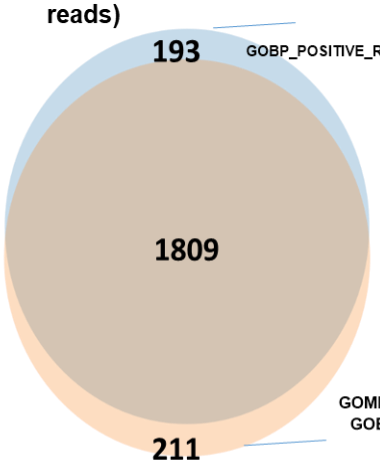
We next wanted to determine if the RNAs decreased in binding to SRSF1 were a result of reduced gene expression. Comparison with our RNAseq data set (Chapter 4) showed no strong correlation with changes in SRSF1 interactions as determined by iCLIP (***Figure 6.14***). In some cases, RNAs interacted with SRSF1 more than

anticipated. For example, they showed a decrease in expression levels in the absence of *Malat1* however showed increased binding to SRSF1 in the absence of *Malat1*. Similarly, some RNAs interacted with SRSF1 less than expected, as some RNAs show increased in gene expression in the absence of *Malat1* but decreased binding to SRSF1.

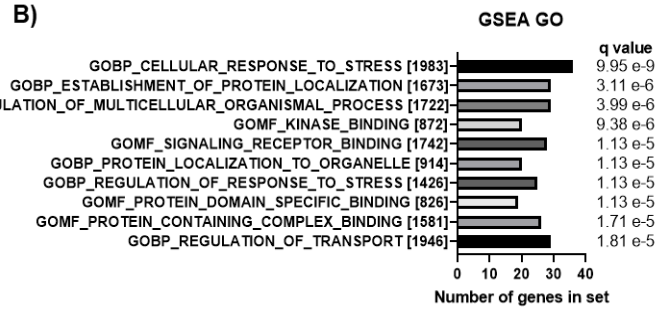
Given that we have previously shown that loss of *Malat1* shows some *in cis-regulatory* effects across chromosome 19 (chapter 4), we wanted to determine if genes that show changes to binding to SRSF1 in the absence of *Malat1* also have a chromosome bias. Analysis of genes which showed increased binding to SRSF1 in the absence of *Malat1* was similarly distributed across the genome **(*Figure 6.15*).** In contrast, analysis of genes which decreased in binding to SRSF1 in the absence of *Malat1* showed some bias toward genes derived from chromosome 19. When normalised to chromosome length chromosome 19 represented the highest number of binding changes in the absence of *Malat1* at both LogFC>1 and LogFC> 0.5. This indicated further *in cis-regulatory* effects for the role of *Malat1* in Th2 cells.

We next manually inspected genes which were identified to show changes in interactions with SRSF1 in the absence of *Malat1 (**Figure 6.16**)*. We chose to inspect genes which were linked to immune cell function and showed some of the biggest changes LogFC in binding. This confirmed changes in binding that were reported at the LogFC level. For example, comparison of SRSF1 binding to *Jak2* in WT and *Malat1*$^{-/-}$Th2 cells showed a decrease in SRSF1 interactions in the absence of *Malat1*. This binding appeared to be reduced across the whole transcript (***Figure 6.16***).

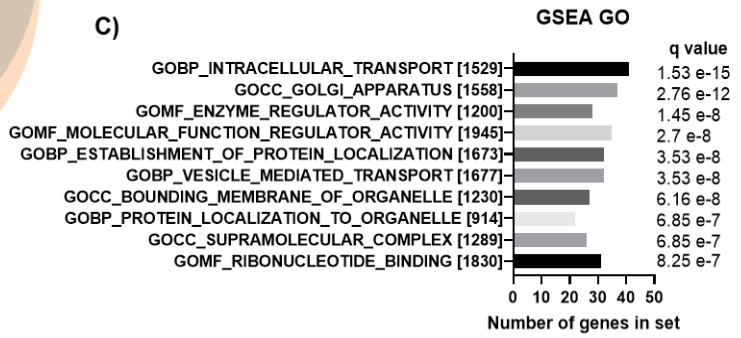
A)



B)



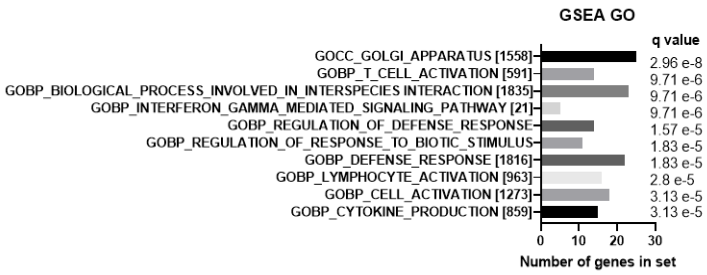
C)



D)

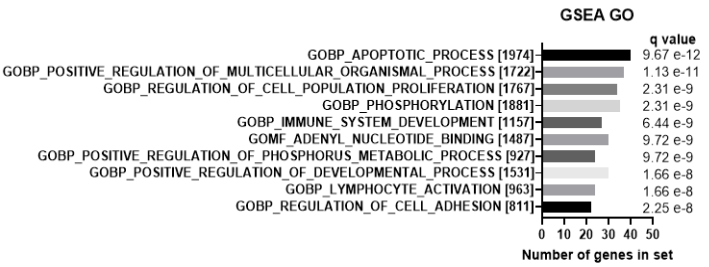
	Decreased binding	Increased binding
LogFC > 0.5	194	126
Log FC>1	50	32
T test sig genes	84	72

E)



Up in *Malat1*$^{-/-}$

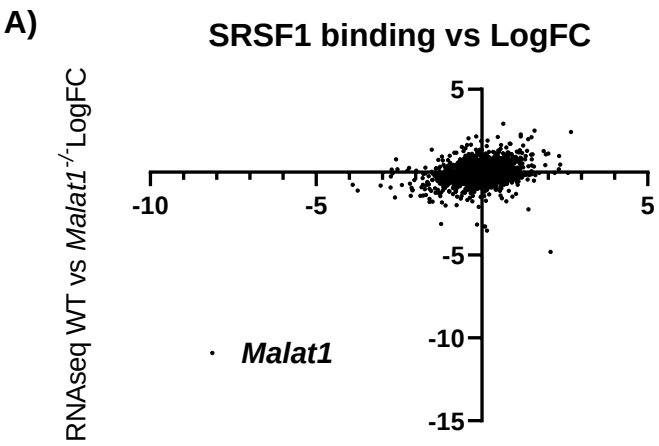
F)



Down in *Malat1*$^{-/-}$

Figure 6.13: Loss of *Malat1* generally decreased SRSF1 interactions.

A) Venn diagram depicting the overlap of the average percentage reads for SRSF1 determined by iCLIP in WT *in vitro* polarised Th2 cells and *Malat1*$^{-/-}$ *in vitro* polarised Th2 cells **B)** GSEA GO analysis of SRSF1 interactions which are unique to WT *in vitro* polarised Th2 cells **C)** GSEA GO analysis of SRSF1 interactions which are unique to *Malat1*$^{-/-}$ *in vitro* polarised Th2 cells **D)** Table depicting the number of genes that increased or decreased binding to SRSF1 and different LogFC changes. An unpaired t-test was used to analyse the significance (n=2 – caution advised). **E)** GSEA GO analysis of genes which increased binding to SRSF1 over LogFC 0.5 when comparing WT and *Malat1*$^{-/-}$ *in vitro* polarised Th2 **F)** GSEA GO analysis of genes which decreased in binding to SRSF1 over LogFC 0.5 when comparing WT and *Malat1*$^{-/-}$ *in vitro* polarisedTh2 cells. A threshold of 0.01% percentage reads and over 100 reads was applied for all comparisons.



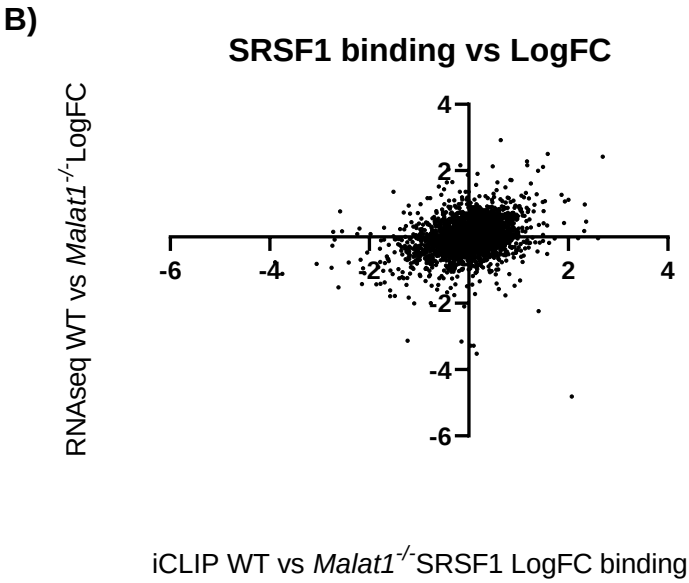


Figure 6.14: Changes in SRSF1 interactions are not influenced by abundance.

A) Comparison of LogFC of changes in gene expression (WT vs *Malat1*$^{-/-}$ Th2 cells) determined by RNAseq with changes in SRSF1 binding (WT vs *Malat1*$^{-/-}$ Th2 cells) determined by iCLIP. **B)** As in A but with *Malat1* removed.

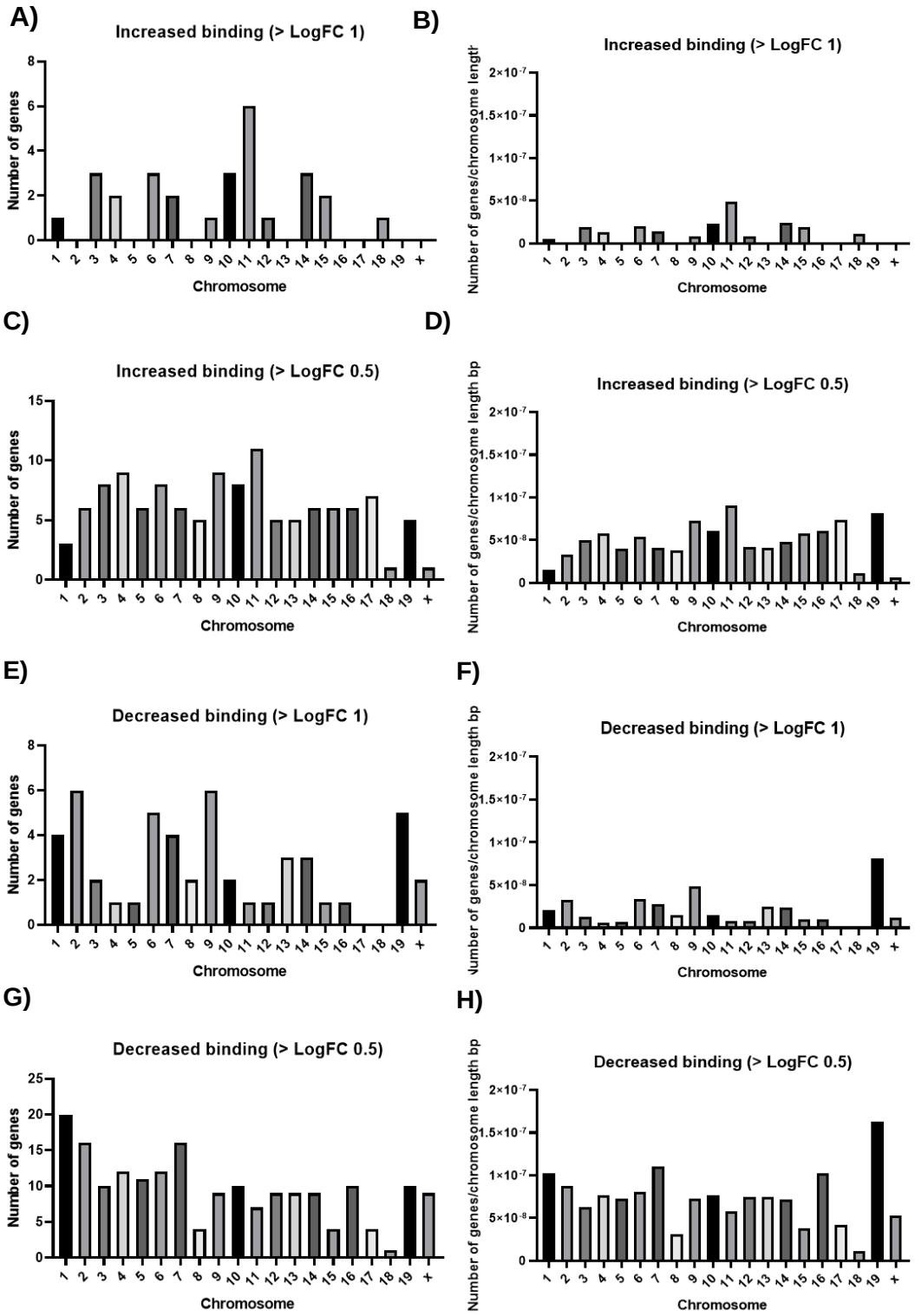

A)
Increased binding (> LogFC 1)
Number of genes
Chromosome
B)
Increased binding (> LogFC 1)
Number of genes/chromosome length bp
Chromosome
C)
Increased binding (> LogFC 0.5)
Number of genes
Chromosome
D)
Increased binding (> LogFC 0.5)
Number of genes/chromosome length bp
Chromosome
E)
Decreased binding (> LogFC 1)
Number of genes
Chromosome
F)
Decreased binding (> LogFC 1)
Number of genes/chromosome length bp
Chromosome
G)
Decreased binding (> LogFC 0.5)
Number of genes
Chromosome
H)
Decreased binding (> LogFC 0.5)
Number of genes/chromosome length bp
Chromosome

Figure 6.15: Loss of *Malat1* altered binding of RNAs to SRSF1 with a bias to those derived from chromosome 19.

Chromosome location of genes bound to SRSF1 which changed in the absence of *Malat1* in *in vitro* polarised Th2 cells **A)** Genes which increased binding LogFC >1 (expressed as a total number of genes), **B)** Genes which increased binding LogFC >1 (normalised to chromosome length) **C)** Genes which increased binding LogFC >0.5 (expressed as a total number of genes), **D)** Genes which increased binding LogFC >0.5(normalised to chromosome length). **E)** Genes which show decreased binding LogFC >1 (expressed as a total number of genes), **F)** Genes which show decreased binding LogFC >1 (normalised to chromosome length) **G)** Genes which decreased binding LogFC >0.5 (expressed as a total number of genes), **H)** Genes which decreased binding LogFC >0.5(normalised to chromosome length).

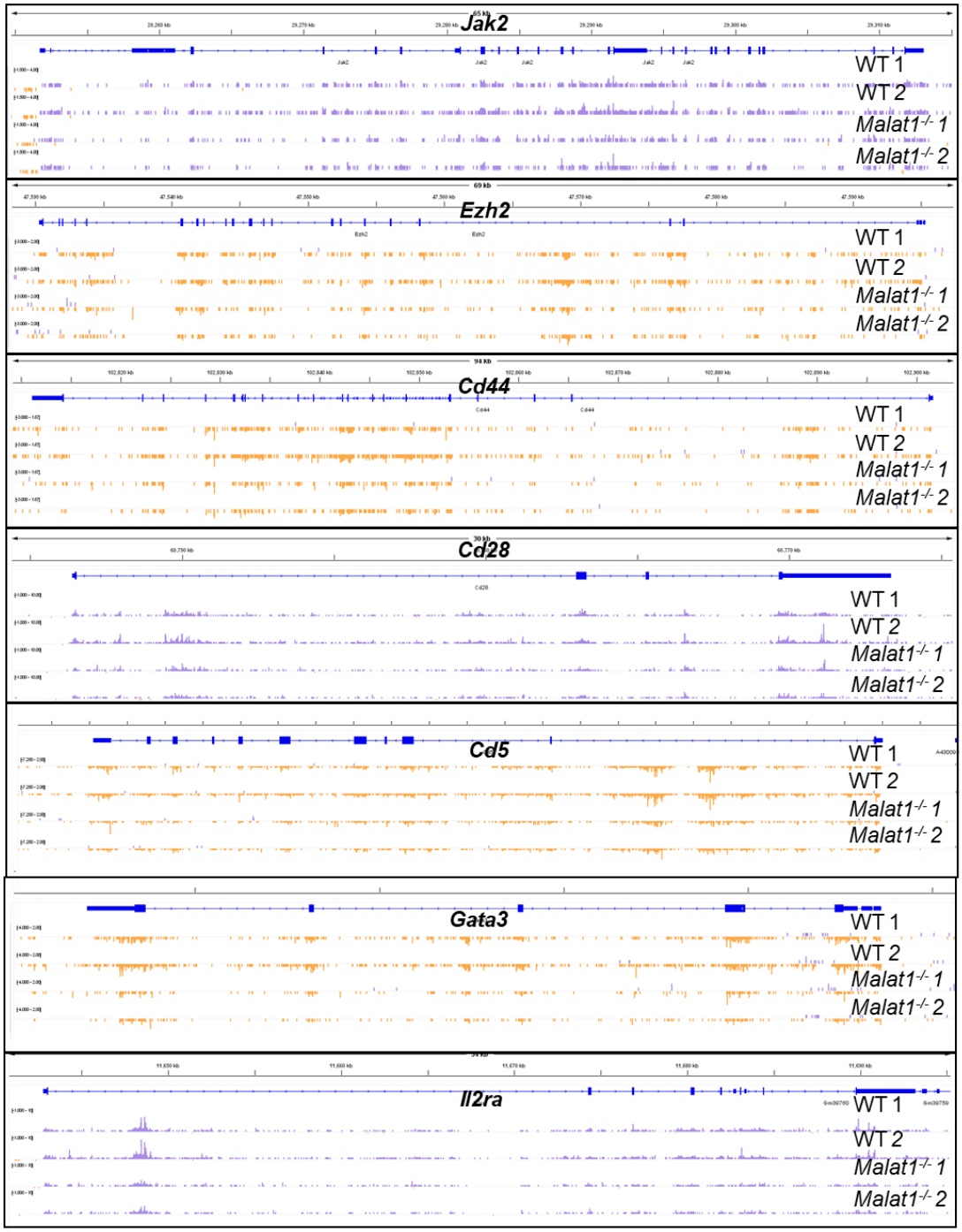


Figure 6.16: Visualisation of changes in SRSF1 binding in the absence of *Malat1.*

Coverage plots for SRSF1 iCLIP in WT and *Malat1*$^{-/-}$ Th2 cells. Tracks for WT Th2 replicate 1, WT Th2 replicate 2, *Malat1*$^{-/-}$ Th2 replicate 1 and *Malat1*$^{-/-}$ Th2 replicate 2 are shown. From top to bottom coverage of the following targets is shown *Jak2, Ezh2,* and *Cd28, Cd5, Gata3, Il2ra.* Tracks were generated in IGV.

To understand why the loss of *Malat1* altered SRSF1 interactions, we next examined SRSF1 localisation in the presence and absence of *Malat1* in *in vitro* polarised Th2 cells. Interestingly, this revealed that in *Malat1*$^{-/-}$ *in vitro* polarised Th2 cells SRSF1 was more cytoplasmic than in WT *in vitro* polarised Th2 cells (***Figure 6.17***). It is likely, that this change in SRSF1 localisation is responsible for some of the changes we observed in SRSF1 interactions.

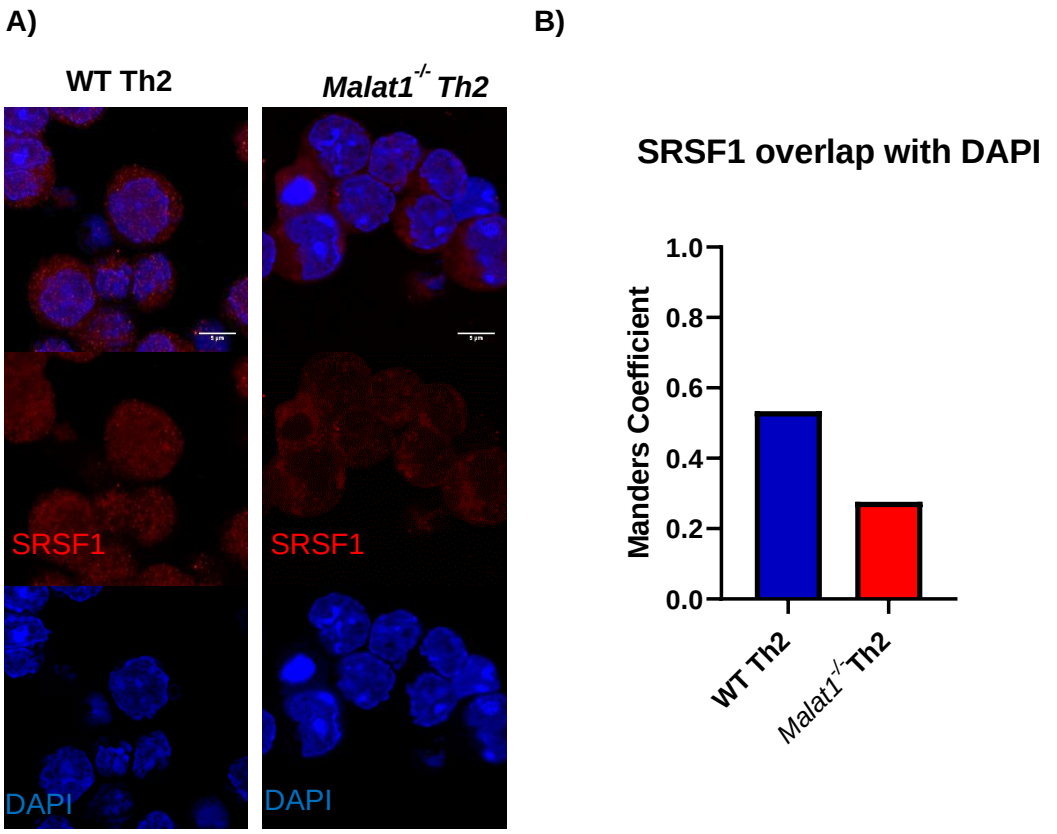


Figure 6.17: *Malat1* retains SRSF1 in the nucleus.

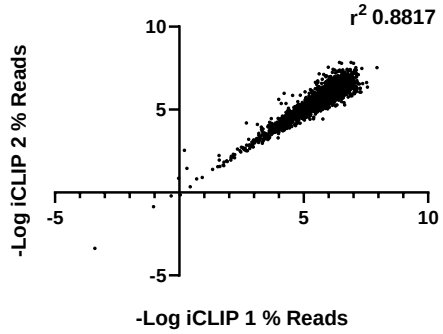
A) Example images of immunofluorescence analysis for SRSF1 (red) and DAPI (blue) in *in vitro* polarised Th2 cells (day 6) samples were derived from female mice. **B)** Quantification of the proportion of overlap of SRSF1 with DAPI in WT and *Malat1*$^{-/-}$ Th2 cells using the ImageJ plugin JaCOP and the manders coefficient. Pilot experiment n=1.

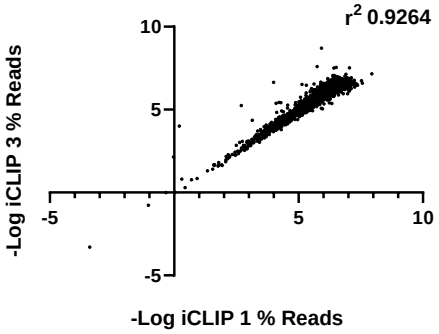
6.2.3 hnRNPA1 interactions in Th2 cells

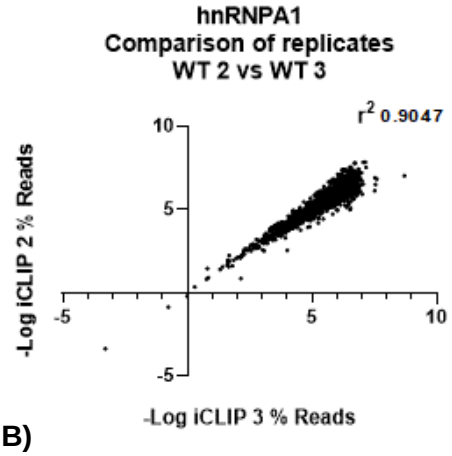
We next examined hnRNPA1 interactions in WT and *Malat1*$^{-/-}$ *in vitro* polarised Th2 cells using iCLIP. Initially, we compared the similarity of iCLIP reads between replicates. We compared the three WT replicates, which all showed a positive correlation, and all replicates correlated similarly with one another. WT replicate 1 and replicate 2 had an r^2 value of 0.8817, WT replicates 1 and 3 had an r^2 value of 0.9264 and WT replicates 2 and 3 had an r^2 of 0.9047 (***Figure 6.18***). Similarly, replicates from *Malat1*$^{-/-}$ Th2 cells strongly correlated with one another. Replicate 1 and 2 had an r^2 of 0.8501, 1 and 3 had an r^2 of 0.9172 and 2 and 3 had an r^2 value of 0.9032 (***Figure 6.18***). These samples did not correlate as strongly as SRSF1 replicates. Of note, Urška Janjoš reported that the SRSF1 antibody appeared to be better for iCLIP than the hnRNPA1 antibody and better libraries were created for SRSF1 (data not shown). This could explain these slight differences between hnRNPA1 replicates compared to SRSF1 replicates.

We then compared the average percentage of reads between WT Th2 and *Malat1*$^{-/-}$ Th2 cells. As with SRSF1 these showed a positive correlation, interestingly, this was similar to the correlation between sample replicates of the same condition (r^2 0.9394) (***Figure 6.19***).

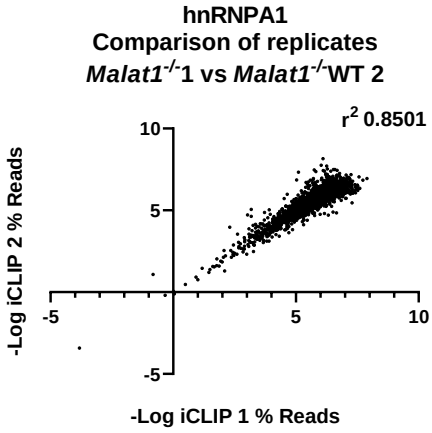
A)

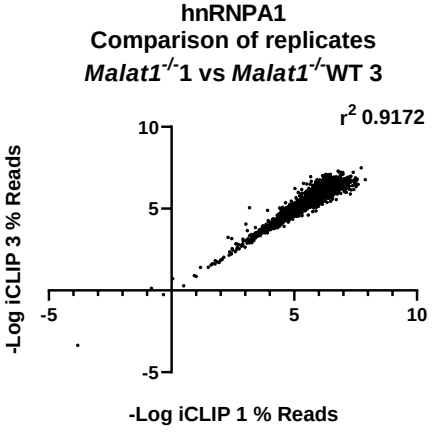

hnRNPA1
Comparison of replicates
WT 1 vs WT 2
r2 0.8817
-Log iCLIP 2 % Reads
-Log iCLIP 1 % Reads


hnRNPA1
Comparison of replicates
WT 1 vs WT 3
r2 0.9264
-Log iCLIP 3 % Reads
-Log iCLIP 1 % Reads


hnRNPA1
Comparison of replicates
WT 2 vs WT 3
r2 0.9047
-Log iCLIP 2 % Reads
-Log iCLIP 3 % Reads

B)


hnRNPA1
Comparison of replicates
Malat1-/-1 vs Malat1-/-WT 2
r2 0.8501
-Log iCLIP 2 % Reads
-Log iCLIP 1 % Reads


hnRNPA1
Comparison of replicates
Malat1-/-1 vs Malat1-/-WT 3
r2 0.9172
-Log iCLIP 3 % Reads
-Log iCLIP 1 % Reads

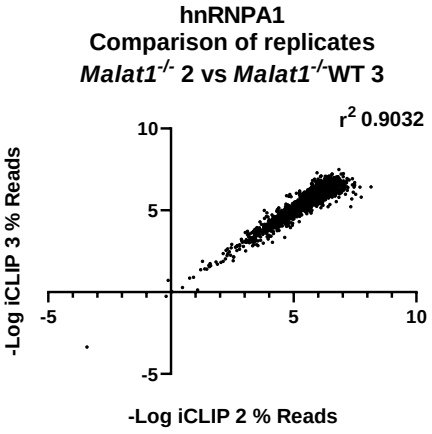

hnRNPA1
Comparison of replicates
Malat1-/- 2 vs Malat1-/-WT 3
r2 0.9032
-Log iCLIP 3 % Reads
-Log iCLIP 2 % Reads

Figure 6.18: hnRNPA1 iCLIP samples positively correlated with one another.

A) Correlation of Log_2 percentage of reads between three UV cross-linked WT Th2 hnRNPA1 iCLIP samples **B)** Correlation of Log_2 percentage of reads between two UV cross-linked *Malat1*$^{-/-}$ Th2 hnRNPA1 iCLIP samples. A threshold of >0.01% reads was applied. n=3 biological replicates representative of 1 independent experiment.

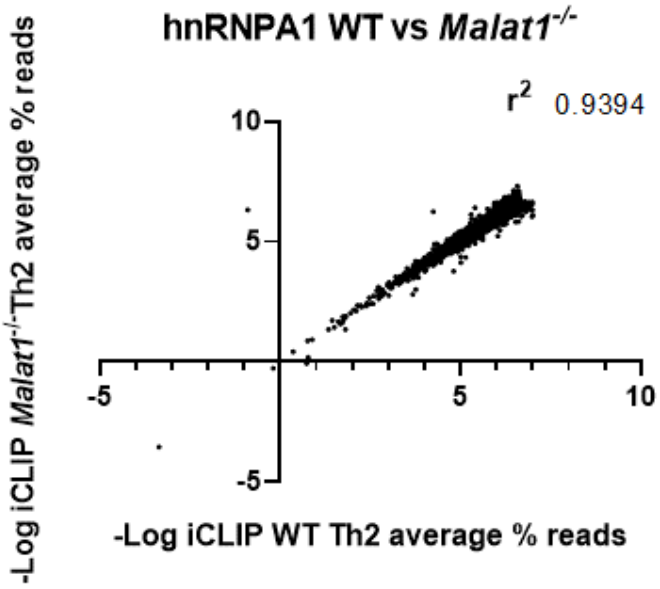


Figure 6.19: hnRNPA1 reads from WT and *Malat1*$^{-/-}$ *in vitro* polarised Th2 cells positively correlate.

Correlation of average Log2 percentage of reads between UV cross-linked WT Th2 hnRNPA1 iCLIP samples and UV cross-linked *Malat1*$^{-/-}$ Th2 iCLIP samples n=3 biological replicates representative of 1 independent experiment. A threshold of >0.01% reads was applied.

We next wanted to determine which RNA regions hnRNPA1 was bound to. Genomic annotation of the bound cDNA sequences revealed that hnRNPA1 is most prominently bound to introns ~72% ***(*Figure *6.20A)***. Only small changes in binding were observed in the absence of *Malat1* with a ~2% drop in ncRNA binding observed in the *Malat1*$^{-/-}$Th2 cells likely representing the loss of *Malat1*. However, when samples derived from WT and *Malat1*$^{-/-}$ Th2 cells were compared, this showed hnRNPA1 increased in binding to mt_rRNA, snoRNA and mt_tRNA in the *Malat1*$^{-/-}$ Th2 cells. A slight decrease in rRNA interactions was also observed in the absence of *Malat1* (***Figure 6.20B***).

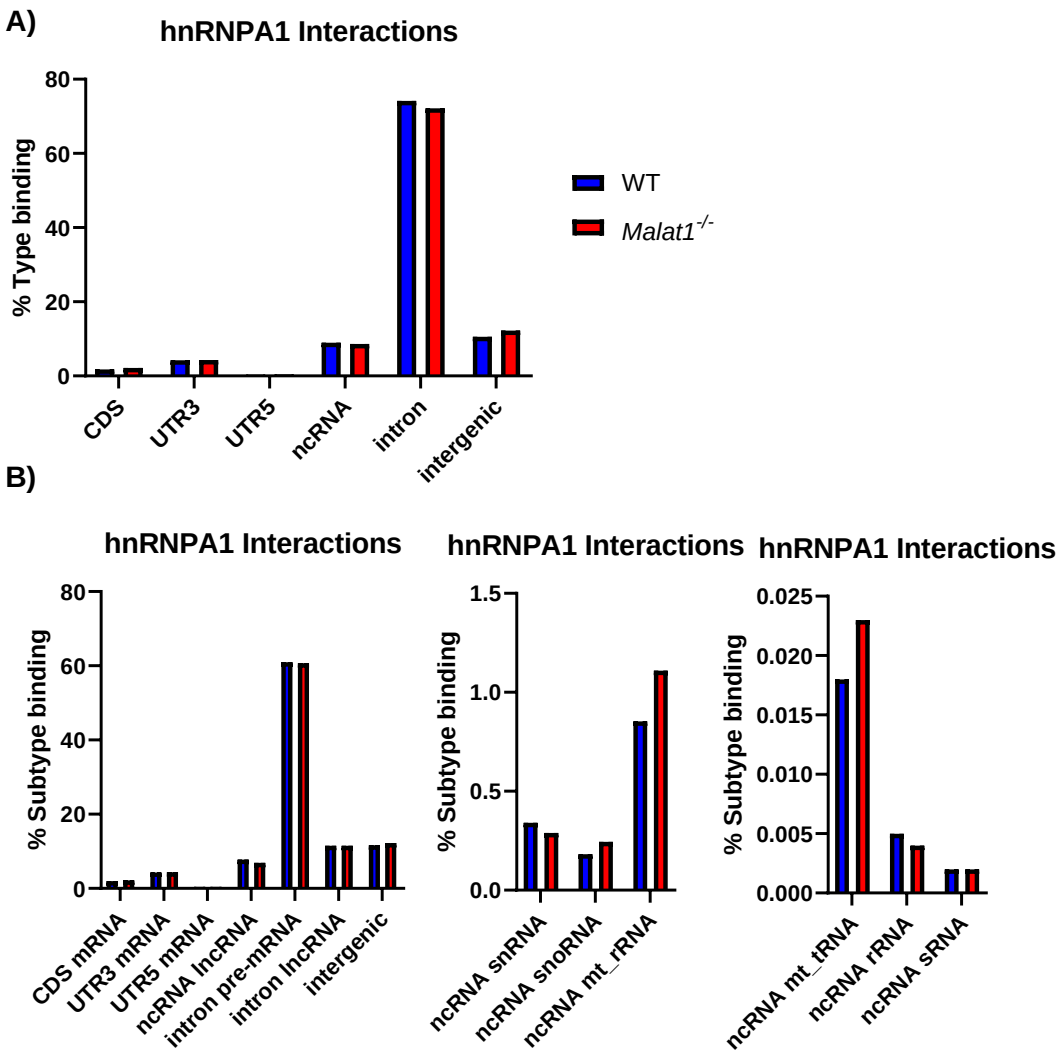


Figure 6.20: hnRNPA1 interacts with intronic cDNA sequences in Th2 cells.

A) The distribution of hnRNPA1 iCLIP crosslinking sites in different genomic regions in WT Th2 (blue) and *Malat1*$^{-/-}$ Th2 cells (red) **B)** The distribution of hnRNPA1 iCLIP crosslinking sites in different genomic regions categorised by subtype in WT Th2 (blue) and *Malat1*$^{-/-}$ Th2 cells (red) – split into three graphs for ease of reading due to scale.

We then examined the top interaction partners for hnRNPA1 derived from WT *in vitro* polarised Th2 cells. *Malat1* represented 1.86% of the total reads for iCLIP and was the top interaction partner in WT cells (excluding intronic RNA reads) (**Figure *6.21*A**). This was a similar but slightly reduced number of percentage reads as that observed for *Malat1* interactions with SRSF1 ~3.04%. Interestingly, we observed similar interactions to those observed with SRSF1, other top interaction partners included *Rn7sk, Arhgap,* and *Inpp4b.* As the top interaction partners of SRSF1 and hnRNPA1 were similar, GSEA and STRING analysis also identified several genes which are involved in T cell differentiation, immune system function, and T cell activation (**Figure *6.21*B/C/D**).

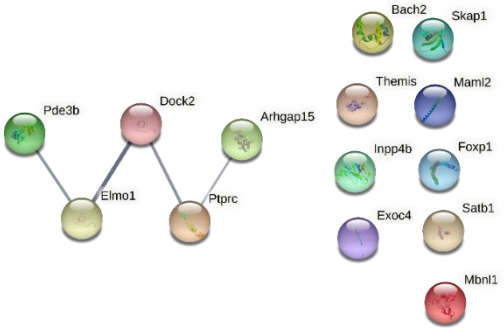
We next inspected the binding pattern of hnRNPA1 across its targets. We found that the hnRNPA1 binding patterns differed depending on the target. In some cases, hnRNPA1 bound discrete locations as with *Rn7sk* and *Inpp4b,* in other cases hnRNPA1 bound ubiquitously across the transcript for *Argap15 (***Figure *6.22*)**. Of note, the binding patterns appeared similar between replicates.

hnRNPA1 WT Th2 cells

A)

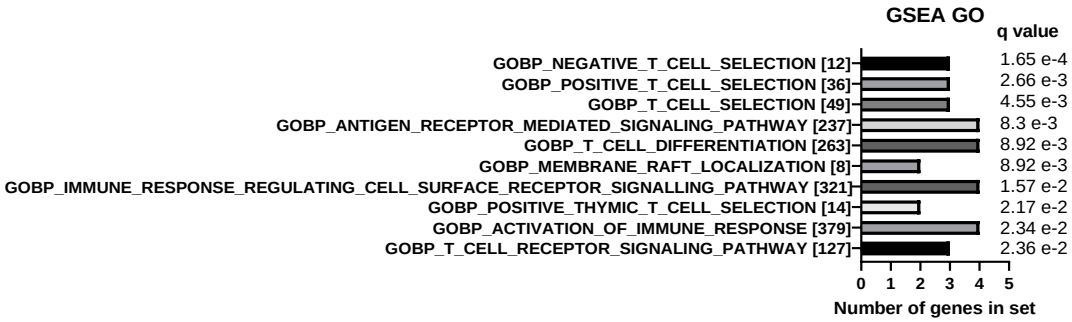
Gene ID	Average % reads WT	Binding in CLIPdb? Mouse	Binding in CLIPdb? Human
Malat1	1.85762	N/A	Yes
Rn7sk	1.137781	N/A	Yes
Arhgap15	1.024078	N/A	Yes
Inpp4b	0.779775	N/A	Yes
Satb1	0.598553	N/A	Yes
mt-Rnr2	0.589625	N/A	Yes
Mbnl1	0.582462	N/A	Yes
Gm42418	0.539405	N/A	No
Ptprc	0.392864	N/A	No
Dock2	0.366279	N/A	Yes
Pde3b	0.352741	N/A	Yes
Skap1	0.325103	N/A	Yes
Bach2	0.313592	N/A	Yes
Foxp1	0.303078	N/A	Yes
Elmo1	0.300336	N/A	Yes
mt-Rnr1	0.297048	N/A	Yes
Themis	0.286241	N/A	Yes
Maml2	0.285074	N/A	Yes
Gm22513	0.282383	N/A	No
Exoc4	0.257672	N/A	Yes

B)



C)

Top 20 binders



D)

Top 200 binders

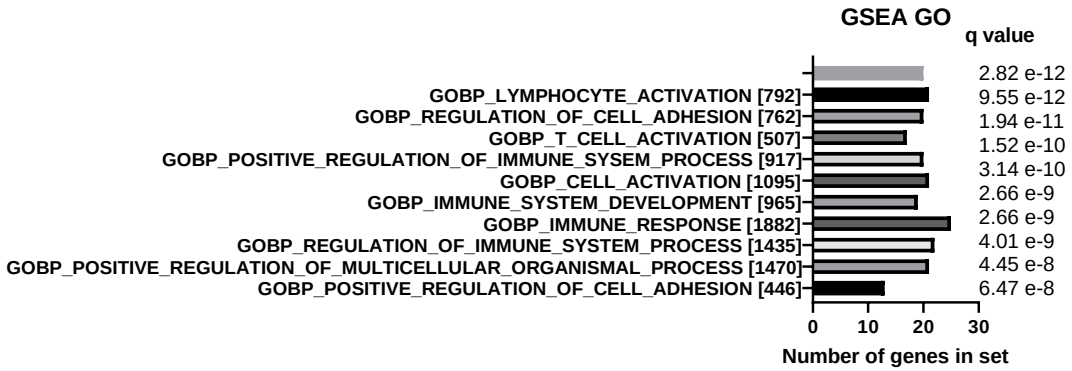


Figure 6.21: hnRNPA1 interactions are linked to T cell function and differentiation in WT Th2 cells.

A) Table depicting the top hnRNPA1 interactions in WT *in vitro* polarised Th2 cells determined by iCLIP. The average % reads are shown. Top binders were compared with CLIPdb to determine if they have previously been identified as interaction partners. **B)** STRING interaction network of the top 20 interaction partners of hnRNPA1 in WT Th2 cells **C)** GSEA GO enrichment analysis of the top 20 hnRNPA1 interaction partners in WT Th2 cells **D)** GSEA GO enrichment analysis of the top 200 hnRNPA1 interaction partners in WT Th2 cells.

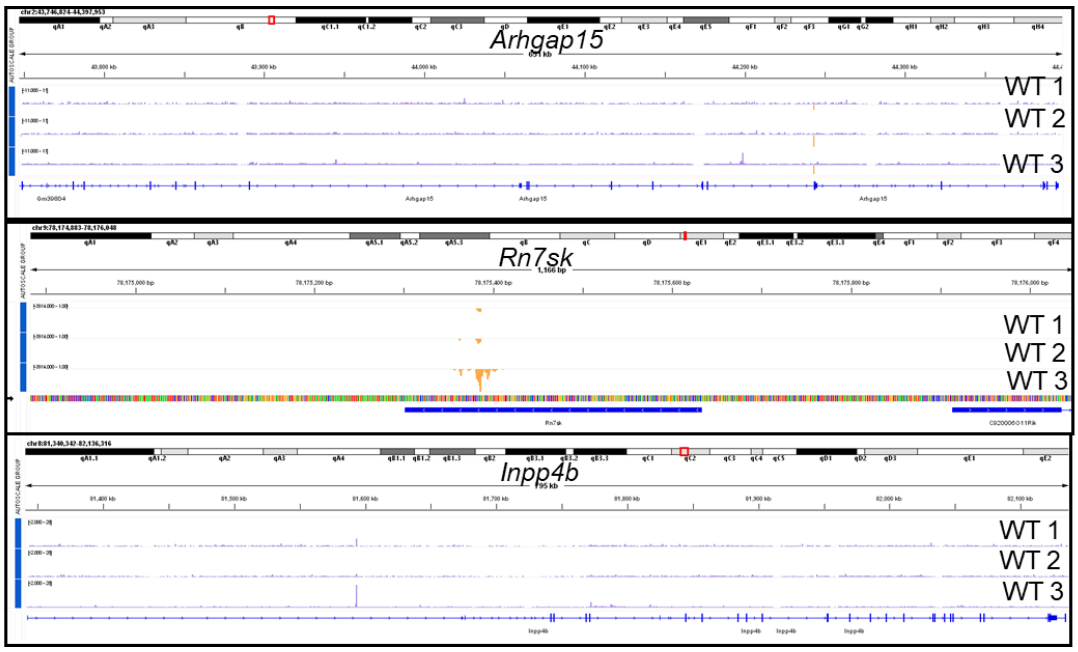


Figure 6.22: hnRNPA1 coverage tracks in WT Th2 cells.

Coverage plots for hnRNPA1 iCLIP in WT Th2 cells. Tracks for WT replicate 1, WT replicates 2 and WT replicate 3 are shown. From top to bottom coverage of the following targets is shown *Arhgap15, Inpp4b and Rn7sk.* Tracks were generated in IGV.

Next, we examined the top interaction partners for hnRNPA1 derived from *Malat1*$^{-/-}$ *in vitro* polarised Th2 cells. As expected, *Malat1* was absent from the list of top interaction partners (***Figure 6.23A***). Similar interaction partners were identified in WT and *Malat1*$^{-/-}$ *in vitro* polarised Th2 cells as 95% of the top 20 interaction partners appeared in both cell types. As expected, given the strong overlap of interaction partners to WT Th2 cells GSEA and STRING analysis also identified that hnRNPA1 interacted with genes that were involved in T cell differentiation, immune system function, and T cell activation (***Figure 6.23B/C/D***).

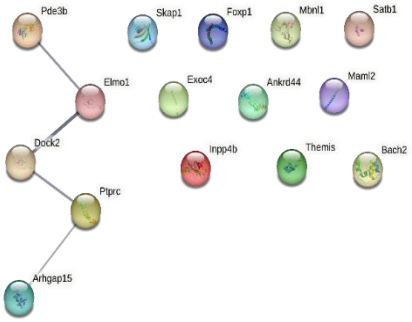
Manual inspection of hnRNPA1 binding revealed that hnRNPA1 binding patterns differed depending on the target. These patterns were similar to what was observed in WT cells. In some cases, hnRNPA1 bound discrete locations as with *Rn7sk* and *Inpp4b,* in other cases hnRNPA1 bound ubiquitously across the transcript for *Argap15* (***Figure 6.24***). As with the other CLIP experiments, the binding patterns appeared similar between replicates. Chromosome analysis of hnRNPA1 interaction partners determined that like SRSF1, hnRNPA1 interacted with genes derived from all chromosomes with no obvious bias and this was similar between WT and *Malat1*$^{-/-}$ conditions (***Figure 6.25***).

A)

hnRNPA1 *Malat1*$^{-/-}$ Th2 cells

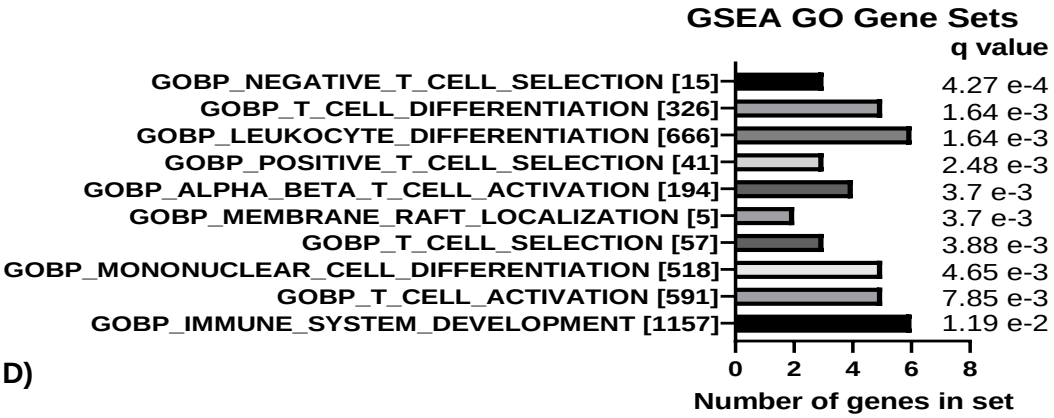
Gene ID	Average % reads WT	Binding in CLIPdb? Mouse	Binding in CLIPdb? Human
Rn7sk	1.200795	N/A	Yes
Arhgap15	0.984699	N/A	Yes
Gm42418	1.049581	N/A	No
mt-Rnr2	0.88537	N/A	Yes
Inpp4b	0.750569	N/A	Yes
Satb1	0.547857	N/A	Yes
Mbnl1	0.52814	N/A	Yes
Ptprc	0.395102	N/A	No
mt-Rnr1	0.393731	N/A	Yes
Dock2	0.37106	N/A	Yes
Skap1	0.333962	N/A	Yes
Bach2	0.32629	N/A	Yes
Gm22513	0.304733	N/A	Yes
Pde3b	0.314491	N/A	Yes
Foxp1	0.300838	N/A	Yes
Elmo1	0.278815	N/A	Yes
Maml2	0.269721	N/A	Yes
Themis	0.273111	N/A	Yes
Exoc4	0.247348	N/A	Yes
Ankrd44	0.229695	N/A	Yes

B)



C)

Top 20 binders



D)

Top 200 binders

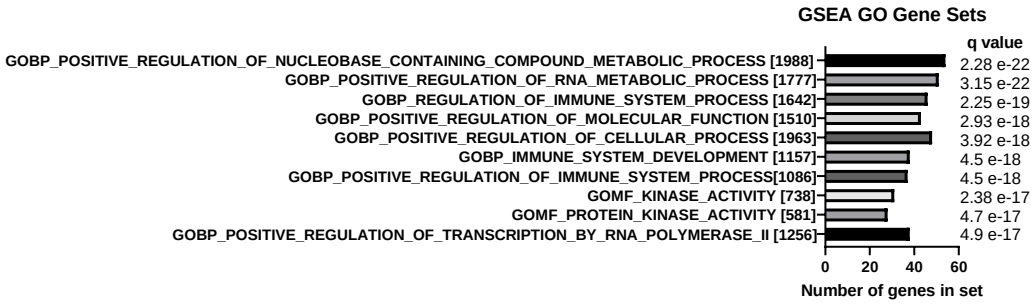


Figure 6.23: hnRNPA1 interactions are linked to T cell function and differentiation in *Malat1*$^{-/-}$ Th2 cells.

A) Table depicting the top hnRNPA1 interactions in *Malat1*$^{-/-}$ *in vitro* polarised Th2 cells determined by iCLIP. The average % reads are shown. Top binders were compared with CLIPdb to determine if they have previously been identified as interaction partners. **B)** STRING interaction network of the top 20 interaction partners of hnRNPA1 in *Malat1*$^{-/-}$ Th2 cells **C)** GSEA GO enrichment analysis of the top 20 hnRNPA1 interaction partners in *Malat1*$^{-/-}$ Th2 cells **D)** GSEA GO enrichment analysis of the top 200 hnRNPA1 interaction partners in *Malat1*$^{-/-}$ Th2 cells.

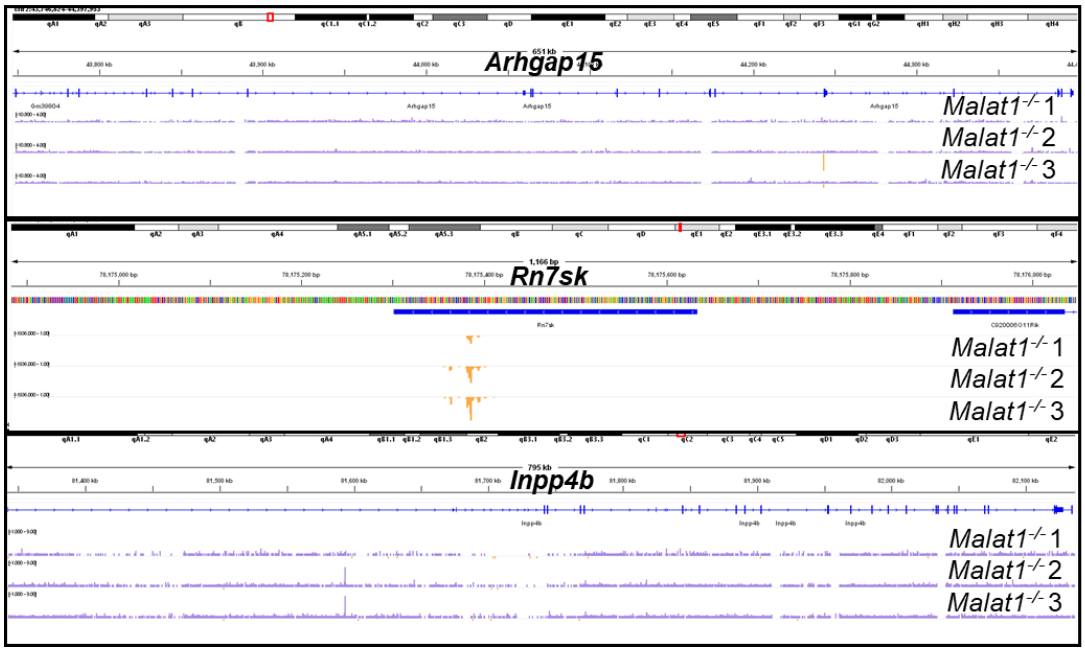


Figure 6.24: hnRNPA1 coverage tracks in *Malat1*$^{-/-}$ Th2 cells.

Coverage plots for hnRNPA1 iCLIP in *Malat1*$^{-/-}$ Th2 cells. Tracks for *Malat1*$^{-/-}$ replicate 1, *Malat1*$^{-/-}$ replicate 2 and *Malat1*$^{-/-}$ replicate 3 are shown. From top to bottom coverage of the following targets is shown *Arhgap15, Rn7sk and Inpp4b.* Tracks were generated in IGV.

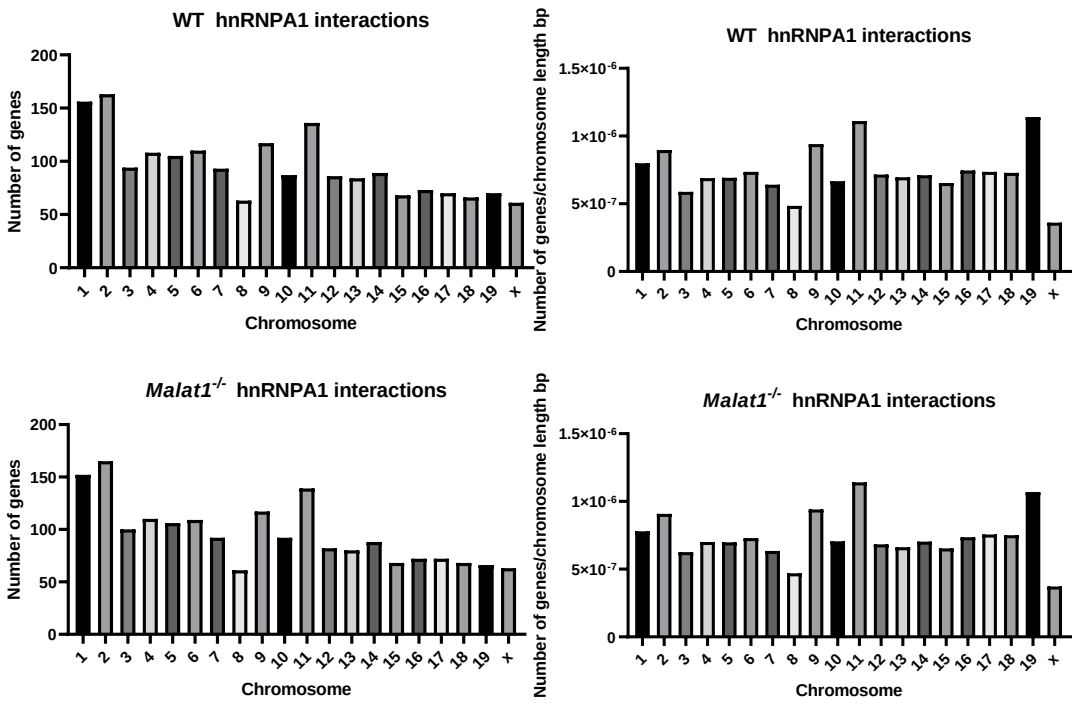


Figure 6.25: hnRNPA1 binds genes derived from chromosomes distributed across the genome.

Chromosome location of genes which bind hnRNPA1 in **A)** WT *in vitro* polarised Th2 cells determined by iCLIP (expressed as the total number of genes) **B)** WT *in vitro* polarised Th2 cells determined by iCLIP (number of genes normalised to chromosome length). **C)** As in A with cells derived from *Malat1*$^{-/-}$ Th2 cells **D)** As in B with cells derived from *Malat1*$^{-/-}$ Th2 cells.

Next, we further investigated our iCLIP dataset and compared it with our transcriptomic dataset described in chapter 4. As shown for SRSF1, binding did not correlate with expression for hnRNPA1 interaction partners. Only a weak negative correlation was observed of with an r^2-0.24 (WT Th2) and -0.241 *(Malat1*$^{-/-}$ Th2) (***Figure 6.26***). Thus, it is likely that other factors such as RNA motifs contribute to the percentage of hnRNPA1 interactions.

Therefore, to determine if RNA motif influenced hnRNPA1 interactions, PEKA was used to determine hnRNPA1 binding preferences in WT and *Malat1*$^{-/-}$ Th2 cells (**Figure *6.27***). As previously mentioned, Urška Janjoš and Miha Modic performed the PEKA analysis. Interestingly, analysis of binding sites across whole genes identified the top k-mers for hnRNPA1 in WT Th2 cells. These motifs were AGGG and GAAA (***Figure 6.27****A*) Similar motifs were observed for hnRNPA1 derived from *Malat1*$^{-/-}$ Th2 cells - TAGG, AGGU and AAGG (***Figure 6.27****A*). Remarkably, several motifs were not specifically defined and were depicted as a series of N. Additionally, the peaks for the analysis appeared noisy, and this suggested that the crosslinking at the site of hnRNPA1 interactions was poor.

As hnRNPA1 most prominently bound introns, we next analysed k-mer motifs across introns (***Figure 6.27****B*). The most prominent k-mer for hnRNPA1 derived from WT Th2 cells was AGAG which represented ~3.5% of k-mers across introns and AAGA which represented ~2.5% of k-mers across introns (***Figure 6.27****B*). Similar AG-rich motifs were observed for hnRNPA1 derived from *Malat1*$^{-/-}$ Th2 cells TAGG, AAGG and AGGG. Together this data set suggested that loss of *Malat1* may have some impact on k-mer binding motifs however G/A rich motifs occurred frequently in both cell types.

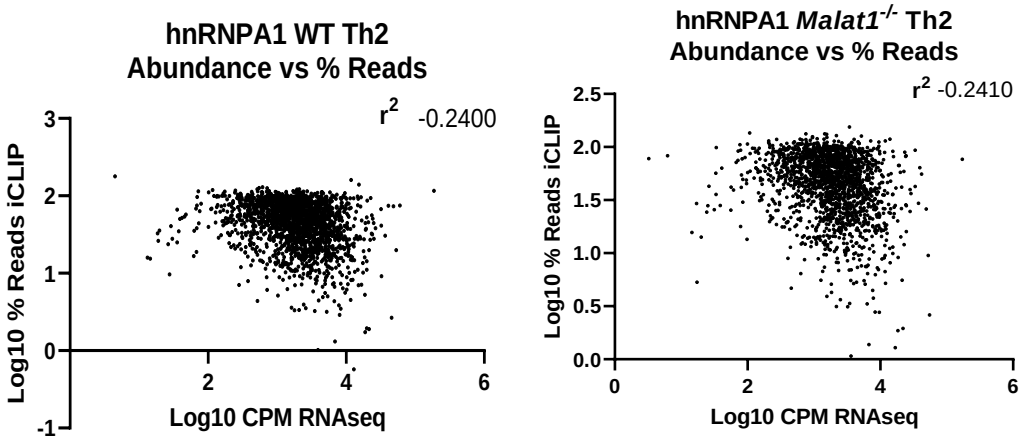


Figure 6.26: hnRNPA1 interactions do not correlate with RNA abundance.

Comparison of Log10 reads derived from hnRNPA1 iCLIP or Log10CPM of reads from RNAseq data from Th2 cells (chapter 4). Comparison is shown for WT in vitro polarised Th2 cells (left) or *Malat1*$^{-/-}$ Th2 cells (right).

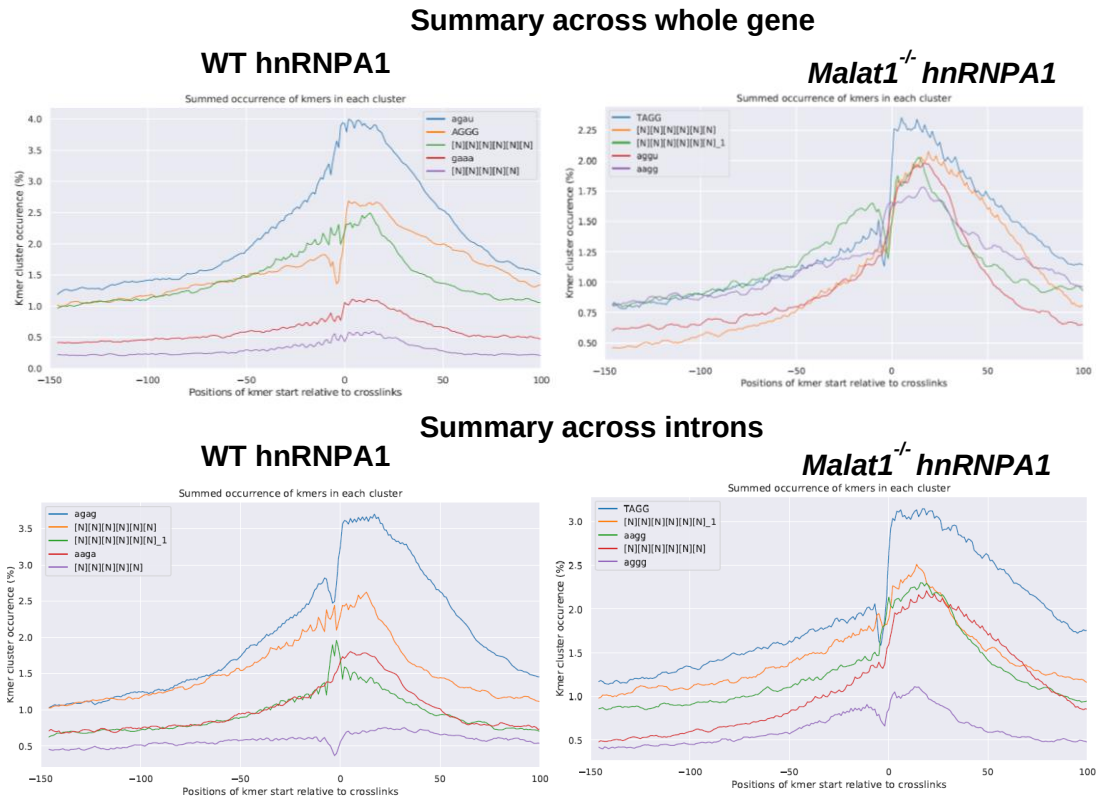


Figure 6.27: hnRNPA1 interacts with AG-rich motifs.

k-mer analysis across all cross-link sites across the whole gene for hnRNPA1 iCLIP in **A)** WT Th2 cells or *Malat1*$^{-/-}$ Th2 cells. k-mer analysis in the vicinity of all cross-link sites across introns for hnRNPA1 iCLIP **B)** WT Th2 cells or B) *Malat1*$^{-/-}$ Th2 cells. The most highly enriched RNA pentamers are depicted as a percentage of occurrence. The position of k-mers is relative to the crosslink site 0. PEKA was used to determine the k-mer frequency.

To assess whether loss of *Malat1* altered hnRNPA1 interactions, we compared iCLIP reads from WT and *Malat1*$^{-/-}$ Th2 cells at a threshold of >100 reads and >0.01% reads. This identified 1,812 reads which were common to hnRNPA1 derived from both WT and *Malat1*$^{-/-}$ Th2 cells **(*Figure 6.28)***. 100 genes were unique to WT Th2 cells at this threshold, and GSEA GO analysis determined that these genes were involved in catalytic complexes, and protein-containing complexes ***(Figure 6.28B***). 106 genes were unique to hnRNPA1 derived from *Malat1*$^{-/-}$ Th2 cells ***(Figure 6.28A***). GSEA GO analysis revealed that these genes were involved in protein localisation, Golgi function and biosynthetic processes cells ***(Figure 6.28C***). However, many of these genes also appear in WT or *Malat1*$^{-/-}$ Th2 cells but represent a low percentage of total reads.

Interestingly, comparison of LogFC of genes which increased or decreased in the absence of *Malat1* revealed a small number of changes. 27 genes increased in the total number of reads at a LogFC 0.5 in the absence of *Malat1* and 24 genes decreased their proportion of the total number of reads ***(Figure 6.28D***). If the LogFC was altered to 1 only 2 genes increased the percentage of reads and 2 genes decreased the percentage of read binding – this included *Malat1*.

As only a small number of genes were identified to alter binding to hnRNPA1 in the absence of *Malat1* these lists were pooled and GSEA GO analysis was performed. This identified carbohydrate metabolic process and protein complex pathways ***(Figure 6.28E***). Collectively, this data suggested that the loss of *Malat1* had a limited impact on the proportion of hnRNPA1 interactions.

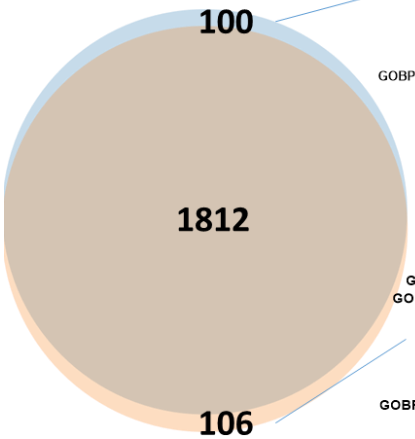
To determine if the RNAs which altered binding to hnRNPA1 were due to changes in gene expression we compared the iCLIP data with our RNAseq dataset

(chapter 4). Comparison with our RNAseq data set showed no strong correlation with changes in hnRNPA1 interactions as determined by iCLIP (**Figure *6.29***). As with SRSF1, some RNAs bound to hnRNPA1 more or less than anticipated based on their abundance (**Figure *6.29***).

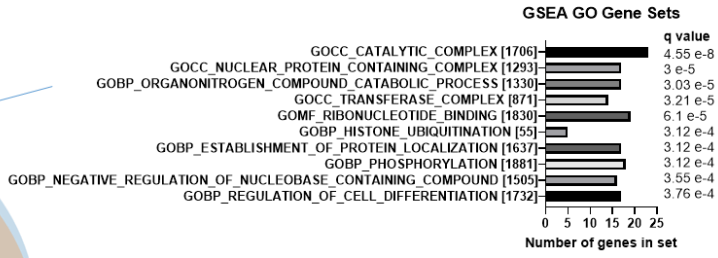
As we are interested in the local effects of *Malat1* we next examined the chromosome location of genes which change their binding to hnRNPA1 in the absence of *Malat1*. Although the number of genes which show altered binding to hnRNPA1 in the absence of *Malat1* is limited to ~40 genes these appear to be mainly located along chromosome 19 (**Figure *6.30***) This provided further evidence for *in cis-regulatory* effects for the role of *Malat1* in Th2 cells.

To confirm the changes in binding to hnRNPA1 in the absence of *Malat1,* genes which showed the highest changes in binding were manually inspected this confirmed the expected changes in binding that were reported at the LogFC level. For example, the loss of *Malat1* reduced binding to *Cd5* but increased binding to *St6galnac3 (***Figure *6.31*)**.

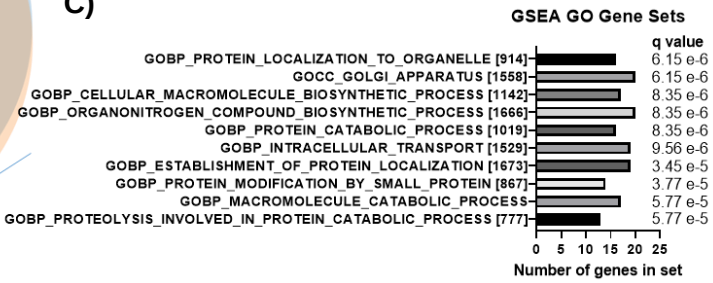
A)



B)



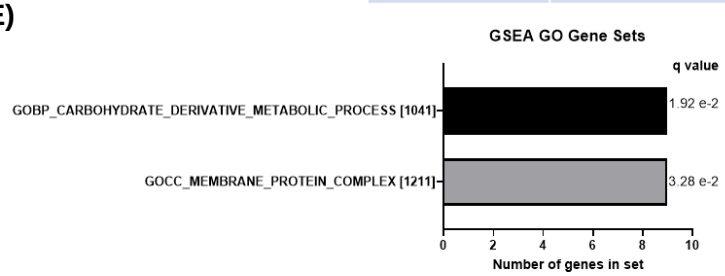
C)



D)

LogFC	Increased binding	Decreased binding
>1/<-1	2	2
>0.5/<0.5	27	24
P<0.01	13	17

E)



Increased/decreased binding FDR >Log+/- 0.5

Figure 6.28: Loss of *Malat1* has minimal impact on hnRNPA1 interactions.

A) Venn diagram depicting the overlap of the average percentage reads for hnRNPA1 determined by iCLIP in WT in vitro polarised Th2 cells and *Malat1*$^{-/-}$ *in vitro* polarised Th2 cells **B)** GSEA GO analysis of hnRNPA1 interactions which are unique to WT *in vitro* polarised Th2 cells **C)** GSEA GO analysis of hnRNPA1 interactions which are unique to Malat1$^{-/-}$ *in vitro* polarised Th2 cells **D)** Table depicting the number of genes that increase or decrease in binding to hnRNPA1 and different LogFC changes. An unpaired t-test was used to analyse the significance. **E)** GSEA GO analysis of genes which increase or decrease in binding to hnRNPA*1* over LogFC 0.5 when comparing WT and *Malat1*$^{-/-}$ in vitro polarised Th2. For all comparisons, a threshold of 0.01% percentage reads and over 100 reads was applied.

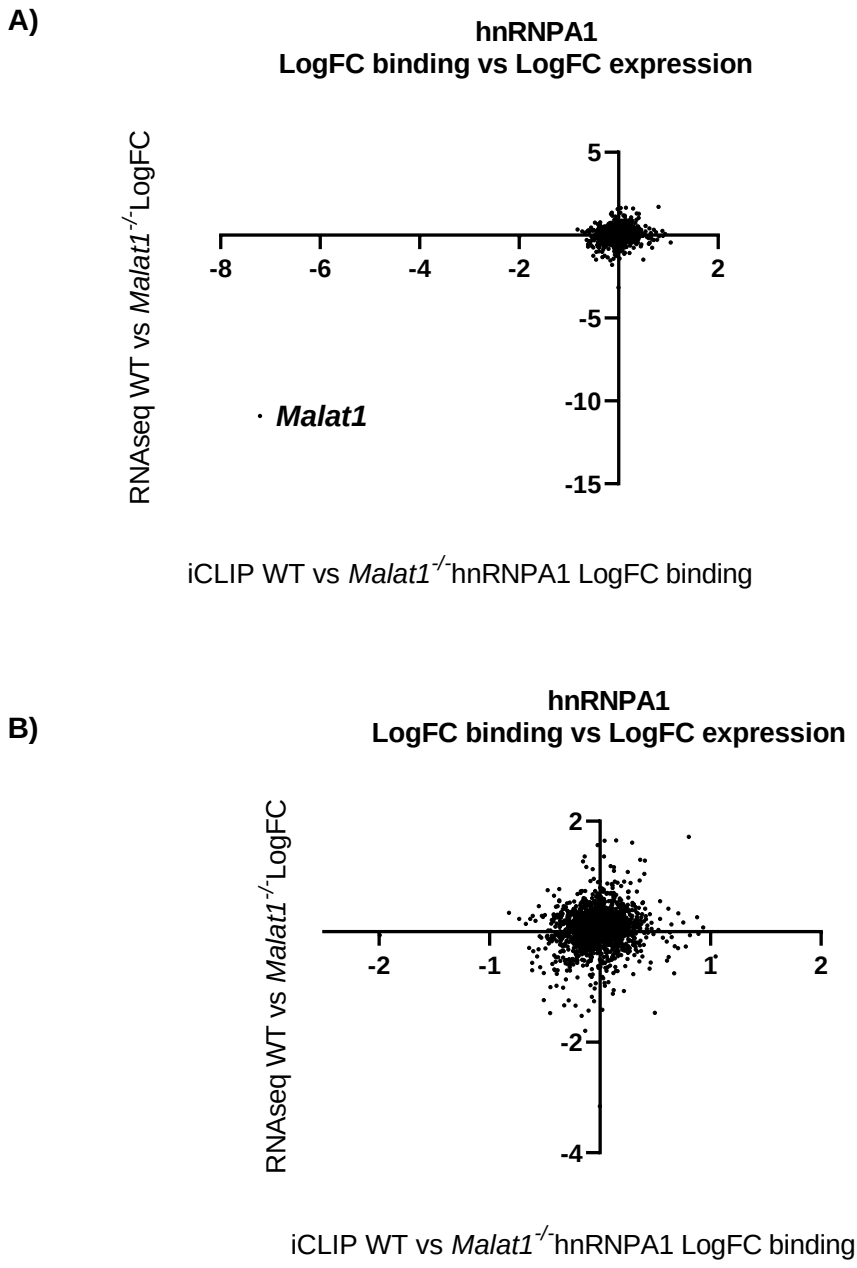


Figure 6.29: Abundance has limited influence on hnRNPA1 interactions.

A) Comparison of LogFC of changes in gene expression (WT vs *Malat1*$^{-/-}$ Th2 cells) determined by RNA-seq with changes hnRNPA1 binding (WT vs *Malat1*$^{-/-}$ Th2 cells) determined by iCLIP. B) As in A but with *Malat1* removed.

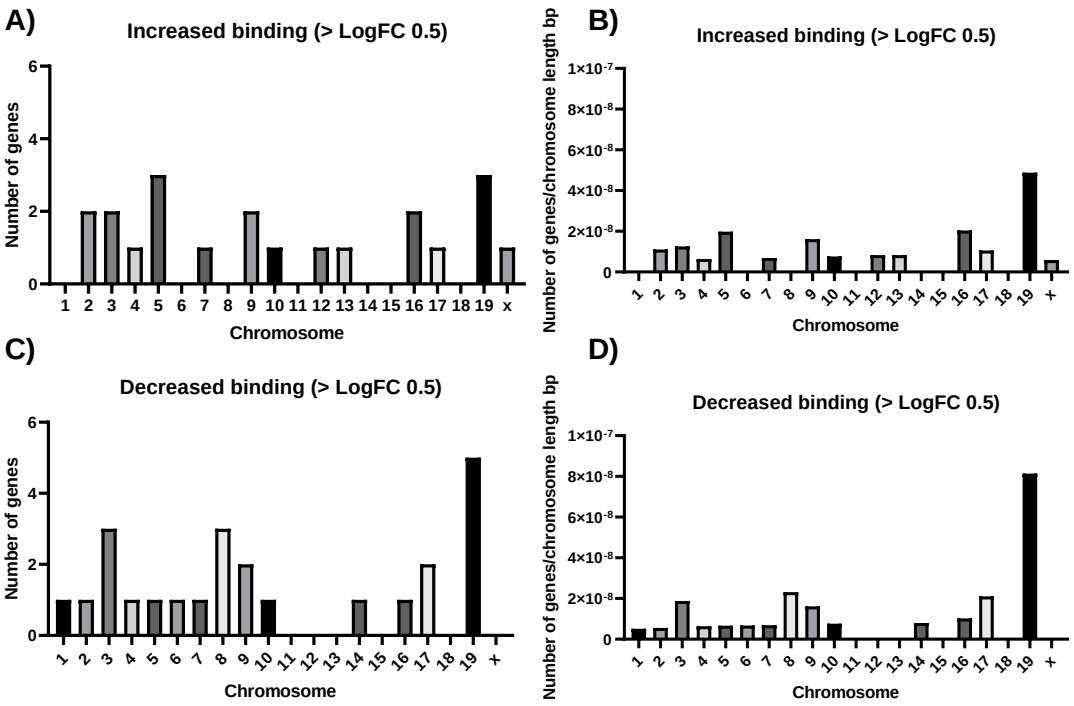


Figure 6.30: Loss of *Malat1* alters binding of genes to hnRNPA1 with a bias to those derived from chromosome 19.

Chromosome location of genes bound to hnRNPA11 which change in the absence of *Malat1* in in vitro polarised Th2 cells **A)** Genes which show increased binding LogFC >0.5 (expressed as a total number of genes), **B)** Genes which show increased binding LogFC >0.5 (normalised to chromosome length) **C)** Genes which show decreased binding LogFC >0.5 (expressed as a total number of genes) **D)** Genes which show decreased binding LogFC >0.5 (normalised to chromosome length).

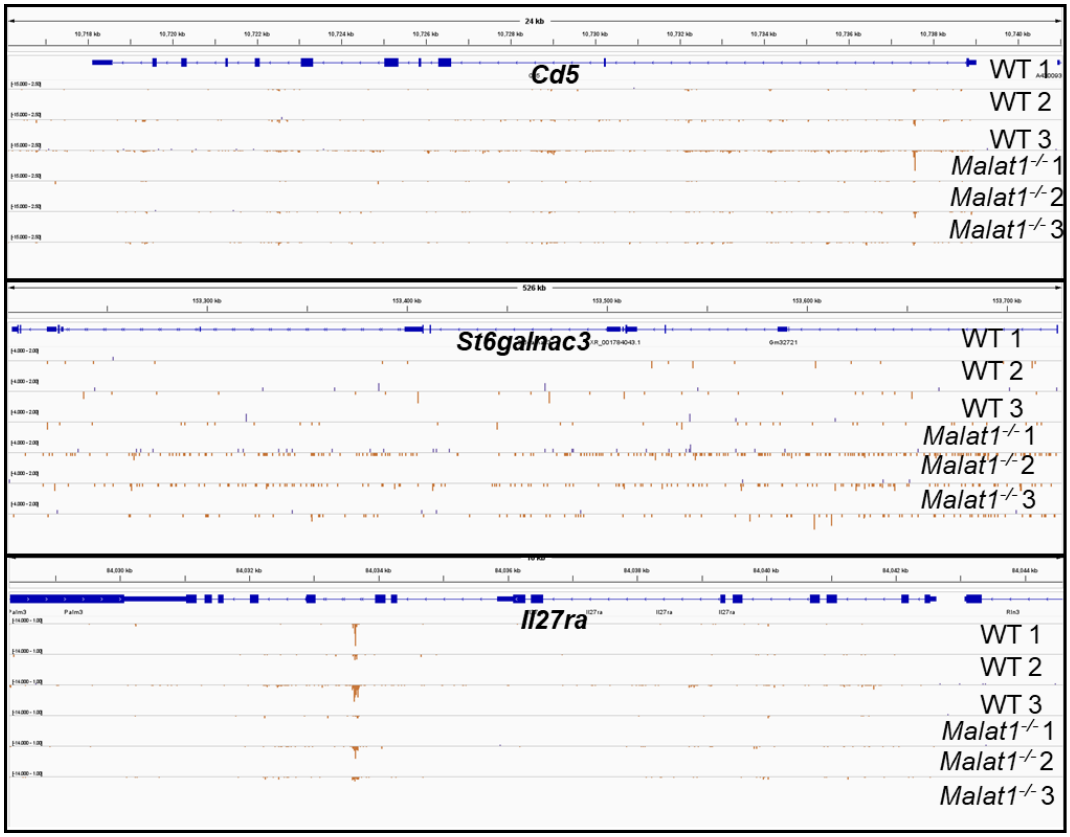


Figure 6.31: Visualisation of changes in hnRNPA1 binding in the absence of *Malat1.*

Coverage plots for hnRNPA11 iCLIP in WT and *Malat1*$^{-/-}$ Th2 cells. Tracks for WT Th2 replicate 1, WT Th2 replicate 2, WT Th2 replicate 3, *Malat1*$^{-/-}$ Th2 replicate 1, *Malat1*$^{-/-}$ Th2 replicate 2, and *Malat1*$^{-/-}$ Th2 replicate 3 are shown. From top to bottom coverage of the following targets is shown *Cd5, St6galnac3,* and *Il27ra.* Tracks were generated in IGV.

As the loss of *Malat1* altered the location of SRSF1, we wanted to determine if *Malat1* also retained hnRNPA1 in the nucleus. Analysis of hnRNPA1 localisation in WT and *Malat1*$^{-/-}$ Th2 cells revealed that hnRNPA1 remained mostly nuclear regardless of the presence of *Malat1 (***Figure *6.32)***. This could explain the smaller number of changes observed for hnRNPA1 binding in the absence of *Malat1* as a similar pool of RNAs will be able to bind hnRNPA1 if it is still retained in the nucleus.

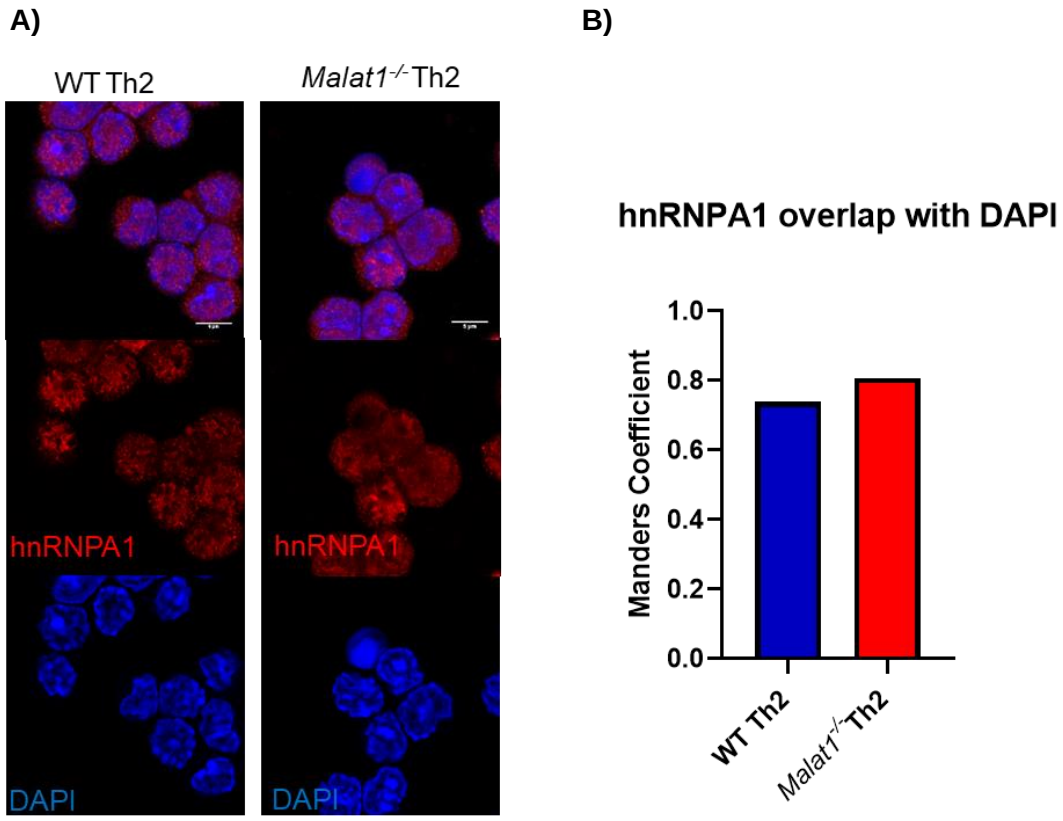


Figure 6.32: *Malat1* does not influence hnRNPA1 localisation.

A) Example images of immunofluorescence analysis for hnRNPA1 (red) and DAPI (blue) in *in vitro* polarised Th2 cells (day 6) samples were derived from female mice. **B)** Quantification of the proportion of overlap of hnRNPA1 with DAPI in WT and *Malat1*$^{-/-}$ Th2 cells using the manders coefficient. Pilot experiment n=1.

6.2.3 SRSF1 and hnRNPA1 bind similar targets in Th2 cells

It is well established that hnRNPA1 and SRSF1 often compete for the binding of the same RNAs. We next compared hnRNPA1 and SRSF1 interactions in the presence and absence of *Malat1*. Initially, we compared binding sites across the *Malat1* transcript. This demonstrated that both SRSF1 and hnRNPA1 bound across the *Malat1* transcript with hot spots at the 5' and 3' ends. There was some overlap of crosslinking sites, particularly at the 3' end (***Figure 6.33***).

Next, we compared the percentage of iCLIP reads from hnRNPA1 and SRSF1 WT Th2 iCLIP experiments. This revealed a positive correlation with an r^2 value of 0.6189 (***Figure 6.34****A*). A similar correlation was observed when comparing the percentage of iCLIP reads from hnRNPA1 and SRSF1 *Malat1*$^{-/-}$ Th2 iCLIP experiments with an r^2 value of 0.5803 (***Figure 6.34***B). Collectively, this suggested that SRSF1 and hnRNPA1 bound RNAs in a similar fashion.

We then wanted to determine the overlap between hnRNPA1 and SRSF1 interaction partners in Th2 cells. Interestingly, the majority of hnRNPA1 and SRSF1 interaction partners overlapped. 79% of hnRNPA1 interactions also appeared in the SRSF1 iCLIP (***Figure 6.35A***). GSEA GO analysis revealed that interactions that were unique to hnRNPA1 were involved in macromolecule catabolic process, protein localisation and intracellular transport (***Figure 6.35B***). GSEA GO analysis of unique SRSF1 interactions revealed these genes to be involved in the regulation of nucleobase-containing compounds, regulation of biosynthetic processes and regulation of gene expression (***Figure 6.35C***).

We then wanted to determine if any genes switch between binding SRSF1 and hnRNPA1 in the absence of *Malat1 (**Figure 6.36**)*. Little to no overlap was observed

between genes that increased binding to one protein and decreased binding to the other and vice versa ***(Figure 6.36)***. Instead, 8 genes increased binding to both SRSF1 and hnRNPA1 in the absence of *Malat1* and 10 genes decreased binding to both SRSF1 and hnRNPA1 in the absence of *Malat1* – this included *Malat1* ***(Figure 6.36)***. Collectively, this highlighted the greater number of changes to SRSF1 binding partners in the absence of *Malat1* than hnRNPA1 and indicated that genes acted similarly in the absence of *Malat1* when interacting with either hnRNPA1 or SRSF1.

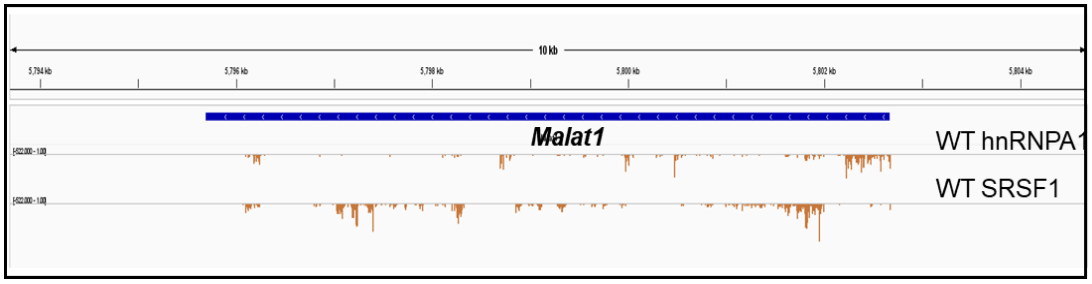
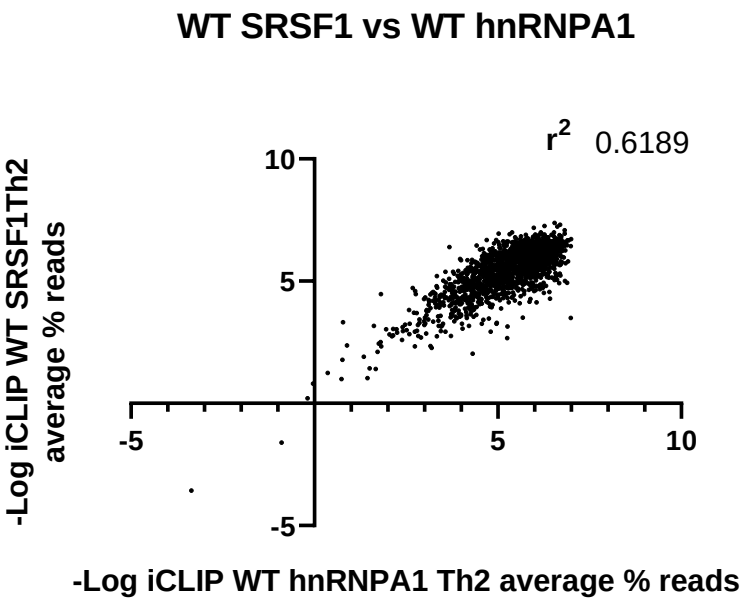


Figure 6.33: SRSF1 and hnRNPA1 bind across the *Malat1* transcript.

Coverage plots for hnRNPA1 and SRSF1 iCLIP in WT Th2 cells across *Malat1*. Representative tracks are shown Tracks were generated in IGV

A)



B)

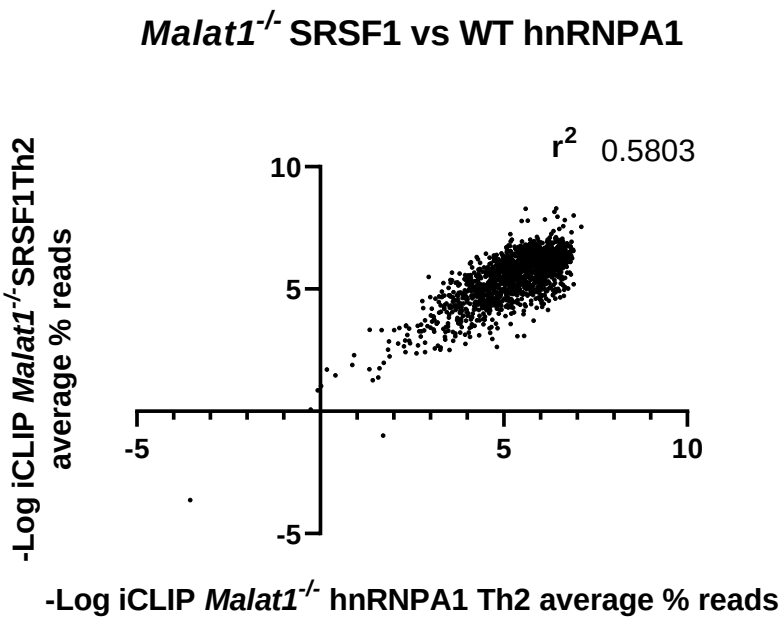


Figure 6.34: SRSF1 and hnRNPA1 iCLIP percentage reads positively correlate.

A) Correlation of average Log_2 percentage of reads between UV cross-linked WT Th2 hnRNPA1 iCLIP samples and UV cross-linked WT Th2 SRSF1 iCLIP samples **B)** Correlation of average Log_2 percentage of reads between UV cross-linked SRSF1 *Malat1*$^{-/-}$ Th2 iCLIP samples and hnRNPA1 *Malat1*$^{-/-}$ Th2 iCLIP samples.

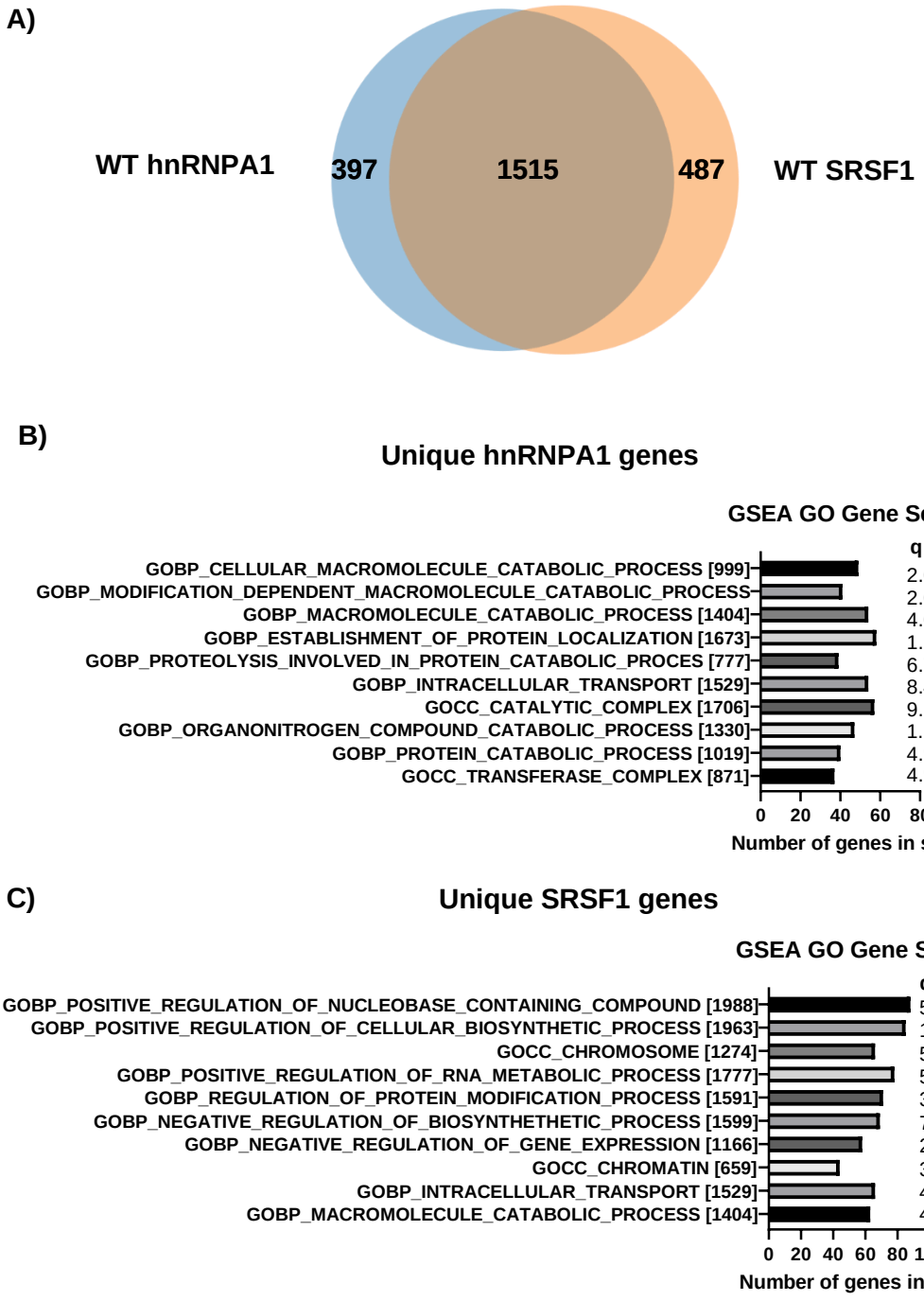


Figure 6.35: SRSF1 and hnRNPA1 bind similar genes.

A) Venn diagram depicting the overlap of WT hnRNPA1 average percentage reads and WT SRSF1 average percentage reads determined by iCLIP at a threshold of >100 reads and >0.01% reads**. B)** GSEA GO analysis of unique hnRNPA1 interactions. **C)** GSEA GO analysis of unique SRSF1 interactions.

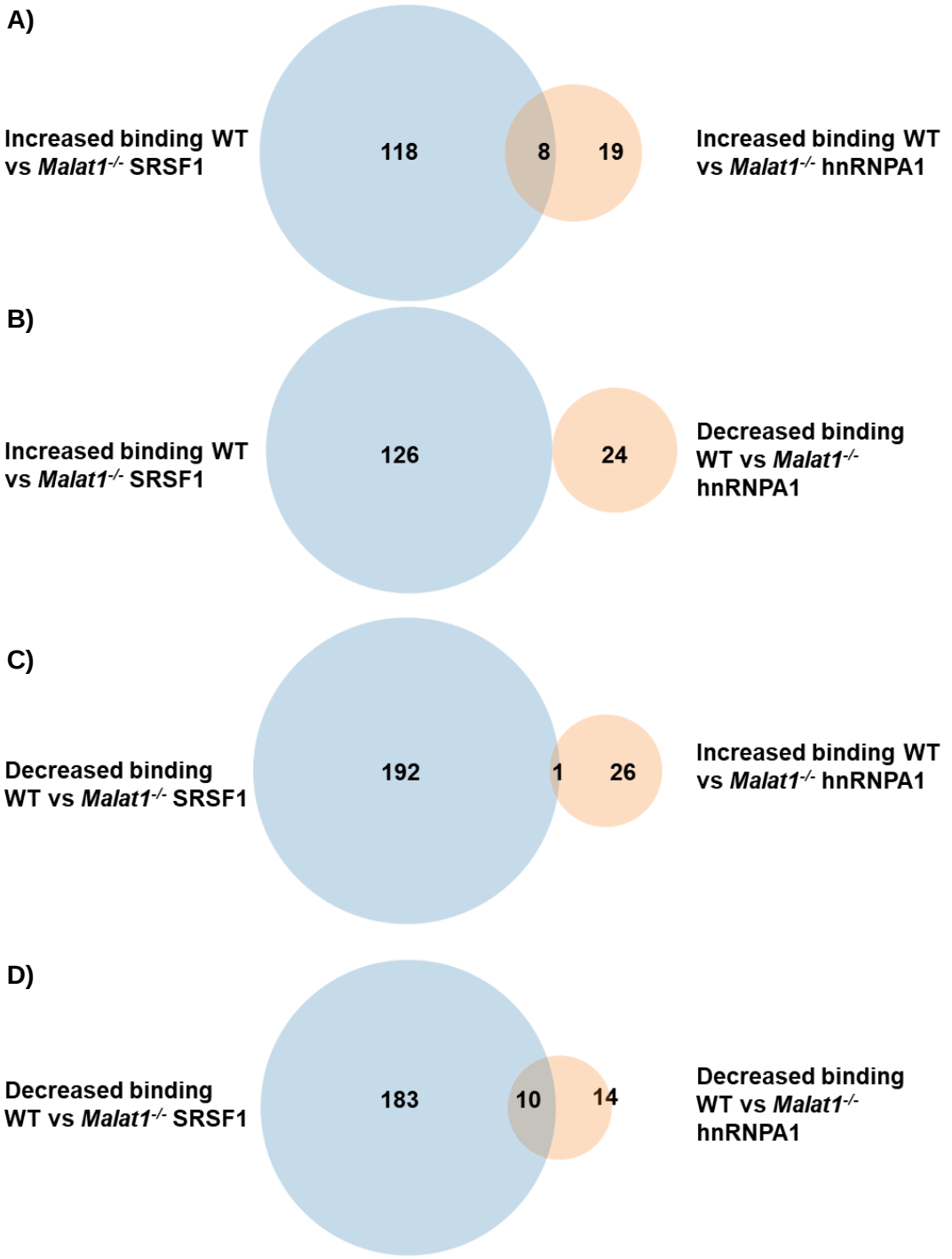


Figure 6.36: Loss of Malat1 influences binding to SRSF1 more than hnRNPA1.

A) Venn diagram depicting the overlap of genes which increase in binding to SRSF1 and hnRNPA1 in the absence of *Malat1.* ***B)*** Venn diagram depicting the overlap of genes which increased in binding to SRSF1 and decreased in binding to hnRNPA1 in the absence of *Malat1* **C)** Venn diagram depicting the overlap of genes which decreased binding to SRSF1 and increased binding to hnRNPA1 in the absence of *Malat1* **D)** Venn diagram depicting the overlap of genes which decreased binding to SRSF1 and decreased binding to hnRNPA1 in the absence of *Malat1.* Changes in binding were at a LogFC >0.5.

6.2.4 Loss of *Malat1* affects AS of SRSF1 and hnRNPA1 targets

Malat1 has previously been linked to the regulation of AS (Arun et al., 2020). We wanted to determine if loss of *Malat1* altered differential transcript usage in Th2 cells and if this could be explained by differences in SRSF1 or hnRNPA1 interactions. Differential transcript usage analysis was performed by Joshua Lee (University of York, York, UK) as described in the methods section. Two approaches were taken to determine differential transcript usage – DRIMseq and DEXseq. The DEXseq tool analyses exon usage whereas DRIMseq was developed specifically to determine differential transcript usage as is based on estimated transcript counts – differential transcript usage is a measure of the relative contribution of one transcript to the overall expression of a gene/the total transcriptional output (Love et al., 2018). Both methods have been shown to be comparable and identify transcripts that have been differentially used rather than specific splicing events (Love et al., 2018).

Initially, we compared the total number of genes which showed differential transcript usage with genes that are differentially expressed in the absence of *Malat1* in Th2 cells. 187 genes were identified to have differential transcript usage by DRIMseq in the absence of Malat1 (***Figure 6.37***). 179 genes were identified to have differential transcript usage by DEXseq. 65 genes showed differential transcript usage in the absence of *Malat1* for both DRIMseq and DEXseq, 13 of which also were differentially expressed at the whole gene level (***Figure 6.37***).

GSEA GO analysis of genes which showed differential transcript usage by DRIMseq identified genes involved in protein conjugation, splicing and catabolic processes (***Figure 6.38****A*). In contrast, genes which showed differential transcript usage by DEXseq identified genes involved in intracellular transport, cell death and apoptosis (***Figure 6.38B***).

Next, we wanted to compare genes which showed differential transcript usage with those that bound SRSF1 or hnRNPA1. 66 of the genes which showed differential transcript usage by DRIMseq or DEXseq upon loss of *Malat1* also bound to SRSF1 in WT *in vitro* polarised Th2 cells (***Figure 6.39A***). 55 genes were found to have differential transcript usage in the absence of *Malat1* by DRIMseq or DEXseq and bound to hnRNPA1 (***Figure 6.39B***). The genes which bound SRSF1 and showed differential transcript usage had considerable overlap ~74.24% with those that showed differential transcript usage and bound to hnRNPA1 (***Figure 6.39A/B***).

Analysis of genes which showed differential transcript usage in the absence of *Malat1* and bound to SRSF1 from WT Th2 cells identified genes involved in apoptosis and regulation of cell death (***Figure 6.40A***). Similarly, genes which showed differential transcript usage in the absence of *Malat1* and interacted with hnRNAP1 in WT Th2 cells were involved in the regulation of cell death, protein-containing complexes, and immune system function (***Figure 6.40B***).

Finally, we plotted genes which had been shown to have differential transcript usage. It is important to note that all genes plotted showed significant differential transcript usage i.e the changes in transcript proportions between WT and *Malat1*$^{-/-}$ Th2 cells was significant $p<0.05$. This can be achieved by the additive effects of changes in transcript usage between conditions. Where the individual transcript showed significant differential transcript usage the p-value is depicted. These graphs visually confirmed differences in transcript usage between WT and *Malat1*$^{-/-}$ Th2 cells this included genes such as *Il27ra, Cd27 and Tardbp (**Figure 6.41**)*.

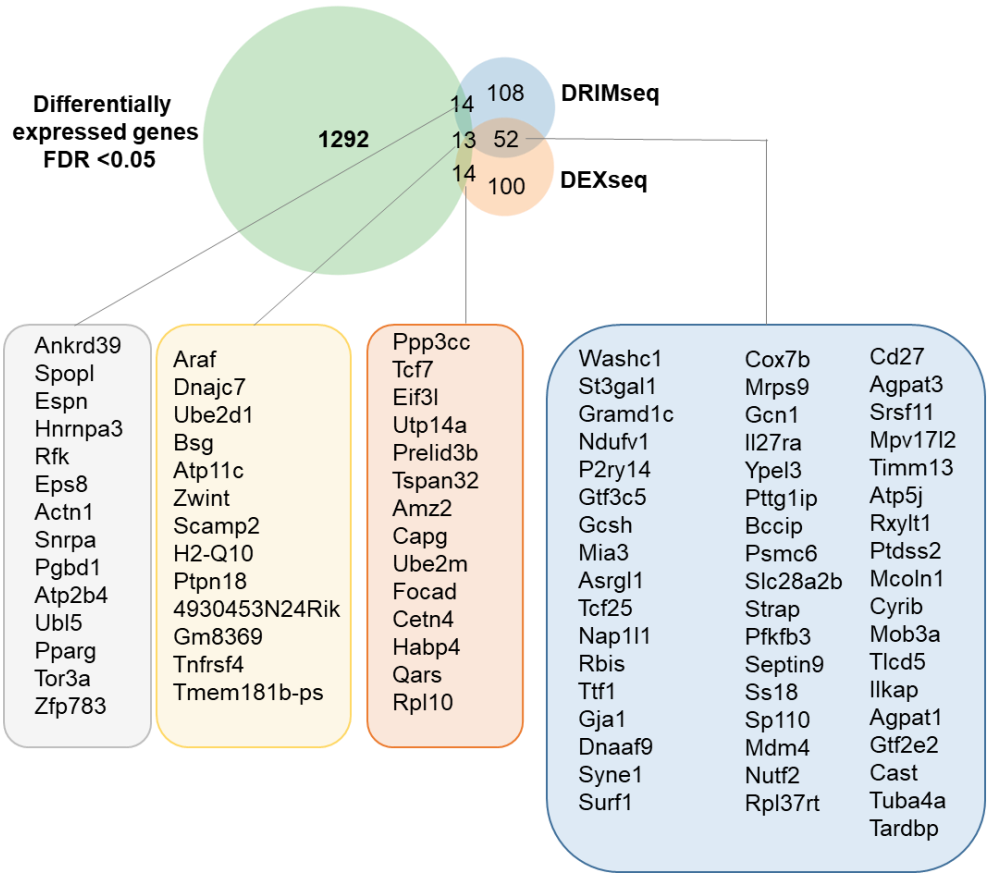


Figure 6.37: Loss of *Malat1* results in differential transcript usage in Th2 cells.

Venn diagram depicting the overlap of genes which are differentially expressed in Th2 cells in the absence of *Malat1* in *in vitro* polarised Th2 cells q <0.05 with genes that show differential transcript usage in the absence of *Malat1* as determined by DRIMseq in *in vitro* polarised Th2 cells q <0.05 and genes that show differential transcript usage in the absence of *Malat1* by DEXseq in *in vitro* polarised Th2 cells q <0.05.

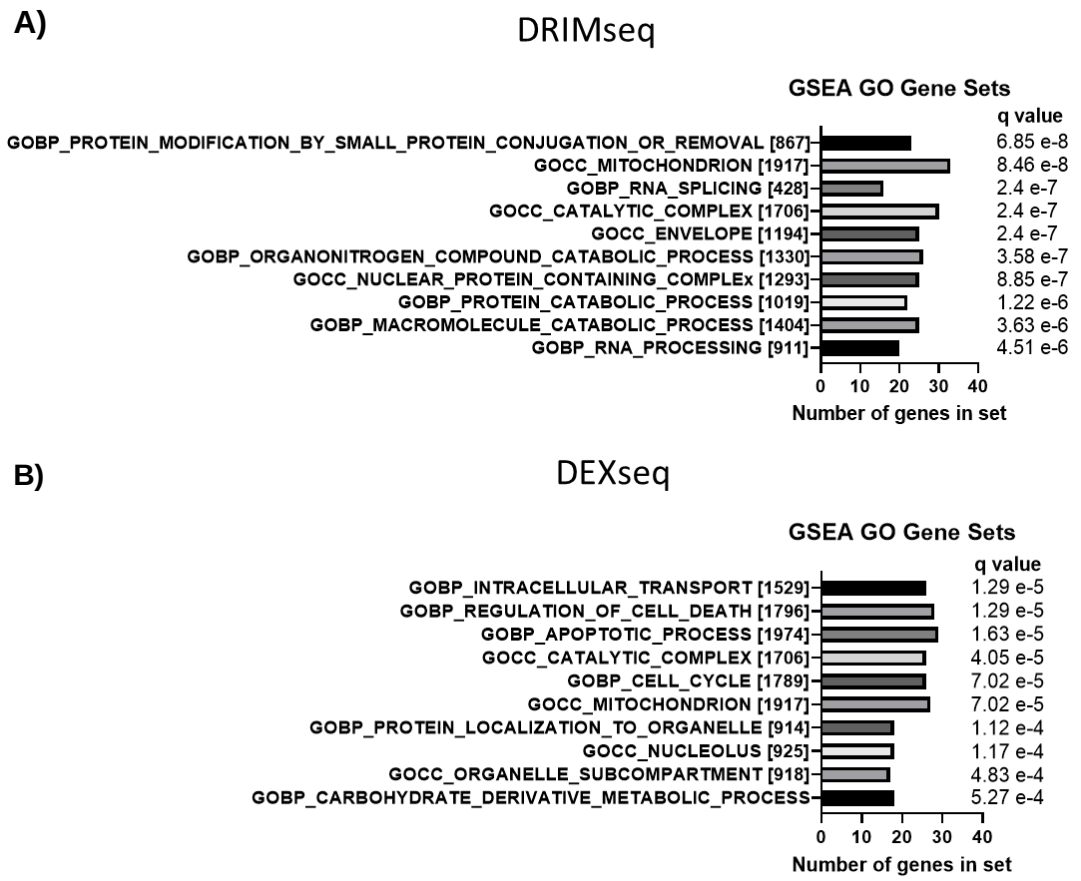


Figure 6.38: Genes which show differential transcript usage in the absence of *Malat1* are involved in the regulation of cell death and splicing.

A) GSEA GO analysis of genes which show differential transcript usage in the absence of *Malat1* as determined by DRIMseq in *in vitro* polarised Th2 cells **B)** GSEA GO analysis of genes which show differential transcript usage in the absence of *Malat1* as determined by DEXseq in *in vitro* polarised Th2 cells.

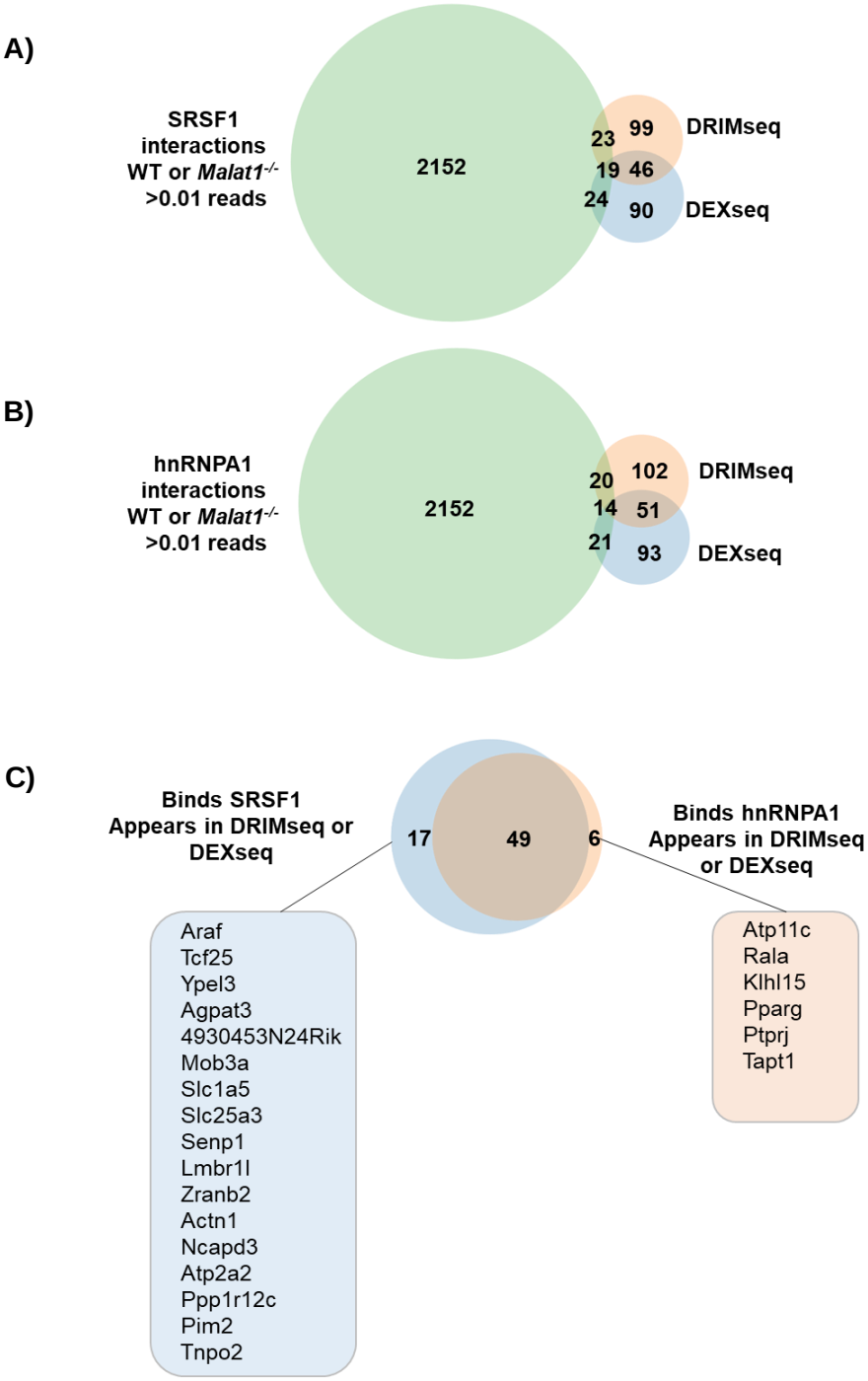
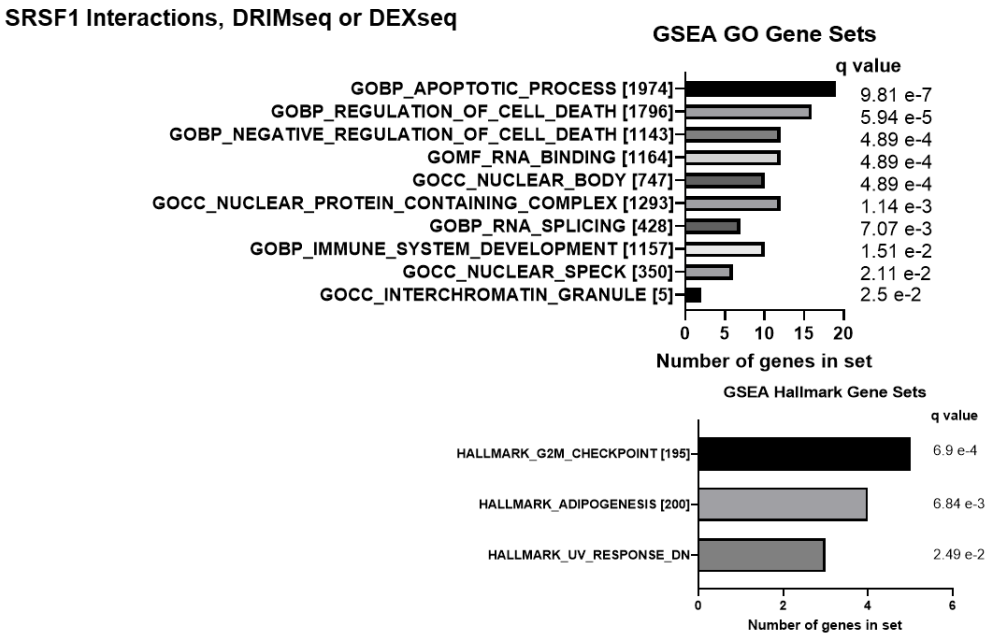


Figure 6.39: Genes have differential transcript usage in the absence of *Malat1* bind hnRNPA1 and SRSF1.

A) Venn diagram depicting the overlap of genes which interact with SRSF1 in WT Th2 cells determined by iCLIP with genes that show differential transcript usage in the absence of *Malat1* through DRIMseq or DEXseq **B)** As in A but for hnRNPA1 interactions in WT Th2 cells determined by iCLIP **C)** Overlap of genes which bind SRSF1 and show differential transcript usage in the absence of *Malat1* with genes that bind hnRNPA1 and show differential transcript usage in the absence of *Malat1.*

A)



B)

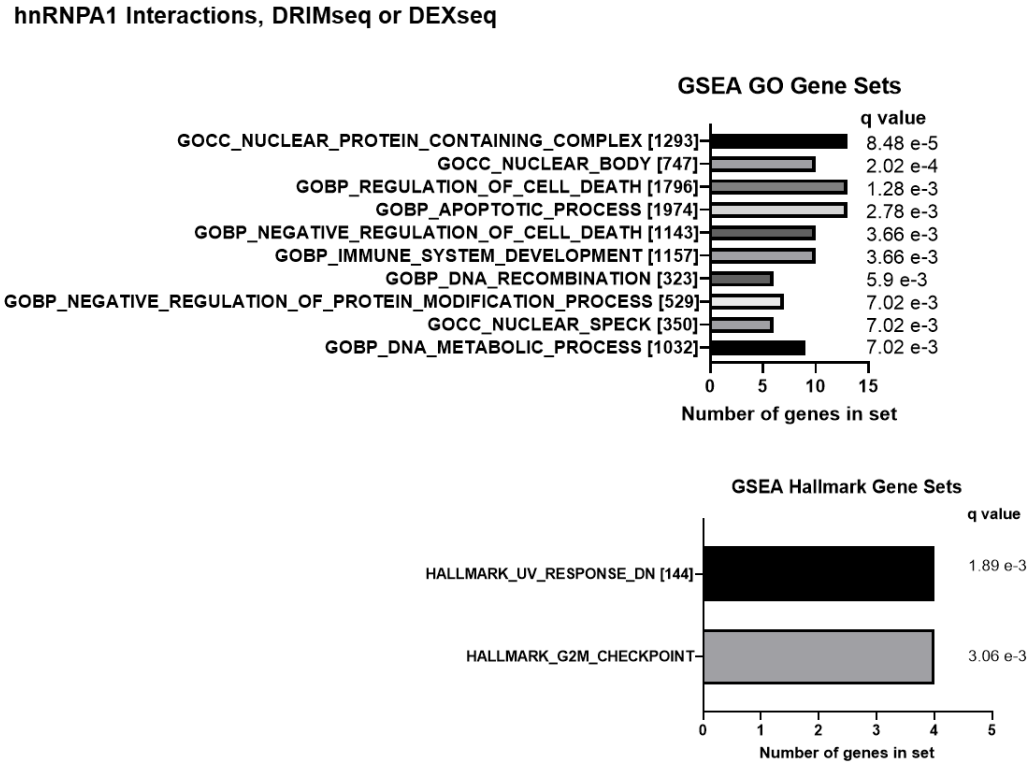


Figure 6.40: GSEA analysis of genes which show differential transcript usage in the absence of *Malat1* and bind RBPs.

A) GSEA GO and hallmark analysis of gene which bind to SRSF1 determined by iCILP in Th2 cells with genes that show differential transcript usage in *Malat1*$^{-/-}$ Th2 cells **B)** As in A but for genes that bind hnRNPA1 as determined by iCLIP.

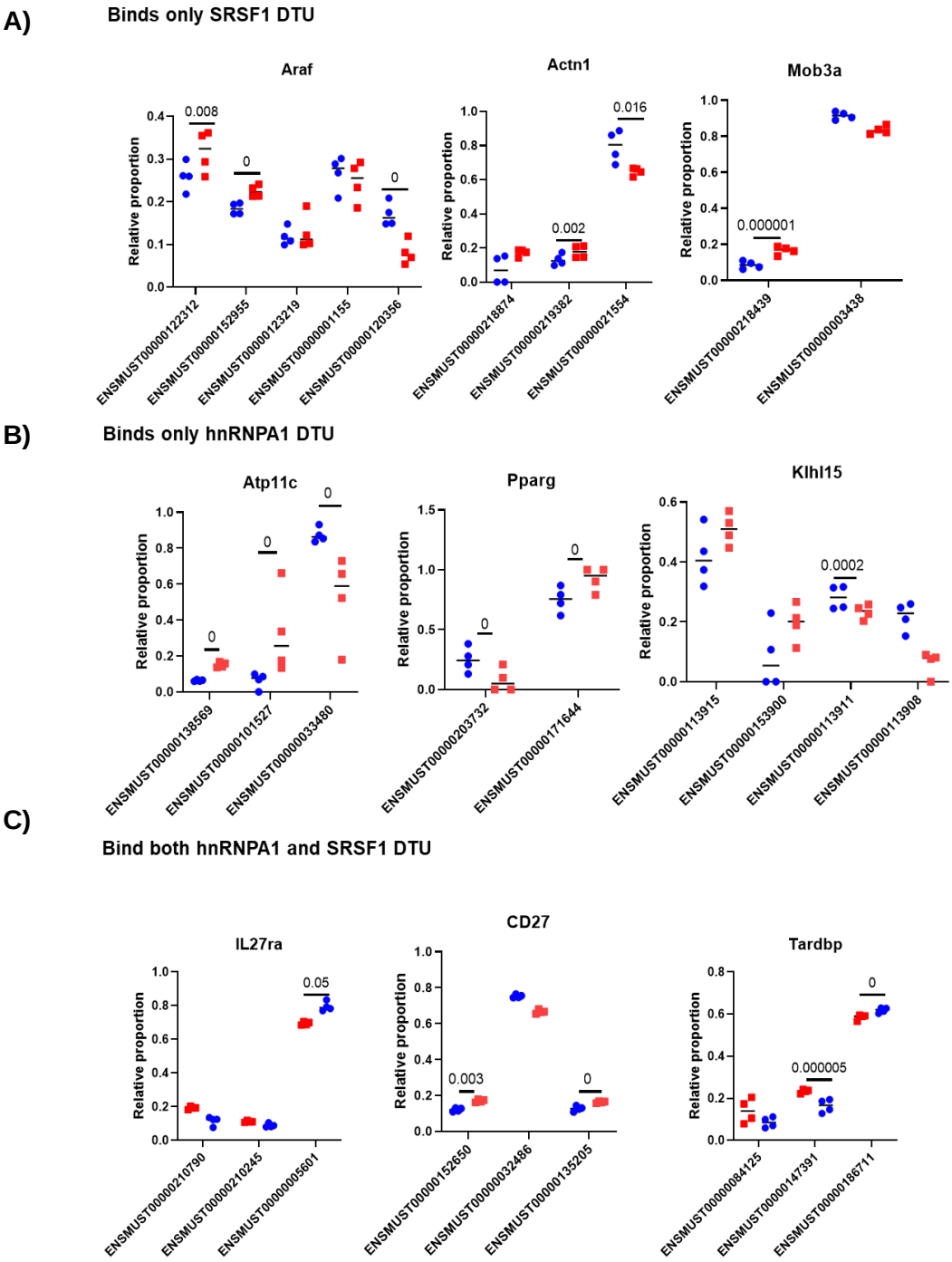


Figure 6.41: Inspection of genes which show differential transcript usage in the absence of *Malat1.*

***A*)** Graphs depicting genes which show differential transcript usage between WT and *Malat1*$^{-/-}$ Th2 cells and also bind SRSF1 in WT Th2 cells– individual transcripts which show significant differential transcript usage are shown **B)** As in A for genes which also bind hnRNPA1 in Th2 cells **C)** As in A but for genes that bind both SRSF1 and hnRNPA1 in Th2 cells. Transcript abundance in WT Th2 cells (blue) or *Malat1-/-* Th2 cells (red).

6.3 Discussion

In this chapter, we aimed to elucidate the relevance of *Malat1*-RBP interactions in $CD4^+$ T cells. We identified SRSF1 and hnRNPA1 RNA interactions in the presence and absence of *Malat1*. We determined that both SRSF1 and hnRNPA1 prominently interacted with RNAs involved in T cell differentiation and function in WT *in vitro* polarised Th2 cells. We also determined that the loss of *Malat1* had a more prominent effect on SRSF1 interactions, as this resulted in a general decrease in RNA binding. Of note, many of the changes in RNA binding in the absence of *Malat1* were localised to chromosome 19 – particularly for hnRNPA1. This suggested a potential *in cis*-regulatory function of *Malat1* through interactions with SRSF1 and hnRNPA1. We also found that RNAs which bound either hnRNPA1 or SRSF1 showed differential transcript usage in the absence of *Malat1*.

We began our studies by characterising the type and subtype interactions of SRSF1 and hnRNPA1. We found that SRSF1 interacted with intronic regions (46.5% of interactions) and coding sequences (22.7% of interactions). This somewhat aligned with a previous study that identified SRSF1 binding sites in thymocytes using CLIP-seq. This study determined that SRSF1 most prominently interacted with coding sequences (50.3% of interactions), and introns (26% of interactions) (Qi et al., 2021). However, the proportion of type-binding sites appears to be influenced by cell type and the relative abundance of other RBPs. For example, one study which used iCLIP to determine SRSF1 interaction partners found that SRSF1 interacted most prominently with intronic regions and that overexpression of hnRNPA1 increased the percentage of intron binding and decreased the proportion of coding sequence binding (Howard et al., 2018). This indicated that SRSF1 interactions in Th2 cells are within previously reported percentages of binding sites. The high proportion of intronic

cDNA sequences indicated that SRSF1 interactions mainly represented nuclear targets. This observation is supported by a previous study that isolated TDP-43 from nuclear and cytoplasmic fractions using iCLIP. Quantification of the iCLIP reads determined that the cytoplasmic version of TDP-43 showed 34% 3' UTR binding in contrast to the nuclear fraction which showed 3.8% 3' UTR binding (Tollervey et al., 2011).

In contrast, we found that hnRNPA1 was most prominently bound to intronic regions. This is similar to previous studies which observed that hnRNPA1 interactions were prominently located in introns in HeLa cells. The authors also determined that hnRNPA1 bound uniformly across introns (Bruun et al., 2016). Because hnRNPA1 mainly interacted with introns this means that exons are free for splice site recognition (Bruun et al., 2016) This is further supported by another iCLIP study which performed iCLIP on quadricep muscle tissue of hnRNPA1 and observed 74% intron binding (Bruun et al., 2016). Similar to SRSF1, the high percentage of intron binding suggests that hnRNPA1 also prominently interacted with nuclear targets (Tollervey et al., 2011).

Studies have indicated that ~85% of the human genome is pervasively transcribed by RNA polymerase II (Pol II) (Jacquier, 2009). Many of these transcripts are well categorised for example mRNAs and some ncRNAs. Despite many updates in the annotation of the human and mouse genomes, a large portion of RNA-seq reads map to unannotated and intergenic regions or the genome. Interestingly, in the CLIP data set many of the reads (in some cases ~18% of total reads) are classified under the broad umbrella term of intergenic RNA. Until recently, the transcriptional origin of these intergenic RNAs has remained unclear. One study used computational methods to identify the source of intergenic transcription and determined that these intergenic

regions are unannotated genes or a result of fuzzy transcription extending beyond the current boundaries of annotated transcripts (Agostini et al., 2021). The largest proportion of intergenic transcripts arose due to fuzzy transcription, this is also known as downstream of gene transcripts, this occurs when Pol II terminates far downstream of the ends of genes this can be a few hundred kilobases downstream of the end of a gene. These RNA products can exert their function *in cis* to their site of transcription and then are rapidly degraded (Agostini et al., 2021). As mapping of these intergenic reads becomes more advanced, it may be possible to incorporate changes in the downstream of transcript interactions with RBPs.

Interestingly, we observed that in the absence of *Malat1* the proportion of mt rRNA, miRNA, snoRNA and snRNA increased interactions with SRSF1. Similarly, snoRNA, mt_rRNA and mt_tRNA interactions with hnRNPA1 increased in the absence of *Malat1*. snoRNAs are a subtype of ncRNAs that reside in the nucleoli (the primary site of ribosome biogenesis) and are important for tRNA and mRNA modifications (Biagioni et al., 2021). They are essential for stabilising rRNA through pseudouridylation and methylation. Little is reported on the relevance of *Malat1*, SRSF1 or hnRNPA1 interactions with snRNAs. However, one possible explanation for the increased binding to SRSF1 and hnRNPA1 is that these proteins become more prevalent in the nucleoli in the absence of *Malat1*– this would likely only be a small fraction of the total RBPs as snoRNA interactions represented less than 0.5% of total interactions in Th2 cells. This could also explain the increased interactions with rRNAs which would reside in the nucleolus for part of their lifetime. Additionally, we see an increase in the proportion of cytoplasmic SRSF1 in the absence of *Malat1* which could explain the increase in rRNA binding. snRNAs are an essential component of the spliceosome, and aid mRNA splicing events (Morais et al., 2021).

The increase in snRNA binding might influence some of the changes in differential gene expression in the absence of *Malat1* at a global level.

Of note, in this study we found that *Malat1* was the top interaction partner for both SRSF1 and hnRNPA1. SRSF1 and hnRNPA1 bound similar parts of the *Malat1* transcript with hotspots of binding localised to the 5' and 3' ends. Neither SRSF1 nor hnRNPA1 bound to the triple helix region of *Malat1*. A previous study, has shown that the *Malat1* triple helix can be essential for interactions with the RBP METTL16 (Brown et al., 2016). However, another study used *in vitro* transcribed fragments of *Malat1* to determine hotspots of interactions. Similar to our study this determined that *Malat1* prominently interacted with RBPs at the 5' and 3' regions (Scherer et al., 2020). Of note, as this study used fragments of *Malat1* and an *in vitro* binding assay to determine binding sites this may not be truly representative of interactions that take place *in vivo.* Nevertheless, the binding sites identified in our study align with previously published data.

PEKA analysis was used to identify k-mers of RBP binding preferences. Previous work has identified CUGGA as key motif for SRSF1 in exonic regions (Das & Krainer, 2014). It appears that typically purine-rich sequences are enriched for example AGAAGAAG has been identified using *RNA Motif Modeler* for SRSF1 CLIP-seq data (X. Wang et al., 2011). This study also predicted that these would likely be single-stranded sequences (Wang et al., 2011). We did not observe this exact motif in our data set for SRSF1 but did observe the enrichment of similar purine-rich motifs such as – GAA, and GGAA or GGAAA. hnRNPA1 has previously been shown to bind the motif UAGGA, GGGAG, AGGG, GAGGG and GGAGG (Lee et al., 2018). We observed similar GA-rich motifs such as GAAA, AGGG and AGAG. However,

the peaks hnRNPA1 k-mers were noisy which indicates that fewer crosslinking events had taken place than that observed for SRSF1.

One factor that is important to keep in mind for our data set is that as iCLIP utilises UV crosslinking this preferentially crosslinks RBPs to uridines and guanosines, although to a lesser extent. Additionally, enzymes used for RNA fragmentation will have preferences in addition to those used for reverse transcription (Hafner et al., 2021). The potential impact of these biases on motif identification for CLIP experiments is poorly understood due to a lack of systemic studies. PEKA aims to reduce the impact of these technical biases; however, it is important to consider that technique and bioinformatics method may impact the k-mer results obtained in this study.

Before we began this study, we speculated that as *Malat1* is such an abundant transcript within the cell and a top interaction partner for both hnRNPA1 and SRSF1, it may act as a shield to protect the RBPs from aberrant interactions -i.e prevent interactions with these that wouldn't arise in a WT state. Consequently, we anticipated that loss of *Malat1* would have caused a general increase in genes binding to SRSF1 and or hnRNPA1. We did not observe a large increase in RNAs bound to SRSF1 or hnRNAP1 in the absence of *Malat1*. Instead, we observed a general decrease in interactions with SRSF1. This suggested that *Malat1* instead may help recruit RNAs to the complex or retain the RBPs within nuclear speckles to enable the interactions to take place. Fewer changes in binding were observed upon loss of *Malat1* when examining hnRNPA1 interactions. As hnRNPA1 and SRSF1 have previously been shown to compete for the same RNAs we also expected to see some RNAs switching interactions between SRSF1 and hnRNPA1 (Howard et al., 2018). We did not observe this, possibly the loss of *Malat1* did not have a big enough impact on the number of

differentially bound RNAs for this to take place in Th2 cells although the ratio of RNAs binding to hnRNPA1 and SRSF1 will have changed slightly.

We also observed that upon loss of *Malat1* genes which bound to SRSF1 and hnRNPA1 changed in binding with a bias towards genes derived from chromosome 19. It is possible that genomic issues are at play at this is a result of generation of the knockout mice. Additionally, as loss of *Malat1* most prominently altered gene expression of those derived from chromosome 19 may have altered the availability of RNAs to bind SRSF1. Finally, some *Malat1* dependent RNA-RNA interactions may be at play which could explain these differences in RNA binding. However, *Malat1* is thought to more commonly interact with RNAs indirectly through proteins (Engreitz et al., 2014).

In this study, we found that SRSF1 and hnRNPA1 interacted with genes that were involved in T cell differentiation and function. This is not unexpected as the experiment took place in a T cell context. Additionally, both SRSF1 and hnRNPA1 have previously been linked to T cell functions as discussed in chapter 5. Of note, in the absence of *Malat1* genes which showed differential binding to SRSF1 were shown to be involved in the regulation of apoptosis and cell death. We observed similar findings in Chapter 4 where the loss of *Malat1* also altered genes involved in the regulation of cell death. As very few changes in interactions occurred in the absence of *Malat1* limited GO pathways were identified and were low in significance. However, it appeared that genes which altered interactions with SRSF1 and hnRNPA1 were commonly derived from chromosome 19, this further suggested *in cis* regulatory functions of *Malat1*. It may be interesting to confirm these interactions using FISH combined with immunofluorescence.

One particularly interesting observation from this dataset is that *Malat1* appears to retain SRSF1 in the nucleus. This is similar to previous findings from Tripathi and colleagues who determined that SRSF1 becomes cytoplasmic in the absence of *Malat1* (Tripathi et al., 2010). It is possible that the increased number of differential binding events to SRSF1 in the absence of *Malat1* could be explained by this movement in contrast to hnRNPA1 which is retained in the nucleus. Additionally, in chapter 5 we showed that SRSF1 was more significantly enriched than hnRNPA1 when identifying *Malat1* interaction partners. Additionally, *Malat1* represented a greater number of total reads for SRSF1 interactions than hnRNPA1 interactions which could also contribute to the greater effect seen for SRSF1 interactions in the absence of *Malat1*.

Tripathi and colleagues have also shown that *Malat1* regulates the levels of phosphorylated forms of SRSF1. They demonstrated that the loss of *Malat1* increased the proportion of dephosphorylated SRSF1 (Tripathi et al., 2010). This is particularly important as both the phosphorylated and dephosphorylated form of SRSF1 is required for AS. For example, hyperphosphorylation of SRSF1 determines the binding of SRSF1 to pre-mRNA and regulates splice site detection, it is also important for localising SRSF1 between nuclear speckles and transcription sites (Cao et al., 1997). We did not examine the phosphorylation status of SRSF1 in our experiments, however, this would be beneficial to determine if SRSF1 is also dephosphorylated in addition to mislocalised in Th2 cells.

It is important to bear in mind this imaging study was a pilot experiment and only represented one sample, SRSF1 localisation may have been affected by other factors such as the stage of the cell cycle. For example, when cells are in telophase,

SRSF1 has minimal overlap with DAPI (Tripathi et al., 2010a). This experiment needs to be repeated to increase our confidence in these findings.

Tripathi and colleagues have previously analysed AS events in the absence of *Malat1*. Microarray analysis determined that 238 genes showed differential transcript usage upon knockdown of *Malat1* (Tripathi et al., 2010). The authors went on to show that these AS events were linked to the cellular levels of SRSF1 (Tripathi et al., 2010a). It would be beneficial if we could recapitulate this experiment and alter the expression of SRSF1 and hnRNPA1 either through knockdown or CRISPR/Cas9 mediated knockout of *Malat1* in primary *in vitro* polarised Th2 cells to determine if these proteins are responsible for the changes in AS observed.

In this experiment, we also observed a similar number of changes in AS in the absence of *Malat1*. Several of these genes were also observed to bind to either SRSF1 or hnRNPA1. This included *Il27ra* which has previously been shown to directly bind to SRSF1 which is responsible for AS of the transcript (Yang et al., 2022). This indicates that loss of *Malat1* may affect AS events through interactions with SRSF1 and hnRNPA1.

As we observed changes in RNA binding to both SRSF1 and hnRNPA1, it is possible that *Malat1* regulates the RNA binding capabilities of its other interaction partners in addition to SRSF1 and hnRNPA1. Although the changes in SRSF1 and hnRNPA1 binding are slight, perhaps potential changes in total RNA-RBP interactions are additive or act synergistically to affect Th cell function. To fully understand the role of *Malat1* in this network of proteins further work is required – perhaps using iCLIP to probe changes in RNA binding of all *Malat1* interaction

parterns in WT and *Malat1*$^{-/-}$ cells could be used or mathematical modelling could be employed to deepen our understanding of these interactions.

6.3.1 Future work

Due to the timing of data arrival, only initial data analysis was performed for this dataset. Here, we identified some potential changes in k-mer motifs between WT and *Malat1*$^{-/-}$ Th2 cells. To build on this and make it easier to determine binding site changes this data could be further analysed by PEKA in the form of a heat map which would show k-mers that are the most enriched and enable easier comparison between WT and *Malat1*$^{-/-}$ cells (Kuret et al., 2022). Additionally, due to time, tracks were manually inspected using IGV software, which is beneficial for determining crosslink sites, however, the reads are not normalised to the library size. Recently, a tool has been developed by the Ule laboratory known as clipplotr which circumvents this issue by normalising the data to the library size in addition to smoothing the crosslink sites which enables easier comparison between experimental conditions (Chakrabarti et al., 2021). It would be beneficial to replot this data using clipplotr.

As previously mentioned, we initially aimed to determine SRSF1 and hnRNPA1 interactions in both naïve and *in vitro* polarised Th2 cells derived from WT and *Malat1*$^{-/-}$ mice. Unfortunately, pilot experiments were unsuccessful due to low protein yields. The Ule lab has recently developed an improved iCLIP (iiCLIP) protocol, this involves the radioactive labelling of RNP complexes, and enhances the efficiency of other iCLIP steps such as library preparation. iiCLIP has been used to identify polypyrimidine tract-binding protein 1 (PTBP1) interactions from 1 million HEK293 cells (Lee et al., 2021.). By using iiCLIP it is possible that a lower cell number could be used to identify RBP interactions in naïve $CD4^+$ T cells.

One particularly interesting finding from our data set is that the loss of *Malat1* results in the movement of SRSF1 to the cytoplasm. We could anticipate that this would show reduced binding to intronic regions of RNAs (as splicing takes place in the nucleus). However, we saw no obvious reduction in the proportion of intron binding. This could be because as SRSF1 moves to the cytoplasm in *Malat1*$^{-/-}$ cells the bound genes move with SRSF1. To understand this observation further, it would be interesting to repeat the iCLIP experiment using a cell fractionation approach where cytoplasmic SRSF1 interaction is compared to nuclear SRSF1 interactions.

To analyse and determine differential gene expression and differential transcript usage we used short read illumina sequencing. Short-read sequencing technologies are inherently limited in their ability to detect long or complex isoforms as they do not sequence the full length of the transcript (Byrne et al., 2017). Bioinformatic methods are used to map the short reads onto the gene but different algorithms can result in different outcomes. Recent advances in the RNA-seq field by Oxford Nanopore Technology have developed long-read RNAseq which can either directly sequence RNA or full-length cDNA transcripts (Byrne et al., 2017). This sequencing technology would be beneficial to use if we wanted to further explore the role of *Malat1* in AS in the future.

6.3.2 Summary

In summary, we found *Malat1*-dependent RNA targets for both RBPs, suggesting that the presence of *Malat1* RNA is necessary for appropriate interaction between these RBPs and some of their target RNAs in T cells. This effect was more prominent for SRSF1 and is at least partly through Malat1-mediated control of the localisation of its RBP partners. Collectively, our results revealed that *Malat1* was an essential orchestrator of RBP function during highly dynamic cellular transitions, with relevance to immune cell function and potential implications in health and pathology.

7. Malat1 regulates IL-10 expression in $CD4^+$ T cells in a sex-specific manner

7.1 Introduction

7.1.1 Sexual dimorphism in immunity

The immune system provides essential protection from a vast and unpredictable number of pathogens. Yet, the immune system needs to maintain a balance between optimal pathogen clearance and minimal tissue damage. Interestingly, this fine balance differs between males and females and is known as a type of sexual dimorphism (Klein & Flanagan, 2016). Consequently, susceptibility to infectious diseases and predisposition to autoimmune disorders differs between the sexes (Márquez et al., 2020). Therefore, sex differences in the immune system should be an important consideration in immune-based studies and treatments.

The significance of sex differences in immunity has recently been highlighted by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic, in which older males were more likely to suffer from severe coronavirus disease 2019 (COVID-19) and a higher mortality rate. This male bias in infection severity is also prominent in other viral infections such as influenza, human immunodeficiency virus 1 (HIV-1), hepatitis C and Zika virus, although it can be challenging to separate this from exposure as reviewed by (Bunders & Altfeld, 2020). In addition, there is an increased prevalence of many parasitic infections in males (particularly helminths/parasitic worms), with males also presenting with a higher parasite burden in comparison to females (Zuk & McKean, 1996). On the whole, studies indicate that females produce stronger immune responses - typically generating double the antibody response to seasonal vaccines when compared to males (Klein & Pekosz, 2014). However, as a consequence, women are more susceptible to developing autoimmune disorders, with 80% of autoimmune cases reported in females (Bunders & Altfeld,

2020). Thus, understanding the underlying mechanisms responsible for these sex differences is essential to be able to treat immune disorders in males and females appropriately.

Several factors contribute to sexual dimorphism in male and female immune systems. These factors are predominantly hormones, chromosome composition and expression of X-linked genes (Bunders & Altfeld, 2020). Sex hormones such as oestrogen and testosterone are responsible for many differences in male and female immunity. An example of this can be seen in COVID-19 patients: angiotensin-converting enzyme 2 (ACE2) is a receptor expressed by numerous tissues which binds SARS-CoV-2 spike protein and enables entry into host cells. Interestingly, oestrogen has been shown to downregulate the expression of ACE2, consequently, ACE2 expression in lung tissue is much lower in women than men (J. Liu et al., 2010). In parallel, expression of the viral entry protein transmembrane serine protease 2 (TMPRSS2) increases in response to androgens. As such it has been observed that males who are treated with androgen deprivation therapy for prostate cancer may be somewhat protected from infection with SARS-CoV-2 (Montopoli et al., 2020). Collectively, these studies demonstrate how sex hormones can contribute to the sexual dimorphism observed in infection susceptibility.

There are numerous examples of sexual dimorphism throughout both arms of the immune system. Unsurprisingly, many differences are observed when comparing $CD4^+$ T cell responses in males and females. Interestingly, total lymphocyte counts are comparable between the sexes. However, the percentage of human αβ T cells is much lower in males than in females (Bouman et al., 2005). Testosterone is known to induce apoptosis in T cells and could be responsible for the lower αβ T cell counts observed in males (Mcmurray et al., 2001). The importance of hormones on Th cell

function is particularly evident in females, as $CD4^+$ T cell activity varies throughout the menstrual cycle and during pregnancy. Pregnancy-associated hormones systemically affect the function of nearly all immune cell types. Of interest, pregnancy hormones are known to create a local Th2 bias, with foetal tissue producing Th2 cytokines including IL-10, IL-4 and IL-5. This Th2 bias can be important in preventing miscarriage (H. Lin et al., 1993). Further understanding of sexual dimorphism in Th cell biology, presents an interesting area of research, in particular, understanding the underlying mechanism of sexual dimorphism would be of great benefit to the field of immunology.

Conflicting studies report on the expression of IL-10 in male and female Th cells. Some studies indicate that there is no difference in IL-10 expression when comparing male and female human Th cells (Bouman et al., 2005). Another study which examined the expression of IL-10 from human female PBMCs throughout various stages of the menstrual cycle found that expression of IL-10 remained stable throughout each menstrual stage, indicating that female hormones do not influence the expression of IL-10 (Faas et al., 2000). Contrastingly, one study which examined serum concentrations of IL-10 in human PBMC cultures found lower levels of IL-10 and IL-4 in male supernatants (Girón-González et al., 2000). In a mouse model of cutaneous melanoma, it was found that only female tumour-bearing mice showed increased expression of serum IL-10 (Surcel et al., 2017). Yet, male murine splenocytes stimulated with anti CD3 have a higher concentration of IL-10 in cell culture supernatants than their female equivalents (Liva & Voskuhl, 2001). Further, if female murine splenocytes were cultured in dihydrotestosterone (DHT) and CD3 this increased expression of IL-10 (Liva & Voskuhl, 2001). Additionally, LPS-challenged male mice produce higher levels of IL-10 than their female counterparts (Marriott et

al., 2006). A complex picture emerges where sex, age and species appear to influence IL-10 expression. The cellular source of IL-10 may also be a determinant of sex-specific effects. More specific studies will deepen our understanding of differences in IL-10 expression in male and female $CD4^+$ T cells.

7.1.2 Sexual dimorphism in lncRNA and RBP biology

The classic example of a sex-specific lncRNA is *Xist*, which plays an essential role in X chromosome inactivation in females (Penny et al., 1996). However, other lncRNAs which play a specific sex-based role remain largely unexplored. One study which carried out microarray analysis of whole blood from ischemic stroke patients found that 299 lncRNAs were dysregulated in male stroke patients compared to just 97 lncRNAs in female stroke patients. However, the functional relevance of these sex-specific lncRNA changes was not examined (Dykstra-Aiello et al., 2016). Other studies have examined sex-specific lncRNAs involved in gonad and ovary development in chickens and Chinese softshell turtles (Zhang et al., 2018; Zou et al., 2020). One study of Chinese softshell turtles found 932 female-specific lncRNAs and 449 male-specific lncRNAs in male or female-specific tissues, the authors postulated that these sex-based differences in lncRNA expression play an important role in sex differentiation (Zhang et al., 2018). Little work has compared sexual dimorphism in lncRNA RBP interactions, however, these may be responsible for some of the sexual dimorphism observed in human and mouse immunity.

Limited work has linked *Malat1* to any sex-specific functions. One study found that *Malat1* expression was significantly higher in male patients than in female patients with colorectal cancer.(Zheng et al., 2014). Additionally, in one study which used scRNA-seq to compare sexual dimorphisms in gene expression across ten different cell types from male and female adult mice, *Malat1* was found to be more

highly expressed in male fibroblasts, heart leukocytes, and heart smooth muscle cells (Lu & Mar, 2020). *Malat1* expression may also be influenced by hormones, as its expression level has been linked to several reproductive system cancers (Zhao et al., 2018). For example, *Malat1* is upregulated in castration-resistant prostate cancer, and its high expression is linked with disease progression (Ren et al., 2013). Additionally, 17β-estradiol treatment of breast cancer cells, inhibited cell proliferation and down regulated *Malat1* expression in a dose dependent manner (Zhao et al., 2014). Although only anecdotal evidence has linked *Malat1* to sex-specific functions it would be interesting to determine if *Malat1* may also have a sex-specific role in an immune context.

Sex-specific gene regulation also contributes to sexual dimorphism phenotypes. Scientists have postulated that epigenetic marks and transcriptomic regulators play a key role in sex-specific gene regulation. RBPs are an important class of post-transcriptional regulators, which control many cellular processes such as pre-mRNA splicing, mRNA nuclear export, mRNA turnover and translation to regulate gene expression (Turner & DÍaz-Muñoz, 2018). One study which examined 1,344 RBPs in the human liver found 45 RBPs to be significantly associated with sex, with ~40% of these proteins more highly expressed in males (Chaturvedi et al., 2015). This study suggested that age and sex-based differences in RBPs are important contributors to differences in liver development in males and females. Little work has compared RBP differences in male and female immune systems. However, given the myriad of sex-based variations in many aspects of immunity and that RBPs are key post-transcriptional regulators it is plausible that sexual dimorphism in RBP composition and expression could contribute to the different male and female phenotypes in the immune system. Some RBPs are also sex-linked for example DEAD-Box Helicase 3

X-Linked (DDX3X) and its Y linked counterpart DEAD-Box Helicase 3 Y Linked (DDX3Y) DDX3X is essential for IFN responses to pathogens in females, as knockout mice are more susceptible to *Listeria monocytogenes* and have a reduced number of lymphocytes. This phenotype could not be rescued with DDX3Y (Szappanos et al., 2018). During $CD4^+$ T cell differentiation thousands of genes are differentially expressed including lncRNAs and RBPs. Thus, $CD4^+$ T cells are an ideal platform for comparing potential differences in RBP and lncRNA functions in males and females.

7.1.3 Hypothesis and aims

The differences between male and female immune responses are striking. Yet, the potential role of lncRNA-RBP interactions in these sex-specific immune responses remains largely unexplored. Little work has reported a sex-specific role for *Malat1* apart from potential differences in IL-10 responses in Th cells between males and females. We wanted to explore the possibile impact of *Malat1* on IL-10 responses in Th cells in males and females. We hypothesised that loss of *Malat1* might result in a different $CD4^+$ T cell phenotype when comparing male and female mice.

To address this hypothesis this chapter aims to:

- Characterise Male and Female WT and *Malat1*$^{-/-}$ *in vitro* $CD4^+$ T cell differentiation
- Compare T cell phenotypes in models of inflammation between Male and Female WT and *Malat1* $^{-/-}$ mice
- Compare *Malat1* interaction partners in male and female $CD4^+$ T cells

7.2 Results

7.2.1 *Malat1* downregulation is a hallmark of $CD4^+$ T cell activation in male and female cells

To gain insight into the role of *Malat1* in male and female $CD4^+$ T cells, we began our studies by using qRT-PCR to determine the levels of *Malat1* expression in different $CD4^+$ T cell states. We observed that after 24 hours of *in vitro* stimulation under Th2 polarising conditions, *Malat1* was significantly downregulated in both female and male $CD4^+$ T cells to a comparable extent (***Figure 7.1A***). Upon further analysis on day 6 of *in vitro* T cell polarisation, we showed that *Malat1* was expressed at lower levels in WT male Th2 cells than their female counterparts (***Figure 7.1B***). Additionally, *Malat1* expression was comparably abolished in both knockout male and female $CD4^+$ T cells.

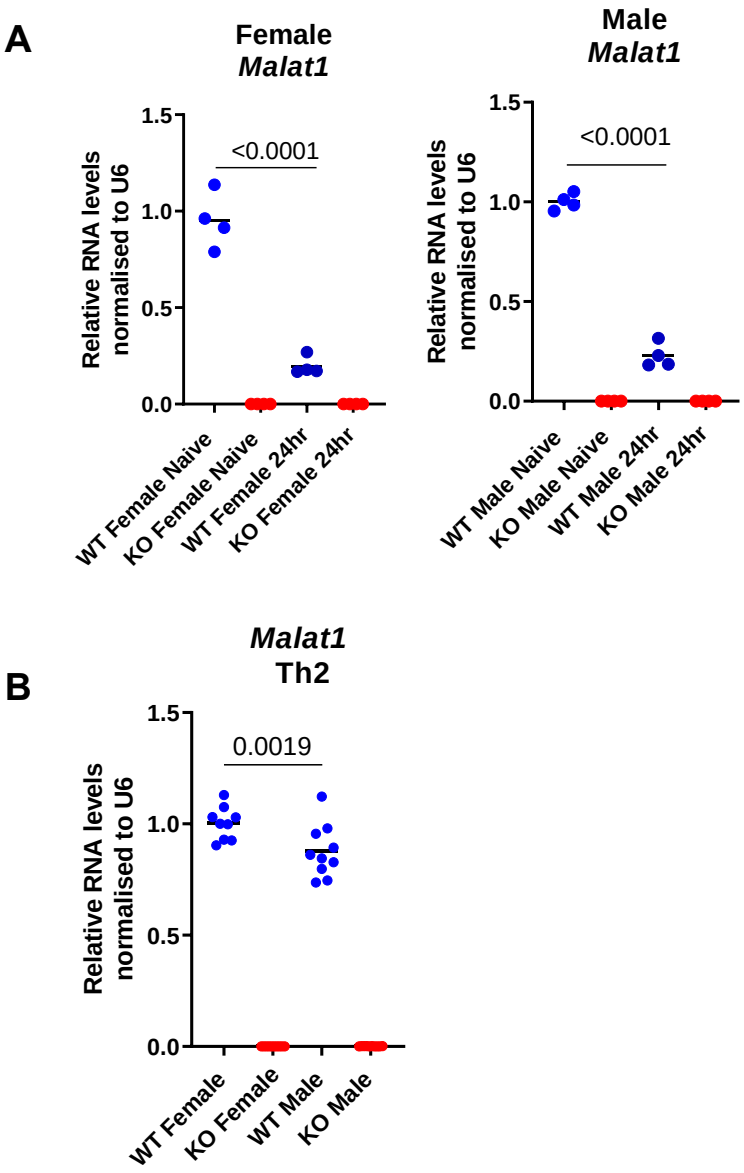


Figure 7.1: *Malat1* is downregulated upon T cell activation in male and female CD4⁺ T cells.

A) RNA levels of *Malat1* in naïve CD4⁺ T cells and after 24 hours of stimulation. RNA levels of *Malat1* in female CD4⁺ T cells are shown on the left WT (blue) *Malat1*$^{-/-}$ (red). RNA levels of *Malat1* in male CD4⁺ T cells are shown on the right WT (blue) *Malat1*$^{-/-}$ (red). **B)** RNA levels of *Malat1* in in vitro polarised Th2 cells (day 6). Male and female cells are labelled, WT (blue) *Malat1*$^{-/-}$ (red). RNA levels are normalised to U6 RNA average expression in naïve cells or WT female cells. RNA levels were determined by qRT-PCR. 24 hours timepoint n=4 biological replicates representative of 2 independent experiments. Day 6 n=9 biological replicates of 2 pooled independent experiments. Data were analysed using a one-way ANOVA with post hoc Tukeys test. p value is displayed. The mean is depicted.

We further investigated these samples and analysed the expression of an early $CD4^+$ T cell activation marker. Notably, loss of *Malat1* significantly reduced the expression of *Cd69* in both male and female $CD4^+$ T cells after 24 hours of stimulation (***Figure 7.2***). This indicated that *Malat1* has similar functions in male and female cells at the early stages of T cell activation, however the kinetics of this have not been explored fully and could differ between male and female $CD4^+$ T cells.

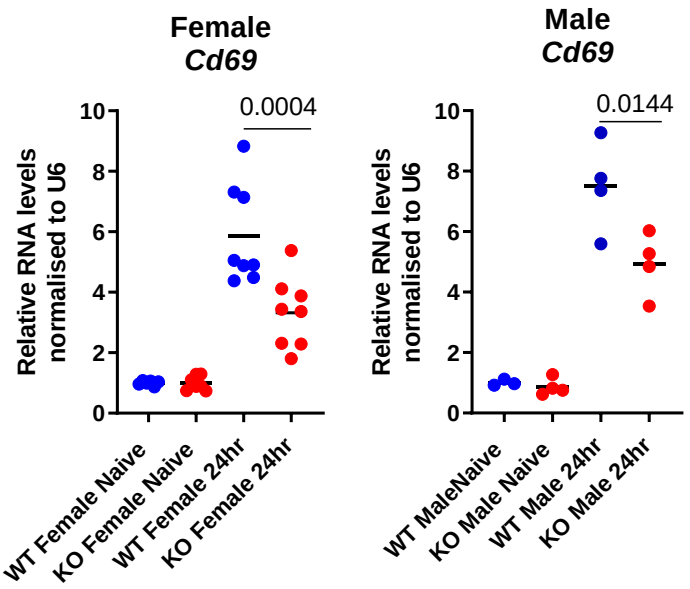


Figure 7.2 CD69 is downregulated in the absence of *Malat1* upon T cell activation in both male and female cells.

RNA levels of *Cd69* in naïve CD4$^+$ T cells and after 24 hours of stimulation. RNA levels of *Cd69* in female CD4$^+$ T cells are shown on the left WT (blue) *Malat1*$^{-/-}$ (red). RNA levels of *Cd69* in male CD4$^+$ T cells are shown on the right WT (blue) *Malat1*$^{-/-}$ (red). RNA levels were normalised to U6 RNA average expression in naïve cells. Expression was determined by qRT-PCR. n=4-8 biological replicates of 1 or 2 independent experiments. Data were analysed using a one-way ANOVA. P value is displayed. The mean is depicted.

7.2.2 Loss of *Malat1* reduced the expression of IL-10 in a female-specific manner *in vitro*

In chapter 3, we demonstrated that loss of *Malat1* suppressed IL-10 expression in female $CD4^+$ T cells. We went on to show that loss of *Malat1* impaired Th cell differentiation which was most pronounced in Th2 cells (chapter 4). We next explored the effects of loss of *Malat1* on IL-10 expression in male and female Th2 cells. We found that following *in vitro* Th2 polarisation, loss of *Malat1* suppressed the expression of IL-10 in female cells. However, this suppression was not observed in male Th2 cells at the protein level (***Figure 7.3***). Interestingly, IL-10 expression was lower in WT males compared to WT females. This indicated that the extent of differentiation was reduced in male $CD4^+$ T cells. However, IL-4 expression was comparable between male and female cells. Additionally, loss of *Malat1* did not affect IL-4 expression in either male or female $CD4^+$ T cells (***Figure 7.3***).

We next analysed IL-10 expression by qRT-PCR in male and female *in vitro* polarised Th2 cells. Similar to protein expression, RNA levels of IL-10 were diminished in female *Malat1*$^{-/-}$ Th2 cells, yet IL-10 was upregulated in male *Malat1*$^{-/-}$ Th2 cells (***Figure 7.4****.)*.

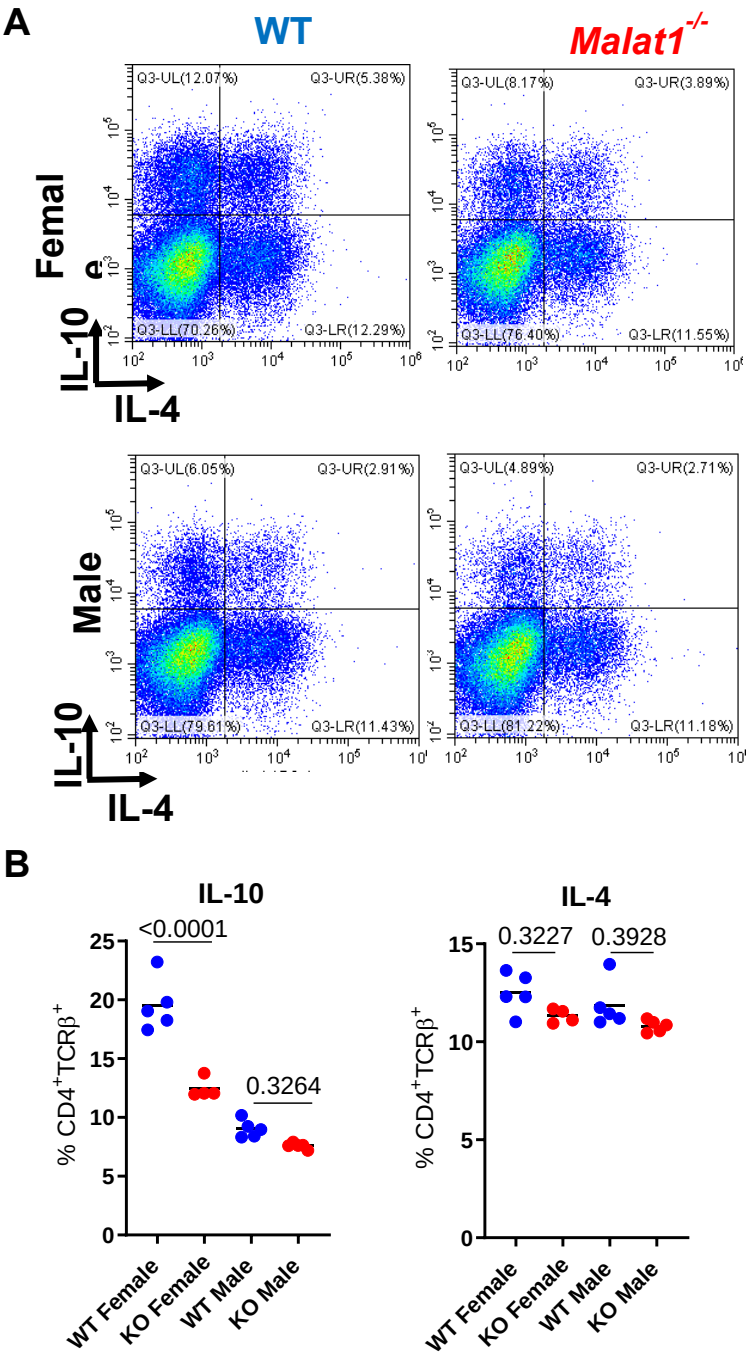


Figure 7.3 Downregulation of IL-10 at the protein level in *Malat1*$^{-/-}$ cells is female-specific *in vitro.* A) Representative FACs plots of IL-10 and IL-4 expression in in vitro polarised Th2 cells (day 6) in WT or *Malat1*$^{-/-}$ cells determined by intracellular cytokine staining. Female cells are shown in the top panel. Male cells are shown in the bottom panel **B)** Quantification of the percentage of live, CD4^{+} TCRβ^{+} cells that express IL-10 or IL-4 in WT (blue) or *Malat1*$^{-/-}$ (red) in vitro polarised Th2 cells (day 6). Male and female cells are annotated. n=5/4 biological replicates representative of 3 independent experiments. Data were analysed using a one-way ANOVA with post hoc Tukeys test. p value is displayed. The mean is depicted.

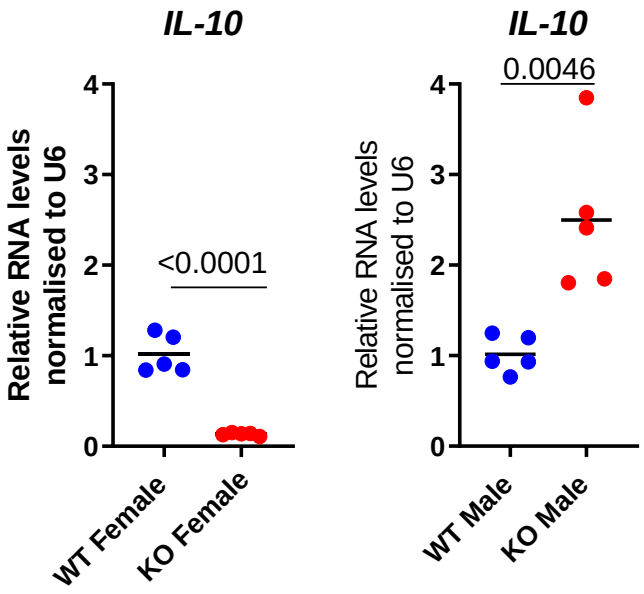


Figure 7.4: Downregulation of IL-10 at the RNA level in *Malat1*$^{-/-}$ cells is female-specific *in vitro.*

RNA levels of IL-10 in Th2 polarised cells (day 6). RNA levels of IL-10 in female CD4^{+} T cells are shown on the left WT (blue) *Malat1*$^{-/-}$ (red). RNA levels of IL-10 in male CD4^{+} T cells are shown on the right WT (blue) *Malat1*$^{-/-}$ (red). RNA levels were normalised to U6 RNA average expression in WT cells. Expression was determined by qRT-PCR. n=5 biological replicates representative of 2 independent experiments. Data were analysed using an unpaired t-test. p value is displayed. The mean is depicted.

7.2.3 Loss of *Malat1* impairs proliferation in male CD4$^+$ T cells *in vitro*

Cell counts were examined at different stages of the *in vitro* T cell differentiation process. Similar cell counts were observed after 24 hours, and 4 days of stimulation. However, loss of *Malat1* failed to expand male Th2 cells between days 4 and 6 of culture (i.e when cultured in IL-2) (***Figure 7.5A***). This suggested that T cell proliferation was impaired. Thus, we next examined cell proliferation in a pilot study using CFSE staining and flow cytometry. WT CD4$^+$ T cells proliferated in response to stimulation in both males and females. However, loss of *Malat1* specifically in male CD4$^+$ T cells impaired proliferation as demonstrated by higher CFSE levels (***Figure 7.5B***).

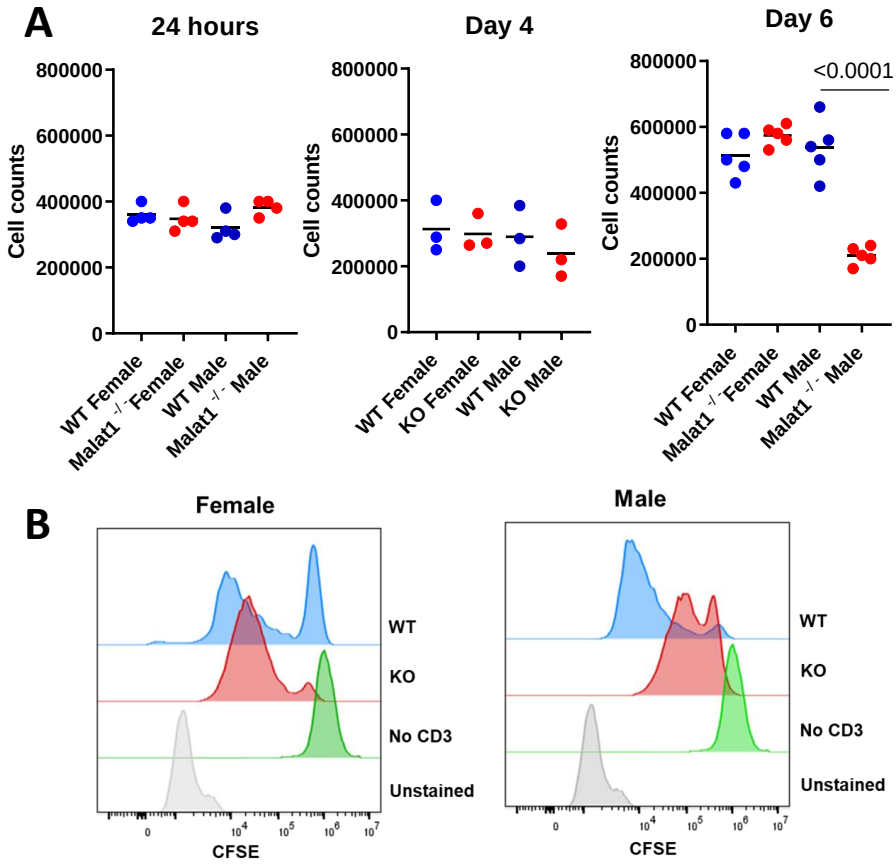


Figure 7.5: Loss of *Malat1* in male CD4⁺ T cells impairs proliferation.

A) Cell counts during *in vitro* T cell polarisation in WT (blue) and *Malat1*$^{-/-}$ (red) cells. Male or female cells are annotated. Cell counts were taken after 24 hours, 4 days and 6 days of culture. n=3-5 representative of 3 independent experiments. Data were analysed by one-way ANOVA with post hoc Tukeys test. p value is displayed. **B)** Representative FACs plots of CD4⁺ T cells derived from WT or *Malat1*$^{-/-}$ mice stimulated for 6 days and stained with CFSE at the start of the experiment. The histograms depict live, single cells CFSE expression. Female cells are shown on the left. Male cells are shown on the right. Representative histograms of n=3 technical replicates of 1 independent experiment.

7.2.4 Loss of *Malat1* reduces IL-10 expression in female $CD4^+$ T cells *in vivo*

Next, to explore the relevance of these findings *in vivo* we used an *S. mansoni* egg injection model of lung inflammation. This model was chosen as it induced a potent Th2 inflammatory response and stimulated $CD4^+$ T cell activation in the lungs and was a relatively short *in vivo* assay (Joyce et al., 2012). A schematic representation of the model is shown in (***Figure 7.6***). First *S. mansoni* eggs are recovered from worm-infected mouse livers. The mice were then primed with eggs through intraperitoneal injection. Next, the mice were challenged with eggs intravenously. The eggs are transported to the lungs via the pulmonary artery and trapped in the lung parenchyma in granulomas. This was a pilot study where male and female samples were processed on different days.

As expected upon egg injection, spleen and lung weights increased in size both in grams and as a percentage of body weight. However, no significant differences were observed between WT and *Malat1*$^{-/-}$ male or female mice ***(Figure 7.7)***. As such, similar cell counts were observed between WT and *Malat1*$^{-/-}$ mice in both the spleen and lung **(*Figure 7.8)*.**

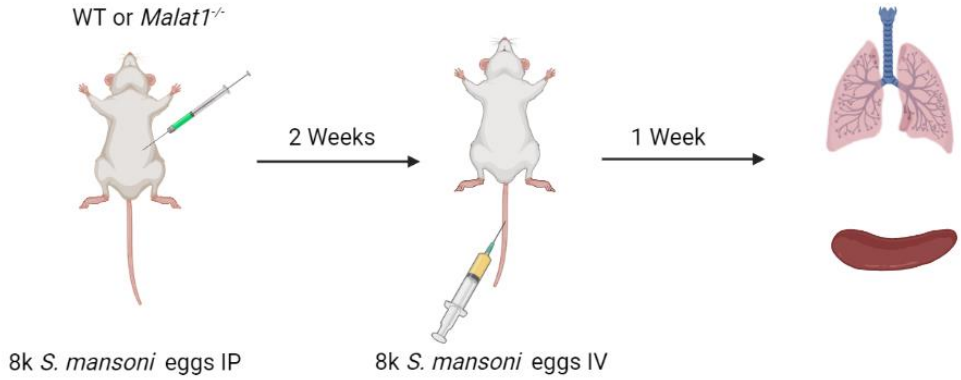


Figure 7.6: Schematic of an *S. mansoni* egg injection model

Diagram depicting *S. mansoni* egg injection model. Briefly, female and male WT and *Malat1*$^{-/-}$ mice (6-12 weeks old), are first primed with 8,000 *S. mansoni* eggs in 200µl of PBS through intraperitoneal injection. After two weeks mice are challenged with 8,000 *S. mansoni* eggs in 200 µl of PBS through intravenous injection. Following the challenge, *S. mansoni* eggs are transported to the lung via the pulmonary artery and become trapped in the lung parenchyma. After one week, mice were sacrificed and tissues were harvested for analysis.

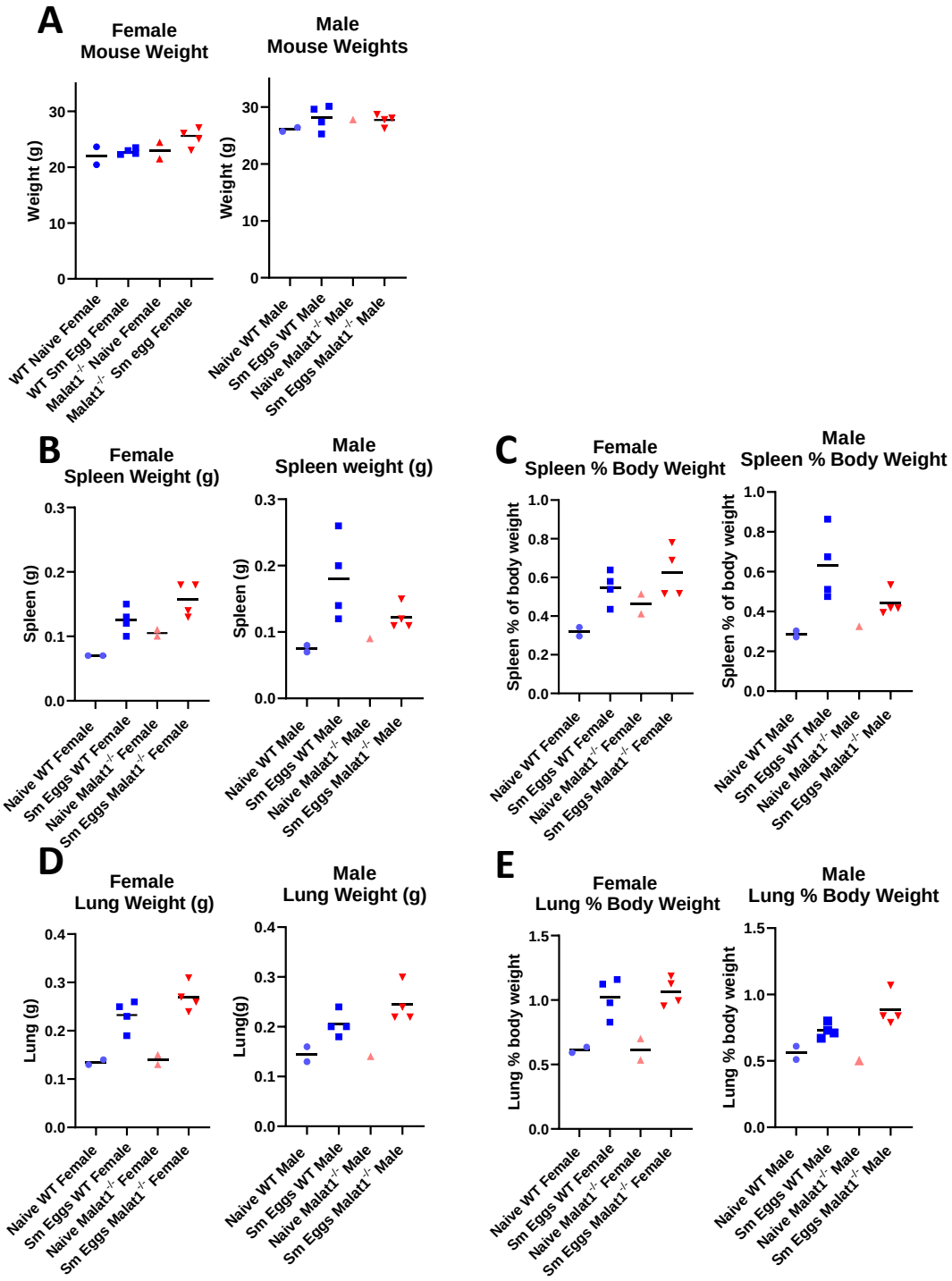

A
Female
Mouse Weight
Male
Mouse Weights
Weight (g)
0
10
20
30
WT Naive Female
WT Sm Egg Female
Malat1 -/- Naive Female
Malat1 -/- Sm egg Female
Naive WT Male
Sm Eggs WT Male
Naive Malat1 -/- Male
Sm Eggs Malat1 -/- Male
B
Female
Spleen Weight (g)
Male
Spleen weight (g)
Spleen (g)
0.0
0.1
0.2
0.3
Naive WT Female
Sm Eggs WT Female
Naive Malat1 -/- Female
Sm Eggs Malat1 -/- Female
Naive WT Male
Sm Eggs WT Male
Naive Malat1 -/- Male
Sm Eggs Malat1 -/- Male
C
Female
Spleen % Body Weight
Male
Spleen % Body Weight
Spleen % of body weight
0.0
0.2
0.4
0.6
0.8
1.0
Naive WT Female
Sm Eggs WT Female
Naive Malat1 -/- Female
Sm Eggs Malat1 -/- Female
Naive WT Male
Sm Eggs WT Male
Naive Malat1 -/- Male
Sm Eggs Malat1 -/- Male
D
Female
Lung Weight (g)
Male
Lung Weight (g)
Lung (g)
Lung(g)
0.0
0.1
0.2
0.3
0.4
Naive WT Female
Sm Eggs WT Female
Naive Malat1 -/- Female
Sm Eggs Malat1 -/- Female
Naive WT Male
Sm Eggs WT Male
Naive Malat1 -/- Male
Sm Eggs Malat1 -/- Male
E
Female
Lung % Body Weight
Male
Lung % Body Weight
Lung % body weight
0.0
0.5
1.0
1.5
Naive WT Female
Sm Eggs WT Female
Naive Malat1 -/- Female
Sm Eggs Malat1 -/- Female
Naive WT Male
Sm Eggs WT Male
Naive Malat1 -/- Male
Sm Eggs Malat1 -/- Male

Figure 7.7: Mouse and organ weight is unchanged in the absence of *Malat1* in an *S. mansoni* egg injection model

A) Body weight of *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice. Female weights are shown on the left. Male weights are shown on the right. **B)** Spleen weight of *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice. Female weights are shown on the left. Male weights are shown on the right. **C)** Spleen weight expressed as a percentage of body weight of *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice. Female weights are shown on the left. Male weights are shown on the right. **D)** Lung weight of *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice. Female weights are shown on the left. Male weights are shown on the right. **E)** Lung weight expressed as a percentage of body weight of *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice. Female weights are shown on the left. Male weights are shown on the right. n=1/2 naïve, n=4 injected biological replicates representative of 1 independent experiment – male and female samples were processed on different days. Data were analysed by one-way ANOVA with post hoc Tukeys. No significant differences were found. The mean is depicted.

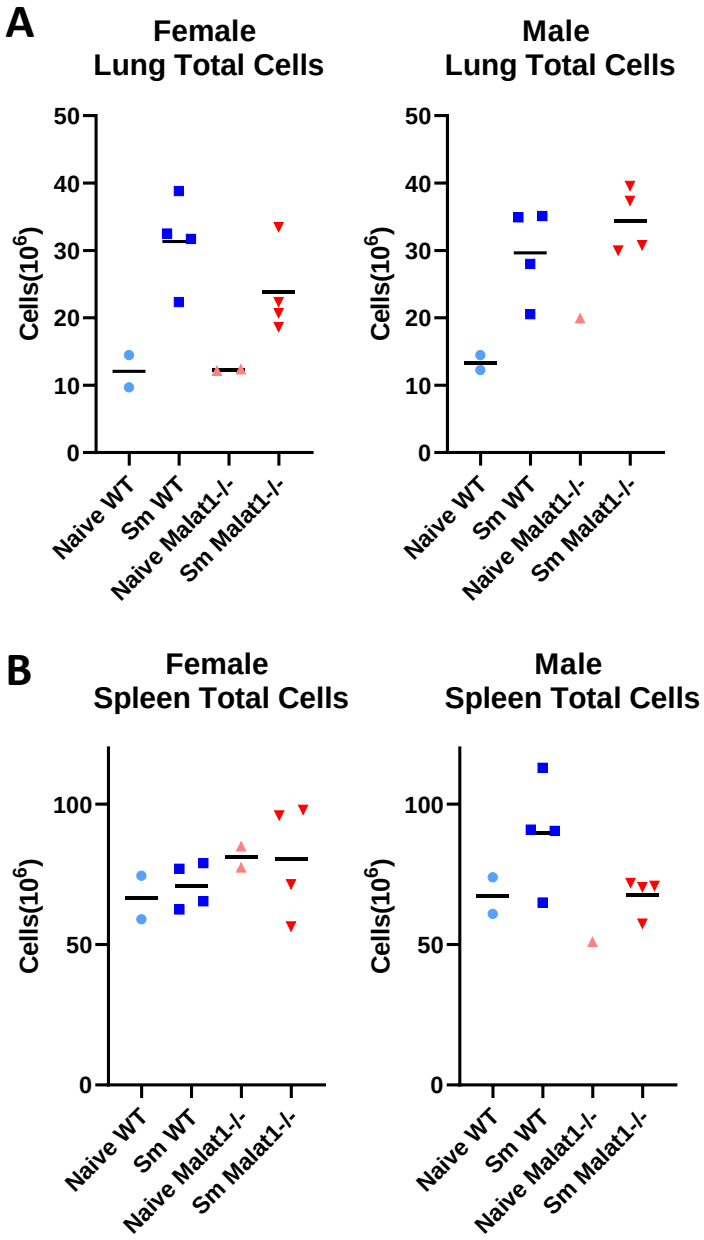


Figure 7.8: Loss of *Malat1* does not alter total lung cell numbers in an *S. mansoni* egg injection model.

A) Total number of live cells in lungs from naïve and *S. mansoni* egg-injected mice. WT (blue) *Malat1*$^{-/-}$ (red). Male and female cells are labelled. **B)** Total number of live cells in spleens from naïve and *S. mansoni* egg-injected mice. WT (blue) *Malat1*$^{-/-}$ (red). Male and female cells are labelled. n=1/2 naïve, n=4 biological replicates injected representative of 1 independent experiment – male and female samples were processed on different days. Data were analysed by one-way ANOVA with post hoc Tukeys. No significant differences were found. The mean is depicted.

We next explored the proportions of different immune populations of naïve and egg-injected mice by flow cytometry. In the lungs of egg-injected mice, the number of $CD4^+$ T cells had expanded in both male and female mice with a similar number of cells observed between WT males and females ***(Figure 7.9A)***. No significant difference in the proportion of $CD4^+$ T cells was observed in the absence of *Malat1*. The percentage of $CD8^+$ T cells and B cells fell in female mouse lungs upon egg injection ***(Figure 7.9B/C).*** Yet, the percentage of $CD8^+$ T cells and B cells remained similar to naïve male mice. This was because $CD8^+$ T cells and B cells expanded specifically in the lungs of male mice upon egg injection, and we saw a greater expansion of eosinophils in females. Notably, Loss of *Malat1* did not affect the proportion of either $CD8^+$ T cells or B cells ***(Figure 7.9)***.

Similarly, upon examination of the proportion of cells in the spleen, no significant differences were observed in the absence of *Malat1 (***Figure *7. 10*).** The percentage and number of $CD4^+$ remained at similar levels before and after egg injection in both male and female mice (**Figure *7. 10*A/C**). Further analysis of the proportion of $CD8^+$ T cells revealed a slight decrease in the percentage in both male and female egg-injected mice compared to naïve (**Figure *7. 10*C**). The percentage of B cells remained consistent in female and male mice after egg injection. However, B cells were approximately twice more prevalent in male mice than female mice in both naïve and egg injection conditions (**Figure *7. 10*C**).

We next analysed the proportion of activated $CD4^+$ T cells in the lungs and spleens of egg-injected mice (**Figure *7.11***). The number and percentage of activated T cells increased upon egg injection. There was no significant difference in the percentage or proportion of activated $CD4^+$ T cells in the absence of *Malat1*.

Lung

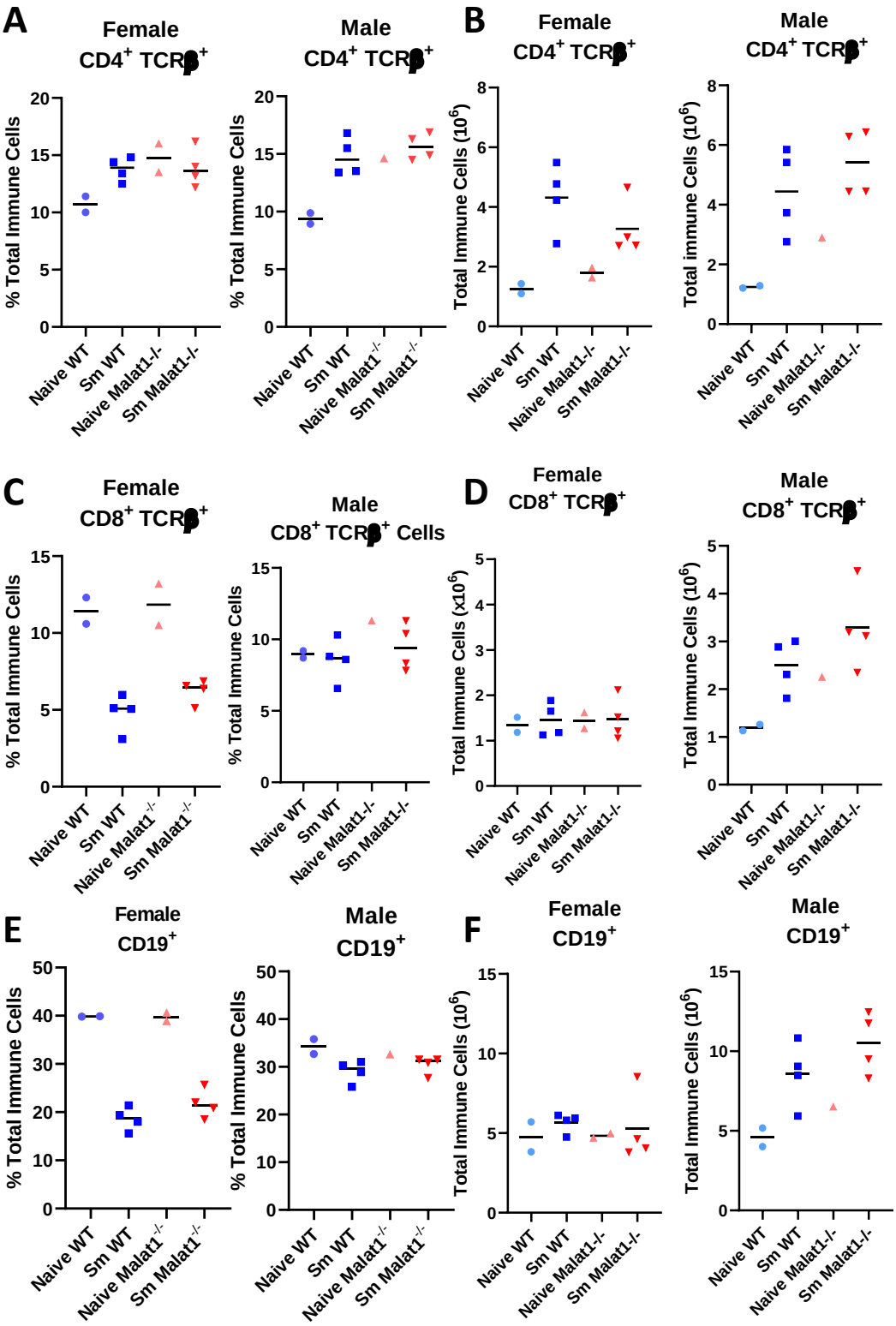

A
Female
CD4+ TCRβ+
% Total Immune Cells
Male
CD4+ TCRβ+
% Total Immune Cells
B
Female
CD4+ TCRβ+
Total Immune Cells (10^6)
Male
CD4+ TCRβ+
Total immune Cells (10^6)
C
Female
CD8+ TCRβ+
% Total Immune Cells
Male
CD8+ TCRβ+ Cells
% Total Immune Cells
D
Female
CD8+ TCRβ+
Total Immune Cells (x10^6)
Male
CD8+ TCRβ+
Total Immune Cells (10^6)
E
Female
CD19+
% Total Immune Cells
Male
CD19+
% Total Immune Cells
F
Female
CD19+
Total Immune Cells (10^6)
Male
CD19+
Total Immune Cells (10^6)
Naive WT
Sm WT
Naive Malat1-/-
Sm Malat1-/-

Figure 7.9: Proportion of CD4$^+$ T cells remains unchanged in the lungs in the absence of *Malat1* in an *S. mansoni* egg injection model.

A) Quantification of the percentage or number **B)** of live, CD4$^+$ TCRβ$^+$ cells of *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice lungs. Female (left) Male (right). **C)** Quantification of the percentage or number **D)** of live, CD8$^+$ TCRβ$^+$ cells of *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice lungs. Female (left) Male (right). **E)** Quantification of the percentage or number **F)** of live, CD19$^+$ cells of *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice lungs. Female data (left) male data (right). Protein levels were determined by flow cytometry. n=1/2 naïve, n=4 injected biological replicates representative of 1 independent experiment – male and female samples were processed on different days. Data were analysed by one-way ANOVA with post hoc Tukeys. No significant differences were found. The mean is depicted.

Spleen

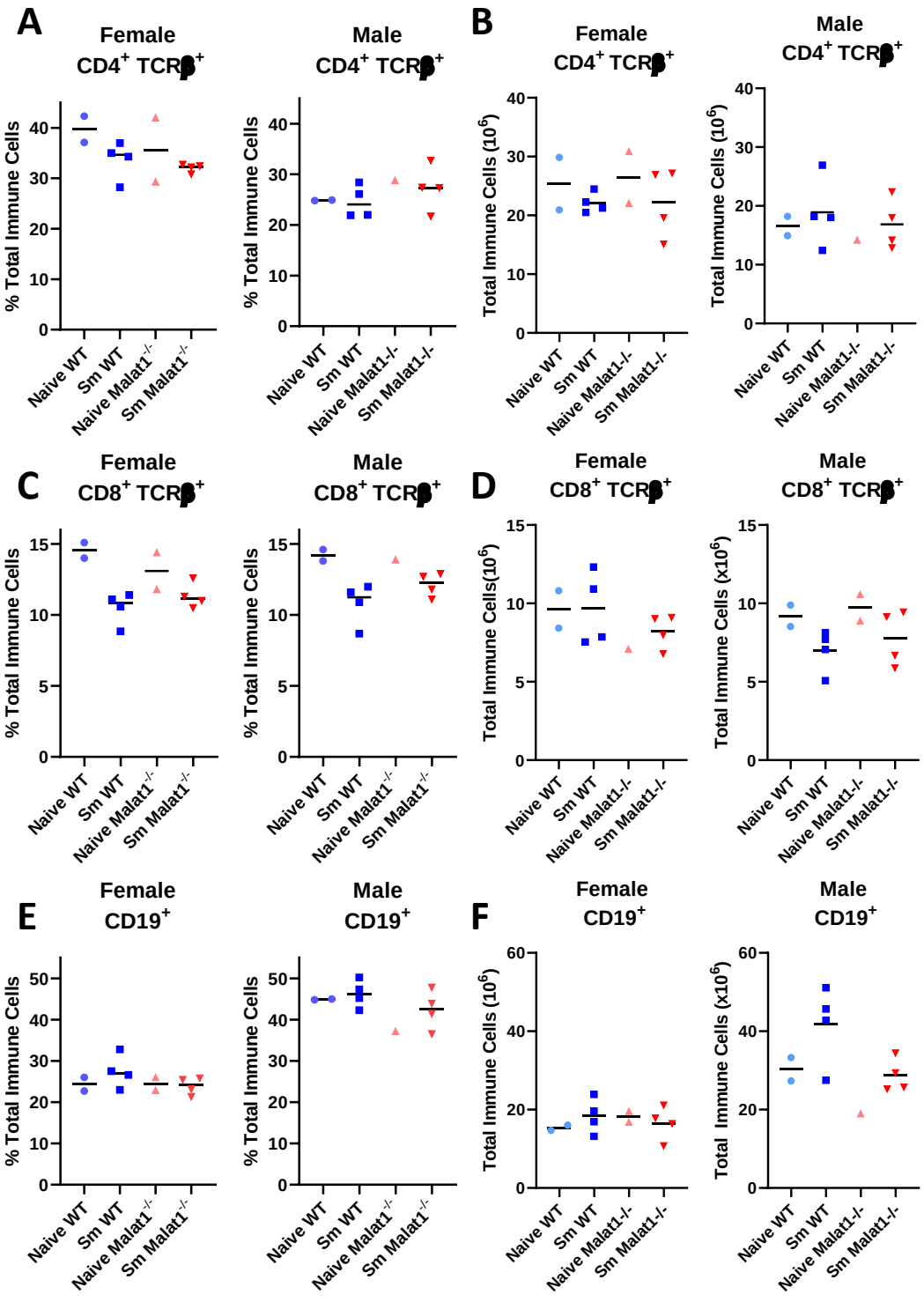

A
Female
CD4+ TCRβ+
% Total Immune Cells
Male
CD4+ TCRβ+
% Total Immune Cells
B
Female
CD4+ TCRβ+
Total Immune Cells (10^6)
Male
CD4+ TCRβ+
Total Immune Cells (10^6)
C
Female
CD8+ TCRβ+
% Total Immune Cells
Male
CD8+ TCRβ+
% Total Immune Cells
D
Female
CD8+ TCRβ+
Total Immune Cells(10^6)
Male
CD8+ TCRβ+
Total Immune Cells (x10^6)
E
Female
CD19+
% Total Immune Cells
Male
CD19+
% Total Immune Cells
F
Female
CD19+
Total Immune Cells (10^6)
Male
CD19+
Total Immune Cells (x10^6)
Naive WT
Sm WT
Naive Malat1-/-
Sm Malat1-/-

Figure 7. 10: Proportion of CD4$^+$ T cells remains unchanged in the spleen in the absence of *Malat1* in an *S. mansoni* egg injection model.

A) Quantification of the percentage or number **B)** of live, CD4$^+$ TCRβ$^+$ cells of *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice spleens. Female (left) Male (right). **C)** Quantification of the percentage or number **D)** of live, CD8$^+$ TCRβ$^+$ cells of *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice spleens. Female (left) Male (right). **E)** Quantification of the percentage or number **F)** of live, CD19$^+$ cells of *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice spleens. Female (left) Male (right). n=1/2 naïve, n=4 injected biological replicates representative of 1 independent experiment – male and female samples were processed on different days. Data were analysed by one-way ANOVA no significant differences were found. The mean is depicted.

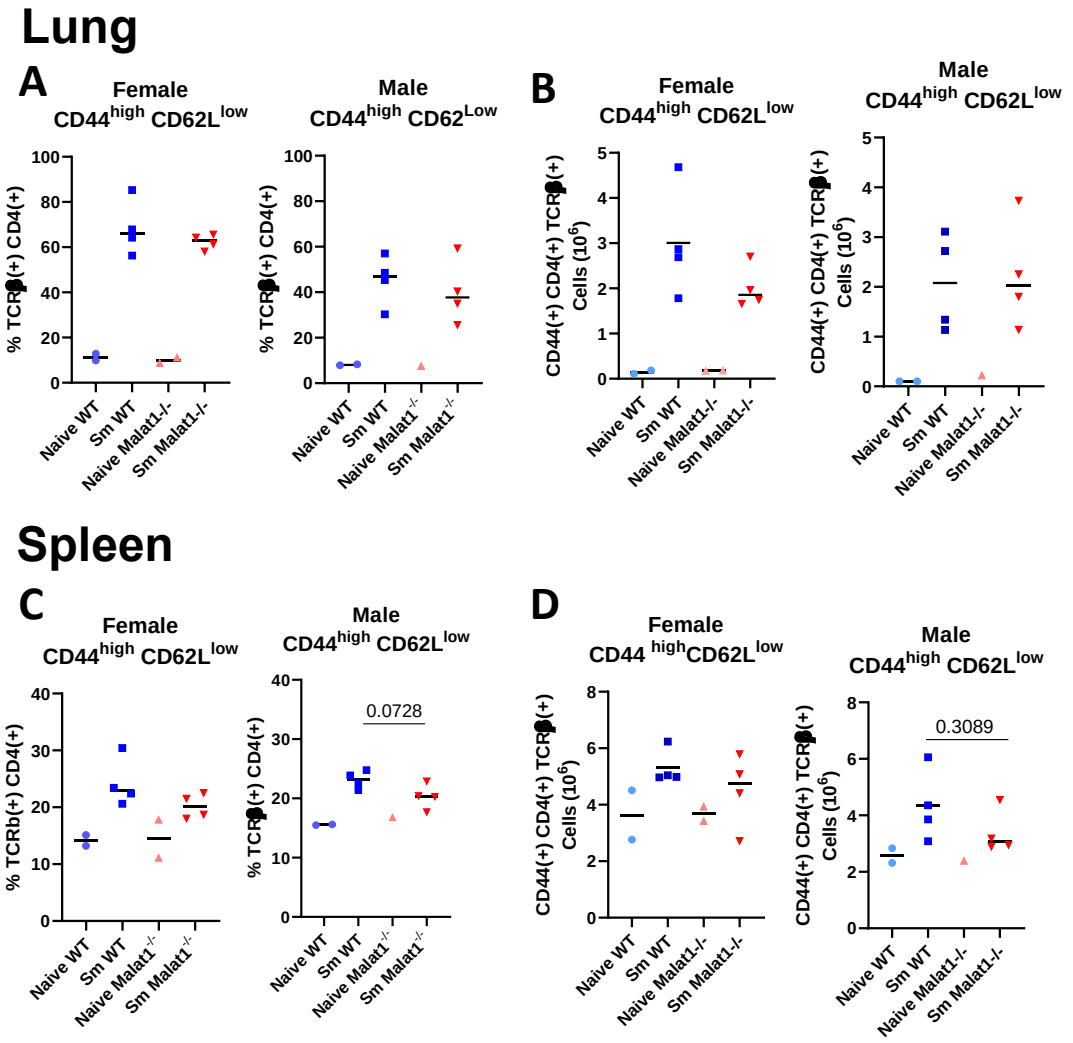


Figure 7.11: Loss of *Malat1* does not impair CD4⁺ T cell activation *in vivo.*

A) Quantification of the percentage or number **B)** of live, CD4⁺ TCRβ⁺ CD44^{high} CD62L^{low} of *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice in the lung. Female data is shown on the left. Male data is shown on the right. **C)** Quantification of the percentage or number **D)** of live, CD4⁺ TCRβ⁺ CD44^{high} CD62L^{low} of *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice in the spleen. Female data are shown on the left. Male data are shown on the right. n=1/2 naïve, n=4 injected biological replicates representative of 1 independent experiment – male and female samples were processed on different days. Data were analysed by one-way ANOVA with post-hoc Tukeys no significant differences were found. The mean is depicted.

To determine if *Malat1* also regulated IL-10 expression in a sex-specific manner *in vivo* we performed intracellular cytokine staining followed by flow cytometry on lungs and spleens from *S. mansoni* egg-injected mice. This demonstrated that loss of *Malat1* reduced IL-10 expression in $CD4^+$ T cells in a sex-specific manner. *Malat1*$^{-/-}$ female mice significantly reduced IL-10 expression as a percentage and number of cells in the lungs (***Figure 7.12)***. This female reduction was borderline significant in the spleen when expressed as a percentage, however, with further repeats performed by Magnus Gwynne, the reduction in IL-10 expression does become significant (data not shown) (***Figure 7.13)***. Interestingly, in male $CD4^+$ T cells in the absence of *Malat1,* the expression of IL-10 increased significantly when expressed as total cell numbers in the lung (***Figure 7.12)***. This reflected similar findings *in vitro* at the RNA level which showed increased IL-10 expression in the absences of *Malat1* in male $CD4^+$ T cells *in vitro.*

IL-4 expression is not significantly altered by loss of *Malat1* in either the spleen or lungs in male and female $CD4^+$ T cells (***Figure 7.12/13)***. Nor is IFNγ significantly altered in the absence of *Malat1,* although in both the spleen and lung the number of IFNγ cells is reduced but not significant (***Figure 7.12/13)***. Together this data showed that *Malat1* regulated IL-10 in a sex-specific manner both *in vitro* and *in vivo* in a Th2 context.

Of note, female WT $CD4^+$ T cells produced approximately double the percentage of IL-10, IL-4 and IFN_γ than their male counterparts. This suggested that female Th cell differentiation is stronger than males in the *S. mansoni* egg injection model.

Lung Intracellular Cytokines

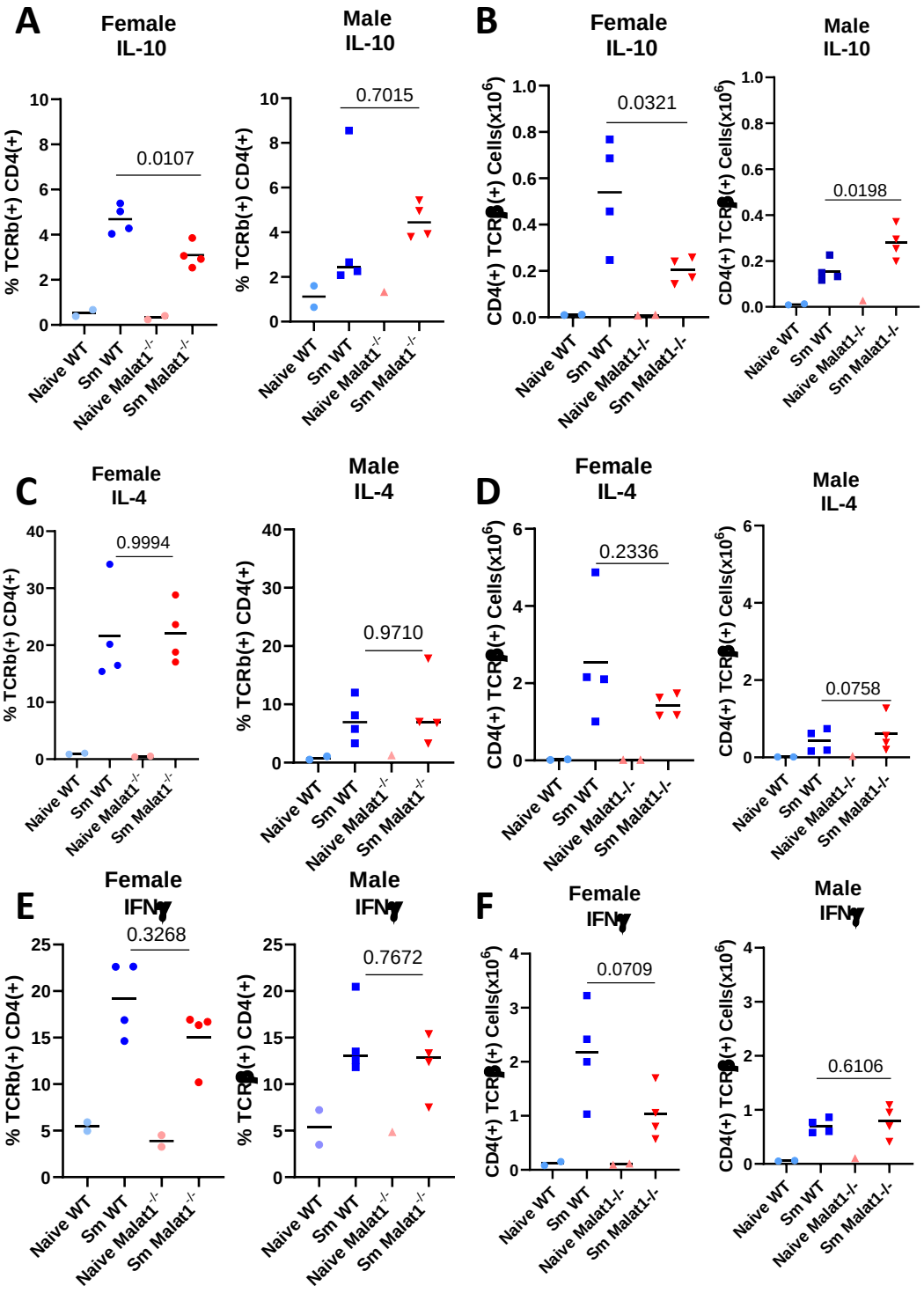


Figure 7.12: IL-10 is downregulated specifically in female *Malat1*$^{-/-}$ CD4^{+} T cells in the lungs of mice of an *S. mansoni* egg injection model.

A) Quantification of the percentage or **B)** number of IL-10^{+} live TCRβ^{+} CD4^{+} cells from *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice lungs, protein levels were determined by intracellular cytokine staining. Female data (left) male data (right). **C)** Quantification of the percentage or **D)** of IL-4^{+} live TCRβ^{+} CD4^{+} cells from *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice lungs, protein levels were determined by intracellular cytokine staining. Female data (left) male data (right). **E)** Quantification of the percentage or **F)** number of IFNγ^{+} live TCRβ^{+} CD4^{+} cells from *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice lungs, protein levels were determined by intracellular cytokine staining. Female data (left) male data (right). n=1/2 naïve, n=4 injected representative of 1 independent experiment – male and female samples were processed on different days. Data were analysed by one-way ANOVA with post hoc Tukeys. p values are displayed. The mean is depicted.

Spleen Intracellular Cytokines

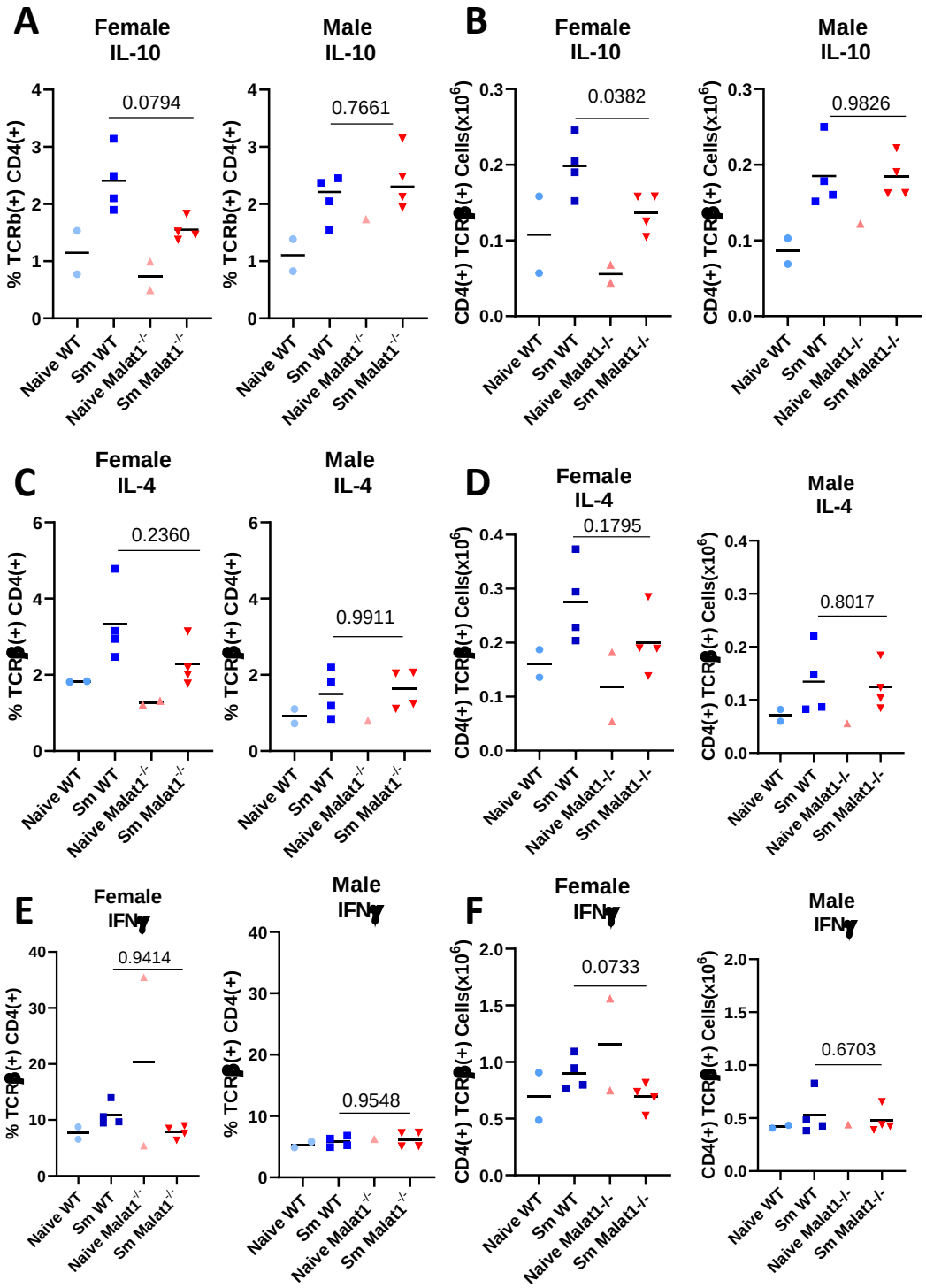


Figure 7.13: IL-10 is downregulated specifically in female *Malat1*$^{-/-}$ CD4^{+} T cells derived from spleens in an *S. mansoni* egg injection model.

A) Quantification of the percentage or **B)** number of IL-10^{+} live TCRβ^{+} CD4^{+} cells from *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice spleens, protein levels were determined by intracellular cytokine staining. Female data (left) male data (right). **C)** Quantification of the percentage or number **D)** of IL-4^{+} live TCRβ^{+} CD4^{+} cells from *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice spleens, protein levels were determined by intracellular cytokine staining. Female data (left) male data (right). **E)** Quantification of the percentage or number **F)** of IFNγ^{+} live TCRβ^{+} CD4^{+} cells from *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice spleens, protein levels were determined by intracellular cytokine staining. Female data (left) male data (right). n=1/2 naïve, n=4 injected biological replicates representative of 1 independent experiment – male and female samples were processed on different days. Data were analysed by one-way ANOVA with post hoc Tukeys. p values are displayed. The mean is depicted.

To determine if loss of *Malat1* altered the proportion or activation of myeloid cells different myeloid populations were examined by flow cytometry. I processed the samples for flow cytometry jointly with Magnus Gwynne. Magnus Gwynne performed the flow cytometry analysis with assistance from Joanna Greenman. In females, loss of *Malat1* significantly increased the percentage of Ly6C^{high} cells (pro-inflammatory monocytes) and decreased the percentage (although non-significant) and the number of Ly6C^{low} cells (patrolling monocytes) upon egg injection (**Figure *7.14***). A higher number of Ly6C^{high} cells were observed in both WT and *Malat1*$^{-/-}$ male mice than their female counterparts. Additionally, a lower number and percentage of Ly6C^{low} cells were found in both WT and *Malat1*$^{-/-}$ egg-injected mice (**Figure *7.14***).

No significant differences were observed when comparing the expression of RELMα and Ym1 in the absence of *Malat1* in Ly6C^{high} cells ***(Figure 7.15)***. However, the percentage of Ly6C^{high} MHCII cells increased in the absence of *Malat1* specifically in male mice ***(Figure 7.15E)***. Loss of *Malat1* reduced the number of RELMα Ly6C^{low} cells specifically in female mice, with a slight but non-significant increase in RELMα ***(Figure 7.16)***. No significant differences were observed when comparing Ym1 or MHCII levels between WT and *Malat1*$^{-/-}$ Ly6C^{low} cells. The percentage of Ym1 Ly6C^{low} cells was elevated when comparing WT females and WT males ***(Figure 7.16)***.

No significant changes in the number or percentage of eosinophils were observed in the absence of *Malat1 (***Figure *7.17)***. Additionally, loss of *Malat1* did not affect the expression of RELMα in either male or female eosinophils upon egg injection. Of note egg injection resulted in a large expansion of eosinophils in both

WT male and female mice, however, this was more pronounced in females *(***Figure 7.17***)*.

Similarly, no differences were observed in the proportion of neutrophils upon loss of *Malat1* in the egg injection model (***Figure 7.18***). However, the proportion of neutrophils was higher in male mice in comparison to female mice (***Figure 7.18)***.

It is important to note, that although we saw some changes in myeloid populations in this experiment these were not reproducible in later experiments performed by Magnus Gwynne. The observation of female specific downregulation of IL-10 was similar between experiments.

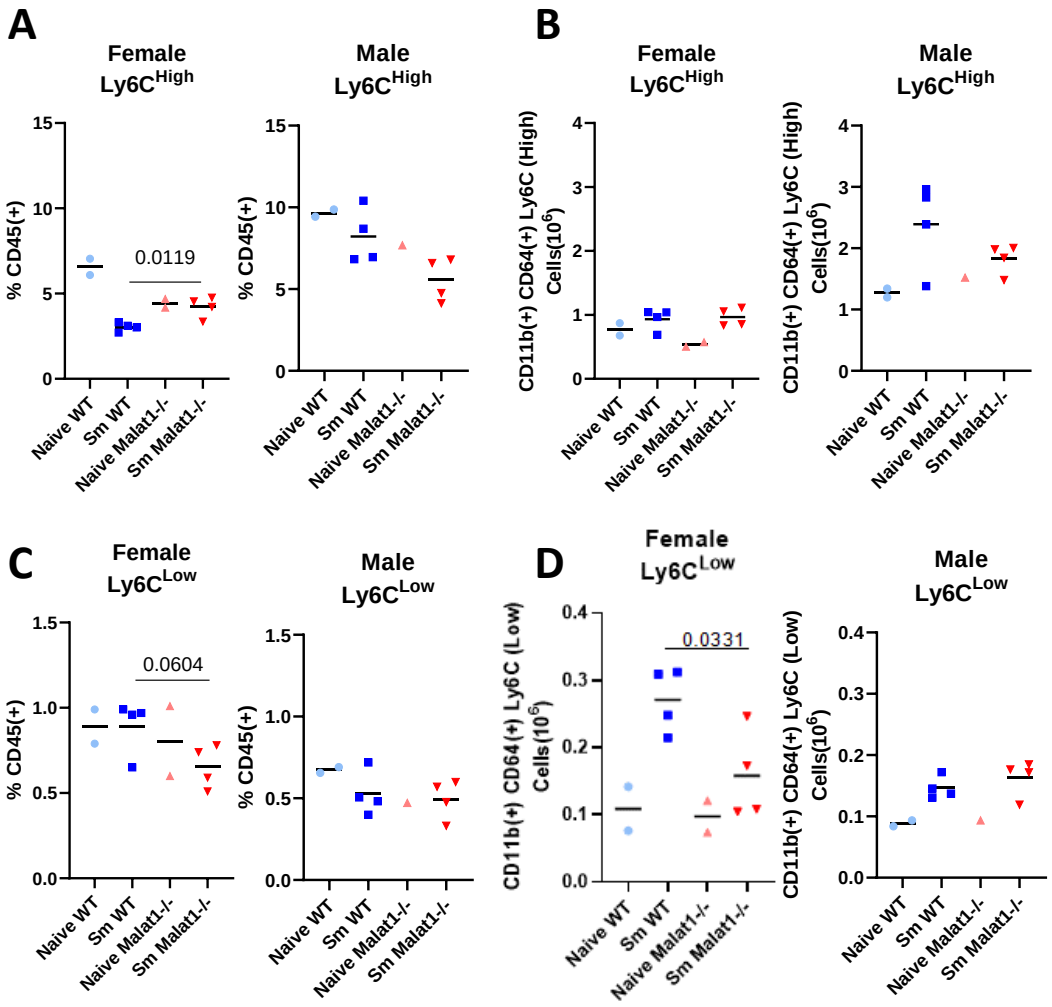


Figure 7.14: Loss of *Malat1* reduces the number of female Ly6C cells in an *S. mansoni* egg injection model.

A) Quantification of the percentage or **B)** number of CD45^{+} live Ly6C^{high} cells from *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice lungs. **C)** Quantification of the percentage or **D)** number of CD45^{+} live Ly6C^{low} cells from *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice lungs protein levels were determined by flow cytometry. Female data are shown on the left, male data are shown on the right of each panel. n=1/2 naïve, n=4 injected biological replicates representative of 1 independent experiment – male and female samples were processed on different days Data were analysed by one-way ANOVA with post hoc Tukeys. p values are displayed. The mean is depicted.

Lung Ly6C high

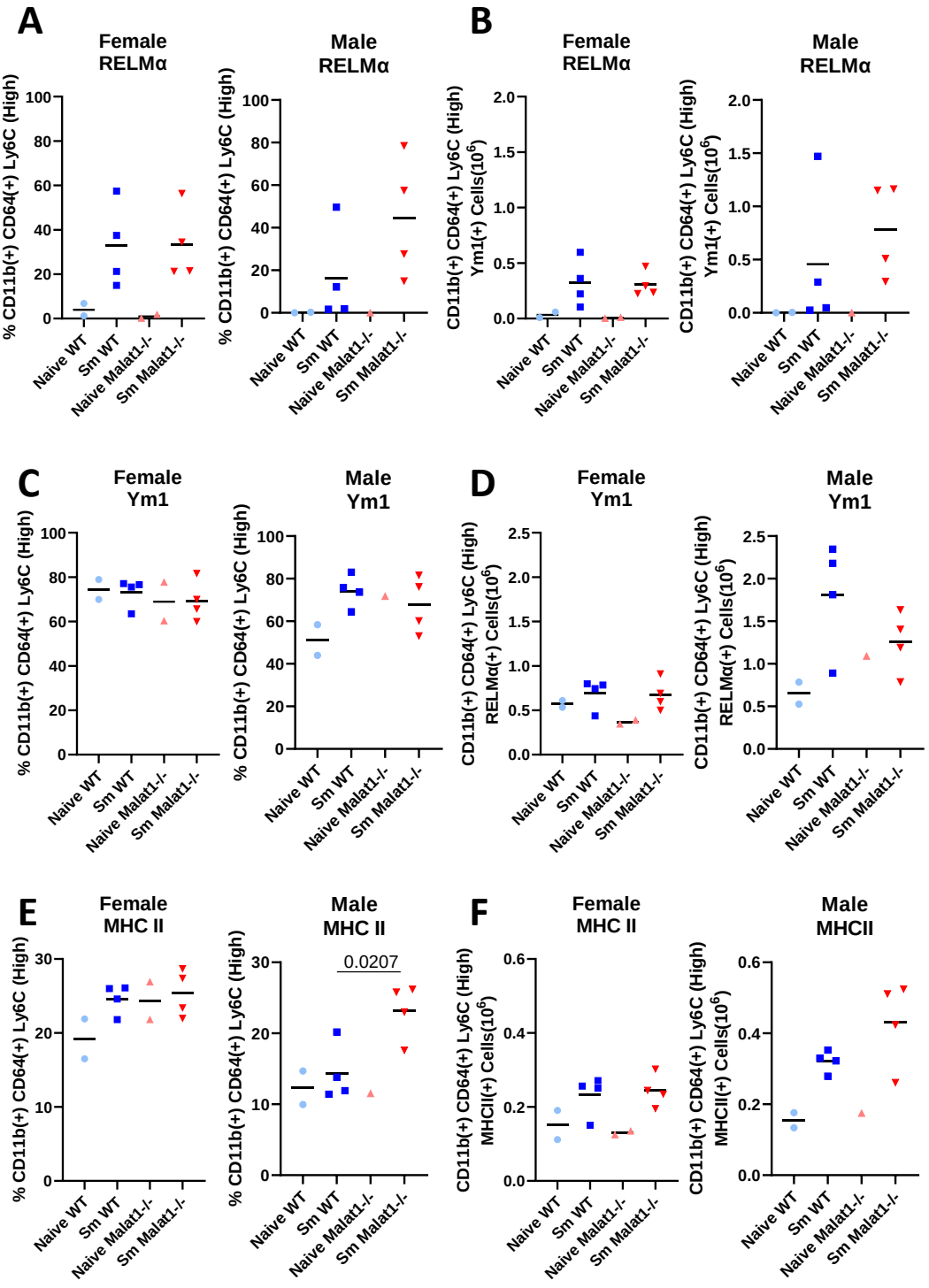


Figure 7.15: Loss of *Malat1* does not alter Ym1 or RELMα expression *in vivo* in Ly6C^{high} cells.

A) Quantification of the percentage or **B)** number of CD45^{+} live Ly6C^{high} Relmα^{+} cells from *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice lungs. **C)** Quantification of the percentage or **D)** number of CD45^{+} live Ly6C^{high} YM1^{+} cells from *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice lungs **E)** Quantification of the percentage or **F)** number of CD45^{+} live Ly6C^{high} MHCII^{+} cells from *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice lungs protein levels were determined by flow cytometry. Female data are shown on the left, male data are shown on the right of each panel. n=1/2 naïve, n=4 injected biological replicates representative of 1 independent experiment – male and female samples were processed on different days. Data were analysed by one-way ANOVA with post hoc Tukeys. p values are displayed where significant. The mean is depicted.

Lung Ly6C Low

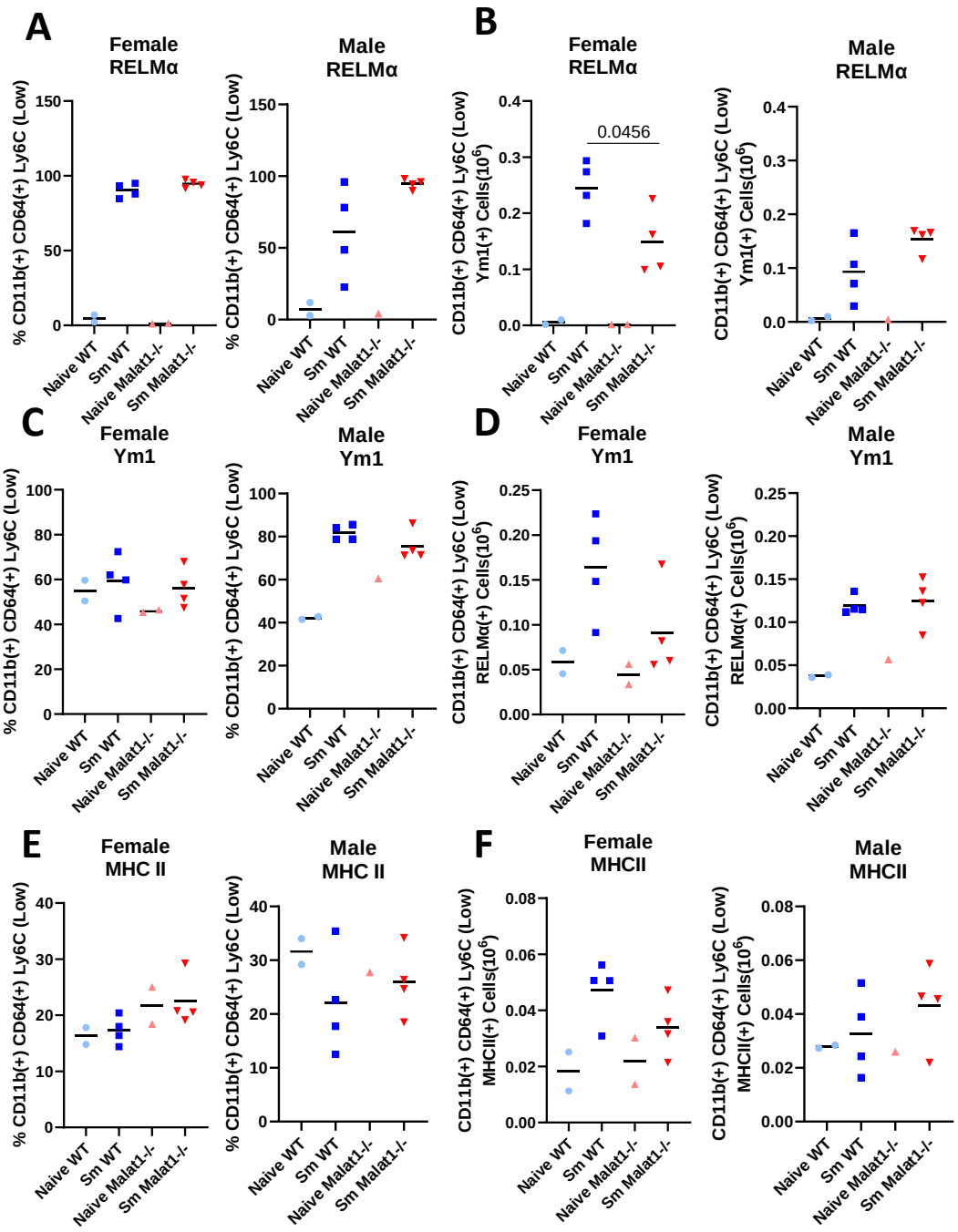


***Figure 7.16*: Loss of *Malat1* does not alter YM1 or RELMa expression in vivo in Ly6C^{low} cells.**

A) Quantification of the percentage or **B)** number of CD45^{+} live Ly6C^{low} Relmα^{+} cells from *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice lungs. **C)** Quantification of the percentage or **D)** number of CD45^{+} live Ly6C^{low} YM1^{+} cells from *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice lungs **E)** Quantification of the percentage or **F)** number of CD45^{+} live Ly6C^{low} MHCII^{+} cells from *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice lungs protein levels were determined by flow cytometry. Female data are shown on the left, male data are shown on the right of each panel. n=1/2 naïve, n=4 injected biological replicates representative of 1 independent experiment – male and female samples were processed on different days. Data were analysed by one-way ANOVA with post hoc Tukeys. p values are displayed where appropriate. The mean is depicted.

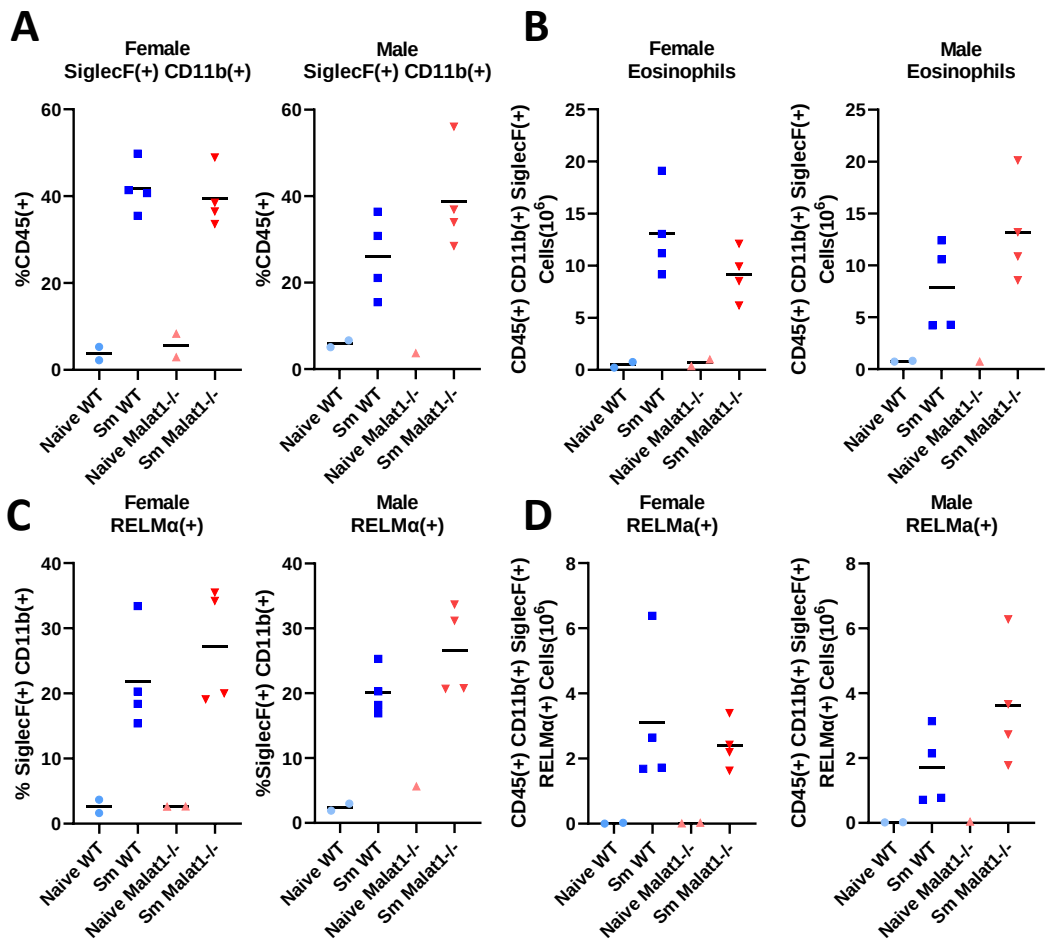


Figure 7.17: Loss of *Malat1* does not alter eosinophil populations *in vivo.*

A) Quantification of the percentage or **B)** number of CD45$^+$ live SiglecF$^+$, CD11b$^+$ cells from *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice lungs. **C)** Quantification of the percentage or **D)** number of CD45$^+$ live SiglecF$^+$, CD11b$^+$ Relmα$^+$ cells from *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice lungs protein levels were determined by flow cytometry. Female data are shown on the left, male data are shown on the right of each panel. n=1/2 naïve, n=4 injected biological replicates representative of 1 independent experiment – male and female samples were processed on different days. Data were analysed by one-way ANOVA with post hoc Tukeys. p values are displayed where significant. The mean is depicted.

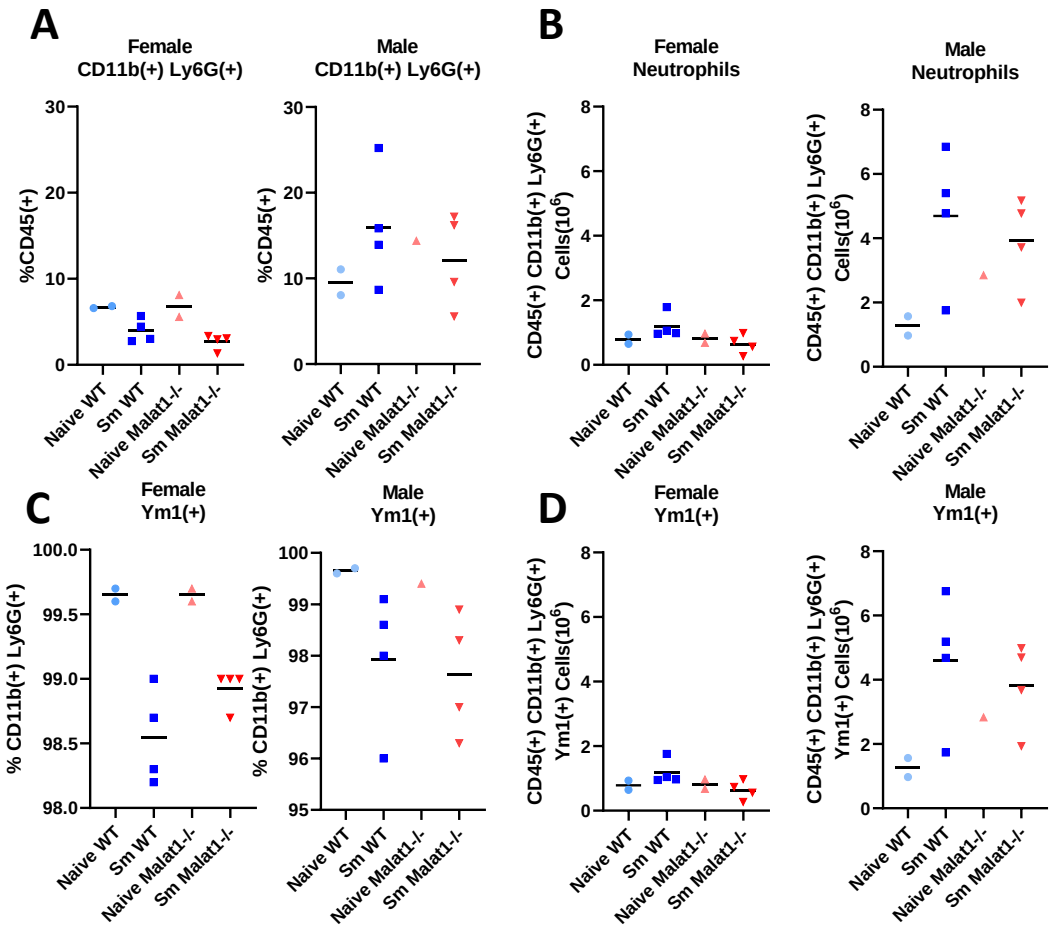


Figure 7.18: Loss of *Malat1* does not alter Neutrophil populations *in vivo.*

A) Quantification of the percentage or **B)** number of CD45$^+$ live CD11b$^+$, Ly6G$^+$ cells from *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice lungs. **C)** Quantification of the percentage or **D)** number of CD45$^+$ live, CD11b$^+$, Ly6G$^+$ Ym1$^+$ cells from *S. mansoni* egg-injected WT (blue) or *Malat1*$^{-/-}$ (red) mice lungs protein levels were determined by flow cytometry. Female data are shown on the left, male data are shown on the right of each panel. n=1/2 naïve, n=4 injected biological replicates representative of 1 independent experiment – male and female samples were processed on different days. Data were analysed by one-way ANOVA with post hoc Tukeys. p values are displayed where significant. The mean is depicted.

7.2.5 *Malat1* binds RBPs in sex-specific hierarchies

Given the above sex-specific *Malat1* effects on Th2 cells we speculated that *Malat1*-protein interactions might also differ between male and female cells. We began exploring this hypothesis by re-analysing our RAP-MS data presented in chapter 5. This identified that *Malat1* had preferential binding partners when comparing male and female naïve $CD4^+$ T cells (**Figure *7.19***). In female naïve $CD4^+$ T cells, *Malat1* showed the most prominent enrichment with SRSF1, SRSF10 and PUF60. In contrast in male naïve $CD4^+$ T cells, *Malat1* is most prominently enriched with MBNL1, ALDOA and TIAL1. When comparing trends in binding. *Malat1* tended to interact more strongly with the SR family of proteins in female cells and the hnRNP family of proteins in male cells. Several proteins were bound with similar enrichment in either male or female naïve $CD4^+$ T cells. This is another interesting finding which highlights the potential role of *Malat1* in regulating sex-specific RBP interactions in naïve $CD4^+$ T cells.

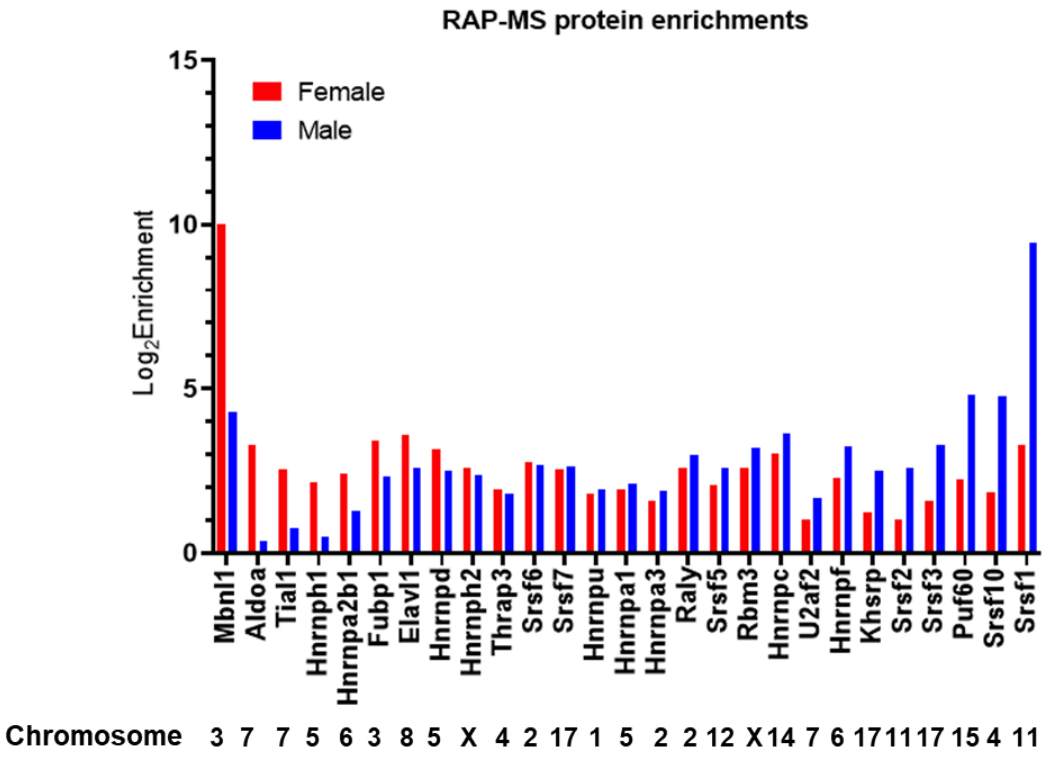


Figure 7.19: *Malat1* binds RBPs in sex-specific hierarchies.

Comparison of Log_2 fold change of cross-linked vs non-cross-linked samples in male and female WT naive $CD4^+$ T cells. Log_2 enrichment females are shown in red. Log_2 enrichment in males is shown in blue. Chromosome location is depicted below.

7.3 Discussion

In this chapter, we have shown for the first time that *Malat1* regulates $CD4^+$ T cell function in a sex-specific manner. This is an interesting finding as *Malat1* is not chromosome linked. *In vitro,* we found that *Malat1* is downregulated in both males and females after 24 hours of stimulation. However, this does not mean that the kinetics of the early stages of $CD4^+$ T cell activation are the same between males and females. Yet, after 6 days of stimulation *Malat1*$^{-/-}$ male $CD4^+$ T cell proliferation was impaired compared to WT and female cells. Further, loss of *Malat1* decreased IL-10 expression only in female $CD4^+$ T cells *in vivo* and *in vitro.* Some sex differences were also observed when comparing WT male and WT female $CD4^+$ T cells.

One possible explanation for the difference in IL-10 expression when comparing WT male and WT female *in vitro* polarised Th2 cells is that WT male cells have lower levels of *Malat1* compared to females at day 6 of *in vitro* $CD4^+$ T cell polarisation. This lower expression of *Malat1* could result in the reduced expression of IL-10 in WT male cells. However, the decrease in *Malat1* levels is only slight and other factors are likely playing a role. Alternatively, this could simply be a reflection of the stronger differentiation observed for female $CD4^+$ T cells. Overexpression of *Malat1* in WT male cells would help determine if *Malat1* levels are linked to IL-10 expression in a sex-specific manner.

The differences in IL-10 expression when comparing *Malat1*$^{-/-}$ male and female cells could be explained by changes in baseline IL-10 expression in WT male and female $CD4^+$ T cells. IL-10 levels were lower in WT male cells compared to WT female $CD4^+$ T cells. We have previously shown that under reduced polarising conditions loss of *Malat1* did not impact IL-10 expression (Hewitson et al., 2020). It

is also a possibility that WT male $CD4^+$ T cells may respond better to different concentrations of polarising antibodies and cytokines, which could be masking the functional role of *Malat1* in regulating IL-10 in male cells. It would be beneficial to repeat these experiments and titrate the concentration of cytokines and antibodies.

This reduced polarising effect in male $CD4^+$ T cells was also observed in the *S. mansoni* egg injection model *in vivo*. Expression of IL-10, IL-4 and IFNγ in $CD4^+$ T cells was lower in WT males than females in both the spleen and lungs. This difference in $CD4^+$ T cell responses in males and females could impact susceptibility to infection. This is reflected in the observation that males are more frequently re-infected with *Schistosoma spp* after praziquantel treatment than females (Trienekens et al., 2020). At a more global level, the intensity and prevalence of infections caused by parasites are higher in male mammals (Klein, 2004).

One benefit of reproducing our findings *in vitro* is that the experiment takes place in the absence of hormones during culture, thus hormones would have to be artificially introduced to influence IL-10 expression. This suggests that *Malat1* may act independently of hormones to influence the expression of IL-10. Additional experiments in which male and female cells are stimulated with oestrogen or testosterone would expand our understanding of the sex-specific effects of *Malat1*. Additionally, *Malat1* levels have not been tested in cells derived from our *in vivo* experiments to see if the expression of *Malat1* was different.

After 24 hours of stimulation under Th2 polarising conditions, we observed that *Malat1* is rapidly downregulated in WT male and WT female cells. Additionally, we showed that loss of *Malat1* reduced the mRNA levels of *Cd69* in both males and females. In female cells, we have previously shown that loss of *Malat1* results in a

differentiation defect (chapter 4). We postulated that in female cells rapid downregulation of *Malat1* is required to prevent overactivation of $CD4^+$ T cells ensuring that activation only takes place at the correct time. Despite similar responses to the loss of *Malat1* at the early stages of activation, IL-10 is not regulated in the same manner in males and females. Perhaps changes at later stages of $CD4^+$ T cell activation are responsible for these differences. As a better reflection of $CD4^+$ T cell activation, CD69 protein levels should be examined, as mRNA levels do not always correlate with protein expression (Jurgens et al., 2021). In addition, other makers such as CD25 could be examined to determine differences in T cell activation. This would expand our understanding of the role of *Malat1* in T cell activation.

We also observed a sex bias for the function of *Malat1* in $CD4^+$ T cell proliferation. *In vitro,* we found that loss of *Malat1* impaired the proliferation of male $CD4^+$ T cells during Th2 differentiation. This effect was most pronounced after culture in IL-2 indicating potential differences in CD25 expression or function between male and female $Malat1^{-/-}$ cells. *Malat1* has been shown to regulate the proliferation of numerous cell types. For example, knockdown of *Malat1* impaired chondrocyte proliferation through AKT3, in addition to inhibiting ovarian cancer cell proliferation (Wang, et al., 2019). The mascRNA and long parent isoform of *Malat1* have been shown to promote proliferation in hepatocellular carcinoma cells (Xie et al., 2021). No obvious sex bias has been reported for the role of *Malat1* in proliferation. This finding was not recapitulated *in vivo* as no significant differences in the proportion of $CD4^+$ T cells or activated $CD4^+$ T cells were observed – however, this is not a direct measurement of T cell proliferation. This is perhaps an artefact of artificial stimulation *in vitro* and not a reflection of antigen-specific T-cell responses which may proliferate

better *in vivo.* Further investigation with repeats of both *in vitro* and *in vivo* experiments would be required.

In this chapter, we used an *S. mansoni* egg injection model to induce a potent Th2 response. In the worm infection of *S. mansoni* IL-10 has been shown to play an important role and act on antigen-presenting cells. When IL-10 is neutralised using antibodies, an increased number of co-stimulatory molecules were found, and enhanced Th1 activation from macrophages was observed (Flores Villanueva et al., 1993). Similarly, exogenous treatment with IL-10 suppressed pulmonary granuloma formation (Flores Villanueva et al., 1994). Interestingly, in IL-$10^{-/-}$ mice infected with *S. mansoni* minimal impact was seen on their Th2 cytokine profile, but had increased expression of IFNγ, IL-2 and TNFα derived from lymphoid cells. Additionally, the knockout mice had larger granulomas in the liver (Wynn et al., 1998).To gain further insight into the full impact of reduced IL-10 expression in this model, granuloma size and formation in the lungs could be examined using microscopy. Of note, although we observed some changes in myeloid populations this did not hold true in later repeat experiments by Magnus Gwynne. Little data has been reported on the impact of IL-10 on myeloid populations in this egg injection model as such further investigation is required.

A particularly interesting observation from this chapter is the differences in *Malat1* RBP interaction partners in male and female naive $CD4^+$ T cells. In chapter 5, we found that *Malat1* interacted with common and unique RBPs in EL4 cells and naïve $CD4^+$ T cells. This suggests that the *Malat1*-protein interactome has some cell type- or activation status-specificity. This is further supported by the observation of sex-specific hierarchies in *Malat1* binding proteins. It is plausible that the sex-specific interactions of *Malat1* are responsible for the sexual dimorphism in $CD4^+$ T cell

function. *Malat1* in male $CD4^+$ T cells interacts more prominently with members of the hnRNP family of proteins. In contrast, *Malat1* in female cells interacts more prominently with members of the SR family of proteins. If these differences in *Malat1* interactions are maintained throughout Th cell differentiation perhaps these could influence effector cell function.

Some sex differences have been reported for SR and hnRNP proteins. One study used genome editing to knockin a nuclear retention signal for SRSF1. Instead of causing embryonic lethality in mice as with SRSF knockouts, $SRSF1^{NRS/NRS}$ mice are smaller than WT, and male mice have immotile sperm (Haward et al., 2021). Additionally, differences in protein isoforms can also be observed in a sex-specific manner. Male and female mice have different ratios of hnRNPD isoforms. When female mice were injected with testosterone this was shown to increase the expression of the p42 and p37 isoforms and decrease the levels of p45 hnRNPD isoforms in the murine submaxillary gland (Sheflin & Spaulding, 2000). It would be interesting to explore this observation further and see if knocking-down these RBPs gave a similar phenotype to male and female *Malat1*$^{-/-}$ Th cells.

Levels of *Malat1* expression are lower in WT males than WT females after 6 days of T cell stimulation, it is possible that differences in the expression level of *Malat1* influence RBP binding levels. *Malat1* expression should be directly compared between naïve WT females and male cells to determine if *Malat1* is consistently lower in WT males. In chapter 5 we showed that protein expression was not directly correlated with *Malat1* binding. However, we did not explore if the expression of the same protein influenced the strength of binding. Perhaps baseline expression of the SR or hnRNP proteins differs between sexes this could be further investigated through western blotting or flow cytometry. Protein isoforms could also be examined by

western blotting. Splice variants, phosphorylation status and localisation of the proteins could also explain some of these differences.

The concept of personalised medicine is that unique aspects of our biology including our immune systems require tailored intervention for improved treatment of disease (Klein & Flanagan, 2016). This chapter has highlighted just one sexual dimorphism in immunity. It is possible that if the long-term goal of personalised medicine is for the most effective treatment of each individual, males and females may have different treatment plans in the future. One area which is particularly lacking in this field is understanding sexual dimorphism in transgender individuals. A limited number of studies have begun to explore this notion, for example, transgender men treated with testosterone increased leukocyte-endothelium interactions and increases the levels of E-selectin, IL-6 and TNFα (Iannantuoni et al., 2021). It was also noted that trans people have an increased risk of COVID-19 exposure and mortality. Societal factors likely contribute to this disparity, as transgender people are more likely to work in COVID-affected industries such as food service (Goldie & Chatterjee, 2021). Additional factors such as age, and ethnicity also have significant impacts on the immune system (Montecino-Rodriguez et al., 2013). Thus, several aspects should be considered in the development of precision medicine. Nevertheless, further understanding sexual dimorphism and the underlying mechanism is of great benefit for precision medicine in immunity.

In this chapter, we have focused on sexual dimorphism in Th2 responses. We have previously shown that *Malat1* regulates IL-10 *in vitro* in Th1 cells, and that loss of *Malat1* affects disease pathology in models of leishmaniasis and malaria (Hewitson et al., 2020). Sexual dimorphism has been reported for both of these diseases, Studies have found that females clear asymptomatic malaria infections faster than males

(Briggs et al., 2020). Additionally, pregnant women are at greater risk for malaria infection, severe disease and mortality (Desai et al., 2007). Further, there is a male bias and incidence of malaria in both school-aged children and adults (Briggs et al., 2020). Sex differences are also observed in *Leishmania* infections. Reports indicate that males have more cases of visceral leishmaniasis in Brazil, East Africa, India and Nepal (Lockard et al., 2019). It would be of great interest to see if the sexual dimorphism of *Malat1* is also observed under type 1 conditions.

Taken together this work presented in this chapter suggests that *Malat1* regulates $CD4^+$ T cell function in a sex-specific manner, potentially through distinct sex-specific interactions with RBPs. We anticipate that the sex bias in lncRNA function extends beyond that of *Malat1* and encompasses other cells of the immune system. Further understanding sex bias in lncRNA function is an exciting area of future research.

8. Concluding Discussion

8.1 Review of aims and summary of findings

This study aimed to understand the role of the lncRNA *Malat1* in $CD4^+$ T cells. A variety of techniques including flow cytometry, qRT-PCR, RAP-MS, RNAseq and iCLIP were used to address the aims of this thesis. In this chapter, I review the Thesis aims and major findings, and suggest potential avenues for further research.

The hypothesis addressed by this research was that: *Malat1* has a non-redundant role in the adaptive immune system through specific $CD4^+$ T cell functions.

Aim 1: Determine *Malat1* expression patterns in WT $CD4^+$ T cells

In chapter 3, we examined the expression levels of *Malat1* through qRT-PCR

Main findings:

- *Malat1* is rapidly downregulated within 24 hours of $CD4^+$ T cell activation
- *Malat1* suppression is sustained throughout *in vitro* Th cell activation
- *Malat1* is more highly expressed in Th2 cells compared to Th1 cells

Aim 2: Compare $CD4^+$ T cell differentiation in WT and *Malat1*$^{-/-}$ cells *in vitro*

In chapter 3, we explored Th cell differentiation *in vitro* using naïve $CD4^+$ T cells derived from WT or *Malat1*$^{-/-}$ mice by flow cytometry and qRT-PCR

Main findings

- Loss of *Malat1* reduced the expression of IL-10 in both Th1 and Th2 cells at the protein and RNA level
- Mechanistically, we found that *Malat1* regulated IL-10 by enhancing the expression of the key transcription factor MAF

Aim 3: Compare $CD4^+$ T cell responses in WT and *Malat1*$^{-/-}$ mice *in vivo*

In chapter 3 we used a mixed bone marrow chimera model followed by a type 1 or type 2 model of infection to examine the role of *Malat1 in vivo*

Main findings:

- When mice were infected with *L. donovani, Malat1*$^{-/-}$ CD4^{+} T cells produced lower levels of IL-10 than WT cells (type 1 model)
- When mice were infected with *S. mansoni Malat1*$^{-/-}$ CD4^{+} T cells produced lower levels of IL-10 than WT cells (type 2 model)

Aim 4: Characterise transcriptomic changes in WT and *Malat1*$^{-/-}$ *in vitro* polarised Th cells

In chapter 4, we used RNA sequencing to determine differentially expressed genes in the absence of *Malat1*

Main findings:

- Loss of *Malat1* in *in vitro* polarised Th2 cells resulted in the greatest number of differentially expressed genes compared to Th1 or naïve cells
- Loss of *Malat1* impaired Th cell differentiation
- Loss of *Malat1* prominently altered gene expression along chromosome 19 which suggested a potential *in cis* regulatory role of *Malat1*

Aim 5: Characterise *Malat1* protein interaction partners in a T cell line and naïve CD4^{+} T cells

In chapter 5, we used RAP-MS to identify *Malat1* protein interaction partners in EL4 cells and naïve CD4^{+} T cells

- *Malat1* prominently interacted with members of the hnRNP and SR family of proteins in EL4 cells
- *Malat1* mainly interacted with hnRNP and SR proteins in naïve $CD4^+$ T cells
- *Malat1* had core binding partners in EL4 cells and naïve $CD4^+$ T cells, however, cell-type and sex specific binding partners were also found.

Aim 6: Characterise SRSF1 and hnRNPA1 interactions in the presence and absence of *Malat1* in *in vitro* polarised Th2 cells.

In chapter 6, we used iCLIP to determine hnRNPA1 and SRSF1 interaction partners in *in vitro* polarised WT and *Malta1*$^{-/-}$Th2 cells

Main findings:

- SRSF1 and hnRNPA1 interacted with genes involved in T cell differentiation and function
- Loss of *Malat1* resulted in a general decrease in RNAs binding to SRSF1 this could be explained by changes in SRSF1 localisation in the absence of *Malat1*
- Genes that bind hnRNPA1 and SRSF1 undergo differential transcript usage in the absence of *Malat1*

Aim 7: Characterise Male and Female WT and *Malat1*$^{-/-}$ in vitro T cell differentiation

In chapter 7, we used flow cytometry and qRT-PCR to examine Th cell differentiation in WT and *Malat1*$^{-/-}$ male and female $CD4^+$ T cells

Main findings:

- We found that *Malat1* is rapidly downregulated in WT male and female $CD4^+$ T cells after stimulation
- *In vitro* loss of *Malat1* impaired male $CD4^+$ T cell proliferation

- Loss of *Malat1* reduced IL-10 expression specifically in female $CD4^+$ Th2 cells

Aim 8: Compare $CD4^+$ T cell phenotypes in models of infection between Male and Female WT and *Malat1*$^{-/-}$ mice

In chapter 7, we used an *S. mansoni* egg injection model to induce a potent Th2 response in the lungs

Main findings:

- Loss of *Malat1* impaired IL-10 expression specifically in female $CD4^+$ T cells
- The number and proportion of activated $CD4^+$ T cells were similar between males and females

Aim 9: Compare *Malat1* interaction partners in male and female $CD4^+$ T cells

In chapter 7, we re-examined the RAP-MS data set in a sex-specific manner.

Main findings:

- *Malat1* bound RBPs in sex-specific hierarchies
- *Malat1* in male $CD4^+$ T cells generally interacted more with hnRNP proteins
- *Malat1* in female $CD4^+$ T cells generally interacted more with SR proteins

In this thesis, we described the link between *Malat1* and the regulation of IL-10 in a $CD4^+$ T cell context. Previous studies have examined the relevance of *Malat1* in a Th cell context; however, they do not always confirm these findings *in vivo*. We are also the first to describe a sexual dimorphism for *Malat1* function in Th cells. Additionally, our study is the first to describe *Malat1* protein interaction partners not only in a T cell context but in naïve primary $CD4^+$ T cells. Prior studies have examined *Malat1* RBP interactions in other cell types (typically cell lines). Given that a large

degree of heterogeneity has been reported for *Malat1* binding partners (see chapter 5) defining the *Malat1* protein interactome in a Th cell context provided a unique insight into the cell-specific functions of *Malat1*. Following this, we determined that the interaction of *Malat1* with SRSF1 is required for some of its interaction by retaining SRSF1 in the nucleus.

8.2 Future work

8.2.1 $CD4^+$ T cell-specific knockout of *Malat1*

In chapter 3, we determined that loss of *Malat1* through either knockout or knockdown models reduced the expression of IL-10 both *in vitro* and *in vivo.* In some cases, it can be challenging to unequivocally attribute the phenotype from the genetic deletion of a lncRNA, as in some cases loss of adjacent regulatory elements is responsible. For example, one study revealed opposite effects when comparing large genetic deletions vs minimal disruption of the lncRNA *Haunt* DNA locus (Yin et al., 2015). We have supported our main findings of *Malat1*-dependent regulation of IL-10 using GapmeRs to knockdown *Malat1*. However, it would be beneficial to overexpress *Malat1* or rescue *Malat1* expression in $CD4^+$ T cells and analyse levels of IL-10 and MAF. This would further improve our confidence in the role of the *Malat1* transcript rather than locus in $CD4^+$ T cell function. Due to time constraints this experiment was not attempted as part of this project.

One limitation of this study is the use of a knockout mouse model which lacks expression of *Malat1* in every tissue. $CD4^+$ T cell-independent mechanisms may be responsible for the changes in IL-10 expression observed. A $CD4^+$ T cell-specific knockout model would be beneficial in confirming these findings. However, the mixed bone marrow chimera adds confidence to the haematopoietic function of *Malat1*.

Additionally, despite the ubiquitous expression of *Malat1*, a strong $CD4^+$ T cell-specific response is observed.

8.2.2 The role of *Malat1* in human $CD4^+$ T cells

Mice are commonly used to decipher biological phenomena. This is because many proteins are well conserved between mice and humans, and only ~300 genes have been shown to be unique to either mice or humans. Mice are particularly beneficial for biological studies as they can be genetically manipulated with relative ease. Despite mice being of great benefit for understanding immune responses, there are some key differences between human and mouse immune systems (Mestas & Hughes, 2004). Differences can be observed across both arms of the immune system and in both immune cell development and response to infection.

A significant difference observed when comparing mouse and human immune systems is the ratio of lymphocytes to neutrophils. Human blood is very neutrophil rich (50-70% neutrophils), contrastingly mouse blood comprises 75-90% lymphocytes and 10-25% neutrophils (Mestas & Hughes, 2004). Notably, several differences between human and mouse T cells have been observed. For example, T cell development in the thymus has been well-characterised in mouse models, which have described both positive and negative selection events (Kumar et al., 2018). It is unclear if the mechanism is exactly the same in humans as insights have mainly been derived from thymus transplantation studies for patients who have a rare disease known as DiGeorge syndrome that lack a functional thymus (Davies et al., 2017). Murine studies have shown that thymic epithelial cells (TEC) are required for positive selection. However, human thymic transplant patients have TECs that are derived from donor cells, yet they produce T cells which can respond to host antigen cells – this suggested that human thymocyte selection may be more permissive than that in mice (Li et al.,

2011). Other differences have also been observed between mouse and human T cells. For example, naïve $CD4^+$ T cells derived from mice have a life span of ~6-10 weeks (Braber et al., 2012). In contrast, naïve CD4+ T cells derived from humans have a life span of between 5 and 10 years (Vrisekoop et al., 2008).

Additionally, some lncRNAs have been shown to act differently in mice and humans. For example, the lncRNA *Fast* is conserved both in genomic location and sequence between mice and humans. However, differences have been observed both for *Fast* localisation and function when comparing species. *Fast* is localised in the cytoplasm in human embryonic stem cells (ESC). In human cells, *Fast* activates WNT signalling by blocking interactions of β-catenin and β-TrCP. (C. J. Guo et al., 2020). In contrast, murine *Fast* is located in the nucleus and does not interact with β-TrCP or alter WNT signalling. This study went on to show that the location of several conserved lncRNAs have different localisation in mouse, human and monkey ESCs which suggested that lncRNAs may act differently across species (Guo et al. 2020). Further investigation is required to fully understand the impact of this observation both in ESCs and in other contexts (C. J. Guo et al., 2020).

Given that there are key differences between mouse and human T cell immunity and that lncRNAs have previously been shown in some cases to act differently across mammalian species it would be beneficial to explore the role of *Malat1* in a human context. This would help determine if *Malat1* also regulates IL-10 in a sex-specific manner in humans.

One way we could examine the role of *Malat1* in a human context is isolate PBMCs from healthy donors and use GapmeRs to knockdown *Malat1* or try to use cells from a human challenge study. It can be challenging to study lncRNAs in

primary human cells for example due to a lack of access to patient samples. Additionally, it is challenging to recapitulate a complex physiological environment through *in vitro* cell culture methods. One method to circumvent this problem is the use of humanised mouse models. For example, a liver-specific humanised mouse model has been used to identify the function of the non-conserved human lncRNA *Linc01018* in the regulation of fatty acid oxidation in the liver (Ruan et al., 2020). Similarly, "human immune system" mice have been developed to study human cells *in vivo*. For example, these may be severe combined immunodeficiency (SCID) mutant mice, which are sub-lethally irradiated and reconstituted with human HSCs (Ruan et al., 2020). Similar models which lack mouse natural killer cells have been developed that enable better engraftment of human cells (Ruan et al., 2020). A model similar to this could be used to explore the function of human *Malat1 in vivo.*

8.2.3 Antigen-specific T cells

All T cells are antigen-specific, consequently, during infection antigen-specific T cells become activated, these cells represent a rare proportion of the total $CD4^+$ T cell population (<0.001% of the total T cell population for many antigens) (Leung et al., 2013). Anti-CD3/CD28 mitogenic stimulation is commonly used to mimic the activation, differentiation and expansion of Th cells *in vitro*. Although this model can be manipulated for example to mimic sub-optimal conditions (chapter 3), the reductionist approach to recapitulate $CD4^+$ T cell responses has limitations. For example, the isolation of naïve $CD4^+$ T cells for *in vitro* culture requires the destruction of lymphoid tissues to create a single-cell suspension. This would destroy the spatial relationship between $CD4^+$ T cells and APCs that occur *in vivo* and removes the cells from interactions with lymphoid stroma (Jenkins et al., 2001). It also does not reflect the movement of $CD4^+$ T cells from secondary lymphoid organs to the sites of

infection. Additionally, *in vitro* stimulation does not recapitulate the entire inflammatory response, as inflammatory mediators can influence the quality of T cell responses to antigens (Jenkins et al., 2001). The use of *in vivo* models coupled with intracellular cytokine staining can identify antigen-specific $CD4^+$ T cells. However, this approach is limited to the cytokine-producing capabilities of the cell.

One method of characterising antigen-specific T cells involves the use of MHC-peptide complexes known as tetramers. A tetramer consists of four peptide-loaded MHC II molecules which have been tetramerised using fluorochrome-labelled streptavidin. They can be used to isolate epitope-specific populations and be isolated using a cell sorter. These can also be used to activate $CD4^+$ T cells *in vitro* (Kurtulus & Hildeman, 2013).

OT-II TCR transgenic mouse models can also be used to study antigen-specific $CD4^+$ T cell responses. OT-II $CD4^+$ T cells express transgenic αβ-TCRs that are specific to chicken ovalbumin (OVA) (Leung et al., 2013). Upon OVA stimulation this induces antigen-specific responses.

One way in which our study could be expanded would be to recapitulate our findings using an antigen-specific $CD4^+$ T cell model system. For example, crossing *Malat1*$^{-/-}$ mice with OT II mice would determine if *Malat1* regulated IL-10 in an antigen-specific manner upon OVA challenge.

8.2.4 Therapeutic potential

Even though ~98% of the human genome is non-coding, the vast majority of current therapeutics target one of ~700 disease-linked proteins (Santos et al., 2017). An expanding number of diseases have been linked to ncRNA. *Malat1* is no exception and plays a role in a diverse range of diseases such as cancer and diabetes (Arun et al.,

2020). The therapeutic potential of targeting *Malat1* is being investigated. In our study, we used GapmeR technology to knockdown *Malat1* in primary $CD4^+$ T cells *in vitro* which reduced IL-10 expression. In a mouse mammary tumour virus PyMT model, the use of *Malat1* targeting GapmeRs reduced tumour growth and metastasis (Arun et al., 2016). Similar studies have used siRNAs to knockdown *Malat1* in other cancer types. Nanocomplex delivery of *Malat1* siRNA *in vivo* improved survival in a model of glioblastoma multiforme (S. Kim et al., 2018).

The discovery of small molecules that bind RNA and have suitable properties for use as a drug has been limited. Early discoveries of RNA binding small molecules had low selectivity and high toxicity. Recently, small molecules which bind *Xist* have been developed (Aguilar et al., 2022). The compound named X1, blocked X chromosome inactivation, suppressed H3K27 trimethylation and inhibited cell differentiation in a female-specific manner. Small molecules which target the 3' triple helix of *Malat1* have also been developed (Donlic et al., 2018). One study which used a 1,500 nt segment of *Malat1* which comprised the triple helix, showed this increased proliferation of colorectal cancer cells (Xu et al., 2011). Additionally, the deletion of this segment decreased the accumulation of the *Malat1* transcript (Brown et al., 2014). Thus, targeting this unique structure is a promising therapeutic target.

One possible extrapolation of our work would be to use *Malat1* targeting therapies to treat infectious diseases where $CD4^+$ T cell derived IL-10 can be a critical determinant of disease outcome. We found that loss of *Malat1* resulted in enhanced immunity and pathogen clearance in models of visceral leishmaniasis (Hewitson et al., 2020). However, targeting *Malat1* would have to provide greater benefit than existing treatments in terms of cost, ease of use, efficacity or administration. One challenge of this approach would be to the ability to deliver *Malat1* GapmeR specifically to $CD4^+$

T cells. Targeted drug delivery systems are being developed such as Aptamer drug conjugates (Gao et al., 2022). Yet, as *Malat1* knockout mice have no overt phenotype, targeting *Malat1* may not cause adverse effects on healthy tissues or other cell types as such cell-specific targeting may not be necessary. Nevertheless, the growing body of pre-clinical data indicates that further investigation of *Malat1* targeting therapeutics could be of great benefit, particularly for cancer treatments.

8.2.5 Other lncRNAs in CD4$^+$ T cells

In this thesis, we have demonstrated that *Malat1* has a non-redundant role in Th cell function. The role of lncRNAs in CD4$^+$ T cells is only beginning to be understood, with a limited number of RNAs functionality tested *in vivo* (West & Lagos, 2019). After just 24 hours of CD4$^+$ T cell stimulation, 120 lncRNAs are differentially expressed and are disturbed evenly across the genome (Hewitson et al., 2020). Perhaps further investigation of a wider panel of lncRNAs would reveal more functionally relevant lncRNAs.

To further profile the relevance of lncRNAs in CD4$^+$ T cells, it could be interesting to determine lncRNA-RBP profiles in Th cells for different lncRNAs. *Malat1* is expressed at a high copy number 5,000-7000 copies per cell, given its high expression this made RAP-MS in primary naïve CD4$^+$ T cells easier to develop. However, RAP-MS has previously been used to identify *Xist* RBP interactions, which are expressed in the region of 100's of copies per cell (McHugh et al., 2015). Additionally, a study examined *Linc-NmR* RBP interactions using RAP-MS, which is a transcript expressed at 10 copies per cell (Gandhi et al., 2020). This suggests that despite the high expression of *Malat1* it would be possible to determine other lncRNA-protein interactions in CD4$^+$ T cells.

One lncRNA which may be of particular interest to study alongside *Malat1* is *Neat1*. In this thesis, we determined that *Malat1* regulated gene expression across chromosome 19 and can affect the expression levels of *Neat1*. This is similar to previous studies which have determined that loss of *Malat1* affects the expression of *Neat1* (Zhang et al., 2012). Additionally, *Neat1* has been shown to play a role in Th cell differentiation as over-expression of *Neat1* increased Th2 activation in human $CD4^+$ T cells and Th17 differentiation (Huang et al. 2021; Karimi et al. 2022). Additionally, both *Neat1* and *Malat1* have been found to be bound to active chromatin sites and co-localise across many loci but with unique binding patterns. Additionally, they have been found to bind to similar RBPs (West et al., 2014a)

Perhaps, *Malat1* and *Neat1* have some co-regulatory functions in $CD4^+$ T cells given they have both been shown to influence T cell function and bind similar genes. It may be interesting to study *Malat1* and *Neat1* together in $CD4^+$ T cells using co-knockdown experiments to see if this has a greater impact on T cell function.

8.2.6 Further profiling sexual dimorphism of *Malat1*

In this thesis, we explored the role of *Malat1* in male and female Th cells and through preliminary experiments determined that *Malat1* regulated IL-10 in a sex-specific manner. One benefit of our study is that some of these experiments took place *in vitro* and thus in the absence of hormones suggesting that *Malat1* may act independently of hormones to regulate IL-10. To further explore this finding, it would be interesting to alter the hormone levels of mice and see if this impacted the role of *Malat1* on IL-10 expression.

In our studies, we tended to find that T cell polarisations were weaker in males than females which may have masked the regulatory role of *Malat1* (Chapter 7). As

discussed in chapter 7, studies have reported conflicting information on the function of $CD4^+$ T cells in males and females. One study which castrated mice resulted in an upregulation of IFN-γ and the proportion of Th1 cells (Kissick et al., 2014; Massa et al., 2017). It would be interesting to see if castration of male mice enhanced Th1 differentiation and reduced IL-10 expression in *Malat1*$^{-/-}$ mice compared to WT mice. However, androgens have also been shown to enhance IL-10 production from $CD4^+$ T cells (Liva & Voskuhl, 2001). Perhaps, treatment of cells with testosterone would enhance the levels of IL-10 in WT male cells to a level where we are able to see differences between WT and *Malat1*$^{-/-}$ $CD4^+$ T cells. Similarly, the Th1/Th2 balance has been shown to be influenced by estrogen changes throughout pregnancy and the menstrual cycle (Eames et al., 2016). Cells could also be treated with 17β-estradiol to explore the possibility that estrogen treatment may make male cells susceptible to the loss of *Malat1*. This could be complemented with experiments that use cells from ovariectomised female mice to see if this alters the role of *Malat1* in female cells.

Other important questions remain – why does *Malat1* interact with different proteins in male and female naïve $CD4^+$ T cells? It would be interesting to expand this to differentiated Th cells to see if these differences remain the same Additional experiments, to examine the total proteome and phosphoproteomics, which would enable us to address the question of whether the expression of the phosphorylation status of *Malat1* interacting proteins influences protein binding in male and female $CD4^+$ T cells.

8.3 Concluding remarks

In summary, *Malat1* is one of the most abundant transcripts in mammals, yet its physiological relevance at the whole organism level remains poorly understood. In this thesis, we have shown that *Malat1* has a non-redundant role in $CD4^+$ T cells. In

female mice, we have characterised that *Malat1* downregulation is a hallmark of $CD4^+$ T cell activation. Complete deletion of *Malat1* results in reduced expression of the anti-inflammatory cytokine IL-10 (regulated by MAF). This resulted in enhanced immunity to infection in models of visceral leishmaniasis and malaria and reduced IL-10 expression in mixed bone marrow chimera models. We found that loss of *Malat1* resulted in transcriptome-wide effects which reflected impaired Th cell differentiation. RAP-MS revealed that *Malat1* interacted with members of the hnRNP and SR family of proteins in a sex-specific manner. We selected SRSF1 and hnRNPA1 as representative members of each family and performed iCLIP. We found that loss of *Malat1* results in a general decrease in RNAs bound by SRSF1 and that *Malat1* is required to retain SRSF1 in the nucleus. Our results show that *Malat1* is an essential orchestrator of RBP function during highly dynamic cellular transitions, with relevance to human health and disease.

Acronyms

Acronym	Abbreviation
ACE2	Angiotensin-converting enzyme 2
AIREs	Adenylate uridylate-rich elements
ALS	Amyotrophic lateral sclerosis
APC	Antigen-presenting cell
AREs	Adenylate uridylate-rich elements
AS	Alternative splicing
ASO	Antisense oligonucleotides
AVMC	Acute viral myocarditis
BCR	B cell receptor
Bcra4	Breast cancer anti-oestrogen resistance 4
BLIMP1	B lymphocyte-induced maturation protein 1
CCR2	C-C chemokine receptor 2
CD	Cluster of differentiation
ceRNA	Competing endogenous RNA
Cerox1	Cytoplasmic endogenous regulator of oxidative phosphorylation 1
CHART	Capture hybridization analysis of RNA targets
ChIP	Chromatin immunoprecipitation
ChIRP	Comprehensive identification of RNA binding proteins by mass spectrometry
CISF	Cytokine synthesis inhibitory factor
CIZ1	Cip1-interacting zinc finger protein
CLIP	Crosslinking immunoprecipitation
ConA	Concanvalin A
COVID-19	Coronavirus disease 2019
Ct	Cycle thresholds
CTL	Cytotoxic lymphocytes

DC	Dendritic cell
DDIT4	DNA-damage inducible transcript 4
DDX3X	DEAD-Box Helicase 3 X-Linked
DDX3Y	DEAD-Box Helicase 3 Y Linked
DHT	Dihydrotestosterone
DLBCL	Diffuse large B-cell lymphoma
DRIP	DNA-RNA immunoprecipitation
DYNLT1B	Dynein light chain Tctex-type 1b
EGFR	Epidermal growth factor receptor
Elavl1/HuR	ELAV like RNA binding protein 1
EMSA	Electrophoretic mobility shift assay
ETS-1	E26 transformation-specific sequence 1
EV	Extracellular vesicle
EZH2	Enhancer of zeste homolog 2
FcR	Fc Receptor
Firre	Functional intergenic repeating RNA element
FTD	Frontotemporal dementia
GATA3	GATA Binding Protein 3
GM-CSF	Granulocyte macrophage colony-stimulating factor
GSEA	Gene set enrichment analysis
HIV-1	Human immunodeficiency virus 1
hnRNP	Heterogeneous nuclear ribonuclear proteins
Hotair	HOX transcript antisense RNA
HSC	Haematopoietic stem cells
HT	Hashimoto's thyroiditis
Hulc	Highly upregulated in liver cancer
HUVEC	Human umbilical vein endothelial cells

iCLIP	Individual-nucleotide resolution UV crosslinking and immunoprecipitation
iDRIP	Identification of direct RNA interacting proteins
IFN	Interferon
iiCLIP	Improved iCLIP
IL-	Interleukin
ImmPRes	Immunological Proteomic Resource
INPP4B	Inositol polyphosphate 4-phosphatase type II
Ips1	Induced by phosphate starvation 1
IQGAP	IQ motif containing GTPase activating protein
IVT	In vitro transcribed
KO	Knockout
KSHV	Kaposi sarcoma-associated -herpesvirus
LCMV	Lymphocytic choriomeningitis virus
lincRNA	Long intergenic non-coding RNA
LLPS	Liquid liquid phase separation
lncRNAs	Long non-coding RNA
LPS	Lipopolysaccharide
MAF	Avian musculoaponeurotic fibrosarcoma oncogene homolog
Malat1	Metastasis associated lung adenocarcinoma transcript 1
MALT1	Mucosa-associated lymphoid tissue lymphoma translocation protein 1
MCL	Mantle cell lymphoma
METTL16	Methyl transferase like 16
MHC	Major histocompatibility complex
miRNA	microRNAs
MS	Mass spectrometry
mTORC1	Mechanistic target of rapamycin complex 1

Mt-RNA	Mitochondrial ribosomal RNA
ncRNA	Non-coding RNA
Neat1	Nuclear paraspeckle assembly transcript 1
Neat2	Nuclear enriched abundant transcript 2
NeST	Nettoie Salmonella pas Theiler's (cleanup Salmonella not Theiler's)
NET	Neutrophil extracellular trap
NF-?B	Nuclear factor kappa light chain enhancer of activated B cells
NFAT	Nuclear factor of activated T cells
NK	Natural killer
Nron	Nuclear repressor of NFAT
NSCLC	Non-small cell lung cancer
OOPs	Orthogonal organic phase separation
PAIR	PNA assisted identification of RBPs
PAMP	Pathogen-associated molecular patterns
PAN	Polyadenylated nuclear
PBS	Phosphate buffered saline
PCBP1	Poly (C) binding protein 1
PD-L1	Programmed death-ligand 1
PDX	Patient derived xenograft
PEKA	Positionally enriched k-mer analysis
PLA	Proximity Ligation Assay
PNA	Peptide nucleic acid
PRC2	Polycomb repressive complex
PRR	Pattern recognition receptors
PTBP1	Polypyrimidine tract-binding protein 1
RAG	Recombinase activating gene
RAP-MS	RNA antisense purification followed by mass spectrometry
RBD	RNA binding domain

RBFOX2	RNA binding protein fox -1, homolog 2
RBP	RNA binding protein
RIC	RNA interactome capture
RIP	RNA immunoprecipitation
RISC	RNA induced silencing complex
Rn7sk	RNA Component Of 7SK Nuclear Ribonucleoprotein
RNA-seq	RNA sequencing
RORγT	RAR-related orphan receptor gamma t
RPM	Rotation per minute
RRM	RNA recognition motif
rRNA	Ribosomal RNA
Sammson	Survival Associated Mitochondrial Melanoma Specific Oncogenic Non-Coding RNA
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SCYL1	SCY1-like pseudokinase 1
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
shRNA	Short hairpin RNA
SILAC	Stable isotope labelling of amino acids in cell culture
siRNA	Small interfering RNA
SLE	Systemic lupus erythematosus
SnoRNA	Small nucleolar
snRNA	Small nuclear
snRNP	Small nuclear ribonuclear proteins
SR	Serine/arginine rich
SRSF1	Serine and arginine-rich splicing factor 1
SRSF2	Serine and arginine-rich splicing factor 2
SRSF3	Serine and arginine-rich splicing factor 3

STAT3	Signal transducer and activator of transcription 3
STAT4	Signal transducer and activator of transcription 4
STAT6	Signal transducer and activator of transcription
Stau1	Stauffen1
SUZ12	Suppressor of zeste 12 protein homolog
Tarid	TCF21 antisense RNA inducing demethylation
T-bet	T-box transcription factor
TCR	T cell receptor
TDP-43	TAR DNA binding protein 43
TdT	Terminal deoxynucleotidyl transferase
TGF	Transforming growth factor
Th	T helper
Tmem181b-ps	Transmembrane protein 181b pseudo gene
TMPRSS2	transmembrane serine protease 2
TRIP	Tandem RNA isolation procedure
tRNA	Transfer ribonucleic acid
TSS	Transcription start sites
TTS	Transcriptional termination sites
UV	Ultraviolet
WDR5	WD repeat-containing protein 5
XCI	X chromosome inactivation
x g	Gravity
Xist	X-inactive specific transcript

References

Agostini, F., Zagalak, J., Attig, J., Ule, J., & Luscombe, N. M. (2021). Intergenic RNA mainly derives from nascent transcripts of known genes. *Genome Biology*, *22*(1). https://doi.org/10.1186/s13059-021-02350-x

Aguilar, R., Spencer, K. B., Kesner, B., Rizvi, N. F., Badmalia, M. D., Mrozowich, T., Mortison, J. D., Rivera, C., Smith, G. F., Burchard, J., Dandliker, P. J., Patel, T. R., Nickbarg, E. B., & Lee, J. T. (2022). Targeting Xist with compounds that disrupt RNA structure and X inactivation. | *Nature* |, *604*. https://doi.org/10.1038/s41586-022-04537-z

Akerman, M., Fregoso, O. I., Das, S., Ruse, C., Jensen, M. A., Pappin, D. J., Zhang, M. Q., & Krainer, A. R. (2015). Differential connectivity of splicing activators and repressors to the human spliceosome. *Genome Biology*, *16*(1). https://doi.org/10.1186/s13059-015-0682-5

Amodio, N., Raimondi, L., Juli, G., Stamato, M. A., Caracciolo, D., Tagliaferri, P., & Tassone, P. (2018). MALAT1: a druggable long non-coding RNA for targeted anti-cancer approaches. *Journal of Hematology & Oncology 2018 11:1*, *11*(1), 1–19. https://doi.org/10.1186/S13045-018-0606-4

Anderson, C. F., Oukka, M., Kuchroo, V. J., & Sacks, D. (2007). CD4+CD25−Foxp3− Th1 cells are the source of IL-10–mediated immune suppression in chronic cutaneous leishmaniasis. *Journal of Experimental Medicine*, *204*(2), 285–297. https://doi.org/10.1084/JEM.20061886

Ansa-Addo, E. A., Huang, H. C., Riesenberg, B., Iamsawat, S., Borucki, D., Nelson, M. H., Nam, J. H., Chung, D., Liu, B., Li, Z., Paulos, C. M., Yu, X. Z., Philpott, C., & Howe, P. H. (2020). RNA binding protein pcbp1 is an intracellular immune checkpoint

for shaping t cell responses in cancer immunity. *Science Advances*, *6*(22), 3865. https://doi.org/10.1126/sciadv.aaz3865

Arab, K., Park, Y. J., Lindroth, A. M., Schäfer, A., Oakes, C., Weichenhan, D., Lukanova, A., Lundin, E., Risch, A., Meister, M., Dienemann, H., Dyckhoff, G., Herold-Mende, C., Grummt, I., Niehrs, C., & Plass, C. (2014). Long noncoding RNA TARID directs demethylation and activation of the tumor suppressor TCF21 via GADD45A. *Molecular Cell*, *55*(4), 604–614. https://doi.org/10.1016/j.molcel.2014.06.031

Araújo, M. I., Hoppe, B. S., Medeiros, M., & Carvalho, E. M. (2004). Schistosoma mansoni infection modulates the immune response against allergic and auto-immune diseases. *Memórias Do Instituto Oswaldo Cruz*, *99*(5), 27–32. https://doi.org/10.1590/S0074-02762004000900005

Arun, G., Aggarwal, D., & Spector, D. L. (2020). MALAT1 long non-coding RNA: Functional implications. In *Non-coding RNA* (Vol. 6, Issue 2, p. 22). MDPI AG. https://doi.org/10.3390/NCRNA6020022

Arun, G., Diermeier, S., Akerman, M., Chang, K.-C., Wilkinson, J. E., Hearn, S., Kim, Y., MacLeod, A. R., Krainer, A. R., Norton, L., Brogi, E., Egeblad, M., & Spector, D. L. (2016). Differentiation of mammary tumors and reduction in metastasis upon *Malat1* lncRNA loss. *Genes & Development*, *30*(1), 34–51. https://doi.org/10.1101/gad.270959.115

Arun, G., Diermeier, S. D., & Spector, D. L. (2018). Therapeutic Targeting of Long Non-Coding RNAs in Cancer. *Trends in Molecular Medicine*, *24*(3), 257–277. https://doi.org/10.1016/j.molmed.2018.01.001

Auweter, S. D., Oberstrass, F. C., & Allain, F. H. T. (2006). Sequence-specific binding of single-stranded RNA: is there a code for recognition? *Nucleic Acids Research*, *34*(17), 4943. https://doi.org/10.1093/NAR/GKL620

Aznaourova, M., Janga, H., Sefried, S., Kaufmann, A., Dorna, J., Volkers, S. M., Georg, P., Lechner, M., Hoppe, J., Dökel, S., Schmerer, N., Gruber, A. D., Linne, U., Bauer, S., Sander, L. E., Schmeck, B., & Schulte, L. N. (2020). Noncoding RNA MaIL1 is an integral component of the TLR4–TRIF pathway. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(16). https://doi.org/10.1073/pnas.1920393117

Bandiera, R., Wagner, R. E., Britto-Borges, T., Dieterich, C., Dietmann, S., Bornelöv, S., & Frye, M. (2021). RN7SK small nuclear RNA controls bidirectional transcription of highly expressed gene pairs in skin. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-26083-4

Barra, J., & Leucci, E. (2017). Probing Long Non-coding RNA-Protein Interactions. *Frontiers in Molecular Biosciences*, *4*, 45. https://doi.org/10.3389/fmolb.2017.00045

Biagioni, A., Tavakol, S., Ahmadirad, N., Zahmatkeshan, M., Magnelli, L., Mandegary, A., Samareh Fekri, H., Asadi, M. H., Mohammadinejad, R., & Ahn, K. S. (2021). Small nucleolar RNA host genes promoting epithelial–mesenchymal transition lead cancer progression and metastasis. In *IUBMB Life* (Vol. 73, Issue 6). https://doi.org/10.1002/iub.2501

Blanchette, M., & Chabot, B. (1999). Modulation of exon skipping by high-affinity hnRNP A1-binding sites and by intron elements that repress splice site utilization. *EMBO Journal*, *18*(7). https://doi.org/10.1093/emboj/18.7.1939

Bond, C. S., & Fox, A. H. (2009). Paraspeckles: Nuclear bodies built on long noncoding RNA. In *Journal of Cell Biology* (Vol. 186, Issue 5, pp. 637–644). The Rockefeller University Press. https://doi.org/10.1083/jcb.200906113

Bouman, A., Jan Heineman, M., & Faas, M. M. (2005). Sex hormones and the immune response in humans. In *Human Reproduction Update* (Vol. 11, Issue 4, pp. 411–423). Oxford Academic. https://doi.org/10.1093/humupd/dmi008

Brajic, A., Franckaert, D., Burton, O., Bornschein, S., Calvanese, A. L., Demeyer, S., Cools, J., Dooley, J., Schlenner, S., & Liston, A. (2018). The Long Non-coding RNA Flatr Anticipates Foxp3 Expression in Regulatory T Cells. *Frontiers in Immunology*, *9*, 1989. https://doi.org/10.3389/fimmu.2018.01989

Briggs, J., Teyssier, N., Nankabirwa, J. I., Rek, J., Jagannathan, P., Arinaitwe, E., Bousema, T., Drakeley, C., Murray, M., Crawford, E., Hathaway, N., Staedke, S. G., Smith, D., Rosenthal, P. J., Kamya, M., Dorsey, G., Rodriguez-Barraquer, I., & Greenhouse, B. (2020). *Sex-based differences in clearance of chronic Plasmodium falciparum infection*. https://doi.org/10.7554/eLife.59872

Brockdorff, N. (2018). Local Tandem Repeat Expansion in Xist RNA as a Model for the Functionalisation of ncRNA. *Non-Coding RNA*, *4*(4). https://doi.org/10.3390/ncrna4040028

Brockdorff, N., Ashworth, A., Kay, G. F., McCabe, V. M., Norris, D. P., Cooper, P. J., Swift, S., & Rastan, S. (1992). The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell*, *71*(3), 515–526.

Brockdorff, N., & Turner, B. M. (2015). Dosage compensation in mammals. *Cold Spring Harbor Perspectives in Biology*, *7*(3), a019406. https://doi.org/10.1101/cshperspect.a019406

Brown, J. A., Bulkley, D., Wang, J., Valenstein, M. L., Yario, T. A., Steitz, T. A., & Steitz, J. A. (2014). Structural insights into the stabilization of MALAT1 noncoding RNA by a bipartite triple helix. *Nature Structural & Molecular Biology*, *21*(7), 633–640. https://doi.org/10.1038/nsmb.2844

Brown, J. A., Kinzig, C. G., Degregorio, S. J., & Steitz, J. A. (2016a). Methyltransferase-like protein 16 binds the 3'-terminal triple helix of MALAT1 long noncoding RNA. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1614759113

Brown, J. A., Kinzig, C. G., Degregorio, S. J., & Steitz, J. A. (2016b). Methyltransferase-like protein 16 binds the 3'-terminal triple helix of MALAT1 long noncoding RNA. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(49), 14013–14018. https://doi.org/10.1073/pnas.1614759113

Brown, J. A., Valenstein, M. L., Yario, T. A., Tycowski, K. T., & Steitz, J. A. (2012). Formation of triple-helical structures by the 3′-end sequences of MALAT1 and MENβ noncoding RNAs. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(47). https://doi.org/10.1073/pnas.1217338109

Bruun, G. H., Doktor, T. K., Borch-Jensen, J., Masuda, A., Krainer, A. R., Ohno, K., & Andresen, B. S. (2016). Global identification of hnRNP A1 binding sites for SSO-based splicing modulation. *BMC Biology*, *14*(1). https://doi.org/10.1186/s12915-016-0279-9

Bunders, M., & Altfeld, M. (2020). Implications of sex differences in immunity for SARS-CoV-2 pathogenesis and design of therapeutic interventions. *Immunity*, *0*(0). https://doi.org/10.1016/j.immuni.2020.08.003

Butler, N., Lamb, T., Stephens, R., Kumar, R., Engwerda, C., & Ng, S. (2019). The Role of IL-10 in Malaria: A Double Edged Sword. *Frontiers in Immunology | Www.Frontiersin.Org*, *10*, 229. https://doi.org/10.3389/fimmu.2019.00229

Byrne, A., Beaudin, A. E., Olsen, H. E., Jain, M., Cole, C., Palmer, T., DuBois, R. M., Forsberg, E. C., Akeson, M., & Vollmers, C. (2017). Nanopore long-read RNAseq reveals widespread transcriptional variation among the surface receptors of individual B cells. *Nature Communications*, *8*. https://doi.org/10.1038/ncomms16027

Cabrera-Rodríguez, R., Pérez-Yanes, S., Montelongo, R., Lorenzo-Salazar, J. M., Estévez-Herrera, J., García-Luis, J., Íñigo-Campos, A., Rubio-Rodríguez, L. A., Muñoz-Barrera, A., Trujillo-González, R., Dorta-Guerra, R., Casado, C., Pernas, M., Blanco, J., Flores, C., & Valenzuela-Fernández, A. (2022). Transactive Response DNA-Binding Protein (TARDBP/TDP-43) Regulates Cell Permissivity to HIV-1 Infection by Acting on HDAC6. *International Journal of Molecular Sciences*, *23*(11), 6180. https://doi.org/10.3390/IJMS23116180

Cáceres, J. F., Screaton, G. R., & Krainer, A. R. (1998). A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm. *Genes and Development*. https://doi.org/10.1101/gad.12.1.55

Cai, L. J., Tu, L., Huang, X. M., Huang, J., Qiu, N., Xie, G. H., Liao, J. X., Du, W., Zhang, Y. Y., & Tian, J. Y. (2020). LncRNA MALAT1 facilitates inflammasome activation via epigenetic suppression of Nrf2 in Parkinson's disease. *Molecular Brain*. https://doi.org/10.1186/s13041-020-00656-8

Cao, W., Jamison, S. F., & Garcia-Blanco, M. A. (1997). Both phosphorylation and dephosphorylation of ASF/SF2 are required for pre-mRNA splicing in vitro. *RNA*, *3*(12).

Carlevaro-Fita, J., Rahim, A., Guigó, R., Vardy, L. A., & Johnson, R. (2016). Cytoplasmic long noncoding RNAs are frequently bound to and degraded at ribosomes in human cells. *RNA*, *22*(6), 867–882. https://doi.org/10.1261/rna.053561.115

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B. M., Strein, C., Davey, N. E., Humphreys, D. T., Preiss, T., Steinmetz, L. M., Krijgsveld, J., & Hentze, M. W. (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell*, *149*(6), 1393–1406. https://doi.org/10.1016/J.CELL.2012.04.031

Castello, A., Frese, C. K., Fischer, B., Järvelin, A. I., Horos, R., Alleaume, A. M., Foehr, S., Curk, T., Krijgsveld, J., & Hentze, M. W. (2017). Identification of RNA-binding domains of RNA-binding proteins in cultured cells on a system-wide scale with RBDmap. *Nature Protocols 2017 12:12*, *12*(12), 2447–2464. https://doi.org/10.1038/nprot.2017.106

Chakrabarti, A. M., Capitanchik, C., Ule, J., & Luscombe, N. M. (2021). *clipplotr-a comparative visualisation and analysis tool for CLIP data*. https://doi.org/10.1101/2021.09.10.459763

Chaplin, D. D. (2010). Overview of the immune response. *Journal of Allergy and Clinical Immunology*, *125*(2 SUPPL. 2), S3. https://doi.org/10.1016/j.jaci.2009.12.980

Chaturvedi, P., Neelamraju, Y., Arif, W., Kalsotra, A., & Janga, S. C. (2015). Uncovering RNA binding proteins associated with age and gender during liver maturation. *Scientific Reports*, *5*. https://doi.org/10.1038/srep09512

Chen, J., Martindale, J. L., Abdelmohsen, K., Kumar, G., Fortina, P. M., Gorospe, M., Rostami, A., & Yu, S. (2020). RNA-Binding Protein HuR Promotes Th17 Cell Differentiation and Can Be Targeted to Reduce Autoimmune Neuroinflammation. *The Journal of Immunology*, *204*(8), 2076–2087. https://doi.org/10.4049/JIMMUNOL.1900769/-/DCSUPPLEMENTAL

Chen, J., Martindale, J. L., Cramer, C., Gorospe, M., Atasoy, U., Drew, P. D., & Yu, S. (2017). The RNA-binding protein HuR contributes to neuroinflammation by promoting C-C chemokine receptor 6 (CCR6) expression on Th17 cells. *The Journal of Biological Chemistry*, *292*(35), 14532–14543. https://doi.org/10.1074/JBC.M117.782771

Chen, Q., Su, Y., He, X., Zhao, W., Wu, C., Zhang, W., Si, X., Dong, B., Zhao, L., Gao, Y., Yang, X., Chen, J., Lu, J., Qiao, X., & Zhang, Y. (2016). Plasma long non-coding RNA MALAT1 is associated with distant metastasis in patients with epithelial ovarian cancer. *Oncology Letters*, *12*(2), 1361–1366. https://doi.org/10.3892/ol.2016.4800

Chen, R., Liu, Y., Zhuang, H., Yang, B., Hei, K., Xiao, M., Hou, C., Gao, H., Zhang, X., Jia, C., Li, L., Li, Y., & Zhang, N. (2017a). Quantitative proteomics reveals that long non-coding RNA MALAT1 interacts with DBC1 to regulate p53 acetylation. *Nucleic Acids Research*, *45*(17), 9947–9959. https://doi.org/10.1093/nar/gkx600

Chen, R., Liu, Y., Zhuang, H., Yang, B., Hei, K., Xiao, M., Hou, C., Gao, H., Zhang, X., Jia, C., Li, L., Li, Y., & Zhang, N. (2017b). Quantitative proteomics reveals that long non-coding RNA MALAT1 interacts with DBC1 to regulate p53 acetylation. *Nucleic Acids Research*, *45*(17), 9947–9959. https://doi.org/10.1093/nar/gkx600

Chen, Y. G., Satpathy, A. T., & Chang, H. Y. (2017). Gene regulation in the immune system by long noncoding RNAs. *Nature Immunology*, *18*(9), 962–972. https://doi.org/10.1038/ni.3771

Cheukfai, L., Wei, Z., Yifan, X., Zhiyong, G., Weiqiang, J., Dongping, W., & Xiaoshun, H. (2017). Expression of lncRNA-AK005641 in Murine CD4+FOXP3+Treg and iTreg Differentiation. *Transplantation*, *101*, S4. https://doi.org/10.1097/01.tp.0000520296.28521.7c

Chu, C., Qu, K., Zhong, F. L., Artandi, S. E., & Chang, H. Y. (2011). Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Molecular Cell*, *44*(4), 667–678. https://doi.org/10.1016/J.MOLCEL.2011.08.027

Chu, C., Zhang, Q. C., Da Rocha, S. T., Flynn, R. A., Bharadwaj, M., Calabrese, J. M., Magnuson, T., Heard, E., & Chang, H. Y. (2015). Systematic discovery of Xist RNA binding proteins. *Cell*, *161*(2), 404. https://doi.org/10.1016/J.CELL.2015.03.025

Chu, C., Zhang, Q. C., da Rocha, S. T., Flynn, R. A., Bharadwaj, M., Calabrese, J. M., Magnuson, T., Heard, E., & Chang, H. Y. (2015). Systematic Discovery of Xist RNA Binding Proteins. *Cell*, *161*(2), 404–416. https://doi.org/10.1016/j.cell.2015.03.025

Chu, H. P., Minajigi, A., Chen, Y., Morris, R., Guh, C. Y., Hsieh, Y. H., Boukhali, M., Haas, W., & Lee, J. T. (2021). iDRiP for the systematic discovery of proteins bound directly to noncoding RNA. *Nature Protocols 2021 16:7*, *16*(7), 3672–3694. https://doi.org/10.1038/s41596-021-00555-9

Clancy, S. (2008). RNA splicing: introns, exons and spliceosome. *Nature Education*, *1*(2008).

Collier, S. P., Collins, P. L., Williams, C. L., Boothby, M. R., & Aune, T. M. (2012). Cutting edge: influence of Tmevpg1, a long intergenic noncoding RNA, on the expression of Ifng by Th1 cells. *Journal of Immunology (Baltimore, Md. : 1950)*, *189*(5), 2084–2088. https://doi.org/10.4049/jimmunol.1200774

Collier, S. P., Henderson, M. A., Tossberg, J. T., & Aune, T. M. (2014). Regulation of the Th1 genomic locus from Ifng through Tmevpg1 by T-bet. *Journal of Immunology (Baltimore, Md. : 1950)*, *193*(8), 3959–3965. https://doi.org/10.4049/jimmunol.1401099

Corley, M., Burns, M. C., & Yeo, G. W. (2020). Molecular Cell How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms. *Molecular Cell*, *78*, 9–29. https://doi.org/10.1016/j.molcel.2020.03.011

Couper, K. N., Blount, D. G., & Riley, E. M. (2008). IL-10: the master regulator of immunity to infection. *Journal of Immunology (Baltimore, Md. : 1950)*, *180*(9), 5771–5777.

Cui, H., Banerjee, S., Guo, S., Xie, N., Ge, J., Jiang, D., Zörnig, M., Thannickal, V. J., & Liu, G. (2019). Long noncoding RNA Malat1 regulates differential activation of macrophages and response to lung injury. *JCI Insight*, *4*(4). https://doi.org/10.1172/JCI.INSIGHT.124522

Da Rocha, S. T., & Heard, E. (2017). Novel players in X inactivation: insights into Xist-mediated gene silencing and chromosome conformation. *Nature Structural & Molecular Biology 2017 24:3*, *24*(3), 197–204. https://doi.org/10.1038/nsmb.3370

Das, S., & Krainer, A. R. (2014). Emerging functions of SRSF1, splicing factor and oncoprotein, in RNA metabolism and cancer. In *Molecular Cancer Research* (Vol. 12, Issue 9). https://doi.org/10.1158/1541-7786.MCR-14-0131

Dassie, J. P., & Giangrande, P. H. (2013). Current progress on aptamer-targeted oligonucleotide therapeutics. In *Therapeutic Delivery* (Vol. 4, Issue 12). https://doi.org/10.4155/tde.13.118

Davies, E. G., Cheung, M., Gilmour, K., Maimaris, J., Curry, J., Furmanski, A., Sebire, N., Halliday, N., Mengrelis, K., Adams, S., Bernatoniene, J., Bremner, R., Browning, M., Devlin, B., Erichsen, H. C., Gaspar, H. B., Hutchison, L., Ip, W., Ifversen, M., … Thrasher, A. J. (2017). Thymus transplantation for complete DiGeorge syndrome: European experience. *Journal of Allergy and Clinical Immunology*, *140*(6). https://doi.org/10.1016/j.jaci.2017.03.020

De Pablos, L. M., Ferreira, T. R., Dowle, A. A., Forrester, S., Parry, E., Newling, K., & Walrad, P. B. (2019). The mRNA-bound Proteome of Leishmania mexicana: Novel genetic insight into an ancient parasite*. *Molecular and Cellular Proteomics*, *18*(7), 1271–1284. https://doi.org/10.1074/mcp.RA118.001307

den Braber, I., Mugwagwa, T., Vrisekoop, N., Westera, L., Mögling, R., Bregje de Boer, A., Willems, N., Schrijver, E. H. R., Spierenburg, G., Gaiser, K., Mul, E., Otto, S. A., Ruiter, A. F. C., Ackermans, M. T., Miedema, F., Borghans, J. A. M., de Boer, R. J., & Tesselaar, K. (2012). Maintenance of Peripheral Naive T Cells Is Sustained by Thymus Output in Mice but Not Humans. *Immunity*, *36*(2). https://doi.org/10.1016/j.immuni.2012.02.006

Desai, M., ter Kuile, F. O., Nosten, F., McGready, R., Asamoa, K., Brabin, B., & Newman, R. D. (2007). Epidemiology and burden of malaria in pregnancy. *The Lancet. Infectious Diseases*, *7*(2), 93–104. https://doi.org/10.1016/S1473-3099(07)70021-X

Díaz-Muñoz, M. D., & Turner, M. (2018). Uncovering the role of RNA-binding proteins in gene expression in the immune system. In *Frontiers in Immunology* (Vol. 9, Issue MAY, p. 1094). Frontiers Media S.A. https://doi.org/10.3389/fimmu.2018.01094

Diederichs, S. (2014). The four dimensions of noncoding RNA conservation. In *Trends in Genetics* (Vol. 30, Issue 4, pp. 121–123). Elsevier Ltd. https://doi.org/10.1016/j.tig.2014.01.004

Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., Xue, C., Marinov, G. K., Khatun, J., Williams, B. A., Zaleski, C., Rozowsky, J., Röder, M., Kokocinski, F., Abdelhamid, R. F., … Gingeras, T. R. (2012). Landscape of transcription in human cells. *Nature*, *489*(7414), 101–108. https://doi.org/10.1038/nature11233

Dong, X., Zhang, J., Zhang, Q., Liang, Z., Xu, Y., Zhao, Y., & Zhang, B. (2022). Cytosolic Nuclear Sensor Dhx9 Controls Medullary Thymic Epithelial Cell Differentiation by p53-Mediated Pathways. *Frontiers in Immunology*, *13*. https://doi.org/10.3389/FIMMU.2022.896472

Donlic, A., Morgan, B. S., Xu, J. L., Liu, A., Roble, C., & Hargrove, A. E. (2018). Discovery of Small Molecule Ligands for MALAT1 by Tuning an RNA-Binding Scaffold. *Angewandte Chemie (International Ed. in English)*, *57*(40), 13242. https://doi.org/10.1002/ANIE.201808823

Dougan, M., Dougan, S., Slisz, J., Firestone, B., Vanneman, M., Draganov, D., Goyal, G., Li, W., Neuberg, D., Blumberg, R., Hacohen, N., Porter, D., Zawel, L., & Dranoff, G. (2010). IAP inhibitors enhance co-stimulation to promote tumor immunity. *The Journal of Experimental Medicine*, *207*(10), 2195. https://doi.org/10.1084/JEM.20101123

Dragomir, M., Chen, B., & Calin, G. A. (2018). Exosomal lncRNAs as new players in cell-to-cell communication. In *Translational Cancer Research* (Vol. 7, Issue Suppl 2, pp. S243–S252). AME Publishing Company. https://doi.org/10.21037/tcr.2017.10.46

Du, J., Wang, Q., Ziegler, S. F., & Zhou, B. (2018). FOXP3 interacts with hnRNPF to modulate pre-mRNA alternative splicing. *The Journal of Biological Chemistry*, *293*(26), 10235–10244. https://doi.org/10.1074/JBC.RA117.001349

Dupage, M., & Bluestone, J. A. (2016). Harnessing the plasticity of CD4+ T cells to treat immune-mediated disease. *Nature Reviews Immunology 2016 16:3*, *16*(3), 149–163. https://doi.org/10.1038/nri.2015.18

Dykstra-Aiello, C., Jickling, G. C., Ander, B. P., Shroff, N., Zhan, X., Liu, D., Hull, H., Orantia, M., Stamova, B. S., & Sharp, F. R. (2016). Altered Expression of Long Noncoding RNAs in Blood after Ischemic Stroke and Proximity to Putative Stroke Risk Loci. *Stroke*, *47*(12), 2896–2903. https://doi.org/10.1161/STROKEAHA.116.013869

Eames, H. L., Corbin, A. L., & Udalova, I. A. (2016). Interferon regulatory factor 5 in human autoimmunity and murine models of autoimmune disease. In *Translational Research* (Vol. 167, Issue 1). https://doi.org/10.1016/j.trsl.2015.06.018

Eißmann, M., Gutschner, T., Hämmerle, M., Günther, S., Caudron-Herger, M., Groß, M., Schirmacher, P., Rippe, K., Braun, T., Zörnig, M., & Diederichs, S. (2012). Loss of the abundant nuclear non-coding RNA MALAT1 is compatible with life and development. *RNA Biology*, *9*(8), 1076–1087. https://doi.org/10.4161/rna.21089

Else, K. J., & Grencis, R. K. (1991). Cellular immune responses to the murine nematode parasite Trichuris muris. I. Differential cytokine production during acute or chronic infection. *Immunology*, *72*(4).

Engreitz, J. M., Sirokman, K., McDonel, P., Shishkin, A. A., Surka, C., Russell, P., Grossman, S. R., Chow, A. Y., Guttman, M., & Lander, E. S. (2014a). RNA-RNA interactions enable specific targeting of noncoding RNAs to nascent pre-mRNAs and chromatin sites. *Cell*. https://doi.org/10.1016/j.cell.2014.08.018

Engreitz, J. M., Sirokman, K., McDonel, P., Shishkin, A. A., Surka, C., Russell, P., Grossman, S. R., Chow, A. Y., Guttman, M., & Lander, E. S. (2014b). RNA-RNA interactions enable specific targeting of noncoding RNAs to nascent pre-mRNAs and chromatin sites. *Cell*, *159*(1), 188–199. https://doi.org/10.1016/j.cell.2014.08.018

Eperon, I. C., Makarova, O. v., Mayeda, A., Munroe, S. H., Cáceres, J. F., Hayward, D. G., & Krainer, A. R. (2000). Selection of Alternative 5′ Splice Sites: Role of U1 snRNP and Models for the Antagonistic Effects of SF2/ASF and hnRNP A1. *Molecular and Cellular Biology*. https://doi.org/10.1128/mcb.20.22.8303-8318.2000

Faas, M., Bouman, A., Moesa, H., Heineman, M. J., De Leij, L., & Schuiling, G. (2000). The immune response during the luteal phase of the ovarian cycle: A Th2-type response? *Fertility and Sterility*, *74*(5), 1008–1013. https://doi.org/10.1016/S0015-0282(00)01553-3

Fattahi, F., Ellis, J. S., Sylvester, M., Bahleda, K., Hietanen, S., Correa, L., Lugogo, N. L., & Atasoy, U. (2022). HuR-Targeted Inhibition Impairs Th2 Proinflammatory Responses in Asthmatic CD4 + T Cells. *Journal of Immunology (Baltimore, Md. : 1950)*, *208*(1), 38–48. https://doi.org/10.4049/JIMMUNOL.2100635

Ferreira, F. M., Palle, P., vom Berg, J., Prajwal, P., Laman, J. D., & Buch, T. (2019). Bone marrow chimeras—a vital tool in basic and translational research. *Journal of Molecular Medicine*, *97*(7), 889–896. https://doi.org/10.1007/S00109-019-01783-Z/TABLES/3

Finley, J. (2015). Reactivation of latently infected HIV-1 viral reservoirs and correction of aberrant alternative splicing in the LMNA gene via AMPK activation: Common mechanism of action linking HIV-1 latency and Hutchinson-Gilford progeria syndrome. *Medical Hypotheses*, *85*(3), 320–332. https://doi.org/10.1016/J.MEHY.2015.06.003

Fiorentino, D. F., Bond, M. W., & Mosmann, T. R. (1989). Two types of mouse T helper cell. IV. Th2 clones secrete a factor that inhibits cytokine production by Th1 clones. *The Journal of Experimental Medicine*, *170*(6), 2081–2095.

Flores Villanueva, P. O., Chikunguwo, S. M., Harris, T. S., & Stadecker, M. J. (1993). Role of IL-10 on antigen-presenting cell function for schistosomal egg-specific monoclonal T helper cell responses in vitro and in vivo. *J Immunol*, *151*(6).

Flores Villanueva, P. O., Reiser, H., & Stadecker, M. J. (1994). Regulation of T helper cell responses in experimental murine schistosomiasis by IL-10. Effect on expression of B7 and B7-2 costimulatory molecules by macrophages. *J Immunol*, *153*(11).

Fluiter, K., Mook, O. R. F., Vreijling, J., Langkjær, N., Højland, T., Wengel, J., & Baas, F. (2009). Filling the gap in LNA antisense oligo gapmers: the effects of unlocked nucleic acid (UNA) and 4′-C-hydroxymethyl-DNA modifications on RNase H recruitment and efficacy of an LNA gapmer. *Molecular BioSystems*, *5*(8), 838–843. https://doi.org/10.1039/B903922H

Franco-Zorrilla, J. M., Valli, A., Todesco, M., Mateos, I., Puga, M. I., Rubio-Somoza, I., Leyva, A., Weigel, D., García, J. A., & Paz-Ares, J. (2007). Target mimicry provides a new mechanism for regulation of microRNA activity. *Nature Genetics*, *39*(8), 1033–1037. https://doi.org/10.1038/ng2079

Gabryšová, L., Alvarez-Martinez, M., Luisier, R., Cox, L. S., Sodenkamp, J., Hosking, C., Pérez-Mazliah, D., Whicher, C., Kannan, Y., Potempa, K., Wu, X., Bhaw, L., Wende, H., Sieweke, M. H., Elgar, G., Wilson, M., Briscoe, J., Metzis, V., Langhorne, J., … O'Garra, A. (2018). c-Maf controls immune responses by regulating disease-specific gene networks and repressing IL-2 in CD4+ T cells. *Nature Immunology*, *19*(5), 497. https://doi.org/10.1038/S41590-018-0083-5

Gandhi, M., Groß, M., Holler, J. M., Coggins, S. A. A., Patil, N., Leupold, J. H., Munschauer, M., Schenone, M., Hartigan, C. R., Allgayer, H., Kim, B., & Diederichs, S. (2020). The lncRNA lincNMR regulates nucleotide metabolism via a YBX1 - RRM2 axis in cancer. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-17007-9

Gao, F., Yin, J., Chen, Y., Guo, C., Hu, H., & Su, J. (2022). *Recent advances in aptamer-based targeted drug delivery systems for cancer therapy*. https://doi.org/10.3389/fbioe.2022.972933

Gao, H., Wang, X., Lin, C., An, Z., Yu, J., Cao, H., Fan, Y., & Liang, X. (2020). Exosomal MALAT1 derived from ox-LDL-treated endothelial cells induce neutrophil extracellular traps to aggravate atherosclerosis. *Biological Chemistry*, *401*(3), 367–376. https://doi.org/10.1515/HSZ-2019-0219

Garitano-Trojaola, A., Agirre, X., Prósper, F., & Fortes, P. (2013). Long non-coding RNAs in haematological malignancies. *International Journal of Molecular Sciences*, *14*(8), 15386–15422. https://doi.org/10.3390/ijms140815386

Gaudreau, M.-C., Heyd, F., Bastien, R., Wilhelm, B., & Möröy, T. (2012). Alternative Splicing Controlled by Heterogeneous Nuclear Ribonucleoprotein L Regulates

Development, Proliferation, and Migration of Thymic Pre-T Cells. *The Journal of Immunology*, *188*(11), 5377–5388. https://doi.org/10.4049/JIMMUNOL.1103142

Gazzaniga, F. S., & Blackburn, E. H. (2014). An antiapoptotic role for telomerase RNA in human immune cells independent of telomere integrity or telomerase enzymatic activity. *Blood*, *124*(25), 3675–3684. https://doi.org/10.1182/blood-2014-06-582254

Gebauer, F., Schwarzl, T., Valcárcel, J., & Hentze, M. W. (2020). RNA-binding proteins in human genetic disease. *Nature Reviews Genetics 2020 22:3*, *22*(3), 185–198. https://doi.org/10.1038/s41576-020-00302-y

Gehring, N. H., Wahle, E., & Fischer, U. (2017). Deciphering the mRNP Code: RNA-Bound Determinants of Post-Transcriptional Gene Regulation. In *Trends in Biochemical Sciences* (Vol. 42, Issue 5, pp. 369–382). Elsevier Ltd. https://doi.org/10.1016/j.tibs.2017.02.004

Gentle, I. E., Moelter, I., Lechler, N., Bambach, S., Vucikuja, S., Häcker, G., & Aichele, P. (2014). Inhibitors of apoptosis proteins (IAPs) are required for effective T-cell expansion/survival during antiviral immunity in mice. *Blood*, *123*(5), 659–668. https://doi.org/10.1182/BLOOD-2013-01-479543

Gerber, A. P. (2021). RNA-Centric Approaches to Profile the RNA-Protein Interaction Landscape on Selected RNAs. *Non-Coding RNA*, *7*(1), 1–15. https://doi.org/10.3390/NCRNA7010011

Geuens, T., Bouhy, D., & Timmerman, V. (2016). The hnRNP family: insights into their role in health and disease. In *Human Genetics*. https://doi.org/10.1007/s00439-016-1683-5

Gibbons, H. R., Shaginurova, G., Kim, L. C., Chapman, N., Spurlock, C. F., Aune, T. M., & Aune, T. M. (2018). Divergent lncRNA GATA3-AS1 Regulates GATA3 Transcription in T-Helper 2 Cells. *Frontiers in Immunology*, *9*, 2512. https://doi.org/10.3389/fimmu.2018.02512

Gil, N., & Ulitsky, I. (2019). Regulation of gene expression by cis-acting long non-coding RNAs. *Nature Reviews Genetics 2019 21:2*, *21*(2), 102–117. https://doi.org/10.1038/s41576-019-0184-5

Girón-González, J. A., Moral, F. J., Elvira, J., García-Gil, D., Guerrero, F., Gavilán, I., & Escobar, L. (2000). Consistent production of a higher T(H)1:T(H)2 cytokine ratio by stimulated T cells in men compared with women. *European Journal of Endocrinology*, *143*(1), 31–36. https://doi.org/10.1530/eje.0.1430031

Glisovic, T., Bachorik, J. L., Yong, J., & Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. In *FEBS Letters* (Vol. 582, Issue 14, pp. 1977–1986). No longer published by Elsevier. https://doi.org/10.1016/j.febslet.2008.03.004

Goldie, P. D., & Chatterjee, I. (2021). Examining the elevated risk of COVID-19 in transgender communities with an intersectional lens. *SN Social Sciences*, *1*(10), 249. https://doi.org/10.1007/S43545-021-00255-X

Gomes, E., & Shorter, J. (2019). The molecular language of membraneless organelles. *Journal of Biological Chemistry*, *294*(18), 7115–7127. https://doi.org/10.1074/JBC.TM118.001192

Gomez, J. A., Wapinski, O. L., Yang, Y. W., Bureau, J.-F., Gopinath, S., Monack, D. M., Chang, H. Y., Brahic, M., & Kirkegaard, K. (2013a). The NeST Long ncRNA Controls Microbial Susceptibility and Epigenetic Activation of the Interferon-γ Locus. *Cell*, *152*(4), 743–754. https://doi.org/10.1016/j.cell.2013.01.015

Gomez, J. A., Wapinski, O. L., Yang, Y. W., Bureau, J.-F., Gopinath, S., Monack, D. M., Chang, H. Y., Brahic, M., & Kirkegaard, K. (2013b). The NeST Long ncRNA Controls Microbial Susceptibility and Epigenetic Activation of the Interferon-γ Locus. *Cell*, *152*(4), 743–754. https://doi.org/10.1016/j.cell.2013.01.015

Gordon, M. A., Babbs, B., Cochrane, D. R., Bitler, B. G., & Richer, J. K. (2019). The long non-coding RNA MALAT1 promotes ovarian cancer progression by regulating RBFOX2-mediated alternative splicing. *Molecular Carcinogenesis*, *58*(2), 196–205. https://doi.org/10.1002/MC.22919

Groff, A. F., Barutcu, A. R., Lewandowski, J. P., & Rinn, J. L. (2018). Enhancers in the Peril lincRNA locus regulate distant but not local genes. *Genome Biology*, *19*(1). https://doi.org/10.1186/S13059-018-1589-8

Gu, W., Jiang, X., Wang, W., Mujagond, P., Liu, J., Mai, Z., Tang, H., li, S., Xiao, H., & Zhao, J. (2022). Super-Enhancer-Associated Long Non-Coding RNA LINC01485 Promotes Osteogenic Differentiation of Human Bone Marrow Mesenchymal Stem Cells by Regulating MiR-619-5p/RUNX2 Axis. *Frontiers in Endocrinology*, *13*. https://doi.org/10.3389/FENDO.2022.846154

Guo, C. J., Ma, X. K., Xing, Y. H., Zheng, C. C., Xu, Y. F., Shan, L., Zhang, J., Wang, S., Wang, Y., Carmichael, G. G., Yang, L., & Chen, L. L. (2020). Distinct Processing of lncRNAs Contributes to Non-conserved Functions in Stem Cells. *Cell*, *181*(3). https://doi.org/10.1016/j.cell.2020.03.006

Guo, W., Lei, W., Yu, D., Ge, Y., Chen, Y., Xue, W., Li, Q., Li, S., Gao, X., & Yao, W. (2017). Involvement of lncRNA-1700040D17Rik in Th17 cell differentiation and the pathogenesis of EAE. *International Immunopharmacology*, *47*, 141–149. https://doi.org/10.1016/j.intimp.2017.03.014

Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., Tsai, M.-C., Hung, T., Argani, P., Rinn, J. L., Wang, Y., Brzoska, P., Kong, B., Li, R., West, R. B., van de Vijver, M. J., Sukumar, S., & Chang, H. Y. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature*, *464*(7291), 1071–1076. https://doi.org/10.1038/nature08975

Hacisuleyman, E., Goff, L. A., Trapnell, C., Williams, A., Henao-Mejia, J., Sun, L., Mcclanahan, P., Hendrickson, D. G., Sauvageau, M., Kelley, D. R., Morse, M., Engreitz, J., Lander, E. S., Guttman, M., Lodish, H. F., Flavell, R., Raj, A., & Rinn, J. L. (2014). *Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre*. https://doi.org/10.1038/nsmb.2764

Haerty, W., & Ponting, C. P. (2013). Mutations within lncRNAs are effectively selected against in fruitfly but not in human. *Genome Biology*, *14*(5), R49. https://doi.org/10.1186/gb-2013-14-5-r49

Hafner, M., Katsantoni, M., Köster, T., Marks, J., Mukherjee, J., Staiger, D., Ule, J., & Zavolan, M. (2021). CLIP and complementary methods. In *Nature Reviews Methods Primers* (Vol. 1, Issue 1). https://doi.org/10.1038/s43586-021-00018-1

Han, J., Shen, L., Zhan, Z., Liu, Y., Zhang, C., Guo, R., Luo, Y., Xie, Z., Feng, Y., & Wu, G. (2021). The long noncoding RNA MALAT1 modulates adipose loss in cancer-associated cachexia by suppressing adipogenesis through PPAR-γ. *Nutrition and Metabolism*. https://doi.org/10.1186/s12986-021-00557-0

Han, S. P., Tang, Y. H., & Smith, R. (2010). Functional diversity of the hnRNPs: Past, present and perspectives. In *Biochemical Journal*. https://doi.org/10.1042/BJ20100396

Han, Y., Liu, Y., Zhang, H., Wang, T., Diao, R., Jiang, Z., Gui, Y., & Cai, Z. (2013). Hsa-miR-125b suppresses bladder cancer development by down-regulating oncogene SIRT7 and oncogenic long noncoding RNA MALAT1. *FEBS Letters*, *587*(23), 3875–3882. https://doi.org/10.1016/j.febslet.2013.10.023

Harris, A. L. (2002). Hypoxia — a key regulatory factor in tumour growth. *Nature Reviews Cancer 2002 2:1*, *2*(1), 38–47. https://doi.org/10.1038/nrc704

Harrow, J., Frankish, A., Gonzalez, J. M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B. L., Barrell, D., Zadissa, A., Searle, S., Barnes, I., Bignell, A., Boychenko, V., Hunt, T., Kay, M., Mukherjee, G., Rajan, J., Despacio-Reyes, G., Saunders, G., … Hubbard, T. J. (2012). GENCODE: The reference human genome annotation for The ENCODE Project. *Genome Research*, *22*(9), 1760–1774. https://doi.org/10.1101/gr.135350.111

Haward, F., Maslon, M. M., Yeyati, P. L., Bellora, N., Hansen, J. N., Aitken, S., Lawson, J., von Kriegsheim, A., Wachten, D., Mill, P., Adams, I. R., & Caceres, J. F. (2021). Nucleo-cytoplasmic shuttling of splicing factor srsf1 is required for development and cilia function. *ELife*, *10*. https://doi.org/10.7554/ELIFE.65104

He, L., Valignat, M., Zhang, L., Gelard, L., Zhang, F., le Guen, V., Audebert, S., Camoin, L., Fossum, E., Bogen, B., Wang, H., Henri, S., Roncagalli, R., Theodoly, O., Liang, Y., Malissen, M., & Malissen, B. (2021). ARHGAP45 controls naïve T- and B-cell entry into lymph nodes and T-cell progenitor thymus seeding. *EMBO Reports*, *22*(4). https://doi.org/10.15252/embr.202052196

Hentze, M. W., Castello, A., Schwarzl, T., & Preiss, T. (2018). A brave new world of RNA-binding proteins. In *Nature Reviews Molecular Cell Biology* (Vol. 19, Issue 5, pp. 327–341). Nature Publishing Group. https://doi.org/10.1038/nrm.2017.130

Hergesheimer, R. C., Chami, A. A., de Assis, D. R., Vourc'h, P., Andres, C. R., Corcia, P., Lanznaster, D., & Blasco, H. (2019). The debated toxic role of aggregated TDP-43 in amyotrophic lateral sclerosis: A resolution in sight? In *Brain* (Vol. 142, Issue 5). https://doi.org/10.1093/brain/awz078

Hewitson, J. P., West, K. A., James, K. R., Rani, G. F., Dey, N., Romano, A., Brown, N., Teichmann, S. A., Kaye, P. M., & Lagos, D. (2020). Malat1 Suppresses Immunity to Infection through Promoting Expression of Maf and IL-10 in Th Cells. *The Journal of Immunology*. https://doi.org/10.4049/JIMMUNOL.1900940

Hoefig, K. P., Reim, A., Gallus, C., Wong, E. H., Behrens, G., Conrad, C., Xu, M., Kifinger, L., Ito-Kureha, T., Defourny, K. A. Y., Geerlof, A., Mautner, J., Hauck, S. M., Baumjohann, D., Feederle, R., Mann, M., Wierer, M., Glasmacher, E., & Heissmeyer, V. (2021). Defining the RBPome of primary T helper cells to elucidate higher-order Roquin-mediated mRNA regulation. *Nature Communications 2021 12:1*, *12*(1), 1–18. https://doi.org/10.1038/s41467-021-25345-5

Hoffman, W., Lakkis, F. G., & Chalasani, G. (2016). B cells, antibodies, and more. *Clinical Journal of the American Society of Nephrology*, *11*(1). https://doi.org/10.2215/CJN.09430915

Howard, J. M., Lin, H., Wallace, A. J., Kim, G., Draper, J. M., Haeussler, M., Katzman, S., Toloue, M., Liu, Y., & Sanford, J. R. (2018). HNRNPA1 promotes recognition of splice site decoys by U2AF2 in vivo. *Genome Research*. https://doi.org/10.1101/gr.229062.117

Hu, G., Tang, Q., Sharma, S., Yu, F., Escobar, T. M., Muljo, S. A., Zhu, J., & Zhao, K. (2013). Expression and regulation of intergenic long noncoding RNAs during T cell

development and differentiation. *Nature Immunology*, *14*(11), 1190–1198. https://doi.org/10.1038/ni.2712

Huang, D., Chen, J., Yang, L., Ouyang, Q., Li, J., Lao, L., Zhao, J., Liu, J., Lu, Y., Xing, Y., Chen, F., Su, F., Yao, H., Liu, Q., Su, S., & Song, E. (2018). NKILA lncRNA promotes tumor immune evasion by sensitizing T cells to activation-induced cell death. *Nature Immunology*, *19*(10), 1112–1125. https://doi.org/10.1038/s41590-018-0207-y

Huang, S., Dong, D., Zhang, Y., Chen, Z., Geng, J., & Zhao, Y. (2019). NEAT1 regulates Th2 cell development by targeting STAT6 for degradation. *Cell Cycle*, *18*(3), 312–319. https://doi.org/10.1080/15384101.2018.1562285

Huang, S., Dong, D., Zhang, Y., Chen, Z., Geng, J., & Zhao, Y. (2021). Long non-coding RNA nuclear paraspeckle assembly transcript 1 promotes activation of T helper 2 cells via inhibiting STAT6 ubiquitination. *Human Cell*, *34*(3). https://doi.org/10.1007/s13577-021-00496-1

Huang, W., & Littman, D. R. (2015). Regulation of RORγt in Inflammatory Lymphoid Cell Differentiation. *Cold Spring Harbor Symposia on Quantitative Biology*, *80*, 257–263. https://doi.org/10.1101/SQB.2015.80.027615

Huppertz, I., Attig, J., D'Ambrogio, A., Easton, L. E., Sibley, C. R., Sugimoto, Y., Tajnik, M., König, J., & Ule, J. (2014). iCLIP: Protein-RNA interactions at nucleotide resolution. *Methods*. https://doi.org/10.1016/j.ymeth.2013.10.011

Huynh, J. P., Lin, C. C., Kimmey, J. M., Jarjour, N. N., Schwarzkopf, E. A., Bradstreet, T. R., Shchukina, I., Shpynov, O., Weaver, C. T., Taneja, R., Artyomov, M. N., Edelson, B. T., & Stallings, C. L. (2018). Bhlhe40 is an essential repressor of IL-10 during

Mycobacterium tuberculosis infection. *Journal of Experimental Medicine*, *215*(7), 1823–1838. https://doi.org/10.1084/jem.20171704

Hwang, S. S., Kim, K., Lee, W., & Lee, G. R. (2012). Aberrant expression of IFN-γ in Th2 cells from Th2 LCR-deficient mice. *Biochemical and Biophysical Research Communications*, *424*(3), 512–518. https://doi.org/10.1016/j.bbrc.2012.06.146

Iannantuoni, F., Salazar, J. D., Martinez de Marañon, A., Bañuls, C., López-Domènech, S., Rocha, M., Hurtado-Murillo, F., Morillas, C., Gómez-Balaguer, M., & Víctor, V. M. (2021). Testosterone administration increases leukocyte-endothelium interactions and inflammation in transgender men. *Fertility and Sterility*, *115*(2), 483–489. https://doi.org/10.1016/J.FERTNSTERT.2020.08.002

Imbeaud, S., Graudens, E., Boulanger, V., Barlet, X., Zaborski, P., Eveno, E., Mueller, O., Schroeder, A., & Auffray, C. (2005). Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. *Nucleic Acids Research*, *33*(6), e56. https://doi.org/10.1093/NAR/GNI054

Indu, S., Sekhar, S. C., Sengottaiyan, J., Kumar, A., Pillai, S. M., Laloraya, M., & Kumar, P. G. (2015). Aberrant expression of Dynein light chain 1 (DYNLT1) is associated with human male factor infertility. *Molecular and Cellular Proteomics*. https://doi.org/10.1074/mcp.M115.050005

Iyer, S. S., & Cheng, G. (2012). Role of interleukin 10 transcriptional regulation in inflammation and autoimmune disease. *Critical Reviews in Immunology*, *32*(1), 23–63.

Jacquier, A. (2009). The complex eukaryotic transcriptome: Unexpected pervasive transcription and novel small RNAs. In *Nature Reviews Genetics* (Vol. 10, Issue 12). https://doi.org/10.1038/nrg2683

Jäger, A., & Kuchroo, V. K. (2010). Effector and regulatory T-cell subsets in autoimmunity and tissue inflammation. *Scandinavian Journal of Immunology*, *72*(3), 173–184. https://doi.org/10.1111/j.1365-3083.2010.02432.x

Jeltsch, K. M., Hu, D., Brenner, S., Zöller, J., Heinz, G. A., Nagel, D., Vogel, K. U., Rehage, N., Warth, S. C., Edelmann, S. L., Gloury, R., Martin, N., Lohs, C., Lech, M., Stehklein, J. E., Geerlof, A., Kremmer, E., Weber, A., Anders, H. J., … Heissmeyer, V. (2014). Cleavage of roquin and regnase-1 by the paracaspase MALT1 releases their cooperatively repressed targets to promote TH17 differentiation. *Nature Immunology*, *15*(11), 1079–1089. https://doi.org/10.1038/ni.3008

Jenkins, M. K., Khoruts, A., Ingulli, E., Mueller, D. L., McSorley, S. J., Lee Reinhardt, R., Itano, A., & Pape, K. A. (2001). In vivo activation of antigen-specific CD4 T cells. In *Annual Review of Immunology* (Vol. 19). https://doi.org/10.1146/annurev.immunol.19.1.23

Jeong, S. (2017). SR proteins: Binders, regulators, and connectors of RNA. In *Molecules and Cells*. https://doi.org/10.14348/molcells.2017.2319

Ji, P., Diederichs, S., Wang, W., Böing, S., Metzger, R., Schneider, P. M., Tidow, N., Brandt, B., Buerger, H., Bulk, E., Thomas, M., Berdel, W. E., Serve, H., & Müller-Tidow, C. (2003). MALAT-1, a novel noncoding RNA and thymosin β4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene*, *22*(39), 8031–8041. https://doi.org/10.1038/sj.onc.1206928

Jiang, R., Tang, J., Chen, Y., Deng, L., Ji, J., Xie, Y., Wang, K., Jia, W., Chu, W.-M., & Sun, B. (2017). The long noncoding RNA lnc-EGFR stimulates T-regulatory cells differentiation thus promoting hepatocellular carcinoma immune evasion. *Nature Communications*, *8*, 15129. https://doi.org/10.1038/ncomms15129

Jin, D., Guo, J., Wu, Y., Du, J., Yang, L., Wang, X., Di, W., Hu, B., An, J., Kong, L., Pan, L., & Su, G. (2019). M6A mRNA methylation initiated by METTL3 directly promotes YAP translation and increases YAP activity by regulating the MALAT1-miR-1914-3p-YAP axis to induce NSCLC drug resistance and metastasis. *Journal of Hematology and Oncology*, *12*(1), 1–22. https://doi.org/10.1186/S13045-019-0830-6/FIGURES/7

Johnsson, P., Lipovich, L., Grandér, D., & Morris, K. V. (2014). Evolutionary conservation of long non-coding RNAs; Sequence, structure, function. In *Biochimica et Biophysica Acta - General Subjects*. https://doi.org/10.1016/j.bbagen.2013.10.035

Joyce, K. L., Morgan, W., Greenberg, R., & Nair, M. G. (2012). Using Eggs from Schistosoma mansoni as an In vivo Model of Helminth-induced Lung Inflammation. *Journal of Visualized Experiments : JoVE*, *64*, 1. https://doi.org/10.3791/3905

Jurgens, A. P., Popović, B., & Wolkers, M. C. (2021). T cells at work: How post-transcriptional mechanisms control T cell homeostasis and activation. *European Journal of Immunology*, *51*(9), 2178–2187. https://doi.org/10.1002/EJI.202049055

Käfer, R., Schmidtke, L., Schrick, K., Montermann, E., Bros, M., Kleinert, H., & Pautz, A. (2019). The RNA-Binding Protein KSRP Modulates Cytokine Expression of CD4 + T Cells. *Journal of Immunology Research*, *2019*. https://doi.org/10.1155/2019/4726532

Kanbar, J. N., Ma, S., Kim, E. S., Kurd, N. S., Tsai, M. S., Tysl, T., Widjaja, C. E., Limary, A. E., Yee, B., He, Z., Hao, Y., Fu, X. D., Yeo, G. W., Huang, W. J., & Chang, J. T. (2022a). The long noncoding RNA Malat1 regulates CD8+ T cell differentiation by mediating epigenetic repression. *The Journal of Experimental Medicine*, *219*(6). https://doi.org/10.1084/JEM.20211756

Kanbar, J. N., Ma, S., Kim, E. S., Kurd, N. S., Tsai, M. S., Tysl, T., Widjaja, C. E., Limary, A. E., Yee, B., He, Z., Hao, Y., Fu, X. D., Yeo, G. W., Huang, W. J., & Chang, J. T. (2022b). The long noncoding RNA Malat1 regulates CD8+ T cell differentiation by mediating epigenetic repression. *Journal of Experimental Medicine*, *219*(6). https://doi.org/10.1084/JEM.20211756/213232

Karimi, E., Azari, H., Tahmasebi, A., Nikpoor, A. R., Negahi, A. A., Sanadgol, N., Shekari, M., & Mousavi, P. (2022). LncRNA-miRNA network analysis across the Th17 cell line reveals biomarker potency of lncRNA NEAT1 and KCNQ1OT1 in multiple sclerosis. *Journal of Cellular and Molecular Medicine*, *26*(8). https://doi.org/10.1111/jcmm.17256

Katsuyama, T., Li, H., Comte, D., Tsokos, G. C., & Moulton, V. R. (2019). Splicing factor SRSF1 controls T cell hyperactivity and systemic autoimmunity. *Journal of Clinical Investigation*. https://doi.org/10.1172/JCI127949

Kim, S. H., Kim, S. H., Yang, W. I., Kim, S. J., & Yoon, S. O. (2017). Association of the long non-coding RNA MALAT1 with the polycomb repressive complex pathway in T and NK cell lymphoma. *Oncotarget*, *8*(19), 31305–31317. https://doi.org/10.18632/oncotarget.15453

Kim, S. S., Harford, J. B., Moghe, M., Rait, A., Pirollo, K. F., & Chang, E. H. (2018). Targeted nanocomplex carrying siRNA against MALAT1 sensitizes glioblastoma to temozolomide. *Nucleic Acids Research*, *46*(3), 1424. https://doi.org/10.1093/NAR/GKX1221

Kirk, J. M., Kim, S. O., Inoue, K., Smola, M. J., Lee, D. M., Schertzer, M. D., Wooten, J. S., Baker, A. R., Sprague, D., Collins, D. W., Horning, C. R., Wang, S., Chen, Q., Weeks, K. M., Mucha, P. J., & Calabrese, J. M. (2018). Functional classification of

long non-coding RNAs by k-mer content. *Nature Genetics*, *50*(10). https://doi.org/10.1038/s41588-018-0207-8

Kissick, H. T., Sanda, M. G., Dunn, L. K., Pellegrini, K. L., On, S. T., Noel, J. K., & Arredouani, M. S. (2014). Androgens alter T-cell immunity by inhibiting T-helper 1 differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(27). https://doi.org/10.1073/pnas.1402468111

Klein, S. L. (2004). Hormonal and immunological mechanisms mediating sex differences in parasite infection. *Parasite Immunology*, *26*(6–7), 247–264. https://doi.org/10.1111/J.0141-9838.2004.00710.X

Klein, S. L., & Flanagan, K. L. (2016). Sex differences in immune responses. In *Nature Reviews Immunology* (Vol. 16, Issue 10, pp. 626–638). Nature Publishing Group. https://doi.org/10.1038/nri.2016.90

Klein, S. L., & Pekosz, A. (2014). Sex-based biology and the rational design of influenza vaccination strategies. *Journal of Infectious Diseases*, *209*(SUPPL. 3). https://doi.org/10.1093/infdis/jiu066

Koh, B. H., Hwang, S. S., Kim, J. Y., Lee, W., Kang, M.-J., Lee, C. G., Park, J.-W., Flavell, R. A., & Lee, G. R. (2010). Th2 LCR is essential for regulation of Th2 cytokine genes and for pathogenesis of allergic asthma. *Proceedings of the National Academy of Sciences*, *107*(23), 10614–10619. https://doi.org/10.1073/pnas.1005383107

Kolls, J., & Sandquist, I. (2018). Update on regulation and effector functions of Th17 cells. In *F1000Research* (Vol. 7). Faculty of 1000 Ltd. https://doi.org/10.12688/f1000research.13020.1

Komeno, Y., Huang, Y.-J., Qiu, J., Lin, L., Xu, Y., Zhou, Y., Chen, L., Monterroza, D. D., Li, H., DeKelver, R. C., Yan, M., Fu, X.-D., & Zhang, D.-E. (2015). SRSF2 Is Essential for Hematopoiesis, and Its Myelodysplastic Syndrome-Related Mutations Dysregulate Alternative Pre-mRNA Splicing. *Molecular and Cellular Biology*. https://doi.org/10.1128/mcb.00202-15

König, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D. J., Luscombe, N. M., & Ule, J. (2010). ICLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature Structural and Molecular Biology*. https://doi.org/10.1038/nsmb.1838

Kopp, F., & Mendell, J. T. (2018). Functional Classification and Experimental Dissection of Long Noncoding RNAs. *Cell*, *172*(3), 393–407. https://doi.org/10.1016/j.cell.2018.01.011

Krainer, A. R., Conway, G. C., & Kozak, D. (1990). Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. *Genes and Development*, *4*(7). https://doi.org/10.1101/gad.4.7.1158

Kratofil, R. M., Kubes, P., & Deniset, J. F. (2017). Monocyte conversion during inflammation and injury. In *Arteriosclerosis, Thrombosis, and Vascular Biology* (Vol. 37, Issue 1). https://doi.org/10.1161/ATVBAHA.116.308198

Kulkarni, S., Lied, A., Kulkarni, V., Rucevic, M., Martin, M. P., Walker-Sperling, V., Anderson, S. K., Ewy, R., Singh, S., Nguyen, H., McLaren, P. J., Viard, M., Naranbhai, V., Zou, C., Lin, Z., Gatanaga, H., Oka, S., Takiguchi, M., Thio, C. L., … Carrington, M. (2019). CCR5AS lncRNA variation differentially regulates CCR5, influencing HIV disease outcome. *Nature Immunology*, *20*(7). https://doi.org/10.1038/s41590-019-0406-1

Kumar, B. v., Connors, T. J., & Farber, D. L. (2018). Human T cell development, localization, and function throughout life. *Immunity*, *48*(2), 202. https://doi.org/10.1016/J.IMMUNI.2018.01.007

Kung, J. T. Y., Colognori, D., & Lee, J. T. (2013). Long noncoding RNAs: Past, present, and future. In *Genetics* (Vol. 193, Issue 3, pp. 651–669). Genetics Society of America. https://doi.org/10.1534/genetics.112.146704

Kuret, K., Amalietti, A. G., Jones, D. M., Capitanchik, C., & Ule, J. (2022). Positional motif analysis reveals the extent of specificity of protein-RNA interactions observed by CLIP. *Genome Biology*, *23*(1), 1–34. https://doi.org/10.1186/S13059-022-02755-2/FIGURES/6

Kurtulus, S., & Hildeman, D. (2013). Assessment of CD4+ and CD8+ T cell responses using MHC class i and II tetramers. *Methods in Molecular Biology, 979*. https://doi.org/10.1007/978-1-62703-290-2_8

Lee, F. C. Y., Chakrabarti, A. M., Hänel, H., Monzón-Casanova, E., Hallegger, M., Militti, C., Capraro, F., Sadée, C., Toolan-Kerr, P., Wilkins, O., Turner, M., König, J., Sibley, C. R., & Ule, J. (n.d.). *An improved iCLIP protocol*. https://doi.org/10.1101/2021.08.27.457890

Lee, F. C. Y., & Ule, J. (2018). Advances in CLIP Technologies for Studies of Protein-RNA Interactions. In *Molecular Cell* (Vol. 69, Issue 3). https://doi.org/10.1016/j.molcel.2018.01.005

Lee, Y. J., Wang, Q., & Rio, D. C. (2018). Coordinate regulation of alternative pre-mRNA splicing events by the human RNA chaperone proteins hnRNPA1 and DDX5. *Genes and Development*, *32*(15–16). https://doi.org/10.1101/gad.316034.118

Lelli, A., Nolan, K. A., Santambrogio, S., Gonçalves, A. F., Schönenberger, M. J., Guinot, A., Frew, I. J., Marti, H. M., Hoogewijs, D., & Wenger, R. H. (2015). Induction of long noncoding RNA MALAT1 in hypoxic mice. *Hypoxia*, *3*, 45–52. https://doi.org/10.2147/HP.S90555

Lerner, M. R., & Argetsinger Steitz, J. (1979). Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(11). https://doi.org/10.1073/pnas.76.11.5495

Leucci, E., Vendramin, R., Spinazzi, M., Laurette, P., Fiers, M., Wouters, J., Radaelli, E., Eyckerman, S., Leonelli, C., Vanderheyden, K., Rogiers, A., Hermans, E., Baatsen, P., Aerts, S., Amant, F., Van Aelst, S., van den Oord, J., de Strooper, B., Davidson, I., … Marine, J.-C. (2016). Melanoma addiction to the long non-coding RNA SAMMSON. *Nature*, *531*(7595), 518–522. https://doi.org/10.1038/nature17161

Leung, S., Smith, D., Myc, A., Morry, J., & Baker, J. R. (2013). OT-II TCR transgenic mice fail to produce anti-ovalbumin antibodies upon vaccination. *Cellular Immunology*, *282*(2). https://doi.org/10.1016/j.cellimm.2012.12.006

Li, B., Li, J., Devlin, B. H., & Markert, M. L. (2011). Thymic microenvironment reconstitution after postnatal human thymus transplantation. *Clinical Immunology*, *140*(3). https://doi.org/10.1016/j.clim.2011.04.004

Li, C., Cui, Y., Liu, L.-F., Ren, W.-B., Li, Q.-Q., Zhou, X., Li, Y.-L., Li, Y., Bai, X.-Y., & Zu, X.-B. (2017). High Expression of Long Noncoding RNA MALAT1 Indicates a Poor Prognosis and Promotes Clinical Progression and Metastasis in Bladder Cancer. *Clinical Genitourinary Cancer*, *15*(5), 570–576. https://doi.org/10.1016/j.clgc.2017.05.001

Li, H., Li, Q., Sun, S., Lei, P., Cai, X., & Shen, G. (2020). Integrated Bioinformatics Analysis Identifies ELAVL1 and APP as Candidate Crucial Genes for Crohn's Disease. *Journal of Immunology Research*, *2020*. https://doi.org/10.1155/2020/3067273

Li, J., Cui, Z., Li, H., Lv, X., Gao, M., Yang, Z., Bi, Y., Zhang, Z., Wang, S., Zhou, B., & Yin, Z. (2018). Clinicopathological and prognostic significance of long noncoding RNA MALAT1 in human cancers: a review and meta-analysis. *Cancer Cell International*, *18*(1), 109. https://doi.org/10.1186/s12935-018-0606-z

Li, J.-Q., Hu, S.-Y., Wang, Z.-Y., Lin, J., Jian, S., Dong, Y.-C., Wu, X.-F., Dai-Lan, & Cao, L.-J. (2016). Long non-coding RNA MEG3 inhibits microRNA-125a-5p expression and induces immune imbalance of Treg/Th17 in immune thrombocytopenic purpura. *Biomedicine & Pharmacotherapy*, *83*, 905–911. https://doi.org/10.1016/j.biopha.2016.07.057

Li, L. J., Chai, Y., Guo, X. J., Chu, S. L., & Zhang, L. S. (2017). The effects of the long non-coding RNA MALAT-1 regulated autophagy-related signaling pathway on chemotherapy resistance in diffuse large B-cell lymphoma. *Biomedicine and Pharmacotherapy*, *89*. https://doi.org/10.1016/j.biopha.2017.02.011

Liang, S. C., Long, A. J., Bennett, F., Whitters, M. J., Karim, R., Collins, M., Goldman, S. J., Dunussi-Joannopoulos, K., Williams, C. M. M., Wright, J. F., & Fouser, L. A. (2007). An IL-17F/A Heterodimer Protein Is Produced by Mouse Th17 Cells and Induces Airway Neutrophil Recruitment. *The Journal of Immunology*, *179*(11). https://doi.org/10.4049/jimmunol.179.11.7791

Liang, Z., & Tang, F. (2020). The potency of lncRNA MALAT1/miR-155/CTLA4 axis in altering Th1/Th2 balance of asthma. *Bioscience Reports*, *40*(2). https://doi.org/10.1042/BSR20190397

Lin, A., & Loré, K. (2017). Granulocytes: New members of the antigen-presenting cell family. In *Frontiers in Immunology* (Vol. 8, Issue DEC). https://doi.org/10.3389/fimmu.2017.01781

Lin, H., Mosmann, T. R., Guilbert, L., Tuntipopipat, S., & Wegmann, T. G. (1993). Synthesis of T helper 2-type cytokines at the maternal-fetal interface. *The Journal of Immunology*, *151*(9).

Liu, J., Ji, H., Zheng, W., Wu, X., Zhu, J. J., Arnold, A. P., & Sandberg, K. (2010). Sex differences in renal angiotensin converting enzyme 2 (ACE2) activity are 17β-oestradiol-dependent and sex chromosome-independent. *Biology of Sex Differences*, *1*(1), 1–11. https://doi.org/10.1186/2042-6410-1-6

Liu, J., Niu, Z., Zhang, R., Peng, Z., Wang, L., Liu, Z., Gao, Y., Pei, H., & Pan, L. (2021). MALAT1 shuttled by extracellular vesicles promotes M1 polarization of macrophages to induce acute pancreatitis via miR-181a-5p/HMGB1 axis. *Journal of Cellular and Molecular Medicine*, *25*(19), 9241–9254. https://doi.org/10.1111/JCMM.16844

Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M., & Pan, T. (2015). N6 -methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature*. https://doi.org/10.1038/nature14234

Liu, T. Y., Chen, Y. C., Jong, Y. J., Tsai, H. J., Lee, C. C., Chang, Y. S., Chang, J. G., & Chang, Y. F. (2017). Muscle developmental defects in heterogeneous nuclear Ribonucleoprotein A1 knockout mice. *Open Biology*. https://doi.org/10.1098/rsob.160303

Liu, W., Wang, Z., Liu, L., Yang, Z., Liu, S., Ma, Z., Liu, Y., Ma, Y., Zhang, L., Zhang, X., Jiang, M., & Cao, X. (2020). LncRNA Malat1 inhibition of TDP43 cleavage suppresses IRF3-initiated antiviral innate immunity. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(38), 23695–23706. https://doi.org/10.1073/PNAS.2003932117

Liu, X., Lin, J., Wu, H., Wang, Y., Xie, L., Wu, J., Qin, H., & Xu, J. (2021). A Novel Long Noncoding RNA lincRNA00892 Activates CD4+ T Cells in Systemic Lupus Erythematosus by Regulating CD40L. *Frontiers in Pharmacology*, *12*. https://doi.org/10.3389/fphar.2021.733902

Liu, X., Tian, N., Huang, Q., Xu, Z., Cheng, H., Liu, X., Li, D., Liang, R., Li, B., & Dai, X. (2021). hnRNPA1 enhances FOXP3 stability to promote the differentiation and functions of regulatory T cells. *FEBS Letters*. https://doi.org/10.1002/1873-3468.14142

Liva, S. M., & Voskuhl, R. R. (2001). Testosterone Acts Directly on CD4 + T Lymphocytes to Increase IL-10 Production . *The Journal of Immunology*, *167*(4), 2060–2067. https://doi.org/10.4049/jimmunol.167.4.2060

Lockard, R. D., Wilson, M. E., & Rodríguez, N. E. (2019). *Sex-Related Differences in Immune Response and Symptomatic Manifestations to Infection with Leishmania Species*. https://doi.org/10.1155/2019/4103819

Locksley, R. M., Heinzel, F. P., Sadick, M. D., Holaday, B. J., & Gardner, K. D. (1987). Murine cutaneous leishmaniasis : Susceptibility correlates with differential expansion of helper T-cell subsets. *Annales de l'Institut Pasteur - Immunology*, *138*(5). https://doi.org/10.1016/S0769-2625(87)80030-2

Love, M. I., Soneson, C., & Patro, R. (2018). Swimming downstream: Statistical analysis of differential transcript usage following Salmon quantification. *F1000Research*, *7*. https://doi.org/10.12688/f1000research.15398.3

Lu, T., & Mar, J. C. (2020). Investigating transcriptome-wide sex dimorphism by multi-level analysis of single-cell RNA sequencing data in ten mouse cell types. *Biology of Sex Differences*, *11*(1). https://doi.org/10.1186/s13293-020-00335-2

Luckheeram, R. V., Zhou, R., Verma, A. D., & Xia, B. (2012). CD4 $^{+}$ T Cells: Differentiation and Functions. *Clinical and Developmental Immunology*, *2012*, 1–12. https://doi.org/10.1155/2012/925135

Lunde, B. M., Moore, C., & Varani, G. (2007). RNA-binding proteins: Modular design for efficient function. In *Nature Reviews Molecular Cell Biology* (Vol. 8, Issue 6, pp. 479–490). Nature Publishing Group. https://doi.org/10.1038/nrm2178

Luo, J., Qu, L., Gao, F., Lin, J., Liu, J., & Lin, A. (2021). LncRNAs: Architectural Scaffolds or More Potential Roles in Phase Separation. In *Frontiers in Genetics* (Vol. 12). https://doi.org/10.3389/fgene.2021.626234

Ma, L., Bajic, V. B., & Zhang, Z. (2013). On the classification of long non-coding RNAs. In *RNA Biology* (Vol. 10, Issue 6, pp. 924–933). Taylor and Francis Inc. https://doi.org/10.4161/rna.24604

Mahata, B., Zhang, X., Kolodziejczyk, A. A., Proserpio, V., Haim-Vilmovsky, L., Taylor, A. E., Hebenstreit, D., Dingler, F. A., Moignard, V., Göttgens, B., Arlt, W., McKenzie, A. N. J., & Teichmann, S. A. (2014). Single-cell RNA sequencing reveals T helper cells synthesizing steroids De Novo to contribute to immune homeostasis. *Cell Reports*, *7*(4), 1130–1142. https://doi.org/10.1016/j.celrep.2014.04.011

Marchese, F. P., Raimondi, I., & Huarte, M. (2017). The multidimensional mechanisms of long noncoding RNA function. In *Genome Biology* (Vol. 18, Issue 1, pp. 1–13). BioMed Central Ltd. https://doi.org/10.1186/s13059-017-1348-2

Márquez, E. J., Chung, C. han, Marches, R., Rossi, R. J., Nehar-Belaid, D., Eroglu, A., Mellert, D. J., Kuchel, G. A., Banchereau, J., & Ucar, D. (2020). Sexual-dimorphism in human immune system aging. *Nature Communications*, *11*(1), 1–17. https://doi.org/10.1038/s41467-020-14396-9

Marriott, I., Bost, K. L., & Huet-Hudson, Y. M. (2006). Sexual dimorphism in expression of receptors for bacterial lipopolysaccharides in murine macrophages: A possible mechanism for gender-based differences in endotoxic shock susceptibility. *Journal of Reproductive Immunology*, *71*(1), 12–27. https://doi.org/10.1016/j.jri.2006.01.004

Martin A. Smith, Tanja Gesell, Peter F. Stadler, J. S. M. (2013). Widespread purifying selection on RNA structure in mammals. *Nucleic Acids Research*, *41*(17), 8220–8236.

Martinez, F. O., & Gordon, S. (2014). The M1 and M2 paradigm of macrophage activation: Time for reassessment. *F1000Prime Reports*, *6*. https://doi.org/10.12703/P6-13

Masliah, G., Barraud, P., & Allain, F. H. T. (2013). RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence. *Cellular and Molecular Life Sciences*, *70*(11), 1875. https://doi.org/10.1007/S00018-012-1119-X

Masoumi, F., Ghorbani, S., Talebi, F., Branton, W. G., Rajaei, S., Power, C., & Noorbakhsh, F. (2019). Malat1 long noncoding RNA regulates inflammation and leukocyte differentiation in experimental autoimmune encephalomyelitis. *Journal of Neuroimmunology*, *328*, 50–59. https://doi.org/10.1016/j.jneuroim.2018.11.013

Massa, M. G., David, C., Jörg, S., Berg, J., Gisevius, B., Hirschberg, S., Linker, R. A., Gold, R., & Haghikia, A. (2017). Testosterone Differentially Affects T Cells and Neurons in Murine and Human Models of Neuroinflammation and Neurodegeneration. *American Journal of Pathology*, *187*(7). https://doi.org/10.1016/j.ajpath.2017.03.006

Matia-González, A. M., Iadevaia, V., & Gerber, A. P. (2017). A versatile tandem RNA isolation procedure to capture in vivo formed mRNA-protein complexes. *Methods (San Diego, Calif.)*, *118–119*, 93–100. https://doi.org/10.1016/J.YMETH.2016.10.005

Mattick, J. S. (2004). RNA regulation: a new genetics? *Nature Reviews Genetics*, *5*(4), 316–323. https://doi.org/10.1038/nrg1321

McHugh, C. A., Chen, C.-K., Chow, A., Surka, C. F., Tran, C., McDonel, P., Pandya-Jones, A., Blanco, M., Burghard, C., Moradian, A., Sweredoski, M. J., Shishkin, A. A., Su, J., Lander, E. S., Hess, S., Plath, K., & Guttman, M. (2015). The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature*, *521*(7551), 232–236. https://doi.org/10.1038/nature14443

McHugh, C. A., & Guttman, M. (2018). RAP-MS: A Method to Identify Proteins that Interact Directly with a Specific RNA Molecule in Cells. In *Methods in molecular biology (Clifton, N.J.)* (Vol. 1649, pp. 473–488). https://doi.org/10.1007/978-1-4939-7213-5_31

McHugh, C. A., Russell, P., & Guttman, M. (2014). Methods for comprehensive experimental identification of RNA-protein interactions. *Genome Biology*, *15*(1), 203. https://doi.org/10.1186/gb4152

Mcmurray, R. W., Suwannaroj, S., Ndebele, K., & Jenkins, J. K. (2001). Differential effects of sex steroids on T and B cells: Modulation of cell cycle phase distribution, apoptosis and bcl-2 protein levels. *Pathobiology*, *69*(1), 44–58. https://doi.org/10.1159/000048757

Mcnamee, E. N., Darlynn, •, Johnson, K., Homann, D., & Clambey, E. T. (2013.). *Hypoxia and hypoxia-inducible factors as regulators of T cell development, differentiation, and function*. https://doi.org/10.1007/s12026-012-8349-8

Meininger, I., Griesbach, R. A., Hu, D., Gehring, T., Seeholzer, T., Bertossi, A., Kranich, J., Oeckinghaus, A., Eitelhuber, A. C., Greczmiel, U., Gewies, A., Schmidt-Supprian, M., Ruland, J., Brocker, T., Heissmeyer, V., Heyd, F., & Krappmann, D. (2016). Alternative splicing of MALT1 controls signalling and activation of CD4+ T cells. *Nature Communications 2016 7:1*, *7*(1), 1–15. https://doi.org/10.1038/ncomms11292

Meng, X., Grötsch, B., Luo, Y., Knaup, K. X., Wiesener, M. S., Chen, X. X., Jantsch, J., Fillatreau, S., Schett, G., & Bozec, A. (2018). Hypoxia-inducible factor-1α is a critical transcription factor for IL-10-producing B cells in autoimmune disease. *Nature Communications*, *9*(1). https://doi.org/10.1038/S41467-017-02683-X

Mestas, J., & Hughes, C. C. W. (2004). Of Mice and Not Men: Differences between Mouse and Human Immunology. *The Journal of Immunology*, *172*(5), 2731–2738. https://doi.org/10.4049/JIMMUNOL.172.5.2731

Minajigi, A., Froberg, J. E., Wei, C., Sunwoo, H., Kesner, B., Colognori, D., Lessing, D., Payer, B., Boukhali, M., Haas, W., & Lee, J. T. (2015). Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science (New York, N.Y.)*, *349*(6245), 1DUIMMY. https://doi.org/10.1126/SCIENCE.AAB2276

Molineros, J. E., Singh, B., Terao, C., Okada, Y., Kaplan, J., McDaniel, B., Akizuki, S., Sun, C., Webb, C. F., Looger, L. L., & Nath, S. K. (2019). Mechanistic Characterization of RASGRP1 Variants Identifies an hnRNP-K-Regulated Transcriptional Enhancer Contributing to SLE Susceptibility. *Frontiers in Immunology*, *10*(MAY). https://doi.org/10.3389/FIMMU.2019.01066

Montecino-Rodriguez, E., Berent-Maoz, B., & Dorshkind, K. (2013). Causes, consequences, and reversal of immune system aging. *The Journal of Clinical Investigation*, *123*(3), 958. https://doi.org/10.1172/JCI64096

Monteleone, G., Monteleone, I., Fina, D., Vavassori, P., del Vecchio Blanco, G., Caruso, R., Tersigni, R., Alessandroni, L., Biancone, L., Naccari, G. C., Macdonald, T. T., & Pallone, F. (2005). Interleukin-21 enhances T-helper cell type I signaling and interferon-γ production in Crohn's disease. *Gastroenterology*, *128*(3). https://doi.org/10.1053/j.gastro.2004.12.042

Montopoli, M., Zumerle, S., Vettor, R., Rugge, M., Zorzi, M., Catapano, C. V., Carbone, G. M., Cavalli, A., Pagano, F., Ragazzi, E., Prayer-Galetti, T., & Alimonti, A. (2020). Androgen-deprivation therapies for prostate cancer and risk of infection by SARS-CoV-2: a population-based study (N = 4532). *Annals of Oncology*, *31*(8), 1040–1045. https://doi.org/10.1016/j.annonc.2020.04.479

Moore, K. W., Vieira, P., Fiorentino, D. F., Trounstine, M. L., Khan, T. A., & Mosmann, T. R. (1990). Homology of cytokine synthesis inhibitory factor (IL-10) to the Epstein-Barr virus gene BCRFI. *Science (New York, N.Y.)*, *248*(4960), 1230–1234.

Morais, P., Adachi, H., & Yu, Y. T. (2021). Spliceosomal snRNA Epitranscriptomics. In *Frontiers in Genetics* (Vol. 12). https://doi.org/10.3389/fgene.2021.652129

Mosmann, T. R., Cherwinski, H., Bond, M. W., Giedlin, M. A., & Coffman, R. L. (1986). Two types of murine helper T cell clone. I. Definition according to profiles of lymphokine activities and secreted proteins. *Journal of Immunology (Baltimore, Md. : 1950)*, *136*(7), 2348–2357.

Mosmann, T. R., & Coffman, R. L. (1989). TH1 and TH2 Cells: Different Patterns of Lymphokine Secretion Lead to Different Functional Properties. *Annual Review of Immunology*, *7*(1), 145–173. https://doi.org/10.1146/annurev.iy.07.040189.001045

Munschauer, M., Nguyen, C. T., Sirokman, K., Hartigan, C. R., Hogstrom, L., Engreitz, J. M., Ulirsch, J. C., Fulco, C. P., Subramanian, V., Chen, J., Schenone, M., Guttman, M., Carr, S. A., & Lander, E. S. (2018). The NORAD lncRNA assembles a topoisomerase complex critical for genome stability. *Nature*, *561*(7721). https://doi.org/10.1038/s41586-018-0453-z

Murphy, M. L., Wille, U., Villegas, E. N., Hunter, C. A., & Farrell, J. P. (2001). IL-10 mediates susceptibility to Leishmania donovani infection. *European Journal of Immunology*, *31*(10), 2848–2856. https://doi.org/10.1002/1521-4141(2001010)31:10<2848::aid-immu2848>3.0.co;2-t

Nagai, S., Kurebayashi, Y., & Koyasu, S. (2013). Role of PI3K/Akt and mTOR complexes in Th17 cell differentiation. *Annals of the New York Academy of Sciences*, *1280*(1), 30–34. https://doi.org/10.1111/nyas.12059

Naganuma, T., & Hirose, T. (2013). Paraspeckle formation during the biogenesis of long non-coding RNAs. *RNA Biology*, *10*(3), 456. https://doi.org/10.4161/RNA.23547

Nakagawa, S., Ip, J. Y., Shioi, G., Tripathi, V., Zong, X., Hirose, T., & Prasanth, K. V. (2012a). Malat1 is not an essential component of nuclear speckles in mice. *RNA (New York, N.Y.)*, *18*(8), 1487–1499. https://doi.org/10.1261/rna.033217.112

Nakagawa, S., Ip, J. Y., Shioi, G., Tripathi, V., Zong, X., Hirose, T., & Prasanth, K. v. (2012b). Malat1 is not an essential component of nuclear speckles in mice. *RNA (New York, N.Y.)*, *18*(8), 1487–1499. https://doi.org/10.1261/rna.033217.112

Nakayama, T., Hirahara, K., Onodera, A., Endo, Y., Hosokawa, H., Shinoda, K., Tumes, D. J., & Okamoto, Y. (2017). Th2 cells in health and disease. In *Annual Review of Immunology* (Vol. 35). https://doi.org/10.1146/annurev-immunol-051116-052350

Okugawa, Y., Toiyama, Y., Hur, K., Toden, S., Saigusa, S., Tanaka, K., Inoue, Y., Mohri, Y., Kusunoki, M., Boland, C. R., & Goel, A. (2014). Metastasis-associated long non-coding RNA drives gastric cancer development and promotes peritoneal metastasis. *Carcinogenesis*, *35*(12), 2731–2739. https://doi.org/10.1093/carcin/bgu200

Opal, S. M., & DePalo, V. A. (2000). Anti-inflammatory cytokines. *Chest*, *117*(4), 1162–1172. https://doi.org/10.1378/CHEST.117.4.1162

Ouyang, W., Rutz, S., Crellin, N. K., Valdez, P. A., & Hymowitz, S. G. (2011). Regulation and Functions of the IL-10 Family of Cytokines in Inflammation and Disease. *Annual Review of Immunology*, *29*(1), 71–109. https://doi.org/10.1146/annurev-immunol-031210-101312

Ozaki, K., Spolski, R., Feng, C. G., Qi, C. F., Cheng, J., Sher, A., Morse, H. C., Liu, C., Schwartzberg, P. L., & Leonard, W. J. (2002). A critical role for IL-21 in regulating immunoglobulin production. *Science*, *298*(5598). https://doi.org/10.1126/science.1077002

Patente, T. A., Pinho, M. P., Oliveira, A. A., Evangelista, G. C. M., Bergami-Santos, P. C., & Barbuto, J. A. M. (2019). Human dendritic cells: Their heterogeneity and clinical application potential in cancer immunotherapy. *Frontiers in Immunology*, *10*(JAN). https://doi.org/10.3389/fimmu.2018.03176

Pedersen, S. F., Collora, J. A., Kim, R. N., Yang, K., Razmi, A., Catalano, A. A., Yeh, Y.-H. J., Mounzer, K., Tebas, P., Montaner, L. J., & Ho, Y.-C. (2022). Inhibition of a Chromatin and Transcription Modulator, SLTM, Increases HIV-1 Reactivation Identified by a CRISPR Inhibition Screen. *Journal of Virology*, *96*(13). https://doi.org/10.1128/JVI.00577-22

Pei, X., Wang, X., & Li, H. (2018). LncRNA SNHG1 regulates the differentiation of Treg cells and affects the immune escape of breast cancer via regulating miR-448/IDO. *International Journal of Biological Macromolecules*, *118*(Pt A), 24–30. https://doi.org/10.1016/j.ijbiomac.2018.06.033

Peng, H., Liu, Y., Tian, J., Ma, J., Tang, X., Rui, K., Tian, X., Mao, C., Lu, L., Xu, H., Jiang, P., & Wang, S. (2016a). The Long Noncoding RNA IFNG-AS1 Promotes T Helper Type 1 Cells Response in Patients with Hashimoto's Thyroiditis. *Scientific Reports*, *5*(1), 17702. https://doi.org/10.1038/srep17702

Peng, H., Liu, Y., Tian, J., Ma, J., Tang, X., Rui, K., Tian, X., Mao, C., Lu, L., Xu, H., Jiang, P., & Wang, S. (2016b). The Long Noncoding RNA IFNG-AS1 Promotes T Helper Type 1 Cells Response in Patients with Hashimoto's Thyroiditis. *Scientific Reports*, *5*(1), 17702. https://doi.org/10.1038/srep17702

Pennington, D. J., Vermijlen, D., Wise, E. L., Clarke, S. L., Tigelaar, R. E., & Hayday, A. C. (2005). The integration of conventional and unconventional T cells that characterizes cell-mediated responses. In *Advances in Immunology* (Vol. 87, pp. 27–59). Academic Press. https://doi.org/10.1016/S0065-2776(05)87002-6

Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S., & Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. *Nature*, *379*(6561), 131–137. https://doi.org/10.1038/379131a0

Pereira, I. T., Spangenberg, L., Cabrera, G., & Dallagiovanna, B. (2019). *Polysome-Associated lncRNAs during Cardiomyogenesis of hESCs*. https://doi.org/10.20944/PREPRINTS201903.0157.V1

Pereira, R. M., Hogan, P. G., Rao, A., & Martinez, G. J. (2017). Transcriptional and epigenetic regulation of T cell hyporesponsiveness. *Journal of Leukocyte Biology*, *102*(3), 601. https://doi.org/10.1189/JLB.2RI0317-097R

Punnonen, J., Aversa, G., Cocks, B. G., McKenzie, A. N. J., Menon, S., Zurawski, G., de Waal Malefyt, R., & de Vries, J. E. (1993). Interleukin 13 induces interleukin 4-independent IgG4 and IgE synthesis and CD23 expression by human B cells. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(8). https://doi.org/10.1073/pnas.90.8.3730

Punnonen, J., & de Vries, J. E. (1994). IL-13 induces proliferation, Ig isotype switching, and Ig synthesis by immature human fetal B cells. *Journal of Immunology (Baltimore, Md. : 1950)*, *152*(3).

Qi, Z., Wang, F., Yu, G., Wang, D., Yao, Y., You, M., Liu, J., Liu, J., Sun, Z., Ji, C., Xue, Y., & Yu, S. (2021). SRSF1 serves as a critical posttranscriptional regulator at the late stage of thymocyte development. *Science Advances*, *7*(16). https://doi.org/10.1126/sciadv.abf0753

Qiu, Y., Wu, Y., Lin, M., Bian, T., Xiao, Y., & Qin, C. (2019). LncRNA-MEG3 functions as a competing endogenous RNA to regulate Treg/Th17 balance in patients with asthma by targeting microRNA-17/ RORγt. *Biomedicine & Pharmacotherapy*, *111*, 386–394. https://doi.org/10.1016/j.biopha.2018.12.080

Qu, D., Sun, W.-W., Li, L., Ma, L., Sun, L., Jin, X., Li, T., Hou, W., & Wang, J.-H. (2019). Long noncoding RNA MALAT1 releases epigenetic silencing of HIV-1 replication by

displacing the polycomb repressive complex 2 from binding to the LTR promoter. *Nucleic Acids Research*, *47*(6), 3013–3027. https://doi.org/10.1093/nar/gkz117

Queiroz, R. M. L., Smith, T., Villanueva, E., Marti-Solano, M., Monti, M., Pizzinga, M., Mirea, D. M., Ramakrishna, M., Harvey, R. F., Dezi, V., Thomas, G. H., Willis, A. E., & Lilley, K. S. (2019). Comprehensive quantitation of RNA-protein interaction dynamics by orthogonal organic phase separation (OOPS). *Nature Biotechnology*, *37*(2), 169. https://doi.org/10.1038/S41587-018-0001-2

Ramanathan, M., Porter, D. F., & Khavari, P. A. (2019). Methods to study RNA–protein interactions. *Nature Methods 2019 16:3*, *16*(3), 225–234. https://doi.org/10.1038/s41592-019-0330-1

Ransohoff, J. D., Wei, Y., & Khavari, P. A. (2018). The functions and unique features of long intergenic non-coding RNA. In *Nature Reviews Molecular Cell Biology* (Vol. 19, Issue 3, pp. 143–157). Nature Publishing Group. https://doi.org/10.1038/nrm.2017.104

Ranzani, V., Rossetti, G., Panzeri, I., Arrigoni, A., Bonnal, R. J. P., Curti, S., Gruarin, P., Provasi, E., Sugliano, E., Marconi, M., De Francesco, R., Geginat, J., Bodega, B., Abrignani, S., & Pagani, M. (2015). The long intergenic noncoding RNA landscape of human lymphocytes highlights the regulation of T cell differentiation by linc-MAF-4. *Nature Immunology*, *16*(3), 318–325. https://doi.org/10.1038/ni.3093

Rao, G. K., Wong, A., Collinge, M., Sarhan, J., Yarovinsky, T. O., Ramgolam, V. S., Gaestel, M., Pardi, R., & Bender, J. R. (2018). T cell LFA-1-induced proinflammatory mRNA stabilization is mediated by the p38 pathway kinase MK2 in a process regulated by hnRNPs C, H1 and K. *PloS One*, *13*(7). https://doi.org/10.1371/JOURNAL.PONE.0201103

Rea, J., Menci, V., Tollis, P., Santini, T., Armaos, A., Garone, M. G., Iberite, F., Cipriano, A., Tartaglia, G. G., Rosa, A., Ballarino, M., Laneve, P., & Caffarelli, E. (2020). HOTAIRM1 regulates neuronal differentiation by modulating NEUROGENIN 2 and the downstream neurogenic cascade. *Cell Death and Disease*, *11*(7). https://doi.org/10.1038/s41419-020-02738-w

Ren, S., Liu, Y., Xu, W., Sun, Y., Lu, J., Wang, F., Wei, M., Shen, J., Hou, J., Gao, X., Xu, C., Huang, J., Zhao, Y., & Sun, Y. (2013). Long noncoding RNA MALAT-1 is a new potential therapeutic target for castration resistant prostate cancer. *Journal of Urology*, *190*(6). https://doi.org/10.1016/j.juro.2013.07.001

Richter, K., Perriard, G., & Oxenius, A. (2013). Reversal of chronic to resolved infection by IL-10 blockade is LCMV strain dependent. *European Journal of Immunology*, *43*(3), 649–654. https://doi.org/10.1002/EJI.201242887

Ridings-Figueroa, R., Stewart, E. R., Nesterova, T. B., Coker, H., Pintacuda, G., Godwin, J., Wilson, R., Haslam, A., Lilley, F., Ruigrok, R., Bageghni, S. A., Albadrani, G., Mansfield, W., Roulson, J.-A., Brockdorff, N., Ainscough, J. F. X., & Coverley, D. (2017). The nuclear matrix protein CIZ1 facilitates localization of Xist RNA to the inactive X-chromosome territory. *Genes & Development*, *31*(9), 876–888. https://doi.org/10.1101/gad.295907.117

Rissland, O. S. (2017). The organization and regulation of mRNA–protein complexes. In *Wiley Interdisciplinary Reviews: RNA* (Vol. 8, Issue 1, p. e1369). Blackwell Publishing Ltd. https://doi.org/10.1002/wrna.1369

Rodgers, S. J., Ooms, L. M., Oorschot, V. M. J., Schittenhelm, R. B., Nguyen, E. v., Hamila, S. A., Rynkiewicz, N., Gurung, R., Eramo, M. J., Sriratana, A., Fedele, C. G., Caramia, F., Loi, S., Kerr, G., Abud, H. E., Ramm, G., Papa, A., Ellisdon, A. M., Daly,

R. J., … Mitchell, C. A. (2021). INPP4B promotes PI3Kα-dependent late endosome formation and Wnt/β-catenin signaling in breast cancer. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-23241-6

Rosales, C. (2018). Neutrophil: A cell with many roles in inflammation or several cell types? In *Frontiers in Physiology* (Vol. 9, Issue FEB). https://doi.org/10.3389/fphys.2018.00113

Ruan, X., Li, P., Chen, Y., Shi, Y., Pirooznia, M., Seifuddin, F., Suemizu, H., Ohnishi, Y., Yoneda, N., Nishiwaki, M., Shepherdson, J., Suresh, A., Singh, K., Ma, Y., Jiang, C. fei, & Cao, H. (2020). In vivo functional analysis of non-conserved human lncRNAs associated with cardiometabolic traits. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-019-13688-z

Santos, R., Ursu, O., Gaulton, A., Patrícia Bento, A., Donadi, R. S., Bologa, C. G., Karlsson, A., Al-Lazikani, B., Hersey, A., Oprea, T. I., & Overington, J. P. (2017). A comprehensive map of molecular drug targets. *Nature Publishing Group*. https://doi.org/10.1038/nrd.2016.230

Saraiva, M., Vieira, P., & O'garra, A. (2019). *Cytokines Focus Biology and therapeutic potential of interleukin-10*. https://doi.org/10.1084/jem.20190418

Scherer, M., Levin, M., Butter, F., & Scheibe, M. (2020). Quantitative proteomics to identify nuclear rna- binding proteins of malat1. *International Journal of Molecular Sciences*, *21*(3). https://doi.org/10.3390/ijms21031166

Schmidt, N., Lareau, C. A., Keshishian, H., Ganskih, S., Schneider, C., Hennig, T., Melanson, R., Werner, S., Wei, Y., Zimmer, M., Ade, J., Kirschner, L., Zielinski, S., Dölken, L., Lander, E. S., Caliskan, N., Fischer, U., Vogel, J., Carr, S. A., …

Munschauer, M. (2021). The SARS-CoV-2 RNA–protein interactome in infected human cells. *Nature Microbiology*, *6*(3). https://doi.org/10.1038/s41564-020-00846-z

Schott, G., Munoz, G. G., Trevino, N., Chen, X., Weirauch, M. T., Gregory, S. G., Bradrick, S. S., & Blanco, M. A. G. (2021). U2AF2 binds IL7R exon 6 ectopically and represses its inclusion. *RNA (New York, N.Y.)*, *27*(5), 571–583. https://doi.org/10.1261/RNA.078279.120

Shaath, H., Vishnubalaji, R., Elango, R., Khattak, S., & Alajez, N. M. (2021). Single-cell long noncoding RNA (lncRNA) transcriptome implicates MALAT1 in triple-negative breast cancer (TNBC) resistance to neoadjuvant chemotherapy. *Cell Death Discovery*. https://doi.org/10.1038/s41420-020-00383-y

Shankarling, G., Cole, B. S., Mallory, M. J., & Lynch, K. W. (2014). Transcriptome-Wide RNA Interaction Profiling Reveals Physical and Functional Targets of hnRNP L in Human T Cells. *Molecular and Cellular Biology*, *34*(1), 71. https://doi.org/10.1128/MCB.00740-13

Sharma, S., Findlay, G. M., Bandukwala, H. S., Oberdoerffer, S., Baust, B., Li, Z., Schmidt, V., Hogan, P. G., Sacks, D. B., & Rao, A. (2011). Dephosphorylation of the nuclear factor of activated T cells (NFAT) transcription factor is regulated by an RNA-protein scaffold complex. *Proceedings of the National Academy of Sciences*, *108*(28), 11381–11386. https://doi.org/10.1073/pnas.1019711108

Sheflin, L. G., & Spaulding, S. W. (2000). Testosterone and dihydrotestosterone regulate AUF1 isoforms in a tissue-specific fashion in the mouse. *American Journal of Physiology. Endocrinology and Metabolism*, *278*(1). https://doi.org/10.1152/AJPENDO.2000.278.1.E50

Shepard, P. J., & Hertel, K. J. (2009). The SR protein family. In *Genome biology*. https://doi.org/10.1186/gb-2009-10-10-242

Shevyrev, D., & Tereshchenko, V. (2020). Treg Heterogeneity, Function, and Homeostasis. In *Frontiers in Immunology* (Vol. 10). https://doi.org/10.3389/fimmu.2019.03100

Shi, L., Magee, P., Fassan, M., Sahoo, S., Leong, H. S., Lee, D., Sellers, R., Brullé-Soumaré, L., Cairo, S., Monteverde, T., Volinia, S., Smith, D. D., di Leva, G., Galuppini, F., Paliouras, A. R., Zeng, K., O'Keefe, R., & Garofalo, M. (2021). A KRAS-responsive long non-coding RNA controls microRNA processing. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-22337-3

Shih, C. H., Chuang, L. L., Tsai, M. H., Chen, L. H., Chuang, E. Y., Lu, T. P., & Lai, L. C. (2021). Hypoxia-Induced MALAT1 Promotes the Proliferation and Migration of Breast Cancer Cells by Sponging MiR-3064-5p. *Frontiers in Oncology*, *11*, 1. https://doi.org/10.3389/FONC.2021.658151

Shui, X., Chen, S., Lin, J., Kong, J., Zhou, C., & Wu, J. (2019). Knockdown of lncRNA NEAT1 inhibits Th17/CD4 $^{+}$ T cell differentiation through reducing the STAT3 protein level. *Journal of Cellular Physiology*, jcp.28811. https://doi.org/10.1002/jcp.28811

Simon, M. D., Wang, C. I., Kharchenko, P. V., West, J. A., Chapman, B. A., Alekseyenko, A. A., Borowsky, M. L., Kuroda, M. I., & Kingston, R. E. (2011). The genomic binding sites of a noncoding RNA. *Proceedings of the National Academy of Sciences*, *108*(51), 20497–20502. https://doi.org/10.1073/pnas.1113536108

Singh, G., Pratt, G., Yeo, G. W., & Moore, M. J. (2015). The Clothes Make the mRNA: Past and Present Trends in mRNP Fashion. *Annual Review of Biochemistry*, *84*(1), 325–354. https://doi.org/10.1146/annurev-biochem-080111-092106

Sirey, T. M., Roberts, K., Haerty, W., Bedoya-Reina, O., Granados, S. R., Tan, J. Y., Li, N., Heather, L. C., Carter, R. N., Cooper, S., Finch, A. J., Wills, J., Morton, N. M., Marques, A. C., & Ponting, C. P. (2019). The long non-coding rna cerox1 is a post transcriptional regulator of mitochondrial complex i catalytic activity. *ELife*, *8*. https://doi.org/10.7554/eLife.45051

Spellberg, B., & Edwards, J. E. (2001). Type 1/type 2 immunity in infectious diseases. In *Clinical Infectious Diseases* (Vol. 32, Issue 1). https://doi.org/10.1086/317537

Spiniello, M., Knoener, R. A., Steinbrink, M. I., Yang, B., Cesnik, A. J., Buxton, K. E., Scalf, M., Jarrard, D. F., & Smith, L. M. (2018). HyPR-MS for Multiplexed Discovery of MALAT1, NEAT1, and NORAD lncRNA Protein Interactomes. *Journal of Proteome Research*, *17*(9), 3022. https://doi.org/10.1021/ACS.JPROTEOME.8B00189

Spits, H. (2002). Development of αβ T cells in the human thymus. In *Nature Reviews Immunology* (Vol. 2, Issue 10, pp. 760–772). https://doi.org/10.1038/nri913

Spurlock, C. F., Tossberg, J. T., Guo, Y., Collier, S. P., Crooke, P. S., Aune, T. M., III, & Aune, T. M. (2015). Expression and functions of long noncoding RNAs during human T helper cell differentiation. *Nature Communications*, *6*, 6932. https://doi.org/10.1038/ncomms7932

Statello, L., Guo, C. J., Chen, L. L., & Huarte, M. (2020). Gene regulation by long non-coding RNAs and its biological functions. In *Nature Reviews Molecular Cell Biology* (pp. 1–23). Nature Research. https://doi.org/10.1038/s41580-020-00315-9

Stellato, C., Gubin, M. M., Magee, J. D., Fang, X., Fan, J., Tartar, D. M., Chen, J., Dahm, G. M., Calaluce, R., Mori, F., Jackson, G. A., Casolaro, V., Franklin, C. L., & Atasoy, U. (2011). Coordinate Regulation of GATA-3 and Th2 Cytokine Gene Expression by

the RNA-Binding Protein HuR . *The Journal of Immunology*, *187*(1), 441–449. https://doi.org/10.4049/jimmunol.1001881

Stewart, E. R., Turner, R. M. L., Newling, K., Ridings-Figueroa, R., Scott, V., Ashton, P. D., Ainscough, J. F. X., & Coverley, D. (2019). Maintenance of epigenetic landscape requires CIZ1 and is corrupted in differentiated fibroblasts in long-term culture. *Nature Communications*. https://doi.org/10.1038/s41467-018-08072-2

Sun, Q., Hao, Q., & Prasanth, K. V. (2018). Nuclear Long Noncoding RNAs: Key Regulators of Gene Expression. In *Trends in Genetics* (Vol. 34, Issue 2, pp. 142–157). Elsevier Ltd. https://doi.org/10.1016/j.tig.2017.11.005

Surcel, M., Constantin, C., Caruntu, C., Zurac, S., & Neagu, M. (2017). Inflammatory Cytokine Pattern Is Sex-Dependent in Mouse Cutaneous Melanoma Experimental Model. *Journal of Immunology Research*, *2017*. https://doi.org/10.1155/2017/9212134

Surh, C. D., & Sprent, J. (1994). T-cell apoptosis detected in situ during positive and negative selection in the thymus. *Nature*, *372*(6501), 100–103. https://doi.org/10.1038/372100a0

Syrett, C. M., Paneru, B., Sandoval-Heglund, D., Wang, J., Banerjee, S., Sindhava, V., Behrens, E. M., Atchison, M., & Anguera, M. C. (2019). Altered X-chromosome inactivation in T cells may promote sex-biased autoimmune diseases. *JCI Insight*, *4*(7). https://doi.org/10.1172/jci.insight.126751

Szappanos, D., Tschismarov, R., Perlot, T., Westermayer, S., Fischer, K., Platanitis, E., Kallinger, F., Novatchkova, M., Lassnig, C., Müller, M., Sexl, V., Bennett, K. L., Foong-Sobis, M., Penninger, J. M., & Decker, T. (2018). The RNA helicase DDX3X

is an essential mediator of innate antimicrobial immunity. *PLoS Pathogens*, *14*(11). https://doi.org/10.1371/journal.ppat.1007397

Szulc, B., & Piasecki, E. (1988). Effects of interferons, interferon inducers and growth factors on phagocytosis measured by quantitative determination of synthetic compound ingested by mouse bone marrow-derived macrophages. *Archivum Immunologiae et Therapiae Experimentalis*, *36*(5).

Techasintana, P., Davis, J. W., Gubin, M. M., Magee, J. D., & Atasoy, U. (2015). Transcriptomic-Wide Discovery of Direct and Indirect HuR RNA Targets in Activated CD4+ T Cells. *PloS One*, *10*(7). https://doi.org/10.1371/JOURNAL.PONE.0129321

Techasintana, P., Ellis, J. S., Glascock, J., Gubin, M. M., Ridenhour, S. E., Magee, J. D., Hart, M. L., Yao, P., Zhou, H., Whitney, M. S., Franklin, C. L., Martindale, J. L., Gorospe, M., Davis, W. J., Fox, P. L., Li, X., & Atasoy, U. (2017). The RNA-Binding Protein HuR Posttranscriptionally Regulates IL-2 Homeostasis and CD4 + Th2 Differentiation . *ImmunoHorizons*, *1*(6), 109–123. https://doi.org/10.4049/immunohorizons.1700017

Tirado-Hurtado, I., Fajardo, W., & Pinto, J. A. (2018). DNA Damage Inducible Transcript 4 Gene: The Switch of the Metabolism as Potential Target in Cancer. *Frontiers in Oncology*, *8*, 106. https://doi.org/10.3389/fonc.2018.00106

Tollervey, J. R., Curk, T., Rogelj, B., Briese, M., Cereda, M., Kayikci, M., König, J., Hortobágyi, T., Nishimura, A. L., Župunski, V., Patani, R., Chandran, S., Rot, G., Zupan, B., Shaw, C. E., & Ule, J. (2011). Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nature Neuroscience*, *14*(4). https://doi.org/10.1038/nn.2778

Trienekens, S. C. M., Faust, C. L., Meginnis, K., Pickering, L., Ericssonid, O., Nankasi, A., Moses, A., Tukahebwa, E. M., & Lamberton, P. H. L. (2020). Impacts of host gender on Schistosoma mansoni risk in rural Uganda—A mixed-methods approach. *PLoS Neglected Tropical Diseases*, *14*(5), 1–19. https://doi.org/10.1371/JOURNAL.PNTD.0008266

Tripathi, V., Ellis, J. D., Shen, Z., Song, D. Y., Pan, Q., Watt, A. T., Freier, S. M., Bennett, C. F., Sharma, A., Bubulya, P. A., Blencowe, B. J., Prasanth, S. G., & Prasanth, K. v. (2010a). The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2010.08.011

Tripathi, V., Ellis, J. D., Shen, Z., Song, D. Y., Pan, Q., Watt, A. T., Freier, S. M., Bennett, C. F., Sharma, A., Bubulya, P. A., Blencowe, B. J., Prasanth, S. G., & Prasanth, K. V. (2010b). The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Molecular Cell*, *39*(6), 925–938. https://doi.org/10.1016/j.molcel.2010.08.011

Tripathi, V., Ellis, J. D., Shen, Z., Song, D. Y., Pan, Q., Watt, A. T., Freier, S. M., Bennett, C. F., Sharma, A., Bubulya, P. A., Blencowe, B. J., Prasanth, S. G., & Prasanth, K. V. (2010c). The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Molecular Cell*, *39*(6), 925–938. https://doi.org/10.1016/j.molcel.2010.08.011

Tsagakis, I., Douka, K., Birds, I., & Aspden, J. L. (2020). Long non-coding RNAs in development and disease: Conservation to mechanisms. *The Journal of Pathology*, path.5405. https://doi.org/10.1002/path.5405

Tsao, H.-W., Tai, T.-S., Tseng, W., Chang, H.-H., Grenningloh, R., Miaw, S.-C., & Ho, I.-C. (2013). Ets-1 facilitates nuclear entry of NFAT proteins and their recruitment to the IL-2 promoter. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(39), 15776–15781. https://doi.org/10.1073/pnas.1304343110

Turner, M., & DÍaz-Muñoz, M. D. (2018). RNA-binding proteins control gene expression and cell fate in the immune system review-article. In *Nature Immunology* (Vol. 19, Issue 2, pp. 120–129). Nature Publishing Group. https://doi.org/10.1038/s41590-017-0028-4

Ulitsky, I. (2016). Evolution to the rescue: Using comparative genomics to understand long non-coding RNAs. In *Nature Reviews Genetics* (Vol. 17, Issue 10). https://doi.org/10.1038/nrg.2016.85

Vieira, P., de Waal-Malefyt, R., Dang, M. N., Johnson, K. E., Kastelein, R., Fiorentino, D. F., deVries, J. E., Roncarolo, M. G., Mosmann, T. R., & Moore, K. W. (1991). Isolation and expression of human cytokine synthesis inhibitory factor cDNA clones: homology to Epstein-Barr virus open reading frame BCRFI. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(4), 1172–1176.

Vigneau, S., Rohrlich, P.-S., Brahic, M., & Bureau, J.-F. (2003a). Tmevpg1, a candidate gene for the control of Theiler's virus persistence, could be implicated in the regulation of gamma interferon. *Journal of Virology*, *77*(10), 5632–5638.

Vigneau, S., Rohrlich, P.-S., Brahic, M., & Bureau, J.-F. (2003b). Tmevpg1, a candidate gene for the control of Theiler's virus persistence, could be implicated in the regulation of gamma interferon. *Journal of Virology*, *77*(10), 5632–5638. https://doi.org/10.1128/jvi.77.10.5632-5638.2003

Volk, H. D., Gruner, S., Falck, P., & von Baehr, R. (1986). The influence of interferon-γ and various phagocytic stimuli on the expression of MHC-class II antigens on human monocytes - relation to the generation of reactive oxygen intermediates. *Immunology Letters*, *13*(4). https://doi.org/10.1016/0165-2478(86)90057-X

Vrisekoop, N., den Braber, I., de Boer, A. B., Ruiter, A. F. C., Ackermans, M. T., van der Crabben, S. N., Schrijver, E. H. R., Spierenburg, G., Sauerwein, H. P., Hazenberg, M. D., de Boer, R. J., Miedema, F., Borghans, J. A. M., & Tesselaar, K. (2008). Sparse production but preferential incorporation of recently produced naïve T cells in the human peripheral pool. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(16). https://doi.org/10.1073/pnas.0709713105

Walker, J. A., & McKenzie, A. N. J. (2018). TH2 cell development and function. In *Nature Reviews Immunology* (Vol. 18, Issue 2, pp. 121–133). Nature Publishing Group. https://doi.org/10.1038/nri.2017.118

Wang, C., Li, Y., Yan, S., Wang, H., Shao, X., Xiao, M., Yang, B., Qin, G., Kong, R., Chen, R., & Zhang, N. (2020). Interactome analysis reveals that lncRNA HULC promotes aerobic glycolysis through LDHA and PKM2. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-16966-3

Wang, D., Ding, L., Wang, L., Zhao, Y., Sun, Z., Karnes, R. J., Zhang, J., & Huang, H. (2015). LncRNA MALAT1 enhances oncogenic activities of EZH2 in castrationresistant prostate cancer. *Oncotarget*. https://doi.org/10.18632/oncotarget.5728

Wang, J., Syrett, C. M., Kramer, M. C., Basu, A., Atchison, M. L., & Anguera, M. C. (2016). Unusual maintenance of X chromosome inactivation predisposes female lymphocytes for increased expression from the inactive X. *Proceedings of the*

National Academy of Sciences of the United States of America, *113*(14), E2029-38. https://doi.org/10.1073/pnas.1520113113

Wang, Q. M., Lian, G. Y., Song, Y., Huang, Y. F., & Gong, Y. (2019). LncRNA MALAT1 promotes tumorigenesis and immune escape of diffuse large B cell lymphoma by sponging miR-195. *Life Sciences*, *231*. https://doi.org/10.1016/j.lfs.2019.03.040

Wang, X., Juan, L., Lv, J., Wang, K., Sanford, J. R., & Liu, Y. (2011). Predicting sequence and structural specificities of RNA binding regions recognized by splicing factor SRSF1. *BMC Genomics*, *12*(SUPPL. 5). https://doi.org/10.1186/1471-2164-12-S5-S8

Wang, X., Liu, Q., & Zhang, B. (2014). Leveraging the complementary nature of RNA-Seq and shotgun proteomics data. *Proteomics*, *14*, 2676–2687. https://doi.org/10.1002/pmic.201400184

Wang, X., Sehgal, L., Jain, N., Khashab, T., Mathur, R., & Samaniego, F. (2016). Lncrna malat1 promotes development of mantle cell lymphoma by associating with ezh2. *Journal of Translational Medicine*, *14*(1). https://doi.org/10.1186/s12967-016-1100-9

Warringa, R. A., Schweizer, R. C., Maikoe, T., Kuijper, P. H., Bruijnzeel, P. L., & Koendermann, L. (1992). Modulation of eosinophil chemotaxis by interleukin-5. *American Journal of Respiratory Cell and Molecular Biology*, *7*(6). https://doi.org/10.1165/ajrcmb/7.6.631

Wei, L., Li, J., Han, Z., Chen, Z., & Zhang, Q. (2019). Silencing of lncRNA MALAT1 Prevents Inflammatory Injury after Lung Transplant Ischemia-Reperfusion by Downregulation of IL-8 via p300. *Molecular Therapy. Nucleic Acids*, *18*, 285–297. https://doi.org/10.1016/J.OMTN.2019.05.009

West, J. A., Davis, C. P., Sunwoo, H., Simon, M. D., Sadreyev, R. I., Wang, P. I., Tolstorukov, M. Y., & Kingston, R. E. (2014a). The Long Noncoding RNAs NEAT1 and MALAT1 Bind Active Chromatin Sites. *Molecular Cell*, *55*(5), 791–802. https://doi.org/10.1016/j.molcel.2014.07.012

West, J. A., Davis, C. P., Sunwoo, H., Simon, M. D., Sadreyev, R. I., Wang, P. I., Tolstorukov, M. Y., & Kingston, R. E. (2014b). The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Molecular Cell*, *55*(5), 791–802. https://doi.org/10.1016/j.molcel.2014.07.012

West, & Lagos. (2019). Long Non-Coding RNA Function in CD4+ T Cells: What We Know and What Next? *Non-Coding RNA*, *5*(3), 43. https://doi.org/10.3390/ncrna5030043

Whisenant, T. C., Peralta, E. R., Aarreberg, L. D., Gao, N. J., Head, S. R., Ordoukhanian, P., Williamson, J. R., & Salomon, D. R. (2015). The Activation-Induced Assembly of an RNA/Protein Interactome Centered on the Splicing Factor U2AF2 Regulates Gene Expression in Human CD4 T Cells. *PLOS ONE*, *10*(12), e0144409. https://doi.org/10.1371/JOURNAL.PONE.0144409

Wilkinson, M. E., Charenton, C., & Nagai, K. (2020). RNA Splicing by the Spliceosome. In *Annual Review of Biochemistry* (Vol. 89). https://doi.org/10.1146/annurev-biochem-091719-064225

Willingham, A. T., Orth, A. P., Batalov, S., Peters, E. C., Wen, B. G., Aza-Blanc, P., Hogenesch, J. B., & Schultz, P. G. (2005). A Strategy for Probing the Function of Noncoding RNAs Finds a Repressor of NFAT. *Science*, *309*(5740), 1570–1573. https://doi.org/10.1126/science.1115901

Wilusz, J. E., JnBaptiste, C. K., Lu, L. Y., Kuhn, C. D., Joshua-Tor, L., & Sharp, P. A. (2012). A triple helix stabilizes the 3′ ends of long noncoding RNAs that lack poly(A) tails. *Genes and Development*, *26*(21). https://doi.org/10.1101/gad.204438.112

Wu, J., Zhang, H., Zheng, Y., Jin, X., Liu, M., Li, S., Zhao, Q., Liu, X., Wang, Y., Shi, M., Zhang, S., Tian, J., Sun, Y., Zhang, M., & Yu, B. (2018). The long noncoding RNA MALAT1 induces tolerogenic dendritic cells and regulatory T cells via miR155/dendritic cell-specific intercellular adhesion molecule-3 grabbing nonintegrin/IL10 axis. *Frontiers in Immunology*, *9*(AUG), 1847. https://doi.org/10.3389/FIMMU.2018.01847/BIBTEX

Wynn, T. A., Cheever, A. W., Williams, M. E., Hieny, S., Caspar, P., Kühn, R., Müller, W., & Sher, A. (1998). IL-10 regulates liver pathology in acute murine Schistosomiasis mansoni but is not required for immune down-modulation of chronic disease. *Journal of Immunology (Baltimore, Md. : 1950)*, *160*(9).

Xie, S. J., Diao, L. T., Cai, N., Zhang, L. T., Xiang, S., Jia, C. C., Qiu, D. B., Liu, C., Sun, Y. J., Lei, H., Hou, Y. R., Tao, S., Hu, Y. X., Xiao, Z. D., & Zhang, Q. (2021). mascRNA and its parent lncRNA MALAT1 promote proliferation and metastasis of hepatocellular carcinoma cells by activating ERK/MAPK signaling pathway. *Cell Death Discovery 2021 7:1*, *7*(1), 1–14. https://doi.org/10.1038/s41420-021-00497-x

Xiong, G., Yang, L., Chen, Y., & Fan, Z. (2015). Linc-POU3F3 promotes cell proliferation in gastric cancer via increasing T-reg distribution. *American Journal of Translational Research*, *7*(11), 2262–2269.

Xu, C., Yang, M., Tian, J., Wang, X., & Li, Z. (2011). MALAT-1: a long non-coding RNA and its important 3’ end functional motif in colorectal cancer metastasis. *International Journal of Oncology*, *39*(1), 169–175. https://doi.org/10.3892/IJO.2011.1007

Xu, S., Sui, S., Zhang, J., Bai, N., Shi, Q., Zhang, G., Gao, S., You, Z., Zhan, C., Liu, F., & Pang, D. (2015). Downregulation of long noncoding RNA MALAT1 induces epithelial-to-mesenchymal transition via the PI3K-AKT pathway in breast cancer. *International Journal of Clinical and Experimental Pathology*, *8*(5), 4881–4891.

Xue, Y., Ke, J., Zhou, X., Chen, Q., Chen, M., Huang, T., Lin, F., & Chen, F. (2022). Knockdown of LncRNA MALAT1 Alleviates Coxsackievirus B3-Induced Acute Viral Myocarditis in Mice via Inhibiting Th17 Cells Differentiation. *Inflammation*, 1–13. https://doi.org/10.1007/S10753-021-01612-X/FIGURES/6

Yang, H., Beutler, B., & Zhang, D. (2022). Emerging roles of spliceosome in cancer and immunity. In *Protein and Cell* (Vol. 13, Issue 8). https://doi.org/10.1007/s13238-021-00856-5

Yao, Y., Guo, W., Chen, J., Guo, P., Yu, G., Liu, J., Wang, F., Liu, J., You, M., Zhao, T., Kang, Y., Ma, X., & Yu, S. (2018). Long noncoding RNA Malat1 is not essential for T cell development and response to LCMV infection. *RNA Biology*, 1–10. https://doi.org/10.1080/15476286.2018.1551705

Yarosh, C. A., Tapescu, I., Thompson, M. G., Qiu, J., Mallory, M. J., Fu, X. D., & Lynch, K. W. (2015). TRAP150 interacts with the RNA-binding domain of PSF and antagonizes splicing of numerous PSF-target genes in T cells. *Nucleic Acids Research*, *43*(18), 9006–9016. https://doi.org/10.1093/NAR/GKV816

Yin, Y., Yan, P., Lu, J., Song, G., Zhu, Y., Li, Z., Zhao, Y., Shen, B., Huang, X., Zhu, H., Orkin, S. H., & Shen, X. (2015). Opposing Roles for the lncRNA Haunt and Its Genomic Locus in Regulating HOXA Gene Activation during Embryonic Stem Cell Differentiation. *Cell Stem Cell*, *16*(5), 504–516. https://doi.org/10.1016/J.STEM.2015.03.007

Yu, S., Tripod, M., Atasoy, U., & Chen, J. (2021). HuR Plays a Positive Role to Strengthen the Signaling Pathways of CD4+ T Cell Activation and Th17 Cell Differentiation. *Journal of Immunology Research*, *2021*. https://doi.org/10.1155/2021/9937243

Zahler, A. M., Damgaard, C. K., Kjems, J., & Caputi, M. (2004). SC35 and Heterogeneous Nuclear Ribonucleoprotein A/B Proteins Bind to a Juxtaposed Exonic Splicing Enhancer/Exonic Splicing Silencer Element to Regulate HIV-1 tat Exon 2 Splicing. *Journal of Biological Chemistry*, *279*(11). https://doi.org/10.1074/jbc.M312743200

Zemmour, D., Pratama, A., Loughhead, S. M., Mathis, D., & Benoist, C. (2017). *Flicr* , a long noncoding RNA, modulates Foxp3 expression and autoimmunity. *Proceedings of the National Academy of Sciences*, *114*(17), E3472–E3480. https://doi.org/10.1073/pnas.1700946114

Zeng, F., Peritz, T., Kannanayakal, T. J., Kilk, K., Eirílksdóttir, E., Langel, U., & Eberwine, J. (2006). A protocol for PAIR: PNA-assisted identification of RNA binding proteins in living cells. *Nature Protocols*, *1*(2), 920–927. https://doi.org/10.1038/NPROT.2006.81

Zerbe, L. K., Pino, I., Pio, R., Cosper, P. F., Dwyer-Nield, L. D., Meyer, A. M., Port, J. D., Montuenga, L. M., & Malkinson, A. M. (2004). Relative amounts of antagonistic splicing factors, hnRNP A1 and ASF/SF2, change during neoplastic lung growth: Implications for pre-mRNA processing. *Molecular Carcinogenesis*. https://doi.org/10.1002/mc.20053

Zhang, B., Arun, G., Mao, Y. S., Lazar, Z., Hung, G., Bhattacharjee, G., Xiao, X., Booth, C. J., Wu, J., Zhang, C., & Spector, D. L. (2012). The lncRNA malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. *Cell Reports*. https://doi.org/10.1016/j.celrep.2012.06.003

Zhang, C., Yang, M., & Ericsson, A. C. (2021). Function of Macrophages in Disease: Current Understanding on Molecular Mechanisms. In *Frontiers in Immunology* (Vol. 12). https://doi.org/10.3389/fimmu.2021.620510

Zhang, F., Liu, G., Li, D., Wei, C., & Hao, J. (2018). DDIT4 and Associated lncDDIT4 Modulate Th17 Differentiation through the DDIT4/TSC/mTOR Pathway. *Journal of Immunology (Baltimore, Md. : 1950)*, *200*(5), 1618–1626. https://doi.org/10.4049/jimmunol.1601689

Zhang, F., Liu, G., Wei, C., Gao, C., & Hao, J. (2017). Linc-MAF-4 regulates T $_{h}$ 1/T $_{h}$ 2 differentiation and is associated with the pathogenesis of multiple sclerosis by targeting MAF. *The FASEB Journal*, *31*(2), 519–525. https://doi.org/10.1096/fj.201600838R

Zhang, H., Nestor, C. E., Zhao, S., Lentini, A., Bohle, B., Benson, M., & Wang, H. (2013). Profiling of human CD4+ T-cell subsets identifies the TH2-specific noncoding RNA GATA3-AS1. *Journal of Allergy and Clinical Immunology*, *132*(4), 1005–1008. https://doi.org/10.1016/j.jaci.2013.05.033

Zhang, J., Yu, P., Zhou, Q., Li, X., Ding, S., Su, S., Zhang, X., Yang, X., Zhou, W., Wan, Q., & Gui, J. F. (2018). Screening and characterisation of sex differentiation-related long non-coding RNAs in Chinese soft-shell turtle (Pelodiscus sinensis). *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-26841-3

Zhang, N., & Bevan, M. J. (2011). CD8+ T Cells: Foot Soldiers of the Immune System. In *Immunity* (Vol. 35, Issue 2). https://doi.org/10.1016/j.immuni.2011.07.010

Zhang, S., Zhao, S., Han, X., Zhang, Y., Jin, X., Yuan, Y., Zhao, X., Luo, Y., Zhou, Y., Gao, Y., Yu, H., Sun, D., Xu, W., Yan, S., Gong, Y., & Li, Y. (2021). Lnc-C2orf63-4-1 Confers VSMC Homeostasis and Prevents Aortic Dissection Formation via

STAT3 Interaction. *Frontiers in Cell and Developmental Biology*, *9*. https://doi.org/10.3389/FCELL.2021.792051

Zhang, W., Xie, M., Shu, M. Di, Steitz, J. A., & DiMaio, D. (2016). A proximity-dependent assay for specific RNA-protein interactions in intact cells. *RNA*. https://doi.org/10.1261/rna.058248.116

Zhang, X., Hamblin, M. H., & Yin, K.-J. (2017). The long noncoding RNA Malat1: Its physiological and pathophysiological functions. *RNA Biology*, *14*(12), 1705–1714. https://doi.org/10.1080/15476286.2017.1358347

Zhang, Y., Gao, L., Ma, S., Ma, J., Wang, Y., Li, S., Hu, X., Han, S., Zhou, M., Zhou, L., & Ding, Z. (2019). MALAT1

-KTN1-EGFR regulatory axis promotes the development of cutaneous squamous cell carcinoma. *Cell Death and Differentiation*. https://doi.org/10.1038/s41418-019-0288-7

Zhang, Y., Wang, F., Chen, G., He, R., & Yang, L. (2019). LncRNA MALAT1 promotes osteoarthritis by modulating miR-150-5p/AKT3 axis. *Cell & Bioscience*, *9*(1). https://doi.org/10.1186/S13578-019-0302-2

Zhao, J., Fan, Y., Wang, K., Ni, X., Gu, J., Lu, H., Lu, Y., Lu, L., Dai, X., & Wang, X. (2015). LncRNA HULC affects the differentiation of Treg in HBV-related liver cirrhosis. *International Immunopharmacology*, *28*(2), 901–905. https://doi.org/10.1016/j.intimp.2015.04.028

Zhao, M., Wang, S., Li, Q., Ji, Q., Guo, P., & Liu, X. (2018). Malat1: A long non-coding RNA highly associated with human cancers (review). In *Oncology Letters* (Vol. 16, Issue 1). https://doi.org/10.3892/ol.2018.8613

Zhao, Y., Li, H., Fang, S., Kang, Y., Wu, W., Hao, Y., Li, Z., Bu, D., Sun, N., Zhang, M. Q., & Chen, R. (2016). NONCODE 2016: an informative and valuable data source of long non-coding RNAs. *Nucleic Acids Research*, *44*(D1), D203-8. https://doi.org/10.1093/nar/gkv1252

Zhao, Z., Chen, C., Liu, Y., & Wu, C. (2014). 17β-Estradiol treatment inhibits breast cell proliferation, migration and invasion by decreasing MALAT-1 RNA level. *Biochemical and Biophysical Research Communications*, *445*(2). https://doi.org/10.1016/j.bbrc.2014.02.006

Zheng, H. T., Shi, D. B., Wang, Y. W., Li, X. X., Xu, Y., Tripathi, P., Gu, W. L., Cai, G. X., & Cai, S. J. (2014). High expression of lncRNA MALAT1 suggests a biomarker of poor prognosis in colorectal cancer. *International Journal of Clinical and Experimental Pathology*, *7*(6).

Zhou, Q., Liu, L., Zhou, J., Chen, Y., Xie, D., Yao, Y., & Cui, D. (2021). Novel Insights Into MALAT1 Function as a MicroRNA Sponge in NSCLC. *Frontiers in Oncology*, *11*, 4426. https://doi.org/10.3389/FONC.2021.758653/BIBTEX

Zhu, Y., Xu, G., Yang, Y. T., Xu, Z., Chen, X., Shi, B., Xie, D., Lu, Z. J., & Wang, P. (2019). POSTAR2: Deciphering the post-Transcriptional regulatory logics. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gky830

Zielinski, J., Kilk, K., Peritz, T., Kannanayakal, T., Miyashiro, K. Y., Eiríksdóttir, E., Jochems, J., Langel, Û., & Eberwine, J. (2006). In vivo identification of ribonucleoprotein-RNA interactions. *Proceedings of the National Academy of*

Sciences of the United States of America, *103*(5), 1557. https://doi.org/10.1073/PNAS.0510611103

Zou, X., Wang, J., Qu, H., Lv, X. H., Shu, D. M., Wang, Y., Ji, J., He, Y. H., Luo, C. L., & Liu, D. W. (2020). Comprehensive analysis of miRNAs, lncRNAs, and mRNAs reveals potential players of sexually dimorphic and left-right asymmetry in chicken gonad during gonadal differentiation. *Poultry Science*, *99*(5), 2696–2707. https://doi.org/10.1016/j.psj.2019.10.019

Zuk, M., & McKean, K. A. (1996). Sex differences in parasite infections: Patterns and processes. *International Journal for Parasitology*, *26*(10), 1009–1024. https://doi.org/10.1016/s0020-7519(96)80001-4

Appendix

Supplementary Table 14: Sequence of primers used for RAP-MS pull down of *Malat1*

Start	Name	Sequence 5′-3′ (All 5′ biotinylated)	Delta start point	Overlap
2	>Primer53	TTTTTACTGCTCAATCTTTTTAATTAAAAACTTATCTGCGATTTCCTCGGGCTGAGTCTCCTGCCTCACGAGCTCAGCTGTGCTGCTCTACGCTGCTCTGCTCTCGCTGCCTGAATGCCT	0	0
47	>Primer183	TTTCGATAAGCTACTCTATTAGCTTAAGTTTAGAGTTCTAATTCTTTTTACTGCTCAATCTTTTTAATTAAAAACTTATCTGCGATTTCCTCGGGCTGAGTCTCCTGCCTCACGAGCTC	45	75
107	>Primer412	GTTTAAATAAGCCTTCATGTTATCTCTTAAGATCTTCTTAGATTATTAAGACTAAGTAATATTTCGATAAGCTACTCTATTAGCTTAAGTTTAGAGTTCTAATTCTTTTTACTGCTCAAT	60	60
137	>Primer95	TACTTTTCTCCTCATTTCCTTTTTCAAACTGTTTAAATAAGCCTTCATGTTATCTCTTAAGATCTTCTTAGATTATTAAGACTAAGTAATATTTCGATAAGCTACTCTATTAGCTTAAGT	30	90
182	>Primer229	AATCTCCTAAACTGCTCTGGTCAGCCTCCATTATACAGTACAAATACTTTTCTCCTCATTTCCTTTTTCAAACTGTTTAAATAAGCCTTCATGTTATCTCTTAAGATCTTCTTAGATTA	45	75
227	>Primer378	ACCTTCAAACTAGAACCTTTTAGAACTTCACAAAACCTCCCTTTACAATCTCCTAAACTGCTCTGGTCAGCCTCCATTATACAGTACAAATACTTTTCTCCTCATTTCCTTTTTCAAACT	45	75
272	>Primer116	ACTTAGATTCTATTTAGGTAACCTTCGTTTTAATCTACAAGGCCGACCTTCAAACTAGAACCTTTTAGAACTTCACAAAACCTCCCTTTACAATCTCCTAAACTGCTCTGGTCAGCCTCC	45	75
317	>Primer288	CCTCATTTTCAAACTATTCTCTACAACTTTACTGTTTTAAATGCCACTTAGATTCTATTTAGGTAACCTTCGTTTTAATCTACAAGGCCGACCTTCAAACTAGAACCTTTTAGAACTTCA	45	75
362	>Primer10	ACGGCCGTCAACTTAACCTACTTTTCTCAATCTTTTAAAACTACACCTCATTTTCAAACTATTCTCTACAACTTTACTGTTTTAAATGCCACTTAGATTCTATTTAGGTAACCTTCGTTT	45	75
407	>Primer169	CTCATGCCTTCAAACGTACATGCGCCAGTCGAAGGATTTTTATAACGGCCGTCAACTTAACCTACTTTTCTCAATCTTTTAAAACTACACCTCATTTTCAAACTATTCTCTACAACTTT	45	75

Start	Name	Sequence 5′-3′ (All 5′ biotinylated)	Delta start point	Overlap
452	>Primer335	CACCATCGCCCGGCTAGCCTAACACTTCCATCTTCCCTGTTTCCAACTCATGCCTTCAAACGTACATGCGCCAGTCGAAGGATTTTTATAACGGCCGTCAACTTAACCTACTTTTCTCAA	45	75
902	>Primer245	GCTTTAAAGATAATTTCCTTCTGACTTTATATCTTATCCAACCTTTTGGCCTCAATCTTATCTTCTTTGCCTATCTTGAATGCTGGCATCCAAAGTTGTCTTAACTTCCATTTTGTTTT	450	-330
947	>Primer374	AAAGACTCTTCTAAACTGCTCGCTCCATCAGAAATTTGAACTTATGGCTTTAAAGATAATTTCCTTCTGACTTTATATCTTATCCAACCTTTTGGCCTCAATCTTATCTTCTTTGCCTAT	45	75
992	>Primer134	CCTAGTAGCTTTGATTGCTAGCTTCAATCTTGTATGTGGCTGTCTAAAGACTCTTCTAAACTGCTCGCTCCATCAGAAATTTGAACTTATGGCTTTAAAGATAATTTCCTTCTGACTTTA	45	75
1037	>Primer295	CTTCTAACTCTTCAAAGGCATTCTGCCTTAACTTTTTACTTCAGTCCTAGTAGCTTTGATTGCTAGCTTCAATCTTGTATGTGGCTGTCTAAAGACTCTTCTAAACTGCTCGCTCCATCA	45	75
1082	>Primer49	TTGTCATCAAGCAAAATTAAGCTACAAGTTAAGGCTTTTAATATTCTTCTAACTCTTCAAAGGCATTCTGCCTTAACTTTTTACTTCAGTCCTAGTAGCTTTGATTGCTAGCTTCAATCT	45	75
1127	>Primer185	CAATGCAAAATGCTGACAATCTTGAAACTGTTATCAAAAGTCCTTTTGTCATCAAGCAAAATTAAGCTACAAGTTAAGGCTTTTAATATTCTTCTAACTCTTCAAAGGCATTCTGCCTTA	45	75
1172	>Primer355	GCCAATGCTAGTCGTTAGGATTTTAAAAGCACCTCAGCTCAAGTCCAATGCAAAATGCTGACAATCTTGAAACTGTTATCAAAAGTCCTTTTGTCATCAAGCAAAATTAAGCTACAAGTT	45	75
1217	>Primer83	TTCCTAGTTCACTGAATGCACTTCTGTGTAGACCTGGGTCAGCTGCCAATGCTAGTCGTTAGGATTTTAAAAGCACCTCAGCTCAAGTCCAATGCAAAATGCTGACAATCTTGAAACTG	45	75
1262	>Primer251	AGCTTCACCAAACTGGCTTCGGGACTCCTGTCTGCCGCTCCTGTCTTCCTAGTTCACTGAATGCACTTCTGTGTAGACCTGGGTCAGCTGCCAATGCTAGTCGTTAGGATTTTAAAAGC	45	75
1307	>Primer405	GCTATCTTCACCATGCACTGCTGCTGCTGGCTCCTCAGTCCTTCCTAGCTTCACCAAACTGGCTTCGGGACTCCTGTCTGCCGCTCCTGTCTTCCTAGTTCACTGAATGCACTTCTGTGT	45	75

Start	Name	Sequence 5′-3′ (All 5′ biotinylated)	Delta start point	Overlap
1352	>Primer143	GCTCCTTCTTCCTAGCTTCCTCCACCGAACCGCACTCTTTCCTGGGCTATCTTCACCATGCACTGCTGCTGCTGGCTCCTCAGTCCTTCCTAGCTTCACCAAACTGGCTTCGGGACTCCT	45	75
1397	>Primer282	CACCATCCTGGAACCCTTTCCCAGCTTCACCACCACATCCGTATGGCTCCTTCTTCCTAGCTTCCTCCACCGAACCGCACTCTTTCCTGGGCTATCTTCACCATGCACTGCTGCTGCTGG	45	75
1442	>Primer14	CAAGCCAAGCCGCCAGCTAGCTTCATCACCAACTCGCTCTCGCTCCACCATCCTGGAACCCTTTCCCAGCTTCACCACCACATCCGTATGGCTCCTTCTTCCTAGCTTCCTCCACCGAAC	45	75
1487	>Primer181	TCTATCCTCTCCACAATGCCTGCTCGCCTCCTCCGCGCAGTTGACAAGCCAAGCCGCCAGCTAGCTTCATCACCAACTCGCTCTCGCTCCACCATCCTGGAACCCTTTCCCAGCTTCAC	45	75
1532	>Primer341	CCTGCAGCCTTTCTTGCACACTGGCATGCTGGTCTAGGAGCCGCTATCTATCCTCTCCACAATGCCTGCTCGCCTCCTCCGCGCAGTTGACAAGCCAAGCCGCCAGCTAGCTTCATCACC	45	75
1577	>Primer59	CCATGTTGCCGACCTCAAGGAATGTTACCGCACCGCATGCTCTCCCTGCAGCCTTTCTTGCACACTGGCATGCTGGTCTAGGAGCCGCTATCTATCCTCTCCACAATGCCTGCTCGCCT	45	75
1622	>Primer250	AAGACAAAGTAAACAAGTTACCATCCAAGTTACAGAAAACCACCACCATGTTGCCGACCTCAAGGAATGTTACCGCACCGCATGCTCTCCCTGCAGCCTTTCTTGCACACTGGCATGCT	45	75
1667	>Primer382	AGATATATCTCTTTACACAGAAGCCTACAACTCCCCCATAACTATTAAGACAAAGTAAACAAGTTACCATCCAAGTTACAGAAAACCACCACCATGTTGCCGACCTCAAGGAATGTTACC	45	75
1712	>Primer107	AAAAAGAAAAACCTACAACACCCGCAAAGGCCTACATACAGCCCCAGATATATCTCTTTACACAGAAGCCTACAACTCCCCCATAACTATTAAGACAAAGTAAACAAGTTACCATCCAAG	45	75
1757	>Primer285	CAGCCCTCAAAAGCTTCTGACAAGATGCAAGAGGACATAACCCTGAAAAAGAAAAACCTACAACACCCGCAAAGGCCTACATACAGCCCCAGATATATCTCTTTACACAGAAGCCTACAA	45	75
1802	>Primer4	AAGACAACTTGCCATCTACCATTCTTCTTTCTGGGCCTTGGCAGTCAGCCCTCAAAAGCTTCTGACAAGATGCAAGAGGACATAACCCTGAAAAAGAAAAACCTACAACACCCGCAAAGG	45	75

Start	Name	Sequence 5′-3′ (All 5′ biotinylated)	Delta start point	Overlap
1847	>Primer198	GGGAGGTTGTGCTGGCTCTACCATTCATTCCCCTCTGAGCGGTTAAAGACAACTTGCCATCTACCATTCTTCTTTCTGGGCCTTGGCAGTCAGCCCTCAAAAGCTTCTGACAAGATGCAA	45	75
1892	>Primer357	GCTTGTGGTAGGTCATCTGTTCAAACTACAACGTCTTACAAAACTGGGAGGTTGTGCTGGCTCTACCATTCATTCCCCTCTGAGCGGTTAAAGACAACTTGCCATCTACCATTCTTCTTT	45	75
1937	>Primer77	CCCAAAAGCACTTTGCCCAATTACCTCCCCTACACAGGAGTGAGGCTTGTGGTAGGTCATCTGTTCAAACTACAACGTCTTACAAAACTGGGAGGTTGTGCTGGCTCTACCATTCATTC	45	75
1982	>Primer231	AGACCTAAGGGGAAAAGAACTCAAAATATATTTTGCCCCCATTCCCCCAAAAGCACTTTGCCCAATTACCTCCCCTACACAGGAGTGAGGCTTGTGGTAGGTCATCTGTTCAAACTACA	45	75
2027	>Primer393	CCTTTTTTCTGGTTCCCTTGAGTCATCTGCCTTTAGGATTCTAGACAGACCTAAGGGGAAAAGAACTCAAAATATATTTTGCCCCCATTCCCCCAAAAGCACTTTGCCCAATTACCTCCC	45	75
2072	>Primer129	CAAACTGTAAACCTGGCGAGCTCTGCTTATCCTGAGAGTGGATTTCCTTTTTTCTGGTTCCCTTGAGTCATCTGCCTTTAGGATTCTAGACAGACCTAAGGGGAAAAGAACTCAAAATAT	45	75
2117	>Primer289	CCTCCACACTCAGTGTGAAAGCTAGCATCCATCCTCTACTTCCTACAAACTGTAAACCTGGCGAGCTCTGCTTATCCTGAGAGTGGATTTCCTTTTTTCTGGTTCCCTTGAGTCATCTGC	45	75
2162	>Primer31	GGAGGGGGAAAGAGTAAACTACCAGCAATTCCGCCATGGCCAGCTCCTCCACACTCAGTGTGAAAGCTAGCATCCATCCTCTACTTCCTACAAACTGTAAACCTGGCGAGCTCTGCTTAT	45	75
2207	>Primer165	ATATTTCAAGTAAAAGTGTTTAGGATTTTACAAATCTCATTAAGGGAGGGGGAAAGAGTAAACTACCAGCAATTCCGCCATGGCCAGCTCCTCCACACTCAGTGTGAAAGCTAGCATCC	45	75
2252	>Primer319	AAAAACGTTTCCCCCACCCACTCCTCCCTGTTAAGACCACTCCCAAATATTTCAAGTAAAAGTGTTTAGGATTTTACAAATCTCATTAAGGGAGGGGGAAAGAGTAAACTACCAGCAATT	45	75
2297	>Primer74	CAGTGTGTCAACACAACTATAGCATCTGTGGAAAATCTTAGAAAAAAAACGTTTCCCCCACCCACTCCTCCCTGTTAAGACCACTCCCAAATATTTCAAGTAAAAGTGTTTAGGATTTT	45	75

Start	Name	Sequence 5′-3′ (All 5′ biotinylated)	Delta start point	Overlap
2342	>Primer256	GAAGGTGTCGTGCCAACAGCATAGCAGTACACGCCTTCTCTAACCCAGTGTGTCAACACAACTATAGCATCTGTGGAAAATCTTAGAAAAAAAACGTTTCCCCCACCCACTCCTCCCTG	45	75
2387	>Primer401	GCAACTGGGAAAACTCTTCCAAGGACAAAAGGCAGCTCCAGTCCCTGAAGGTGTCGTGCCAACAGCATAGCAGTACACGCCTTCTCTAACCCAGTGTGTCAACACAACTATAGCATCTGT	45	75
2432	>Primer136	CCTGAGGTGACTGTGAACCAAAGCCGCACTGTGCTGACTTCAGCGGCAACTGGGAAAACTCTTCCAAGGACAAAAGGCAGCTCCAGTCCCTGAAGGTGTCGTGCCAACAGCATAGCAGTA	45	75
2477	>Primer267	CTTAACTTCAACCTCTGGCCTAGCCAAGCTCCTGAGGTTCTCCTGAGGTGACTGTGAACCAAAGCCGCACTGTGCTGACTTCAGCGGCAACTGGGAAAACTCTTCCAAGGACAAAA	45	75
2522	>Primer51	TTTTACCCCTCCCCTTTAATGAAATATTTTAAATCACGGTGCTGTAAAACTTAACTTCAACCTCTGGCCTAGCCAAGCTCCTGAGGTTCTCCTGAGGTGACTGTGAACCAAAGCCGCACT	45	75
2567	>Primer188	CCTAACACCCCCACCCACCCAAACACAAGGCCACAGCCAACTAAGTTTTACCCCTCCCCTTTAATGAAATATTTTAAATCACGGTGCTGTAAAACTTAACTTCAACCTCTGGCCTAGCCA	45	75
2612	>Primer361	TAAGTTCTCTGGTATGATTATCTGAAATCATAAACTAAACAATTACCTAACACCCCCACCCACCCAAACACAAGGCCACAGCCAACTAAGTTTTACCCCTCCCCTTTAATGAAATATTTT	45	75
2657	>Primer94	TACTTGCCAACTTGAAAGCTGAGATTTCCTGTTTTTCCAAATATTTAAGTTCTCTGGTATGATTATCTGAAATCATAAACTAAACAATTACCTAACACCCCCACCCACCCAAACACAAGG	45	75
2702	>Primer226	AAAAAAAGAAAAAGGAAAAAAGAAGCAAAAACTGGATTGGGAGTTACTTGCCAACTTGAAAGCTGAGATTTCCTGTTTTTCCAAATATTTAAGTTCTCTGGTATGATTATCTGAAATCA	45	75
2747	>Primer421	TTCGAGGGACCGGCAGAGGAACCAACCTTCCTTAGCTGCCCGCCTCAAAAAAAGAAAAAGGAAAAAAGAAGCAAAAACTGGATTGGGAGTTACTTGCCAACTTGAAAGCTGAGATTTCCT	45	75
2792	>Primer123	ATCACATCATCCCAGTGGACCAGACCAACCCCCAAGCCCTACGCTTTCGAGGGACCGGCAGAGGAACCAACCTTCCTTAGCTGCCCGCCTCAAAAAAAGAAAAAGGAAAAAAGAAGCAAA	45	75

Start	Name	Sequence 5′-3′ (All 5′ biotinylated)	Delta start point	Overlap
2837	>Primer280	CAATCTATATTCATCCAACAGCTTCAGAAGAGTCCCCACTGTAGCATCACATCATCCCAGTGGACCAGACCAACCCCCAAGCCCTACGCTTTCGAGGGACCGGCAGAGGAACCAACCTTC	45	75
2882	>Primer9	ACATTAAGTCACCTGAAAAAAATTTCAAAAGAGAACCACACACTACAATCTATATTCATCCAACAGCTTCAGAAGAGTCCCCACTGTAGCATCACATCATCCCAGTGGACCAGACCAACC	45	75
2927	>Primer199	GGTCATTAAAGCCACTTCCTTTGTTCCTATAGTAGTTATTAAGATACATTAAGTCACCTGAAAAAAATTTCAAAAGAGAACCACACACTACAATCTATATTCATCCAACAGCTTCAGAAG	45	75
2972	>Primer348	CTTACTTGATAATATAAAAGCTATCACCAGAAGAAATTCCTTCAGGGTCATTAAAGCCACTTCCTTTGTTCCTATAGTAGTTATTAAGATACATTAAGTCACCTGAAAAAAATTTCAAAA	45	75
3017	>Primer92	TACACTAGGAAAAAGACTTGCTTATACAAAACTGAGATAGTATCTCTTACTTGATAATATAAAAGCTATCACCAGAAGAAATTCCTTCAGGGTCATTAAAGCCACTTCCTTTGTTCCTAT	45	75
3062	>Primer244	CATACCTTTACATCTTGTTTAGTCACAAGGAAAATCATTTCTCCTACACTAGGAAAAAGACTTGCTTATACAAAACTGAGATAGTATCTCTTACTTGATAATATAAAAGCTATCACCAG	45	75
3107	>Primer376	AAGTTACAAATAAACACAAGTATACAATGCACAAGAAGAAAAAAAGCATACCTTTACATCTTGTTTAGTCACAAGGAAAATCATTTCTCCTACACTAGGAAAAAGACTTGCTTATACAAA	45	75
3152	>Primer135	CCTCTAAAAGACATTCAGGCTGAATTATCATAATTCTTAAATTATAAGTTACAAATAAACACAAGTATACAATGCACAAGAAGAAAAAAAGCATACCTTTACATCTTGTTTAGTCACAAG	45	75
3197	>Primer307	TACAGAAAAAAAAAGGTTTCCCCTCCCTCATCAACAAAAGCCCACCCTCTAAAAGACATTCAGGCTGAATTATCATAATTCTTAAATTATAAGTTACAAATAAACACAAGTATACAATGC	45	75
3242	>Primer26	CTGCCAGGCTGGTTATGACTCAGATGGTGTTATCTGAAAAAGGTCTACAGAAAAAAAAAGGTTTCCCCTCCCTCATCAACAAAAGCCCACCCTCTAAAAGACATTCAGGCTGAATTATCA	45	75
3287	>Primer170	CTCATTCACCAAGGAGCTGCTCCCTCTGCATCTACGTCATCACACTGCCAGGCTGGTTATGACTCAGATGGTGTTATCTGAAAAAGGTCTACAGAAAAAAAAAGGTTTCCCCTCCCTCA	45	75

Start	Name	Sequence 5′-3′ (All 5′ biotinylated)	Delta start point	Overlap
3332	>Primer332	ATTCCCCATGGAGACATGACATTATTTTTTCTGCCTTTACTTATCACTCATTCACCAAGGAGCTGCTCCCTCTGCATCTACGTCATCACACTGCCAGGCTGGTTATGACTCAGATGGTGT	45	75
3377	>Primer72	ATATGCAGCTTTTCATCAGTAGGAACAATCTCTGGCTCATGCTCATTCCCCATGGAGACATGACATTATTTTTTCTGCCTTTACTTATCACTCATTCACCAAGGAGCTGCTCCCTCTGC	45	75
3422	>Primer265	TGCCATAACTTTATACTGGTTGCTTTCATTTGCTTAAATTTTTGCATATGCAGCTTTTCATCAGTAGGAACAATCTCTGGCTCATGCTCATTCCCCATGGAGACATGACATTATTTTTT	45	75
3467	>Primer423	TTTTGTGGATAAAGCTTGGTAGATAAGCCATAACTTTTAAAGGTATTGCCATAACTTTATACTGGTTGCTTTCATTTGCTTAAATTTTTGCATATGCAGCTTTTCATCAGTAGGAACAAT	45	75
3512	>Primer158	TTGAGATGAACATTTGATCTTCACTGTTTTTCATCAATTCTTTACTTTTGTGGATAAAGCTTGGTAGATAAGCCATAACTTTTAAAGGTATTGCCATAACTTTATACTGGTTGCTTTCAT	45	75
3557	>Primer303	GCAGCATTTTCACTTCATTTCTATTCTGCTTTTGTAAAAGCAGTTTTGAGATGAACATTTGATCTTCACTGTTTTTCATCAATTCTTTACTTTTGTGGATAAAGCTTGGTAGATAAGCCA	45	75
3602	>Primer19	CCACTCATCTCAACAAGCTCAGCTTCTTTTTACTCCAGGCTTAATGCAGCATTTTCACTTCATTTCTATTCTGCTTTTGTAAAAGCAGTTTTGAGATGAACATTTGATCTTCACTGTTTT	45	75
3647	>Primer184	CAAACGAAACATTGGCACACTGCACCGCCTCGCAGCCGCTCGATCCCACTCATCTCAACAAGCTCAGCTTCTTTTTACTCCAGGCTTAATGCAGCATTTTCACTTCATTTCTATTCTGCT	45	75
3692	>Primer358	GGAATGAAGCAACTCTTCTGCTTATGAAGAGAAACCTGTCTGAGGCAAACGAAACATTGGCACACTGCACCGCCTCGCAGCCGCTCGATCCCACTCATCTCAACAAGCTCAGCTTCTTTT	45	75
3737	>Primer93	TACACTTATCTGTCAACAGCAGTCTGCTGTTTCCTGCTCCGAGATGGAATGAAGCAACTCTTCTGCTTATGAAGAGAAACCTGTCTGAGGCAAACGAAACATTGGCACACTGCACCGCCT	45	75
3782	>Primer213	AAGGCACTTATCTATCCCTAACATGCAATACTGCAGATCCAAGTTACACTTATCTGTCAACAGCAGTCTGCTGTTTCCTGCTCCGAGATGGAATGAAGCAACTCTTCTGCTTATGAAGA	45	75

Start	Name	Sequence 5′-3′ (All 5′ biotinylated)	Delta start point	Overlap
3827	>Primer415	TGCAAACATTCAACAGCTCTACAGGTCTTTTTGGAAAAAGAGAAAAAAGGCACTTATCTATCCCTAACATGCAATACTGCAGATCCAAGTTACACTTATCTGTCAACAGCAGTCTGCTGT	45	75
3872	>Primer125	ATGGGAAAACTACTCAAAATTCTGAACTGCCTAAGAGGGGCCAGCTGCAAACATTCAACAGCTCTACAGGTCTTTTTGGAAAAAGAGAAAAAAGGCACTTATCTATCCCTAACATGCAAT	45	75
3917	>Primer275	ATGAAAGCCCATCGGTGCAAGGCTTAGGAATTTTTAAGAGGCTGGATGGGAAAACTACTCAAAATTCTGAACTGCCTAAGAGGGGCCAGCTGCAAACATTCAACAGCTCTACAGGTCTTT	45	75
3962	>Primer50	TTTGTGTTGAAGTTTACGATGCAAAAGCCTATTAGCTATCCCATCATGAAAGCCCATCGGTGCAAGGCTTAGGAATTTTTAAGAGGCTGGATGGGAAAACTACTCAAAATTCTGAACTGC	45	75
4007	>Primer182	TCTTGTAAGCAATGTAATTAAAGTAGGCATTAATCATGTAGGCTTTTGTGTTGAAGTTTACGATGCAAAAGCCTATTAGCTATCCCATCATGAAAGCCCATCGGTGCAAGGCTTAGGAA	45	75
4052	>Primer328	ACATAATGATCCCTTTCATGGGGTCTTCAAGATAAAGATTCCTTAATCTTGTAAGCAATGTAATTAAAGTAGGCATTAATCATGTAGGCTTTTGTGTTGAAGTTTACGATGCAAAAGCCT	45	75
4097	>Primer101	TGGAGCACATTTAAATTTTAGCAATATGAACATCTAATTTTCAGCACATAATGATCCCTTTCATGGGGTCTTCAAGATAAAGATTCCTTAATCTTGTAAGCAATGTAATTAAAGTAGGCA	45	75
4142	>Primer225	TTTTATTAATTTGTATAATTTAATGATTTTAAGCACAAGTACATTGGAGCACATTTAAATTTTAGCAATATGAACATCTAATTTTCAGCACATAATGATCCCTTTCATGGGGTCTTCAA	45	75
4187	>Primer375	AAATAAGGAGACAGCCTTCTAAATACATACATTCTCTAGTGAAGTATTTTATTAATTTGTATAATTTAATGATTTTAAGCACAAGTACATTGGAGCACATTTAAATTTTAGCAATATGAA	45	75
4232	>Primer128	CAAAATTGCCCACACTAACTACAGACAACAAACAAGACTTTATTTAAATAAGGAGACAGCCTTCTAAATACATACATTCTCTAGTGAAGTATTTTATTAATTTGTATAATTTAATGATTT	45	75
4277	>Primer314	TGTTCGAAGTCAAGTTTCTGAAAAGATTAGAGAAGAACATCCCCCCAAAATTGCCCACACTAACTACAGACAACAAACAAGACTTTATTTAAATAAGGAGACAGCCTTCTAAATACATAC	45	75

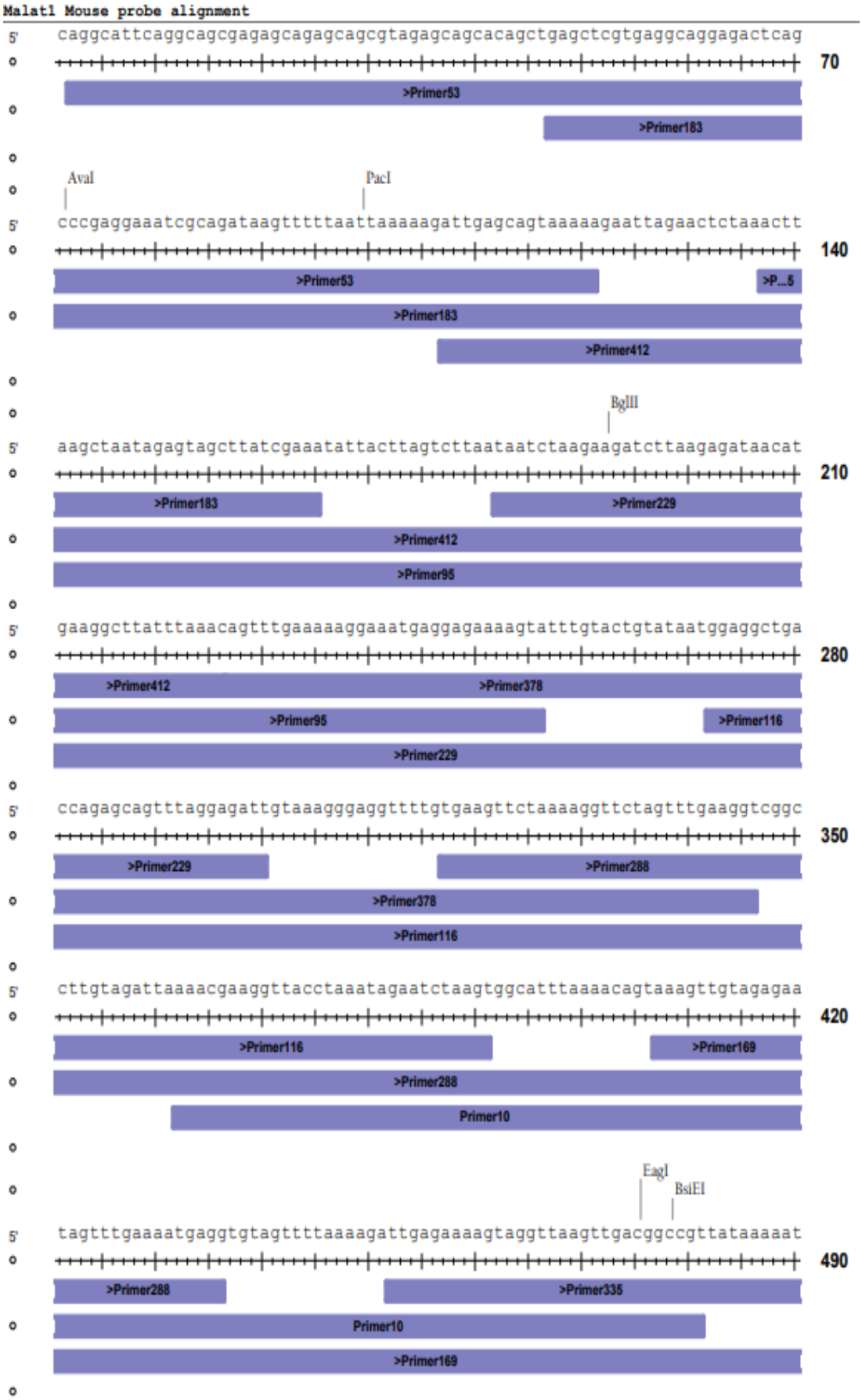
Start	Name	Sequence 5′-3′ (All 5′ biotinylated)	Delta start point	Overlap
4322	>Primer47	TTAATTTCGGTCTTCTGGCTCAAATCCTGATCTGGTCCACTTAAGTGTTCGAAGTCAAGTTTCTGAAAAGATTAGAGAAGAACATCCCCCCAAAATTGCCCACACTAACTACAGACAACA	45	75
4367	>Primer190	CTCATCACTGCATGGAGAATAAAATTTGTCTTTCCTGCCTTAAAGTTAATTTCGGTCTTCTGGCTCAAATCCTGATCTGGTCCACTTAAGTGTTCGAAGTCAAGTTTCTGAAAAGATTAG	45	75
4412	>Primer366	TCCTTAGTTAGACGGCCTCTATGCCAGGCCTGCAATTATTAAATGCTCATCACTGCATGGAGAATAAAATTTGTCTTTCCTGCCTTAAAGTTAATTTCGGTCTTCTGGCTCAAATCCTGA	45	75
4457	>Primer60	CCTTTTACTCTGACCATCATCTCCCACCTGCCTAAGGTACTTAGTCCTTAGTTAGACGGCCTCTATGCCAGGCCTGCAATTATTAAATGCTCATCACTGCATGGAGAATAAAATTTGTC	45	75
4502	>Primer260	GGTCAAATTAGACCCCTGACTTTCTGGAAACAAAATATGTAGTTACCTTTTACTCTGACCATCATCTCCCACCTGCCTAAGGTACTTAGTCCTTAGTTAGACGGCCTCTATGCCAGGCC	45	75
4547	>Primer422	TTTGGAAATGTGGGGGAAAAGTGTCTTACCCTAGATGTTTAGCCATGGTCAAATTAGACCCCTGACTTTCTGGAAACAAAATATGTAGTTACCTTTTACTCTGACCATCATCTCCCACCT	45	75
4592	>Primer146	TAAAGTGGATGAGATGATCGTAAGCATTTAAACTCAACATGCATATTTGGAAATGTGGGGGAAAAGTGTCTTACCCTAGATGTTTAGCCATGGTCAAATTAGACCCCTGACTTTCTGGAA	45	75
4637	>Primer281	CACATGCCTGACCCCACTCGTGGCTCAAGTGAGGTGACAAAAGGCTAAAGTGGATGAGATGATCGTAAGCATTTAAACTCAACATGCATATTTGGAAATGTGGGGGAAAAGTGTCTTACC	45	75
4682	>Primer17	CATGAAGGATGAAATGAGGCTCTGCAAAGGAAAACTCTTTAAACCCACATGCCTGACCCCACTCGTGGCTCAAGTGAGGTGACAAAAGGCTAAAGTGGATGAGATGATCGTAAGCATTTA	45	75
4727	>Primer187	CAGAAGACAGAGGCAAGCGCTTATATGCAAAGTCCTGAGCAGCTCCATGAAGGATGAAATGAGGCTCTGCAAAGGAAAACTCTTTAAACCCACATGCCTGACCCCACTCGTGGCTCAAGT	45	75
4772	>Primer322	AACAAACTCACTGCAAGGTCTCACATCACACACTCACTAGCAGAACAGAAGACAGAGGCAAGCGCTTATATGCAAAGTCCTGAGCAGCTCCATGAAGGATGAAATGAGGCTCTGCAAAGG	45	75

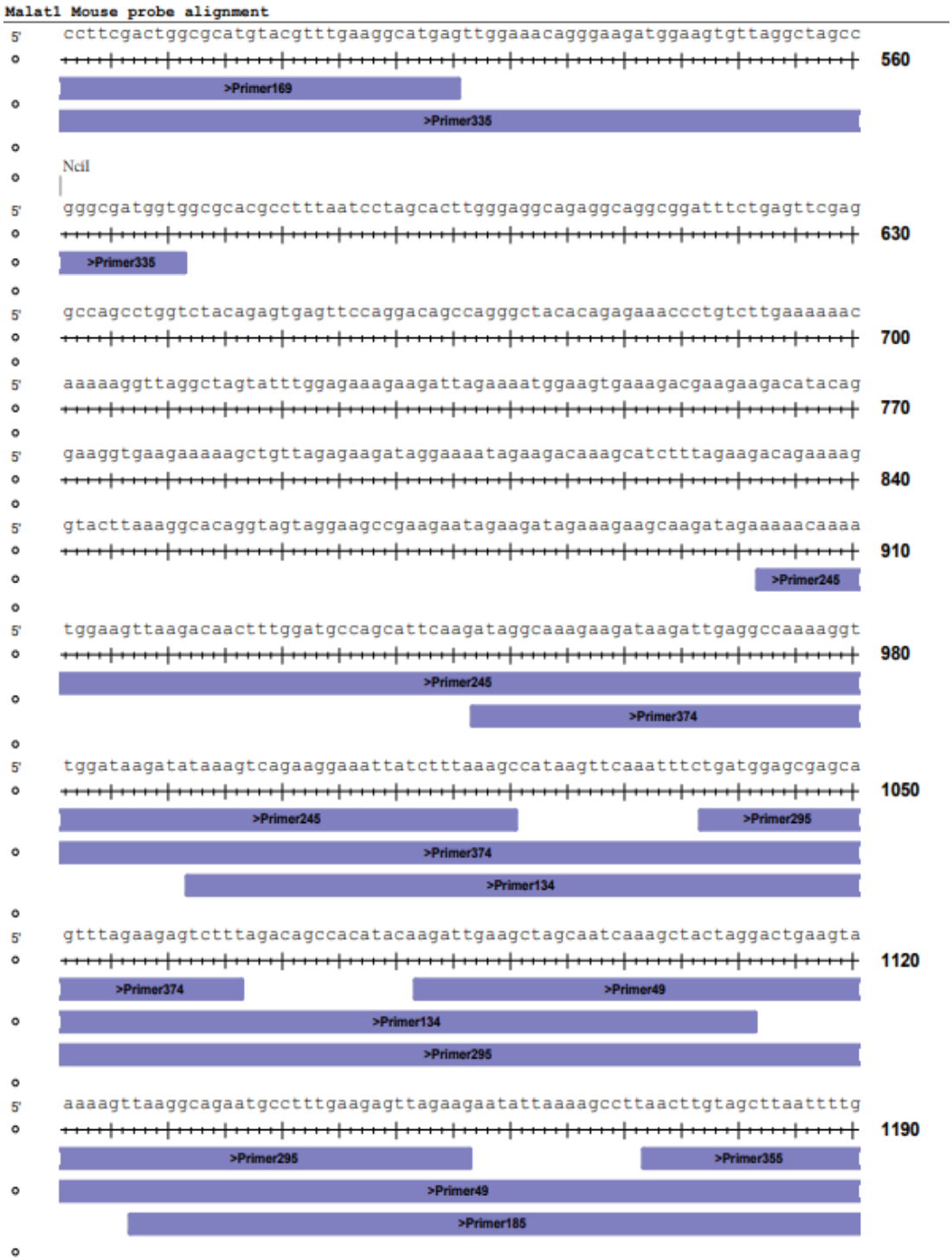
Start	Name	Sequence 5′-3′ (All 5′ biotinylated)	Delta start point	Overlap
4817	>Primer63	GAACAAGTAAGCCCCATCCCCCCCTCCCTCCACATTCCAGGAAAAACAAACTCACTGCAAGGTCTCACATCACACACTCACTAGCAGAACAGAAGACAGAGGCAAGCGCTTATATGCAA	45	75
4862	>Primer241	AAGTCAAGACACCTGCATTCTGTGTGGTCTGTAAAAAAAAAGCTAGAACAAGTAAGCCCCATCCCCCCCTCCCTCCACATTCCAGGAAAAACAAACTCACTGCAAGGTCTCACATCACA	45	75
4907	>Primer408	GGAACAGTACTGCACATATTACTTGCCAAAGAACAGACATGACCTGAAGTCAAGACACCTGCATTCTGTGTGGTCTGTAAAAAAAAAGCTAGAACAAGTAAGCCCCATCCCCCCCTCCCT	45	75
4952	>Primer124	ATCATACTCCAGTCGCGTCACAATGCATTCTAATAGCAGCAGATTGGAACAGTACTGCACATATTACTTGCCAAAGAACAGACATGACCTGAAGTCAAGACACCTGCATTCTGTGTGGTC	45	75
4997	>Primer279	CAACAACCACTACTCCAAACACTTGGGGAAACACAACTTTCTTTAATCATACTCCAGTCGCGTCACAATGCATTCTAATAGCAGCAGATTGGAACAGTACTGCACATATTACTTGCCAAA	45	75
5042	>Primer28	GAGCCATTTCCTCAACACTCAGCCTGTTACTCATGGCTTTTCCTCCAACAACCACTACTCCAAACACTTGGGGAAACACAACTTTCTTTAATCATACTCCAGTCGCGTCACAATGCATTC	45	75
5087	>Primer197	GGGACTCGGCTCCAATCACAAACACGGGTTACTTAAAGCTGCAGAGAGCCATTTCCTCAACACTCAGCCTGTTACTCATGGCTTTTCCTCCAACAACCACTACTCCAAACACTTGGGGAA	45	75
5132	>Primer364	TCACCAGAAATGAACAAAAACATTTACCTAAGGCAGCACAGCAAAGGGACTCGGCTCCAATCACAAACACGGGTTACTTAAAGCTGCAGAGAGCCATTTCCTCAACACTCAGCCTGTTAC	45	75
5177	>Primer90	GTTGAATCTGGAAGAGACTAAAGGCTTCAGTGCTCCCAACCCCCCTCACCAGAAATGAACAAAAACATTTACCTAAGGCAGCACAGCAAAGGGACTCGGCTCCAATCACAAACACGGGTT	45	75
5222	>Primer237	TTCAAGAATGTTGCTTGTCTGATTTATTTCTTGTCAGATTTTAAGTTGAATCTGGAAGAGACTAAAGGCTTCAGTGCTCCCAACCCCCCTCACCAGAAATGAACAAAAACATTTACCTA	45	75
5267	>Primer399	GACCACGGAACTGTTCAAAACATTTCCACTTGCCAGTTAAAATTTCTTCAAGAATGTTGCTTGTCTGATTTATTTCTTGTCAGATTTTAAGTTGAATCTGGAAGAGACTAAAGGCTTCAG	45	75

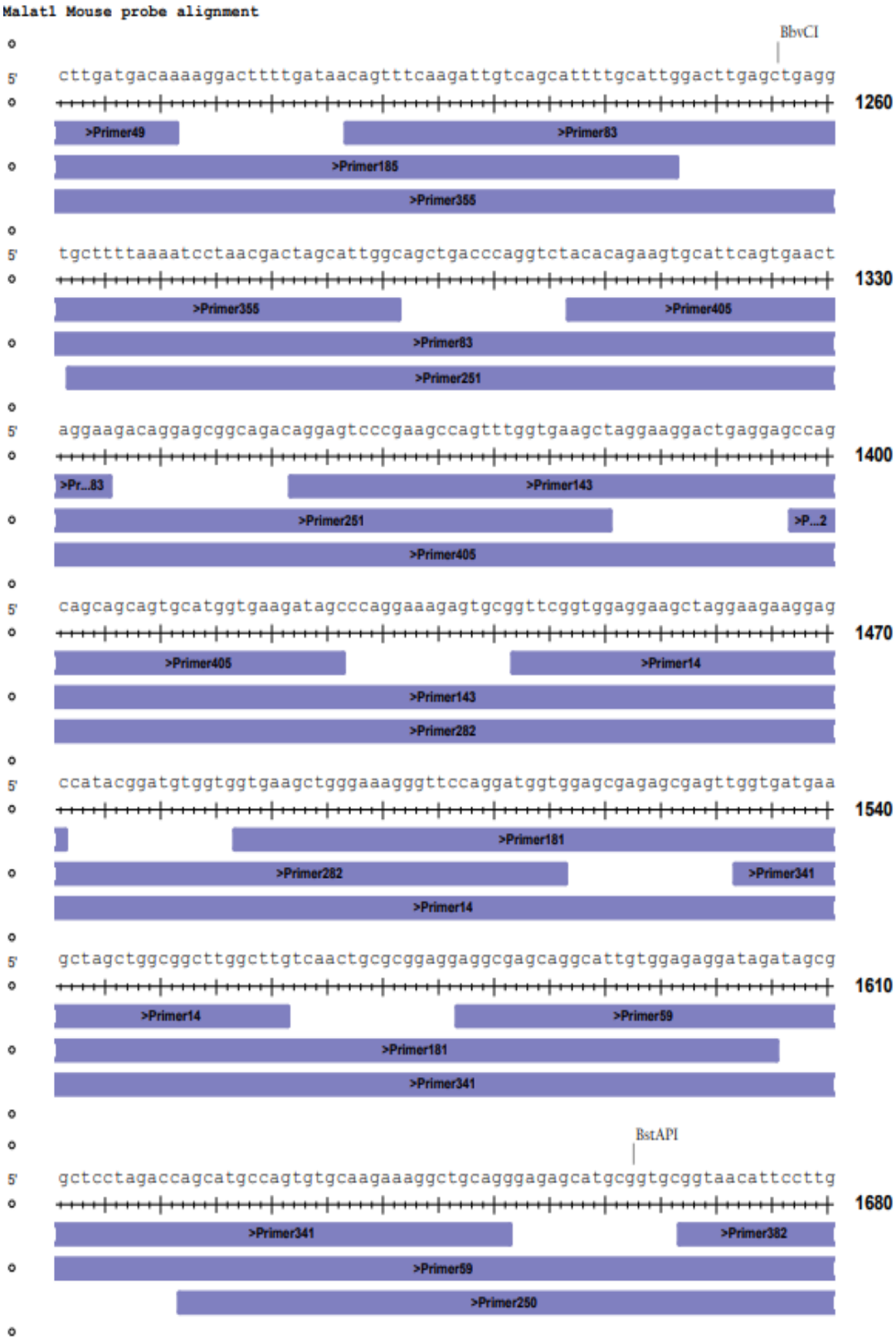
Start	Name	Sequence 5′-3′ (All 5′ biotinylated)	Delta start point	Overlap
5312	>Primer130	CAAGGGAGGGGAGAGAGAACACCTACACAAAGATAATGCACTAAAGACCACGGAACTGTTCAAAACATTTCCACTTGCCAGTTAAAATTTCTTCAAGAATGTTGCTTGTCTGATTTATTT	45	75
5357	>Primer308	TAGATGTCTGCTGTTGTCAATGTTCCTGCATGTAAGAATTAAGACCAAGGGAGGGGAGAGAGAACACCTACACAAAGATAATGCACTAAAGACCACGGAACTGTTCAAAACATTTCCACT	45	75
5417	>Primer105	TTTAAAGTAAATGGGCTATTTTTCCTTACTGGGTCTGGATTCTCTGGCCCCTTGAATAGATAGATGTCTGCTGTTGTCAATGTTCCTGCATGTAAGAATTAAGACCAAGGGAGGGGAGAG	60	60
5582	>Primer220	GCTAATCTTAAACAAGAAAAGGCTCGTTCACCTGTTGTCCTCATTTTGTCCACTGGTGAATTTCAATTCAACTGGAAGCTCCTTCTACAGTCTGAAGTACATCTGAAAGAACTAGGGCT	165	-45
5627	>Primer416	TGCTCCAGCATGAAGCTGGAGAGGATTCAACACTAGATTACCAGTAGCTAATCTTAAACAAGAAAAGGCTCGTTCACCTGTTGTCCTCATTTTGTCCACTGGTGAATTTCAATTCAACTG	45	75
5672	>Primer113	ACCTCACCCCTCCACCCCAAGGCCAACATTACATCACATGCTAGCTGCTCCAGCATGAAGCTGGAGAGGATTCAACACTAGATTACCAGTAGCTAATCTTAAACAAGAAAAGGCTCGTTC	45	75
5717	>Primer315	TTAGTGAGGGGTACCTGAAAAATCTTAAAAAAAGGCTTAGCGCCCACCTCACCCCTCCACCCCAAGGCCAACATTACATCACATGCTAGCTGCTCCAGCATGAAGCTGGAGAGGATTCAA	45	75
5762	>Primer18	CCACACAGGAAGGCTCCGCTGTCCTACATTAAGCCTTCAGTGCCTTTAGTGAGGGGTACCTGAAAAATCTTAAAAAAAGGCTTAGCGCCCACCTCACCCCTCCACCCCAAGGCCAACATT	45	75
5807	>Primer203	TGATACCACTTTGGTCTCGATACAATACTGCTTGCTTGATTCTTGCCACACAGGAAGGCTCCGCTGTCCTACATTAAGCCTTCAGTGCCTTTAGTGAGGGGTACCTGAAAAATCTTAAAA	45	75
5852	>Primer329	AGGTGTTACGGTAGGGTAGTCCCCACTGCTAATCAAAACCGACCATGATACCACTTTGGTCTCGATACAATACTGCTTGCTTGATTCTTGCCACACAGGAAGGCTCCGCTGTCCTACATT	45	75
5897	>Primer76	CCAGGGCCTCTCAAGTATTTTCTTTGGATGCTTCAATTCCAACAAGGTGTTACGGTAGGGTAGTCCCCACTGCTAATCAAAACCGACCATGATACCACTTTGGTCTCGATACAATACTG	45	75

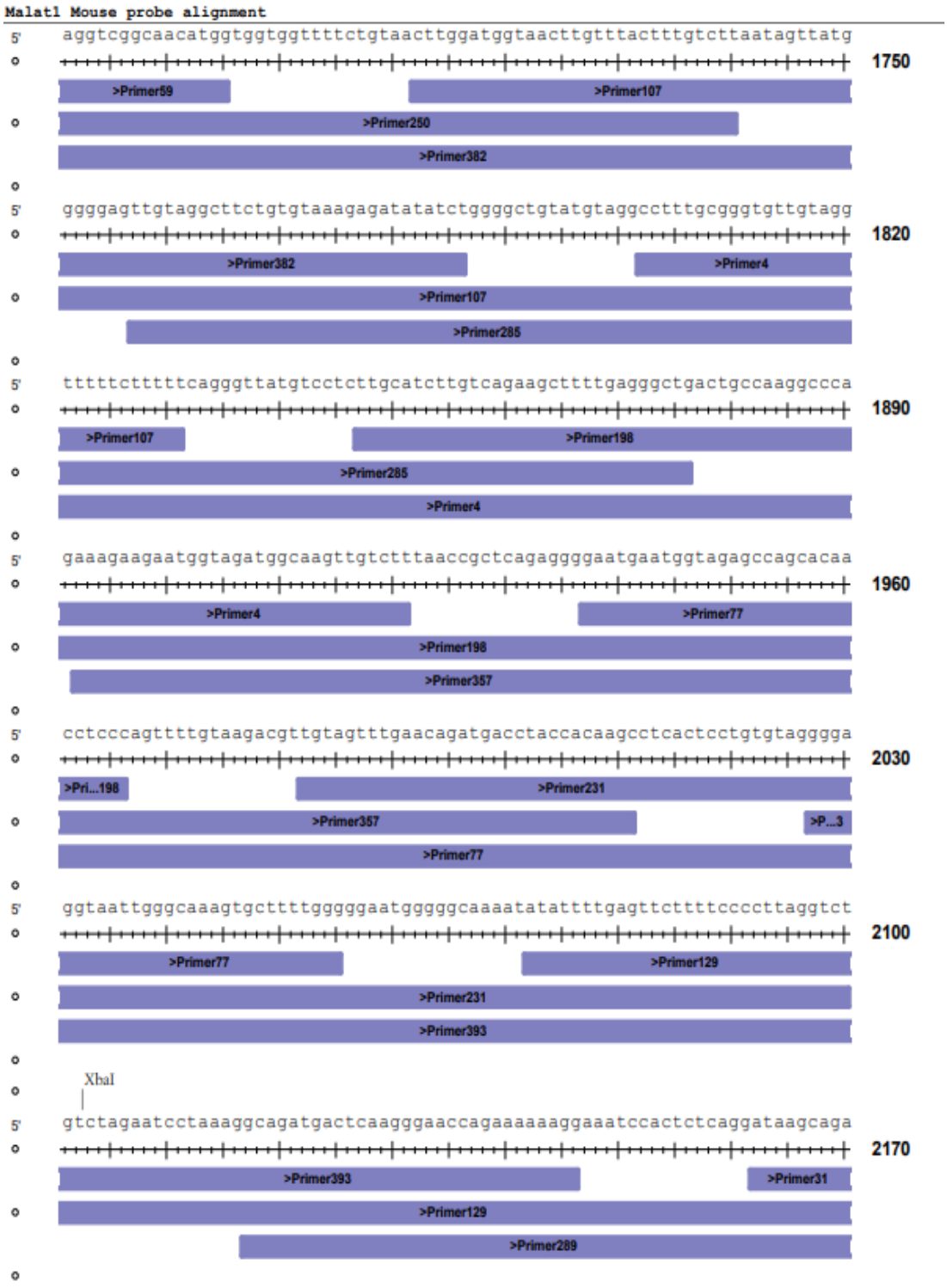
Start	Name	Sequence 5′-3′ (All 5′ biotinylated)	Delta start point	Overlap
5942	>Primer214	ACCGCTGCTATAAAAACAGCCTTTTTTCCAGATGTTAAAACAAGCCCAGGGCCTCTCAAGTATTTTCTTTGGATGCTTCAATTCCAACAAGGTGTTACGGTAGGGTAGTCCCCACTGCT	45	75
5987	>Primer391	CCCCTTTTCCCTCCCCTGCAAGCACAACTTGAGGTTTGGGCTGGTAACCGCTGCTATAAAAACAGCCTTTTTTCCAGATGTTAAAACAAGCCCAGGGCCTCTCAAGTATTTTCTTTGGAT	45	75
6032	>Primer144	GTAACAGGATTCTGGAAAAGCTGGGGAAACTGGTTGCCCGCTTTCCCCCTTTTCCCTCCCCTGCAAGCACAACTTGAGGTTTGGGCTGGTAACCGCTGCTATAAAAACAGCCTTTTTTCC	45	75
6077	>Primer287	CCCATGGTGGCGATGTGGCAGAGAAATCACTTGTGGGGAGACCTTGTAACAGGATTCTGGAAAAGCTGGGGAAACTGGTTGCCCGCTTTCCCCCTTTTCCCTCCCCTGCAAGCACAACTT	45	75
6122	>Primer30	GCTGCATCAAGGTGAGGGGTGAAGGGTCTGTGATTAGGCCAAAGGCCCATGGTGGCGATGTGGCAGAGAAATCACTTGTGGGGAGACCTTGTAACAGGATTCTGGAAAAGCTGGGGAAAC	45	75
6167	>Primer168	CCTTGAAACCGATATGCAACGTGACCTCAAGGATCCAGCTACTGGCTGCATCAAGGTGAGGGGTGAAGGGTCTGTGATTAGGCCAAAGGCCCATGGTGGCGATGTGGCAGAGAAATCAC	45	75
6212	>Primer354	GATGGCCTTTTCTGGTGCAACCCACAGGACCTTGGCACCATGGTTACCTTGAAACCGATATGCAACGTGACCTCAAGGATCCAGCTACTGGCTGCATCAAGGTGAGGGGTGAAGGGTCTG	45	75
6257	>Primer98	TCTTAGCTATGGTTTTAATGTTAAATTACAGGCAAGGGGAAAATTGATGGCCTTTTCTGGTGCAACCCACAGGACCTTGGCACCATGGTTACCTTGAAACCGATATGCAACGTGACCTCA	45	75
6302	>Primer234	CCATACTGGTTTGTTTACTCTGCATAGGTGCTATGTATAAAACATCTTAGCTATGGTTTTAATGTTAAATTACAGGCAAGGGGAAAATTGATGGCCTTTTCTGGTGCAACCCACAGGAC	45	75
6347	>Primer394	CGACTTCCTACATTCCCACCCAGCACTGGTATCAAACATACTATACCCATACTGGTTTGTTTACTCTGCATAGGTGCTATGTATAAAACATCTTAGCTATGGTTTTAATGTTAAATTACA	45	75
6392	>Primer153	TGCAATCCCACCCCAACAACTTCCTACAAAGGCTTGCTTTTCATCCGACTTCCTACATTCCCACCCAGCACTGGTATCAAACATACTATACCCATACTGGTTTGTTTACTCTGCATAGGT	45	75

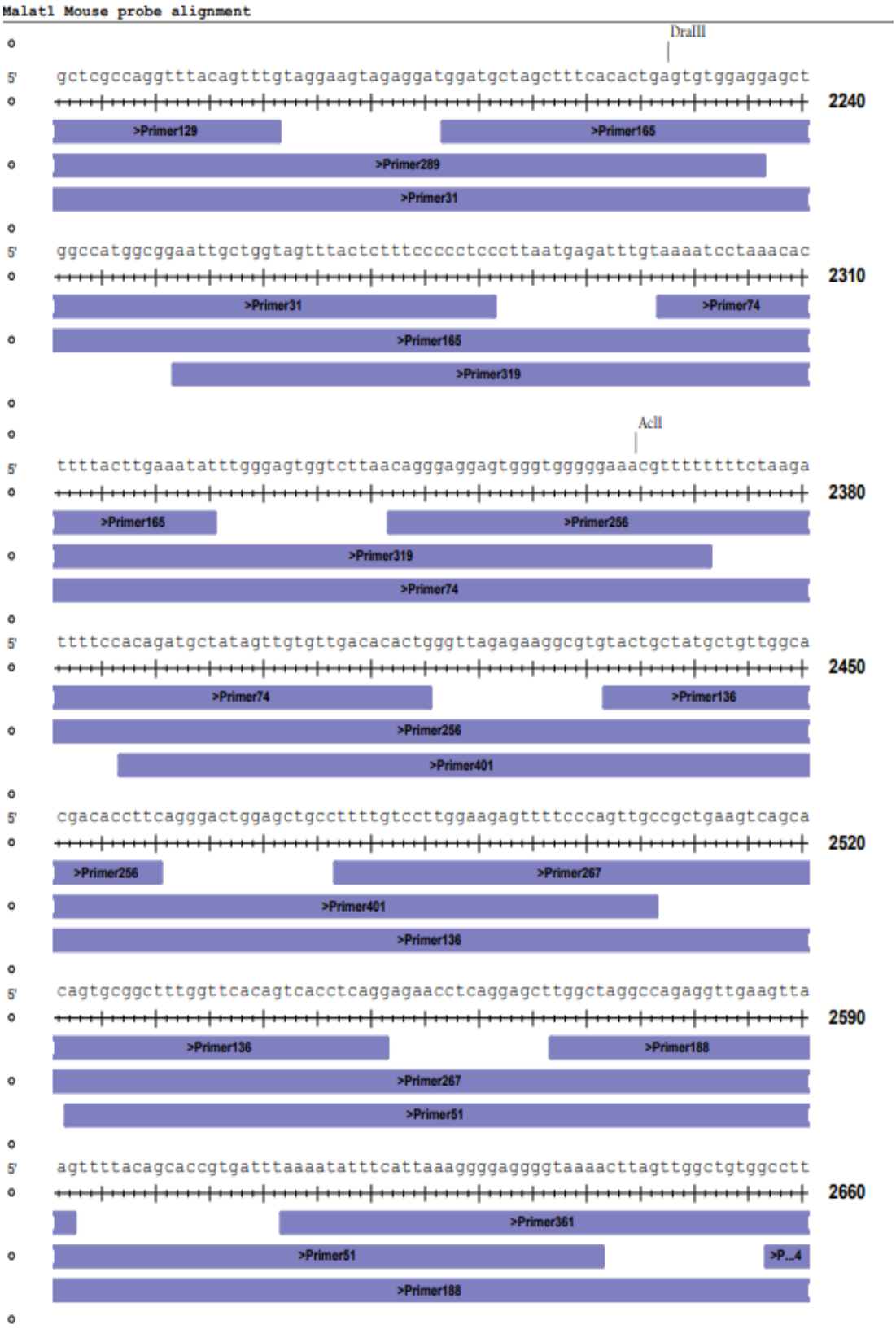
Start	Name	Sequence 5′-3′ (All 5′ biotinylated)	Delta start point	Overlap
6437	>Primer286	CCAAGTCTGTTATGTCCACCTGAAAAAGTCTTAGCAGAGAATTTTTGCAATCCCACCCCAACAACTTCCTACAAAGGCTTGCTTTTCATCCGACTTCCTACATTCCCACCCAGCACTGGT	45	75
6482	>Primer40	TACAACCCTACTGACGAATCTGCTTCCACTAAGATGCTAGCTTGGCCAAGTCTGTTATGTCCACCTGAAAAAGTCTTAGCAGAGAATTTTTGCAATCCCACCCCAACAACTTCCTACAAA	45	75
6527	>Primer167	CCTGAGAAAACAAAAGGTTGTTTTCTCAGGAAAAGAAAAACCTTTACAACCCTACTGACGAATCTGCTTCCACTAAGATGCTAGCTTGGCCAAGTCTGTTATGTCCACCTGAAAAAGTC	45	75
6572	>Primer371	TTTTGCTTTTTTTTTTTTAAAGCTAGGGAAAGGCCAAAAAGCAAAACCTGAGAAAACAAAAGGTTGTTTTCTCAGGAAAAGAAAAACCTTTACAACCCTACTGACGAATCTGCTTCCACT	45	75
6617	>Primer62	CTTGAACCCCGTCCTGGAAACCAGGAGTGCCAGCCACCAGCGTCTTTTGCTTTTTTTTTTTTAAAGCTAGGGAAAGGCCAAAAAGCAAAACCTGAGAAAACAAAAGGTTGTTTTCTCAG	45	75
6662	>Primer261	GTAATGGCCTCATGATATAAGAGTCAAGCAAAGACACCGCAGGGACTTGAACCCCGTCCTGGAAACCAGGAGTGCCAGCCACCAGCGTCTTTTGCTTTTTTTTTTTTAAAGCTAGGGAA	45	75
6707	>Primer384	ATATCAGCTACCATACCCAGAGCCTTTAGAACCCTCCAAGAAAAATGTAATGGCCTCATGATATAAGAGTCAAGCAAAGACACCGCAGGGACTTGAACCCCGTCCTGGAAACCAGGAGTG	45	75
6752	>Primer155	TTAATTATCAAGAGTTCAACACTGAGGCTGGGGAGTGTTCCAGTGATATCAGCTACCATACCCAGAGCCTTTAGAACCCTCCAAGAAAAATGTAATGGCCTCATGATATAAGAGTCAAGC	45	75
6797	>Primer284	CAGAGGTAATAAGACGAATTGGGCATAACCTGAAAGACAATGCAGTTAATTATCAAGAGTTCAACACTGAGGCTGGGGAGTGTTCCAGTGATATCAGCTACCATACCCAGAGCCTTTAGA	45	75
6842	>Primer42	TCATAAAATCAAAGTATTCAATAAATAGTAGGAGGTGTGTCGACTCAGAGGTAATAAGACGAATTGGGCATAACCTGAAAGACAATGCAGTTAATTATCAAGAGTTCAACACTGAGGCTG	45	75
6983	>Primer411	GGGTCTGGATTCTCTGGCCCCTTGAATAGATAGATGTCTGCTGTTGTCAATGTTCCTGCATGTAAGAATTAAGACCAAGGGAGGGGAGAGAGAACACCTACACAAAGATAATGCACTAAA	141	-21

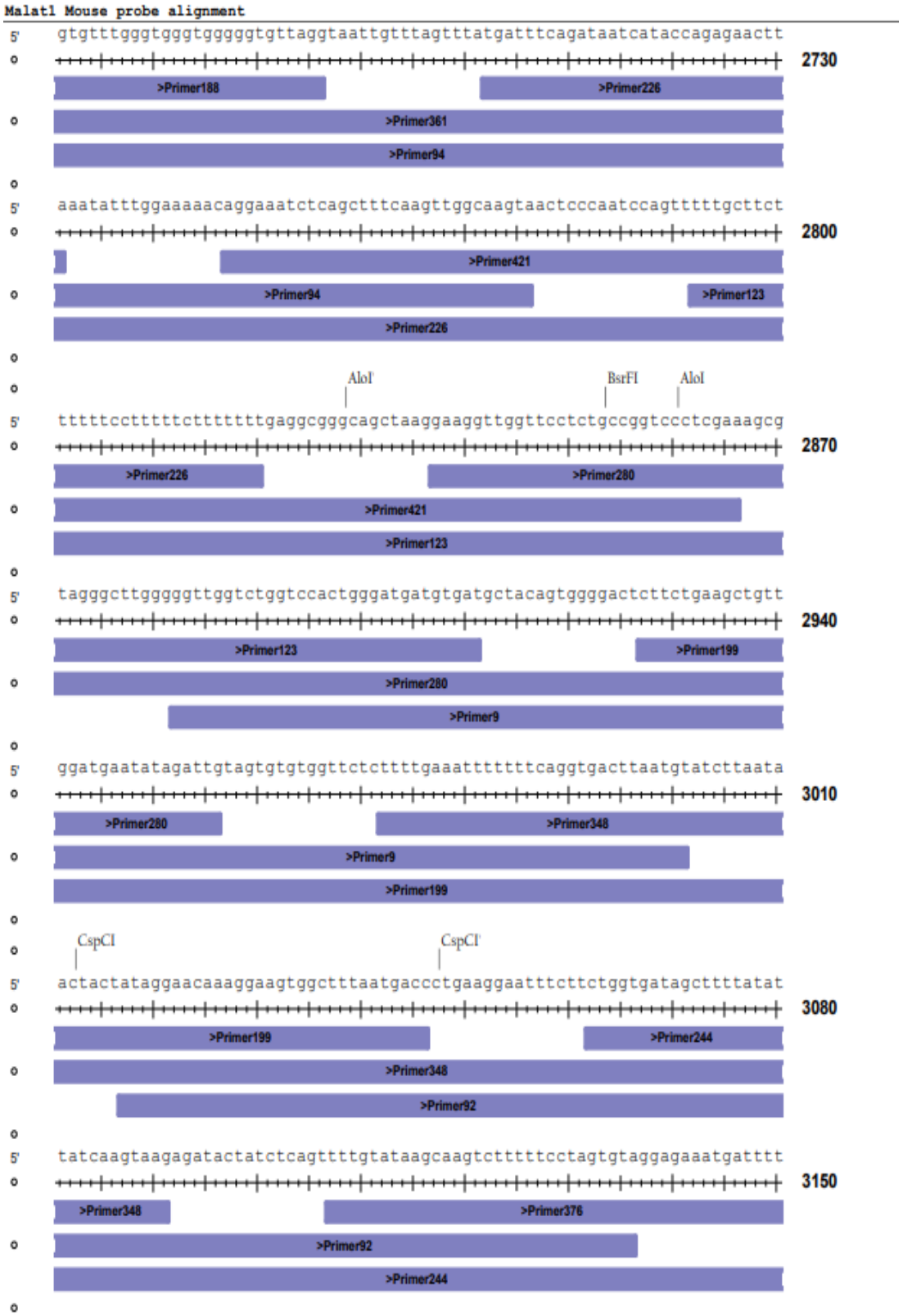

Malat1 Mouse probe alignment
caggcattcaggcagcgagagcagagcagcgtagagcagcacagctgagctcgtgaggcaggagactcag
70
>Primer53
>Primer183
AvaI
PacI
cccgaggaaatcgcagataagtttttaattaaaaagattgagcagtaaaaagaattagaactctaaacttt
140
>Primer53
>P...5
>Primer183
>Primer412
BglII
aagctaatagagtagcttatcgaaatattacttagtcttaataatctaagaagatcttaagagataacat
210
>Primer183
>Primer229
>Primer412
>Primer95
gaaggcttatttaaacagtttgaaaaaggaaatgaggagaaaagtatttgtactgtataatggaggctga
280
>Primer412
>Primer378
>Primer95
>Primer116
>Primer229
ccagagcagtttaggagattgtaaagggaggttttgtgaagttctaaaaggttctagtttgaaggtcggc
350
>Primer229
>Primer288
>Primer378
>Primer116
cttgtagattaaaacgaaggttacctaaatagaatctaagtggcatttaaaacagtaaagttgtagagaa
420
>Primer116
>Primer169
>Primer288
Primer10
EagI
BsiEI
tagtttgaaaatgaggtgtagttttaaaagattgagaaaagtaggttaagttgacggccgttataaaaat
490
>Primer288
>Primer335
Primer10
>Primer169

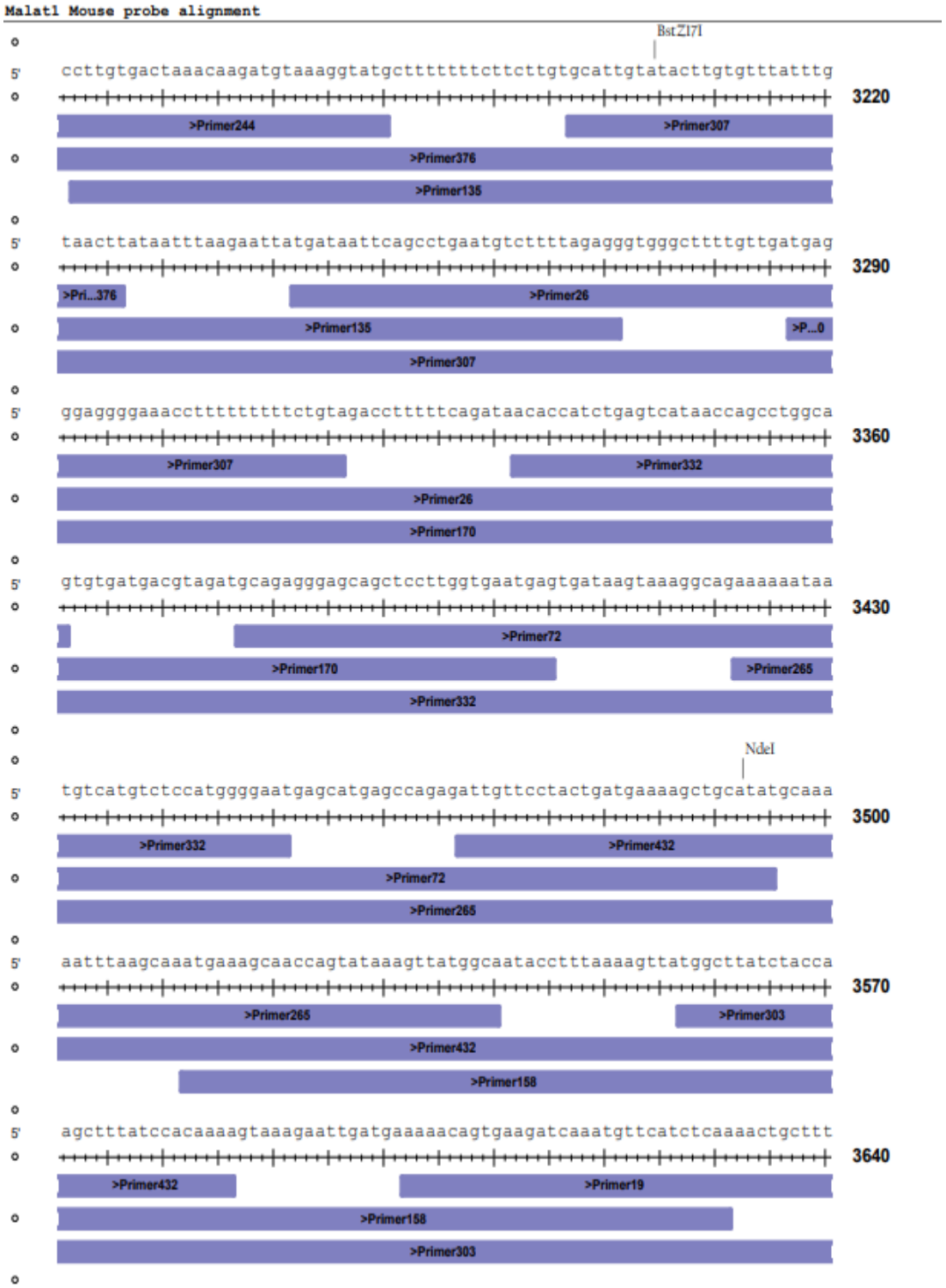

Malat1 Mouse probe alignment
5'
ccttcgactggcgcatgtacgtttgaaggcatgagttggaaacagggaagatggaagtgttaggctagcc
560
>Primer169
>Primer335
NciI
gggcgatggtggcgcacgcctttaatcctagcacttgggaggcagaggcaggcggatttctgagttcgag
630
>Primer335
gccagcctggtctacagagtgagttccaggacagccagggctacacagagaaaccctgtcttgaaaaaac
700
aaaaaggttaggctagtatttggagaaagaagattagaaaatggaagtgaaagacgaagaagacatacag
770
gaaggtgaagaaaaagctgttagagaagataggaaaatagaagacaaagcatctttagaagacagaaaag
840
gtacttaaaggcacaggtagtaggaagccgaagaatagaagatagaaagaagcaagatagaaaaacaaaa
910
>Primer245
tggaagttaagacaactttggatgccagcattcaagataggcaaagaagataagattgaggccaaaaggt
980
>Primer245
>Primer374
tggataagatataaagtcagaaggaaattatctttaaagccataagttcaaatttctgatggagcgagca
1050
>Primer245
>Primer295
>Primer374
>Primer134
gtttagaagagtctttagacagccacatacaagattgaagctagcaatcaaagctactaggactgaagta
1120
>Primer374
>Primer49
>Primer134
>Primer295
aaaagttaaggcagaatgcctttgaagagttagaagaatattaaaagccttaacttgtagcttaattttg
1190
>Primer295
>Primer355
>Primer49
>Primer185

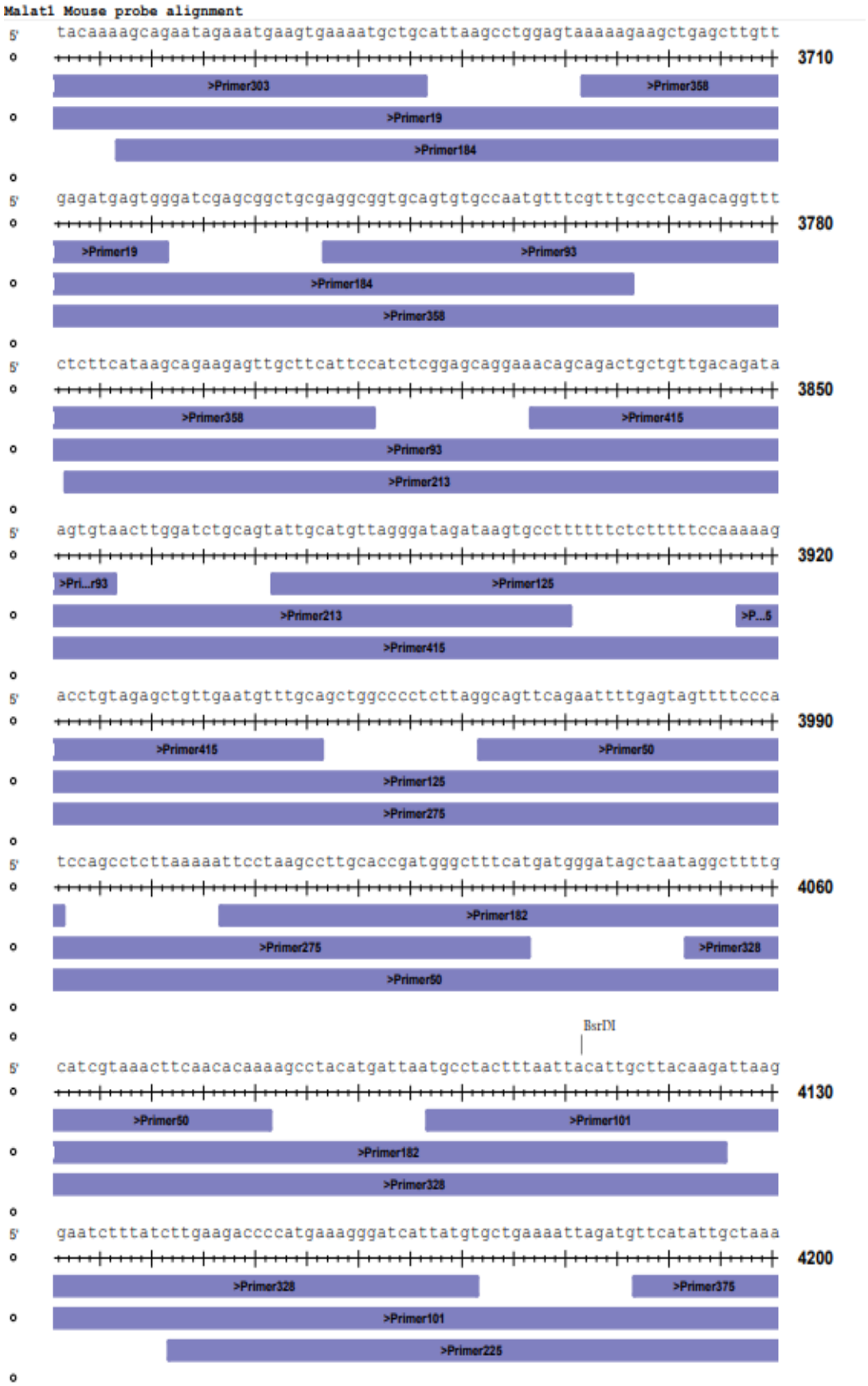

Malat1 Mouse probe alignment
BbvCI
5'
cttgatgacaaaaggacttttgataacagtttcaagattgtcagcattttgcattggacttgagctgagg
1260
>Primer49
>Primer83
>Primer185
>Primer355
5'
tgcttttaaaatcctaacgactagcattggcagctgacccaggtctacacagaagtgcattcagtgaact
1330
>Primer355
>Primer405
>Primer83
>Primer251
5'
aggaagacaggagcggcagacaggagtcccgaagccagtttggtgaagctaggaaggactgaggagccag
1400
>Pr...83
>Primer143
>Primer251
>P...2
>Primer405
5'
cagcagcagtgcatggtgaagatagcccaggaaagagtgcggttcggtggaggaagctaggaagaaggag
1470
>Primer405
>Primer14
>Primer143
>Primer282
5'
ccatacggatgtggtggtgaagctgggaaagggttccaggatggtggagcgagagcgagttggtgatgaa
1540
>Primer181
>Primer282
>Primer341
>Primer14
5'
gctagctggcggcttggcttgtcaactgcgcggaggaggcgagcaggcattgtggagaggatagatagcg
1610
>Primer14
>Primer59
>Primer181
>Primer341
BstAPI
5'
gctcctagaccagcatgccagtgtgcaagaaaggctgcagggagagcatgcggtgcggtaacattccttg
1680
>Primer341
>Primer382
>Primer59
>Primer250

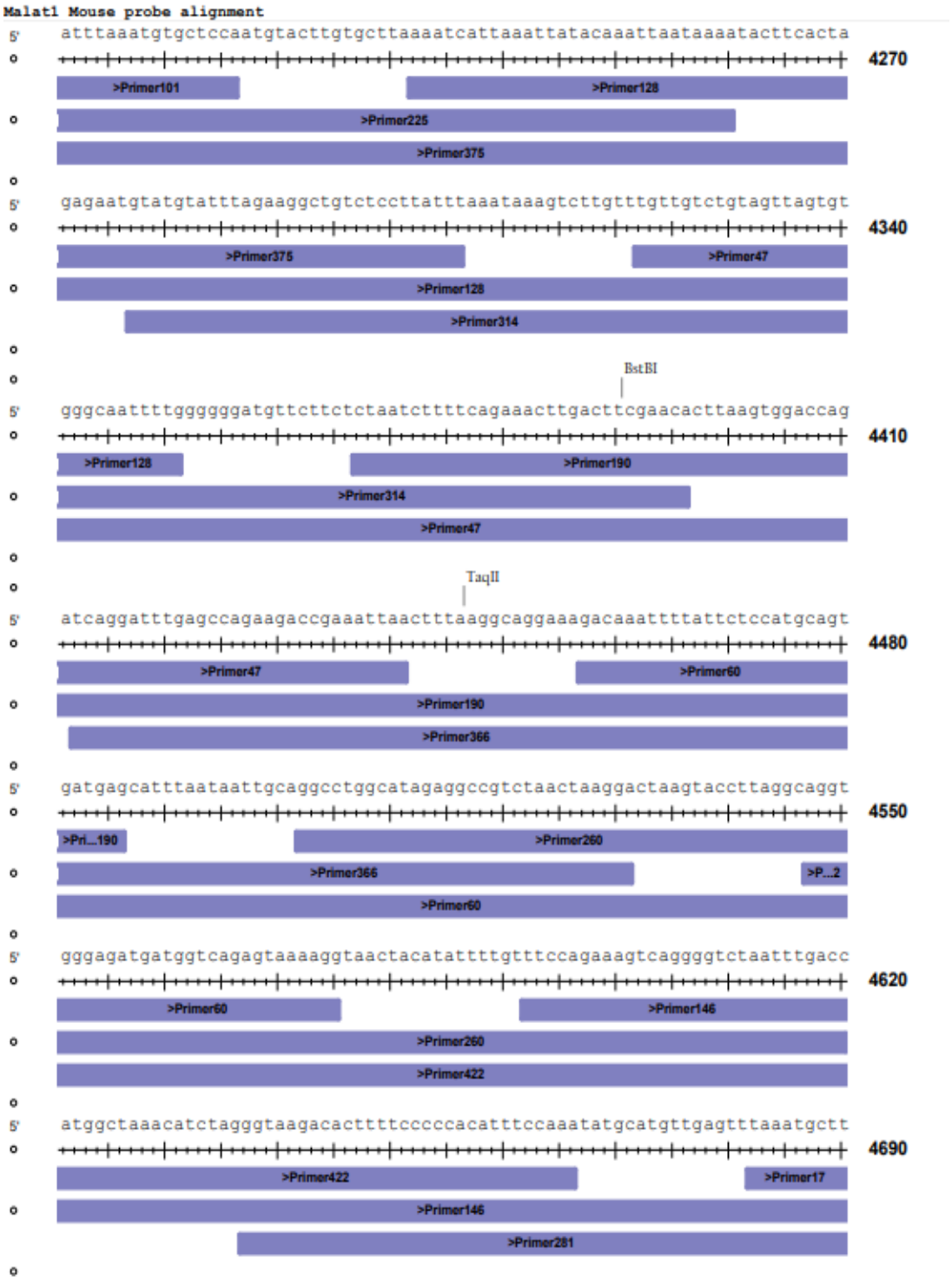

Malat1 Mouse probe alignment
5'
aggtcggcaacatggtggtggttttctgtaacttggatggtaacttgtttactttgtcttaatagttatg
1750
>Primer59
>Primer107
>Primer250
>Primer382
5'
ggggagttgtaggcttctgtgtaaagagatatatctggggctgtatgtaggcctttgcgggtgttgtagg
1820
>Primer382
>Primer4
>Primer107
>Primer285
5'
ttttctttttcagggttatgtcctcttgcatcttgtcagaagcttttgagggctgactgccaaggccca
1890
>Primer107
>Primer198
>Primer285
>Primer4
5'
gaaagaagaatggtagatggcaagttgtctttaaccgctcagaggggaatgaatggtagagccagcacaa
1960
>Primer4
>Primer77
>Primer198
>Primer357
5'
cctcccagttttgtaagacgttgtagtttgaacagatgacctaccacaagcctcactcctgtgtagggga
2030
>Pri...198
>Primer231
>Primer357
>P...3
>Primer77
5'
ggtaattgggcaaagtgcttttgggggaatgggggcaaaatatattttgagttcttttccccttaggtct
2100
>Primer77
>Primer129
>Primer231
>Primer393
XbaI
5'
gtctagaatcctaaaggcagatgactcaagggaaccagaaaaaaggaaatccactctcaggataagcaga
2170
>Primer393
>Primer31
>Primer129
>Primer289

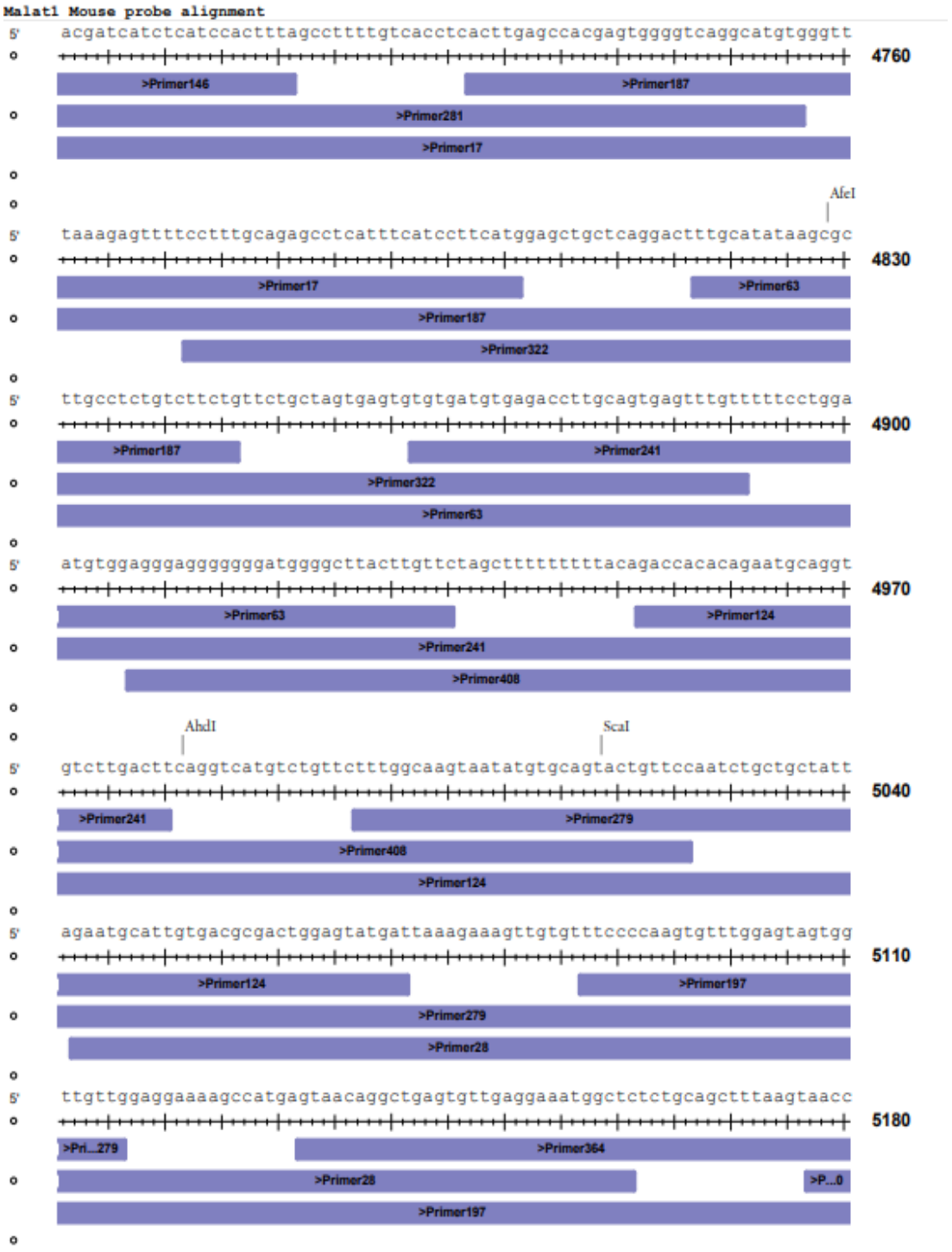

Malat1 Mouse probe alignment
DraIII
gctcgccaggtttacagtttgtaggaagtagaggatggatgctagctttcacactgagtgtggaggagct
2240
>Primer129
>Primer165
>Primer289
>Primer31
ggccatggcggaattgctggtagtttactctttccccctcccttaatgagatttgtaaaatcctaaacac
2310
>Primer31
>Primer74
>Primer165
>Primer319
AclI
ttttacttgaaatatttgggagtggtcttaacagggaggagtgggtgggggaaacgtttttttctaaga
2380
>Primer165
>Primer256
>Primer319
>Primer74
ttttccacagatgctatagttgtgttgacacactgggttagagaaggcgtgtactgctatgctgttggca
2450
>Primer74
>Primer136
>Primer256
>Primer401
cgacaccttcagggactggagctgccttttgtccttggaagagttttcccagttgccgctgaagtcagca
2520
>Primer256
>Primer267
>Primer401
>Primer136
cagtgcggctttggttcacagtcacctcaggagaacctcaggagcttggctaggccagaggttgaagtta
2590
>Primer136
>Primer188
>Primer267
>Primer51
agttttacagcaccgtgatttaaaatatttcattaaaggggaggggtaaaacttagttggctgtggcctt
2660
>Primer361
>Primer51
>P...4
>Primer188

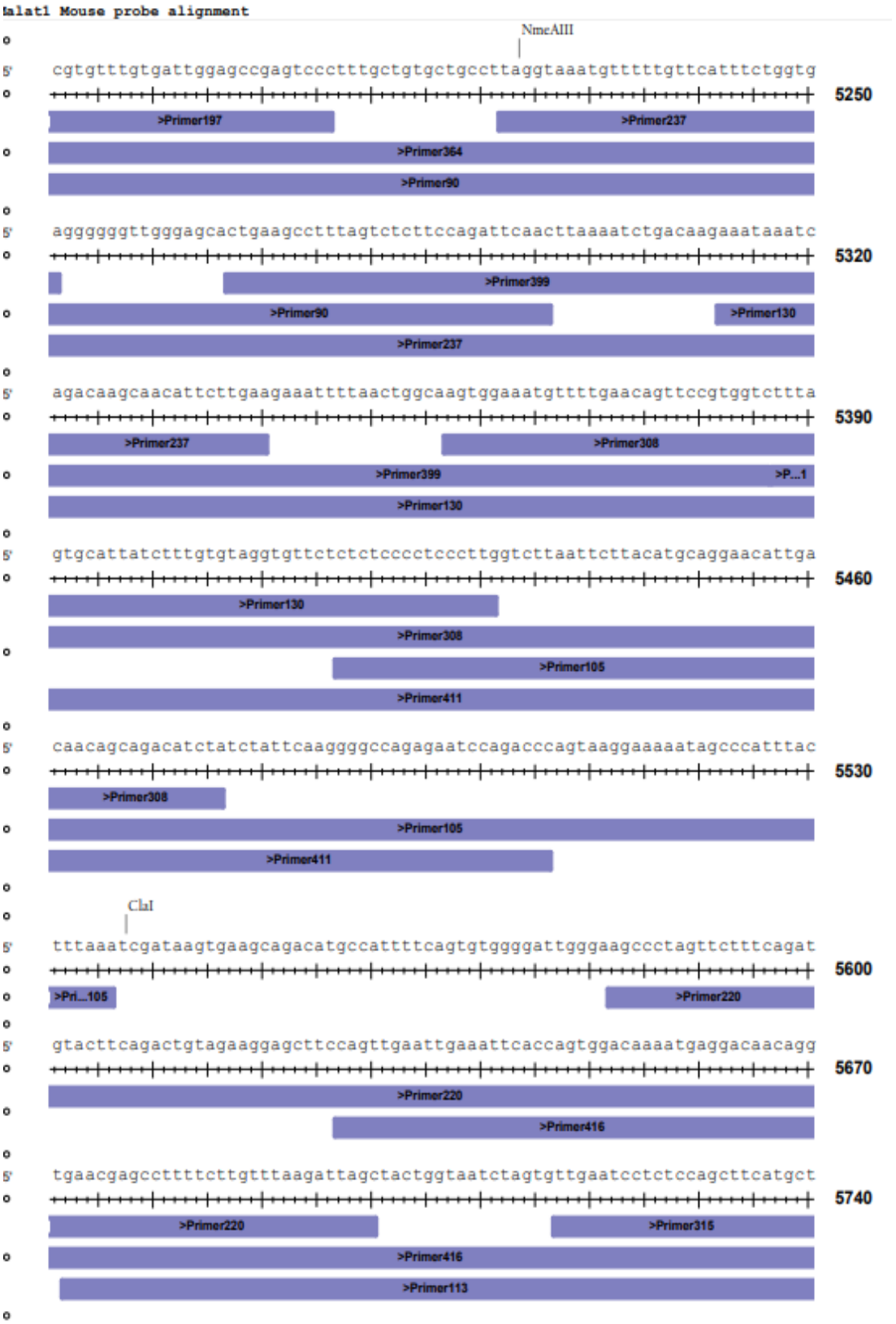

Malat1 Mouse probe alignment
5'
gtgtttgggtgggtgggggtgttaggtaattgtttagtttatgatttcagataatcataccagagaactt
2730
>Primer188
>Primer226
>Primer361
>Primer94
5'
aaatatttggaaaaacaggaaatctcagctttcaagttggcaagtaactcccaatccagtttttgcttct
2800
>Primer421
>Primer94
>Primer123
>Primer226
AloI
BsrFI
AloI
5'
tttttccttttttctttttttgaggcgggcagctaaggaaggttggttcctctgccggtccctcgaaagcg
2870
>Primer226
>Primer280
>Primer421
>Primer123
5'
tagggcttgggggttggtctggtccactgggatgatgtgatgctacagtggggactcttctgaagctgtt
2940
>Primer123
>Primer199
>Primer280
>Primer9
5'
ggatgaatatagattgtagtgtgtggttctcttttgaaatttttttcaggtgacttaatgtatcttaata
3010
>Primer280
>Primer348
>Primer9
>Primer199
CspCI
CspCI
5'
actactataggaacaaaggaagtggctttaatgaccctgaaggaatttcttctggtgatagcttttatat
3080
>Primer199
>Primer244
>Primer348
>Primer92
5'
tatcaagtaagagatactatctcagttttgtataagcaagtctttttcctagtgtaggagaaatgatttt
3150
>Primer348
>Primer376
>Primer92
>Primer244

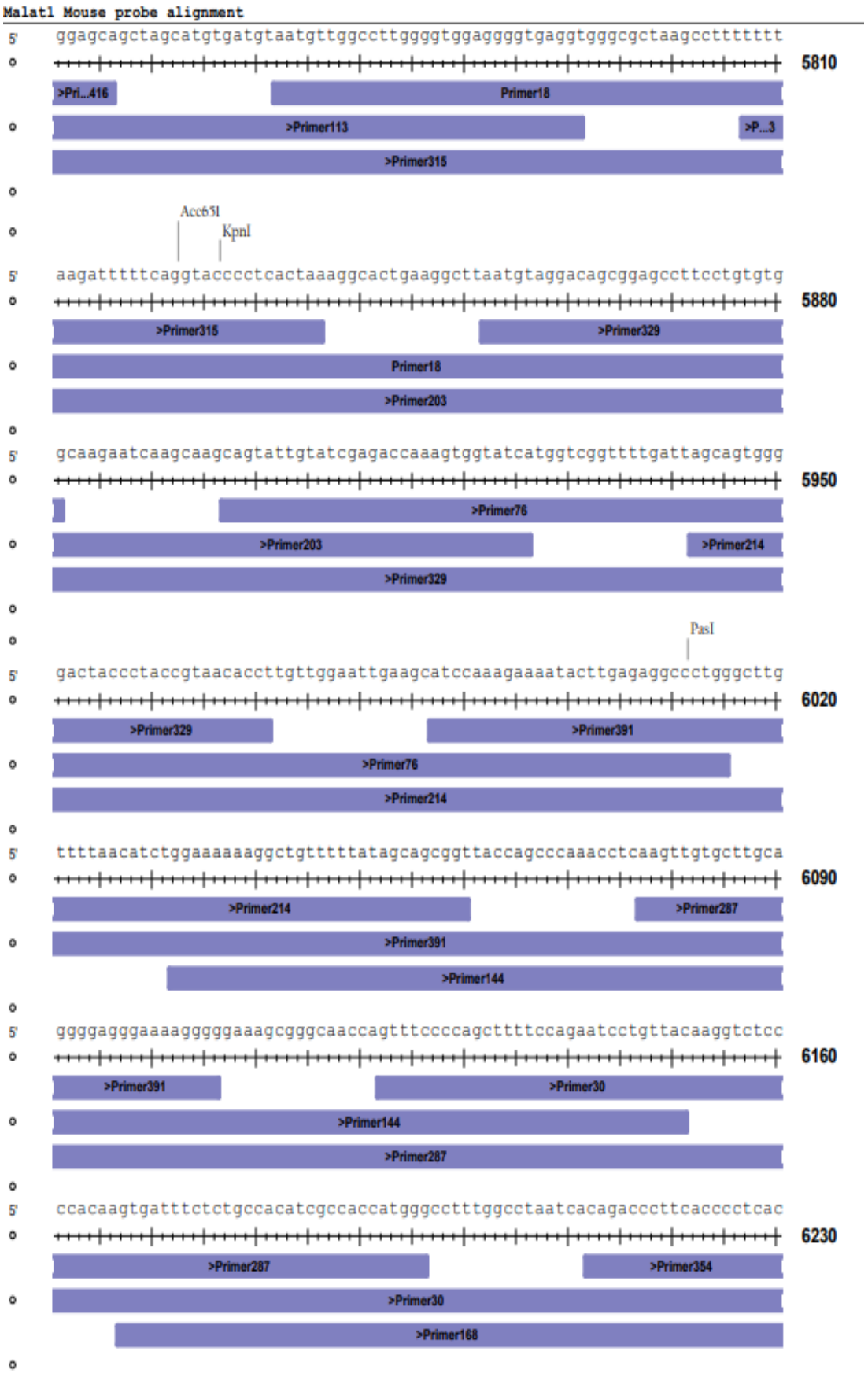

Malat1 Mouse probe alignment
BstZ17I
5' ccttgtgactaaacaagatgtaaaggtatgctttttttcttcttgtgcattgtatacttgtgtttatttg 3220
>Primer244
>Primer307
>Primer376
>Primer135
5' taacttataatttaagaattatgataattcagcctgaatgtcttttagagggtgggcttttgttgatgag 3290
>Pri...376
>Primer26
>Primer135
>P...0
>Primer307
5' ggaggggaaacctttttttttctgtagacctttttcagataacaccatctgagtcataaccagcctggca 3360
>Primer307
>Primer332
>Primer26
>Primer170
5' gtgtgatgacgtagatgcagagggagcagctccttggtgaatgagtgataagtaaaggcagaaaaaataa 3430
>Primer72
>Primer170
>Primer265
>Primer332
NdeI
5' tgtcatgtctccatggggaatgagcatgagccagagattgttcctactgatgaaaagctgcatatgcaaa 3500
>Primer332
>Primer432
>Primer72
>Primer265
5' aatttaagcaaatgaaagcaaccagtataaagttatggcaatacctttaaaagttatggcttatctacca 3570
>Primer265
>Primer303
>Primer432
>Primer158
5' agctttatccacaaaagtaaagaattgatgaaaaacagtgaagatcaaatgttcatctcaaaactgcttt 3640
>Primer432
>Primer19
>Primer158
>Primer303

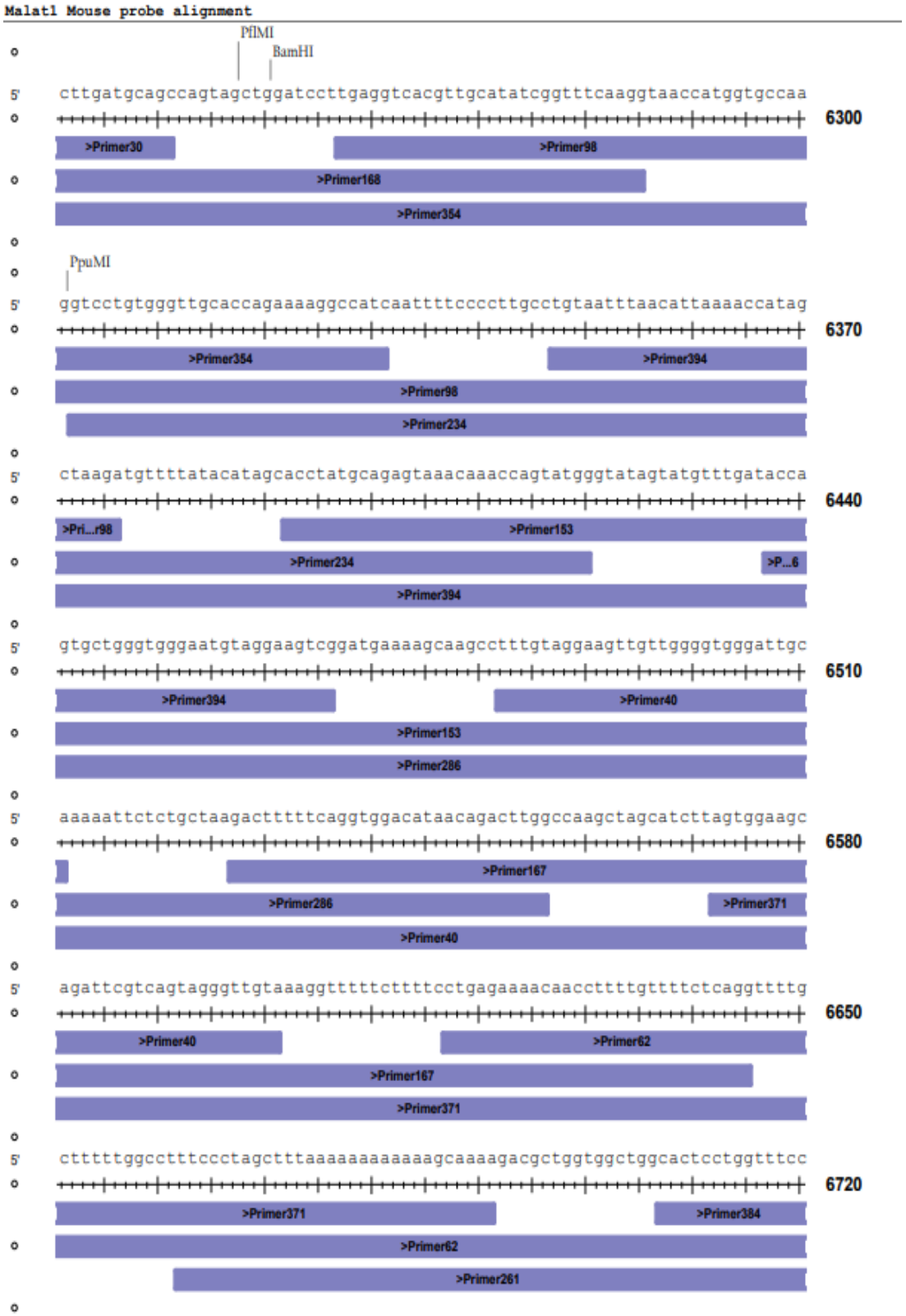

Malat1 Mouse probe alignment
5'
tacaaaagcagaatagaaatgaagtgaaaatgctgcattaagcctggagtaaaaagaagctgagcttgtt
3710
>Primer303
>Primer358
>Primer19
>Primer184
5'
gagatgagtgggatcgagcggctgcgaggcggtgcagtgtgccaatgtttcgtttgcctcagacaggttt
3780
>Primer19
>Primer93
>Primer184
>Primer358
5'
ctcttcataagcagaagagttgcttcattccatctcggagcaggaaacagcagactgctgttgacagata
3850
>Primer358
>Primer415
>Primer93
>Primer213
5'
agtgtaacttggatctgcagtattgcatgttagggatagataagtgccttttttctcttttccaaaaag
3920
>Pri...r93
>Primer125
>Primer213
>P...5
>Primer415
5'
acctgtagagctgttgaatgtttgcagctggcccctcttaggcagttcagaattttgagtagttttccca
3990
>Primer415
>Primer50
>Primer125
>Primer275
5'
tccagcctcttaaaaattcctaagccttgcaccgatgggctttcatgatgggatagctaataggcttttg
4060
>Primer182
>Primer275
>Primer328
>Primer50
BsrDI
5'
catcgtaaacttcaacacaaaagcctacatgattaatgcctactttaattacattgcttacaagattaag
4130
>Primer50
>Primer101
>Primer182
>Primer328
5'
gaatctttatcttgaagaccccatgaaagggatcattatgtgctgaaaattagatgttcatattgctaaa
4200
>Primer328
>Primer375
>Primer101
>Primer225

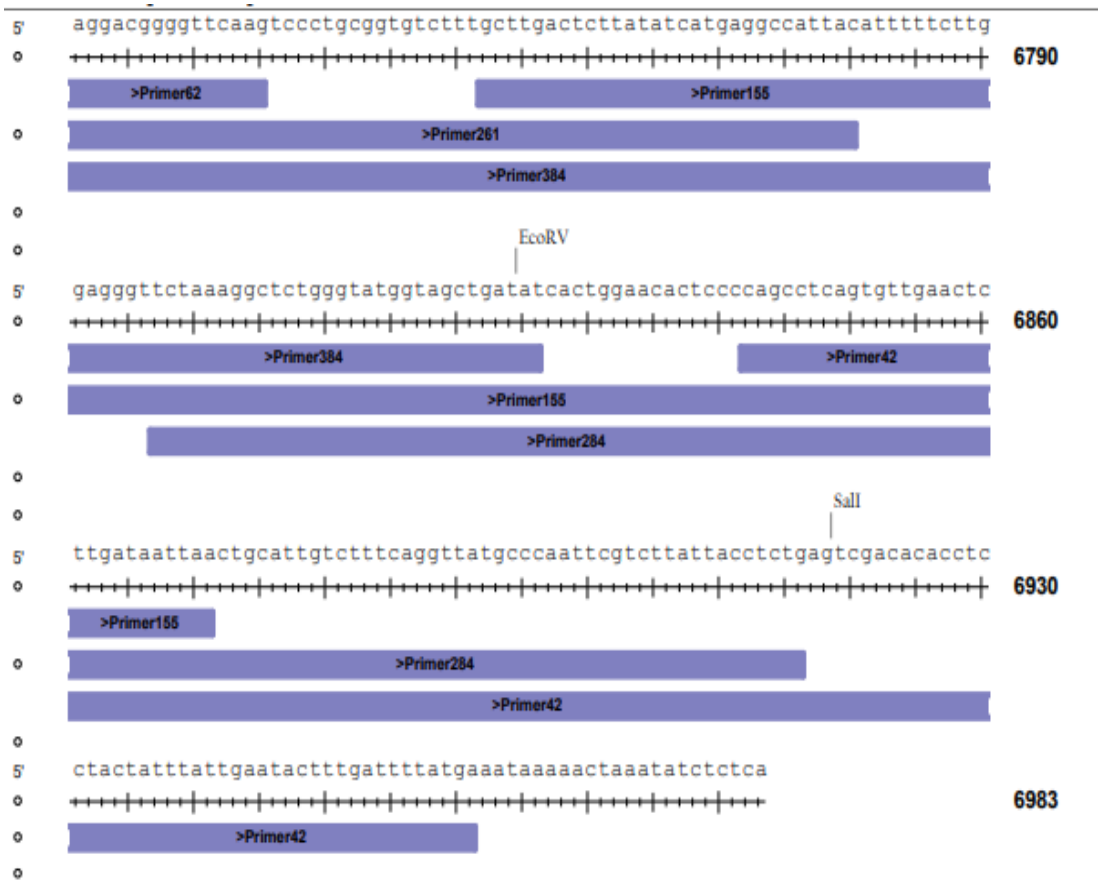

Malat1 Mouse probe alignment
5'
atttaaatgtgctccaatgtacttgtgcttaaaatcattaaattatacaaattaataaaatacttcacta
4270
>Primer101
>Primer128
>Primer225
>Primer375
5'
gagaatgtatgtatttagaaggctgtctccttatttaaataaagtcttgtttgttgtctgtagttagtgt
4340
>Primer375
>Primer47
>Primer128
>Primer314
BstBI
5'
gggcaattttgggggggatgttcttctctaatcttttcagaaacttgacttcgaacacttaagtggaccag
4410
>Primer128
>Primer190
>Primer314
>Primer47
TaqII
5'
atcaggatttgagccagaagaccgaaattaactttaaggcaggaaagacaaattttattctccatgcagt
4480
>Primer47
>Primer60
>Primer190
>Primer366
5'
gatgagcatttaataattgcaggcctggcatagaggccgtctaactaaggactaagtaccttaggcaggt
4550
>Pri...190
>Primer260
>Primer366
>P...2
>Primer60
5'
gggagatgatggtcagagtaaaaggtaactacatattttgtttccagaaagtcaggggtctaatttgacc
4620
>Primer60
>Primer146
>Primer260
>Primer422
5'
atggctaaacatctagggtaagacacttttccccccacatttccaaatatgcatgttgagtttaaatgctt
4690
>Primer422
>Primer17
>Primer146
>Primer281


Malat1 Mouse probe alignment
5'
acgatcatctcatccactttagccttttgtcacctcacttgagccacgagtggggtcaggcatgtgggtt
4760
>Primer146
>Primer187
>Primer281
>Primer17
AfeI
5'
taaagagttttcctttgcagagcctcatttcatccttcatggagctgctcaggactttgcatataagcgc
4830
>Primer17
>Primer63
>Primer187
>Primer322
5'
ttgcctctgtcttctgttctgctagtgagtgtgtgatgtgagaccttgcagtgagtttgtttttcctgga
4900
>Primer187
>Primer241
>Primer322
>Primer63
5'
atgtggagggaggggggatggggcttacttgttctagctttttttttacagaccacacagaatgcaggt
4970
>Primer63
>Primer124
>Primer241
>Primer408
AhdI
ScaI
5'
gtcttgacttcaggtcatgtctgttctttggcaagtaatatgtgcagtactgttccaatctgctgctatt
5040
>Primer241
>Primer279
>Primer408
>Primer124
5'
agaatgcattgtgacgcgactggagtatgattaaagaaagttgtgtttccccaagtgtttggagtagtgg
5110
>Primer124
>Primer197
>Primer279
>Primer28
5'
ttgttggaggaaaagccatgagtaacaggctgagtgttgaggaaatggctctctgcagctttaagtaacc
5180
>Pri...279
>Primer364
>Primer28
>P...0
>Primer197


Malat1 Mouse probe alignment
NmeAIII
5'
cgtgtttgtgattggagccgagtccctttgctgtgctgccttaggtaaatgtttttgttcatttctggtg
5250
>Primer197
>Primer237
>Primer364
>Primer90
5'
agggggggttgggagcactgaagcctttagtctcttccagattcaacttaaaatctgacaagaaataaatc
5320
>Primer399
>Primer90
>Primer130
>Primer237
5'
agacaagcaacattcttgaagaaattttaactggcaagtggaaatgttttgaacagttccgtggtctttta
5390
>Primer237
>Primer308
>Primer399
>P...1
>Primer130
5'
gtgcattatctttgtgtaggtgttctctctcccctcccttggtcttaattcttacatgcaggaacattga
5460
>Primer130
>Primer308
>Primer105
>Primer411
5'
caacagcagacatctatctattcaaggggccagagaatccagacccagtaaggaaaaatagcccatttac
5530
>Primer308
>Primer105
>Primer411
ClaI
5'
tttaaatcgataagtgaagcagacatgccattttcagtgtggggattgggaagccctagttctttcagat
5600
>Pri...105
>Primer220
5'
gtacttcagactgtagaaggagcttccagttgaattgaaattcaccagtggacaaaatgaggacaacagg
5670
>Primer220
>Primer416
5'
tgaacgagccttttcttgtttaagattagctactggtaatctagtgttgaatcctctccagcttcatgct
5740
>Primer220
>Primer315
>Primer416
>Primer113

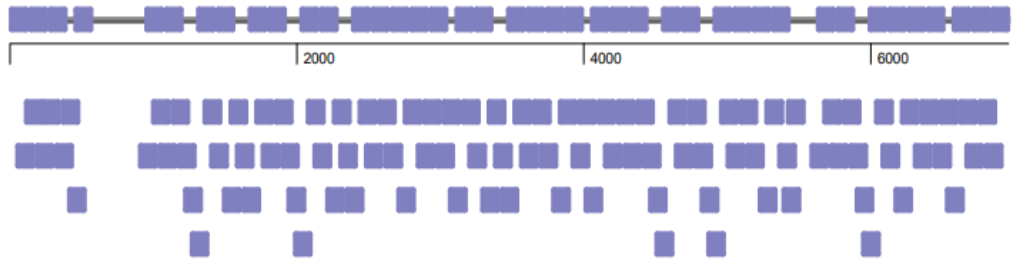

Malat1 Mouse probe alignment
5'
ggagcagctagcatgtgatgtaatgttggccttggggtggaggggtgaggtgggcgctaagccttttttt
5810
>Pri...416
Primer18
>Primer113
>P...3
>Primer315
Acc65I
KpnI
5'
aagatttttcaggtacccctcactaaaggcactgaaggcttaatgtaggacagcggagccttcctgtgtg
5880
>Primer315
>Primer329
Primer18
>Primer203
5'
gcaagaatcaagcaagcagtattgtatcgagaccaaagtggtatcatggtcggttttgattagcagtggg
5950
>Primer76
>Primer203
>Primer214
>Primer329
PasI
5'
gactaccctaccgtaacaccttgttggaattgaagcatccaaagaaaatacttgagaggccctgggcttg
6020
>Primer329
>Primer391
>Primer76
>Primer214
5'
ttttaacatctggaaaaaaggctgtttttatagcagcggttaccagcccaaacctcaagttgtgcttgca
6090
>Primer214
>Primer287
>Primer391
>Primer144
5'
ggggagggaaaaggggggaaagcgggcaaccagtttccccagcttttccagaatcctgttacaaggtctcc
6160
>Primer391
>Primer30
>Primer144
>Primer287
5'
ccacaagtgatttctctgccacatcgccaccatgggcctttggcctaatcacagacccttcacccctcac
6230
>Primer287
>Primer354
>Primer30
>Primer168


Malat1 Mouse probe alignment
PflMI
BamHI
5' cttgatgcagccagtagctggatccttgaggtcacgttgcatatcggtttcaaggtaaccatggtgccaa
6300
>Primer30
>Primer98
>Primer168
>Primer354
PpuMI
5' ggtcctgtgggttgcaccagaaaaggccatcaattttccccttgcctgtaatttaacattaaaaccatag
6370
>Primer354
>Primer394
>Primer98
>Primer234
5' ctaagatgttttatacatagcacctatgcagagtaaacaaaccagtatgggtatagtatgtttgatacca
6440
>Pri...r98
>Primer153
>Primer234
>P...6
>Primer394
5' gtgctgggtgggaatgtaggaagtcggatgaaaagcaagcctttgtaggaagttgttggggtgggattgc
6510
>Primer394
>Primer40
>Primer153
>Primer286
5' aaaaattctctgctaagacttttcaggtggacataacagacttggccaagctagcatcttagtggaagc
6580
>Primer167
>Primer286
>Primer371
>Primer40
5' agattcgtcagtagggttgtaaaggtttttctttttcctgagaaaacaacctttgttttctcaggttttg
6650
>Primer40
>Primer62
>Primer167
>Primer371
5' ctttttggcctttccctagctttaaaaaaaaaaaagcaaaagacgctggtggctggcactcctggtttcc
6720
>Primer371
>Primer384
>Primer62
>Primer261



Supplementary Figure 1: Schematic representation of *Malat1* probe binding sites across the transcript.

Primer binding sites were mapped to the *Malat1* transcript usingn Lasergene software.



Malat1

Probe

Supplementary Figure 2: *Malat1* Schematic representation of *Malat1* probe binding sites across the transcript – zoomed out

Primer binding sites were mapped to the *Malat1* transcript using Lasergene software

Supplementary table 2: List of genes in coherent or incoherent regulatory loops described in Figure 4.3.

Genes controlled by *Malat*1 in $CD4^+$ T cell differentiation correlate with *Malat1* at the single-cell level. Genes which are in coherent loops are genes which correlate positively with *Malat1* are down-regulated upon loss of *Malat1* and vice versa and genes which are in incoherent loops are genes which correlate positively with *Malat1* are up-regulated upon loss of *Malat1* and vice versa.

Gene	**Coherent/Incoherent**
Mlh1	Coherent Upregulated
Utp15	Coherent Upregulated
Scrib	Coherent Upregulated
Usp45	Coherent Upregulated
Slc35e2	Coherent Upregulated
Zfp512b	Coherent Upregulated
Ppid	Coherent Upregulated
Setdb2	Coherent Upregulated
Gatc	Coherent Upregulated
Gpx1	Coherent Upregulated
Ift122	Coherent Upregulated
Las1l	Coherent Upregulated
Mat2a	Coherent Upregulated
Pttg1	Coherent Upregulated
Exosc7	Coherent Upregulated
Ankrd39	Coherent Upregulated
Chd3	Coherent Upregulated
4930481A15Rik	Coherent Upregulated
Malat1	Coherent Downregulated
Emilin2	Coherent Downregulated
Ppp1r3b	Coherent Downregulated
Asb2	Coherent Downregulated
Hilpda	Coherent Downregulated
Ero1l	Coherent Downregulated
Sdc3	Coherent Downregulated
Ctla2a	Coherent Downregulated
Vldlr	Coherent Downregulated
Il12rb2	Coherent Downregulated
Sox5	Coherent Downregulated
Eif4ebp1	Coherent Downregulated
Bhlhe41	Coherent Downregulated

Pim1	Coherent Downregulated
Bcl2	Coherent Downregulated
Chd7	Coherent Downregulated
R3hdm1	Coherent Downregulated
Wipi1	Coherent Downregulated
Bend5	Coherent Downregulated
Nfil3	Coherent Downregulated
Icos	Coherent Downregulated
Mxi1	Coherent Downregulated
Slc2a3	Coherent Downregulated
Nabp1	Coherent Downregulated
Scd2	Coherent Downregulated
Jarid2	Coherent Downregulated
Lpin1	Coherent Downregulated
Adarb1	Coherent Downregulated
Sdf4	Coherent Downregulated
Sema4f	Coherent Downregulated
Spsb1	Coherent Downregulated
Ndufv3	Coherent Downregulated
Tnfaip3	Coherent Downregulated
P4ha1	Coherent Downregulated
Csrnp1	Coherent Downregulated
Abhd18	Coherent Downregulated
Map3k1	Coherent Downregulated
Sdf2l1	Coherent Downregulated
Slc25a5	Coherent Downregulated
Psat1	Coherent Downregulated
Mcm2	Coherent Downregulated
Gins1	Coherent Downregulated
Fignl1	Coherent Downregulated
Impdh2	Coherent Downregulated
Poc1a	Coherent Downregulated
Cd320	Coherent Downregulated
Psmc3ip	Coherent Downregulated
Mybl2	Coherent Downregulated
Hsp90ab1	Coherent Downregulated
Ak6	Coherent Downregulated
Stmn1	Coherent Downregulated
Tkt	Coherent Downregulated
Ran	Coherent Downregulated
Pdk3	Coherent Downregulated
Dut	Coherent Downregulated
Rfc4	Coherent Downregulated
Hnrnpa1	Coherent Downregulated
Eef1g	Coherent Downregulated
Snrpd1	Coherent Downregulated

Tipin	Coherent Downregulated
Mcm3	Coherent Downregulated
Shcbp1	Coherent Downregulated
Mad2l2	Coherent Downregulated
Ptma	Coherent Downregulated
Rfc5	Coherent Downregulated
Lrrc59	Coherent Downregulated
Psmd8	Coherent Downregulated
Rps2	Coherent Downregulated
Cox6a1	Coherent Downregulated
Hsp90aa1	Coherent Downregulated
Tuba1b	Coherent Downregulated
Ruvbl2	Coherent Downregulated
Banf1	Coherent Downregulated
Hspa8	Coherent Downregulated
Ndc1	Coherent Downregulated
Dnmt1	Coherent Downregulated
Tyms	Coherent Downregulated
Txn1	Coherent Downregulated
Uchl5	Coherent Downregulated
Atp5b	Coherent Downregulated
Ints10	Coherent Downregulated
Ywhae	Coherent Downregulated
Kif4	Coherent Downregulated
Gtf2f1	Coherent Downregulated
Rbbp7	Coherent Downregulated
Nhp2	Coherent Downregulated
Npm1	Coherent Downregulated
Ndufa12	Coherent Downregulated
Mki67	Coherent Downregulated
Alyref	Coherent Downregulated
Ppia	Coherent Downregulated
Rfc2	Coherent Downregulated
Stoml2	Coherent Downregulated
Ccnb2	Coherent Downregulated
Ndufb6	Coherent Downregulated
Hnrnpa3	Coherent Downregulated
Psmb2	Coherent Downregulated
Cdk4	Coherent Downregulated
Ppil1	Coherent Downregulated
Hint1	Coherent Downregulated
Hmgb2	Coherent Downregulated
Kars	Coherent Downregulated
Atp5o	Coherent Downregulated
Sf3a1	Coherent Downregulated
Rbm3	Coherent Downregulated

Cct2	Coherent Downregulated
Eif3l	Coherent Downregulated
Eif2d	Coherent Downregulated
Lap3	Coherent Downregulated
Pa2g4	Coherent Downregulated
Hmgb1	Coherent Downregulated
Tagln2	Coherent Downregulated
Ncl	Coherent Downregulated
Mrpl35	Coherent Downregulated
Cct5	Coherent Downregulated
Anp32b	Coherent Downregulated
Mrps25	Coherent Downregulated
Sdhd	Coherent Downregulated
Atp5g2	Coherent Downregulated
Cct7	Coherent Downregulated
Tsnax	Coherent Downregulated
Arid3a	Coherent Downregulated
Csnk2b	Coherent Downregulated
Nedd8	Coherent Downregulated
Mdh2	Coherent Downregulated
1700017B05Rik	Coherent Downregulated
Eif2s3x	Coherent Downregulated
Metap2	Coherent Downregulated
Gm8186	Coherent Downregulated
Rnps1	Coherent Downregulated
Mrpl34	Coherent Downregulated
Cbx3	Coherent Downregulated
Rpl35a	Coherent Downregulated
Polr3c	Coherent Downregulated
Sdf2l1	Incoherent Downregulated
Slc25a5	Incoherent Downregulated
Psat1	Incoherent Downregulated
Mcm2	Incoherent Downregulated
Gins1	Incoherent Downregulated
Fignl1	Incoherent Downregulated
Impdh2	Incoherent Downregulated
Poc1a	Incoherent Downregulated
Cd320	Incoherent Downregulated
Psmc3ip	Incoherent Downregulated
Mybl2	Incoherent Downregulated
Hsp90ab1	Incoherent Downregulated
Ak6	Incoherent Downregulated
Stmn1	Incoherent Downregulated
Tkt	Incoherent Downregulated
Ran	Incoherent Downregulated
Pdk3	Incoherent Downregulated

Dut	Incoherent Downregulated
Rfc4	Incoherent Downregulated
Hnrnpa1	Incoherent Downregulated
Eef1g	Incoherent Downregulated
Snrpd1	Incoherent Downregulated
Tipin	Incoherent Downregulated
Mcm3	Incoherent Downregulated
Shcbp1	Incoherent Downregulated
Mad2l2	Incoherent Downregulated
Ptma	Incoherent Downregulated
Rfc5	Incoherent Downregulated
Lrrc59	Incoherent Downregulated
Psmd8	Incoherent Downregulated
Rps2	Incoherent Downregulated
Cox6a1	Incoherent Downregulated
Hsp90aa1	Incoherent Downregulated
Tuba1b	Incoherent Downregulated
Ruvbl2	Incoherent Downregulated
Banf1	Incoherent Downregulated
Hspa8	Incoherent Downregulated
Ndc1	Incoherent Downregulated
Dnmt1	Incoherent Downregulated
Tyms	Incoherent Downregulated
Txn1	Incoherent Downregulated
Uchl5	Incoherent Downregulated
Atp5b	Incoherent Downregulated
Ints10	Incoherent Downregulated
Ywhae	Incoherent Downregulated
Kif4	Incoherent Downregulated
Gtf2f1	Incoherent Downregulated
Rbbp7	Incoherent Downregulated
Nhp2	Incoherent Downregulated
Npm1	Incoherent Downregulated
Ndufa12	Incoherent Downregulated
Mki67	Incoherent Downregulated
Alyref	Incoherent Downregulated
Ppia	Incoherent Downregulated
Rfc2	Incoherent Downregulated
Stoml2	Incoherent Downregulated
Ccnb2	Incoherent Downregulated
Ndufb6	Incoherent Downregulated
Hnrnpa3	Incoherent Downregulated
Psmb2	Incoherent Downregulated
Cdk4	Incoherent Downregulated
Ppil1	Incoherent Downregulated
Hint1	Incoherent Downregulated

Hmgb2	Incoherent Downregulated
Kars	Incoherent Downregulated
Atp5o	Incoherent Downregulated
Sf3a1	Incoherent Downregulated
Rbm3	Incoherent Downregulated
Cct2	Incoherent Downregulated
Eif3l	Incoherent Downregulated
Eif2d	Incoherent Downregulated
Lap3	Incoherent Downregulated
Pa2g4	Incoherent Downregulated
Hmgb1	Incoherent Downregulated
Tagln2	Incoherent Downregulated
Ncl	Incoherent Downregulated
Mrpl35	Incoherent Downregulated
Cct5	Incoherent Downregulated
Anp32b	Incoherent Downregulated
Mrps25	Incoherent Downregulated
Sdhd	Incoherent Downregulated
Atp5g2	Incoherent Downregulated
Cct7	Incoherent Downregulated
Tsnax	Incoherent Downregulated
Arid3a	Incoherent Downregulated
Csnk2b	Incoherent Downregulated
Nedd8	Incoherent Downregulated
Mdh2	Incoherent Downregulated
1700017B05Rik	Incoherent Downregulated
Eif2s3x	Incoherent Downregulated
Metap2	Incoherent Downregulated
Gm8186	Incoherent Downregulated
Rnps1	Incoherent Downregulated
Mrpl34	Incoherent Downregulated
Cbx3	Incoherent Downregulated
Rpl35a	Incoherent Downregulated
Polr3c	Incoherent Downregulated
Polr3c	Incoherent Downregulated
Hmbox1	Incoherent Upregulated
Rb1cc1	Incoherent Upregulated
Smarca2	Incoherent Upregulated
AW549877	Incoherent Upregulated
Glcci1	Incoherent Upregulated
Il18bp	Incoherent Upregulated
Btg1	Incoherent Upregulated
Adat1	Incoherent Upregulated