Assessing a potential *in cis* regulatory function of the lncRNA *Malat1* in CD4$^+$ T cells

Euan John Mitchell

MSc by Research

University of York

Biology

September 2018

Abstract

The long non-coding RNA (lncRNA) Metastasis Associated Lung Adenocarcinoma Transcript 1 (*Malat1*) has been indicated to have a variety of cellular functions, from cancer cell growth and survival to regulating pre-mRNA splicing. In several *Malat1* knockout (KO) models, *Malat1* neighbouring genes showed significant expression changes, but these effects were not consistent across all *Malat1* KO mouse models. This implies potential and tissue specific *in cis* regulation of these genes by *Malat1*. Here, we assessed if *Malat1* regulated its surrounding genes *in cis* in $CD4^+$ T cells, as previous experiments on these cells showed significant downregulation of *Malat1* upon T cell activation. This was achieved by analysing primary naïve mouse $CD4^+$ cells *in vitro* under Th0, Th1 and Th2 differentiation conditions at two different time points. Our results showed significant expression changes in *Neat1*, Scyl1 and Map3k11 in *Malat1* KO cells under specific activation conditions (e.g. naïve, 4 day Th0 and 6 day Th1 cultures). To assess the cell type specificity of these effect, mouse tail fibroblasts (MTFs) were isolated and analysed – only Scyl1 was differentially expressed in *Malat1* KO MTFs. We also assessed potential *in trans* effects of *Malat1*. It has been shown that *Malat1* binds the mRNA of several RNA binding proteins – amongst those, we found that *Malat1* regulates U2af1 in specific $CD4^+$ cell types. This data indicates that gene regulation in T helper cells by *Malat1* does not occur uniformly across all subsets, but is dependent on the type and duration of activation. For the *Malat1* neighbouring genes, this may be via transcription factors binding to the *Malat1* locus and acting on nearby enhancer and promoter sequences.

Contents

List of tables

List of figures

Acknowledgements

Thanks to my supervisors, Dr Dimitris Lagos and Dr James Hewitson, for all their patient advice and support over the past year, as well as the other members and students of the Lagos and Hewitson lab groups.

Authors 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 this, or any other, University. All sources are acknowledged as References.

Introduction

Long non-coding RNAs

In recent years, an increasing number of studies have highlighted the key role of non-coding RNAs, transcribed regions of the genome which are not protein coding, in genomic regulation (Carninci 2009). In particular, long non-coding RNAs (lncRNAs), non-coding transcripts longer than 200bp, are increasingly considered to have key roles in both transcriptional and epigenetic regulation, such as protein complex assembly scaffolds, transcriptional coregulators or regulators of mRNA processing (Ernst and Morton 2013, Kung *et al* 2013). As this tremendous diversity of function makes it difficult to clearly classify lncRNAs, they are often categorised based upon their relative genomic location, such as genic or intergenic transcripts located less or more than 5kb from protein coding genes respectively (Luo *et al* 2016) and intronic or antisense transcripts located within introns or on opposing strands of other genes (Kung *et al* 2013). However, while these categories are commonly used by researchers, they are not necessarily related to the functions of the lncRNAs (Kung *et al* 2013).

LncRNA regulatory functions can occur either *in cis*, directly mediating expression of neighbouring genes, or *in trans*, acting on distal genes and proteins. *In cis* regulation directly mediates expression of neighbouring genes through a variety of mechanisms, such as *COLDAIR* transcription targeting repressive chromatin modifications to the FLC locus during vernalisation, *Tsix* transcription directly inhibiting transcription of *Xist* or *Xist* binding to targets across the X chromosome to recruit repressive PRC2 to genomic loci (Heo and Sung 2011, Sado *et al* 2006, Brockdorff 2013). Conversely, *in trans* regulatory lncRNAs influence expression of genes on other chromosomes, acting either generally, such as inhibition of RNA polymerase 2 phosphorylation by the 7SK transcript to regulate transcriptional elongation, or specifically at certain loci, such as targeting of repressive chromatin markers to the *HOXD* cluster by the lncRNA *HOTAIR* (Peterlin *et al* 2012, Rin *et al* 2007). Though at present it is unclear which form of activity is more common among lncRNAs, some recent studies have suggested a greater prevalence of *in cis* regulatory functions among specific subsets of lncRNAs, with genic transcripts showing particularly strong association with *in cis* regulation (Yan *et al* 2017, Luo *et al* 2016).

Malat1

One of the most widely studied lncRNAs is Metastasis Associated Lung Adenocarcinoma Transcript 1 (*Malat1*), an intergenic transcript around 6.9kb long transcribed from chromosome 19 in mice and chromosome 11 in humans (Wilusz *et al* 2008, Zhang *et al* 2012). *Malat1* is processed from its initial transcript by RNAse P to produce both a small, cytoplasmic tRNA like sequence called *mascRNA* and the remaining *Malat1* lncRNA, which is subsequently localised to alternative splicing factor clusters in

the nucleus known as nuclear speckles (Wilusz *et al* 2008, Hutchinson *et al* 2007). The lncRNA was originally identified in lung cancer cell lines, where it was found to be a strong predictor of tumour metastasis and survival (Ji *et al* 2003) and has since been found to be strongly associated with growth and proliferation of a variety of other cancer cell lines, including ovarian, colorectal and squamous cell tumours (Lin *et al* 2018, Yang *et al* 2015, Zhou *et al* 2015). As such, there has been considerable research into *Malat1*, looking both at proteins and RNAs which interact with *Malat1* and at potential cellular functions for the transcript to determine the biological role of *Malat1* (Engreitz *et al* 2014, Chen *et al* 2017).

In addition to its apparent importance in cancer cell proliferation, *Malat1* has several unusual features compared to other lncRNAs which have made it a target for the aforementioned functional studies**.** Firstly, the transcript is relatively stable, with the 3' sequence containing an expression and nuclear retention element that promotes folding into a bipartite triple helix structure, protecting the 3' end of the transcript from degradation by RNAses (Brown *et al* 2014). Furthermore**,** *Malat1* has an unusually highly conserved primary sequence across 20 mammalian species, with analysis of 12 different primate species identifying *Malat1* as the most conserved sequence in a 120kb region and around 69% sequence identity observed between the human and mouse sequences (Ma *et al* 2015, Ji *et al* 2003, Figure 1). Finally, the lncRNA is very abundantly expressed across a range of tissues, and in some tissues is expressed at a level similar to some housekeeping genes, such as *Actb* and *Gapdh* (Ma *et al* 2015, Zhang *et al* 2012). Taken together, this high conservation, abundant expression and functional importance in cancer lines indicate that *Malat1* has some key housekeeping function in normal cell biology, a possibility supported by some studies showing a strong association of *Malat1* with pre-mRNA splicing factors (Tripathi *et al* 2010, Engreitz *et al* 2014).

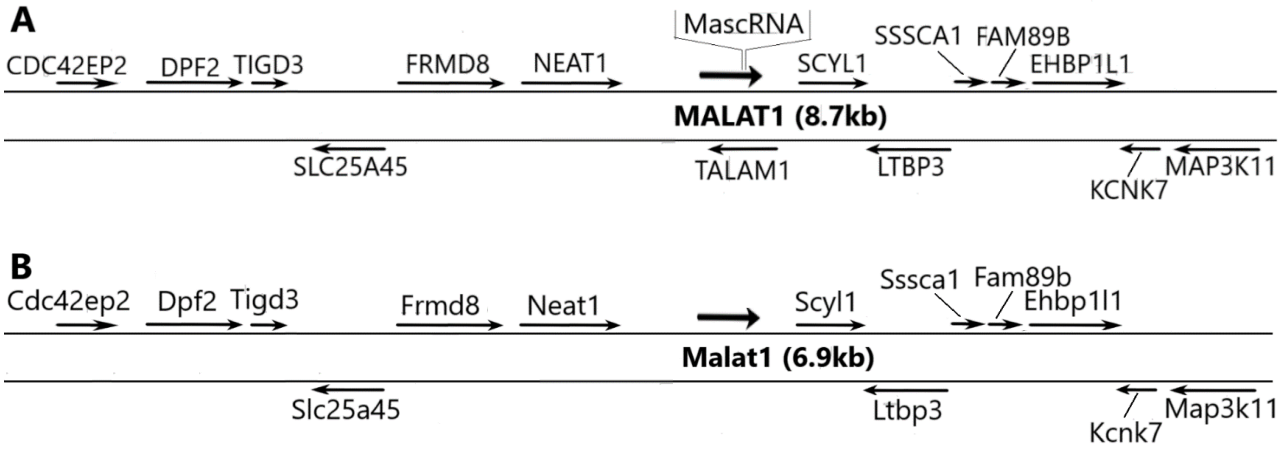

Figure 1: Simplified diagram of the *Malat1* locus and surrounding genes in the (A) human and (B) mouse genomes, showing approximate positions, directions of transcription and relative lengths of each gene

However, this apparent functional importance is contrasted by the results of 3 different papers generating *Malat1* knockout (KO) mouse models by different methods (Zhang *et al* 2012, Nakagawa *et al* 2012, Eiβmann *et al* 2012), as none of these models showed any significant differences in growth, viability, fertility, histology or development compared to wild type. Furthermore, while one *Malat1* KO model showed significantly impaired metastasis of human lung and breast cancer lines (Arun *et al* 2016), another displayed no significant effect of *Malat1* KO on proliferation of A549 or HLE cancer lines, despite its association with growth and metastasis in both the A549 line and liver carcinoma cells in other studies (Eiβmann *et al* 2012, Gutschner *et al* 2013, Lai *et al* 2010). Additionally, while *Malat1* was indicated to have significant effects on pre-mRNA splicing in previous knockdown (KD) models, no significant differences in the number or localisation of nuclear speckles was observed in these KO models, nor any differences in the levels or phosphorylation status of pre-mRNA splicing factors associated with these structures (Zhang *et al* 2012, Nakagawa *et al* 2012, Tripathi *et al* 2010). Several explanations were proposed for these results, such as differing results from knockdown and knockout models, functional relevance only under abnormal conditions or functional compensation by other lncRNAs (Zhang *et al* 2012, Nakagawa *et al* 2012). The possibility of functional compensation seems likely, given the lack of significant *Malat1* knockout effects on lung adenocarcinoma cells despite its strong associations with metastasis of such cancer lines (Eiβmann *et al* 2012, Ji *et al* 2003) – however, these possible solutions to the differences in *Malat1* functional data are yet to be fully assessed.

A notable result from Zhang *et al* (2012) was that, of 12 genes which showed significant expression changes in the KO's, 5 were located immediately around the *Malat1* locus, with a further 7 genes surrounding *Malat1* showing significant expression changes upon further analysis of the knockouts. The paper interpreted this as transcription of *Malat1* regulating these surrounding genes *in cis*, a theory consistent with a recently identified example of *in cis* regulation at the *Malat1* locus by its antisense transcript *Talam1* (Zong *et al* 2016). However, the data from the other knockout papers were inconsistent with this theory – though neither tested the same range of surrounding genes as Zhang *et al* (2012), Eiβmann *et al* (2012) found no significant change in *Neat1* expression in the knockout cells, while Nakagawa *et al* (2012) observed significant downregulation of *Neat1* with *Malat1* KO, rather than upregulation. While this may arise from the different systems of *Malat1* KO used for each study, it may also be that this regulation of surrounding genes only occurs in specific cell types, given both that these papers used different tissue types for expression analysis of these surrounding genes and that Nakagawa *et al* (2012) observed significantly reduced *Neat1* expression only occurring in specific *Malat1* KO tissues. As such, this possible *in cis* regulation by *Malat1* warrants further study.

CD4$^+$ T cells

Previous data from our research group (Hewitson *et al*, unpublished) indicated *Malat1* to have some functional importance in CD4$^+$ T cells, a subset of T-lymphocytes which have key functions in activation and regulation of adaptive immune responses through secretion of specific cytokines. CD4$^+$ cells differentiate from naïve cells into a wide variety of subtypes with different immune functions, including Th1 cells, which mediate intracellular pathogen defence and some autoimmune diseases, Th2 cells, which regulate extracellular parasite responses and a number of allergic diseases, and Th17 cells, a more recently identified line required for responses to extracellular bacteria and fungi (Kurts 2007), with regulatory T cells (Tregs) acting to suppress immunopathological effects of these Th cells (Luckheeram *et al* 2012). Differentiation of these different CD4$^+$ cell types is dependent on extracellular cytokine signalling, coordinating with T cell receptor (TCR) activation to trigger downstream signalling cascades which activate expression of different master regulators for each cell type, such as T-bet, GATA3 and RORγt for Th1, Th2 and Th17 cells respectively (Zhu *et al* 2010). These, in turn, promote inhibition of signalling pathways from other CD4$^+$ cell types and stimulate the expression of distinct effector cytokines for their respective cell types, such as IFNγ in Th1 cells, IL-4, IL-5 and IL-13 in Th2 cells, IL-17 and IL-21 in Th17 cells and IL-35 and TGF-β in Tregs (Luckheeram *et al* 2012).

Given that numerous changes in histone and chromatin modifications have been identified during differentiation of these Th cell types, CD4$^+$ differentiation has been proposed as an effective model for studying mechanisms of epigenetic regulation and their function in cell differentiation (Russ *et al* 2013). This includes assessing lncRNA functions in T cell differentiation – although studies of these functions in both the general immune system and CD4$^+$ T cells specifically have only recently begun, expression analysis of these Th cell types revealed expression of numerous intergenic lncRNA clusters to be highly specific to different stages of CD4$^+$ differentiation and between the CD4$^+$ subtypes, showing far greater cell type specificity than mRNAs (Hu *et al* 2013). Furthermore, many of these differentially expressed lncRNA genes were found to be located near lineage specific mRNA genes, with altered lncRNA expression shown to coincide with corresponding expression changes in nearby protein coding genes (Aune *et al* 2016, Xia *et al* 2014). This suggests *in cis* regulation of these genes by the lncRNAs, with several transcripts identified which were found to have functions in regulating CD4$^+$ lineage specific genes on the same chromosome (Xia *et al* 2014, Aune *et al* 2016). Similarly, genome wide association studies have found robust immune cell expression of lncRNAs in loci strongly associated with autoimmune diseases, further highlighting the functional importance of lncRNAs in immune system and T cell signalling and function (Hrdlickova *et al* 2014). As such, CD4$^+$ T cells are a key model for further studies of lncRNA functions, including functions and interactions of *Malat1*.

Hypothesis and aims

Based on these results and the data on the genes surrounding *Malat1* from Nakagawa *et* al (2012), and Zhang *et al* (2012), we hypothesise that *Malat1* regulates surrounding genes *in cis*, potentially in a cell type specific manner. As such, this project aimed:

1. To determine if immortalised cell lines could be used as a viable model instead of primary $CD4^+$ T cells by comparing the effects of mitogenic stimulation of immortalised cell lines on *Malat1* expression to previous $CD4^+$ expression data on *Malat1* (Hewitson *et al*, unpublished).
2. To test for evidence of *in cis* regulation of neighbouring genes by *Malat1* by comparing their expression in a model of wild type (Wt) and *Malat1* knockout (KO) murine $CD4^+$ T cells.
3. To determine if any such regulation was cell type specific by assessing gene expression in mouse tail fibroblast (MTF) cultures for both genotypes.
4. To assess if there was any evidence for *in trans* regulation by *Malat1* in $CD4^+$ T cells by comparing Wt and *Malat1* KO expression levels of genes known to interact with *Malat1*.
5. To test if *Malat1* KO had any effect on the growth and proliferation of MTF cultures in order to compare any such effects to those observed in previous studies of *Malat1* depleted cells.

Methods

Malat1 knockout mice

The *Malat1* knockout cells used for this study were derived from the KO model developed by Nakagawa *et al* (2012). Briefly, a *Malat1* targeting vector was used to insert a *LacZ/PolyA* cassette immediately downstream of the *Malat1* transcriptional start site, thus preventing transcription of the *Malat1* gene body from this site. Nakagawa *et al* (2012) showed effective knockout of *Malat1* across numerous different tissues using this system, although low level residual expression of a 3.2kb transcript was observed in neuronal cells. This was suggested to be due to the knockout cassette triggering artificial activation of an internal promoter sequence within *Malat1* in specific tissues, leading to transcription of a truncated lncRNA lacking the 5' end of the transcript. Wt control mice were C57BL/6, and female mice aged 6-12 weeks were used for all experiments.

Immortalised cell line culture

Jurkat (ATCC TIB-152) and EL4 (ATCC TIB-39) cell lines were maintained in 25ml of complete CD4 media (1640 RPMI with 10% Fetal Calf Serum (FCS), 100units/ml Penicillin, 100µg/ml Streptomycin and 2mM L-Glutamine). To set up stimulated plates, the cells were counted and a suspension of $1x10^6$ cells/ml set up and seeded onto 6 well plates at 1ml/well before being stimulated with either 5µg/ml Phytohaemagglutinin (PHA) and 1µg/ml Phorbol 12-Myristate 13-Acetate (PMA) for Jurkat cultures or various concentrations of Concavalin A (ConA) for EL4 (10, 5, 1 and 0.1 µg/ml). For the EL4 cultures, given that different concentrations of ConA were applied to test which would be best for any further EL4 cultures, 0.1µg/ml PHA was used as a positive control of T cell stimulation, as this mitogen had been shown by our experiments to successfully stimulate growth of Jurkat cells. N.B. These cell lines are not regularly validated, save for annual mycoplasma screening, and as such should be considered with some caution.

EL4 culture treatment	0.1µg/ml ConA	1µg/ml ConA	5µg/ml ConA	10µg/ml ConA	0.1µg/ml PMA	Unstimulated control
Concentration (ng/µl)	465.1	592.3	322.5	72.1	112.4	467.4
260:230	2.03	2.08	2.08	2.00	2.03	2.08
260:280	0.61	0.81	1.66	0.57	0.51	1.98

Table 1: Example of data from Nanodrop analysis of EL4 culture RNA samples

CD4$^+$ T cell differentiation and culture

In order to harvest and purify naïve primary mouse CD4$^+$ cells (this was carried out by James Hewitson), spleen and lymph nodes (axillary, inguinal, brachial, mesenteric) were taken from groups of Wt and *Malat1* KO mice (3 in each group) and used to create single cell suspensions by dissociation through a 70µm cell strainer. The suspensions were then centrifuged for 5 mins at 405x *g*, resuspended in 7ml Ammonium-Chloride-Potassium (ACK) lysing buffer to lyse red blood cells (5 mins, room temperature) and washed twice by centrifuging and resuspending in complete CD4 media, with cell counts performed using a haemocytometer and Trypan Blue (HyClone) to identify and exclude any dead cells. Following the second wash, the cells were incubated for 5 mins with 3µl CD4 microbeads (Miltenyi Biotech, L3T4) per 100 million cells, before being washed again and the CD4$^+$ cells separated using Magnetic Activated Cell Sorting (MACS). This involved running the sample through a LS column (Miltenyi Biotech) placed in a MidiMACS cell separator (Miltenyi Biotech) to bind the microbeads and thus the CD4$^+$ cells labelled with them, with the column subsequently washed 3 times with 5mls of complete CD4 media and removed from the magnetic separator**.** The CD4$^+$ cell fraction was then eluted using 5ml of media, before being centrifuged and resuspended once more.

These CD4$^+$ cells were subsequently sorted using a MoFlo Astrios EQ cell sorter (Beckman Coulter) – following MACS enrichment, 85µg/ml rat IgG were added to the cells to prevent nonspecific antigen binding and solutions of staining antibodies prepared from stocks as follows: 0.2mg/ml CD4 PerCP5.5 (Biolegend, clone RM4-5) diluted 1/200, 0.5mg/ml CD44 FITC (Biolegend, clone IM7) diluted 1/400, 0.2mg/ml CD62L PE (Biolegend, clone MEL-14) diluted 1/400, 0.2mg/ml CD8α APC (Biolegend, clone 53-6.7) diluted 1/400, 0.2mg/ml CDIIb APC (Biolegend, clone M1/7) diluted 1/400 and 0.5mg/ml MHCII A700 (Biolegend, clone M5/114.152) diluted 1/400. 50µl of this combined antibody mixture was added to the 450µl cell suspension, with single stain controls also carried out to allow correct compensation, and the resulting mixtures incubated on ice for 20-30 minutes. 900µl of media was then added to each solution to wash off excess antibodies, with the CD4 suspension then centrifuged at 405x *g* for 5 mins (4^{o}C), resuspended in 1ml of complete CD4 media, run through a 70µm cell strainer to remove any small clumps and transferred to Flow Assisted Cell Sorting (FACS) tubes. The single stains were then run through the cell sorter to determine the correct compensation to be applied for background fluorescence, before using this data as parameters to sort live, individual, CD4$^+$ naïve cells from both the Wt and KO cell suspensions (i.e. cells which are positive for CD4 and CD62 and negative for CD44, CD8α, CDIIb and MHCII).

Samples of these naïve cells were removed and lysed for further analysis and the rest cultured in 96 well plates, with 3 replicate cultures of 1x10^6 cells prepared for each treatment. These cultures were

stimulated with 10µg/ml platebound anti-CD3 (Biolegend, clone 145-2C11) and 2µg/ml soluble anti-CD28 (Biolegend, clone 37.51) antibodies with further lineage specific stimulation (Th0 = no further stimulation; Th1 = 15ng/ml recombinant IL-12 (Peprotech) and 5µg/ml anti IL-4 (Biolegend, clone 11B11); Th2 = 5µg/ml anti IFNγ (Biolegend, clone XMG1.2) and 5ng/ml recombinant IL-4 (Peprotech)) to induce differentiation of specific $CD4^+$ cell types. These cells were cultured for 4 days, after which half the cells were lysed and the remainder were allowed to rest and expand in 10units/ml IL-2 (Peprotech) to promote further *in vitro* culture growth and survival, and maintained for a further 2 days before harvesting (Figure 2).

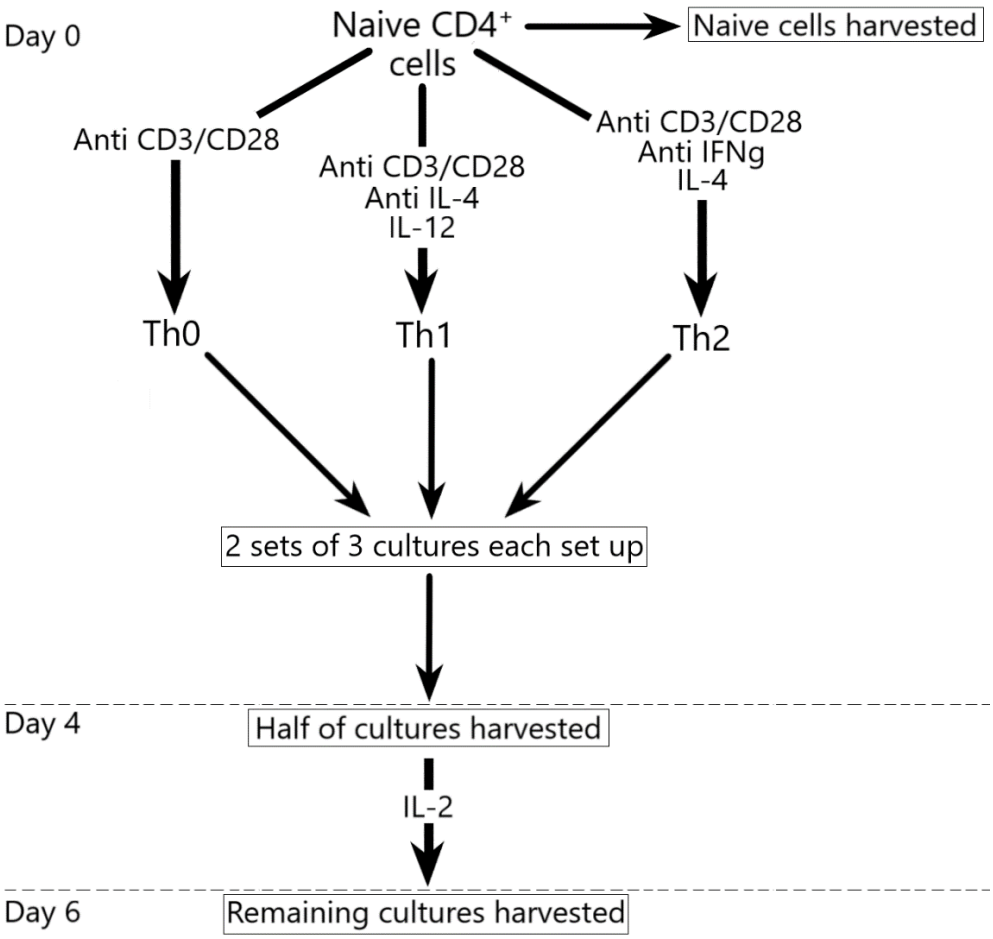

Figure 2: Simplified diagram of primary $CD4^+$ T cell stimulation and culture

Mouse tail fibroblast extraction and culture

Fibroblast cultures were isolated from mouse tails using the protocol detailed by Khan and Gasser (2016). Specifically, tails were harvested from 3 Wt and 3 *Malat1* KO mice, washed in 70% ethanol and air dried for 5 minutes before being placed in dishes containing 10ml of complete fibroblast media (1640 RPMI with 10% Fetal Calf Serum (FCS), 100units/ml Penicillin, 100µg/ml Streptomycin, 2mM L-Glutamine and 100µM L-Asparagine. Each tail was then cut into 2mm segments with sterile scissors

and transferred to a cryotube, and a mixture of 2.5mg/ml Collagenase D and 20mg/ml Pronase was added to a total volume of 2ml. These tubes were then incubated at 37^0C for 90 minutes, after which the tail segments were placed in a 70µm cell strainer and ground for 5 minutes into a further 10ml of media. The cell suspensions were then transferred to Falcon tubes and rinsed twice by centrifuging at 580x *g* for 5 minutes (4°C), removing the supernatant and resuspending in 10ml of complete fibroblast media. The cells were then transferred to 10cm culture dishes for incubation (37°C, 5% CO_2), with the media changed every 2 days for the first week and every 3 days afterwards.

Cultures were passaged upon reaching approximately 70-80% confluence – after removing the media and washing with 5ml of PBS (Gibco), each culture was incubated with 2ml of 0.05% Trypsin-EDTA solution (ThermoFisher Scientific) for 5 mins (37°C), following which 6ml of complete fibroblast media was added for neutralisation and the cells transferred to a 15ml Falcon tube. This was then centrifuged for 5 minutes at 450x *g*, following which the supernatant was removed and the cell pellet resuspended in 1ml of complete fibroblast media for cell counts using a haemocytometer, with 10µl of the cell stain Trypan Blue (HyClone) added to 10µl samples for easier identification of live and dead cells. The cells were then either split evenly between 2 new culture dishes or $2x10^6$ cells used to seed new dishes.

Fibroblast growth assays

To set up growth assays from the MTF cultures, the cells were extracted and counted as described above, then diluted with complete fibroblast media to $1x10^5$ cells/ml – for assays seeded with smaller cell numbers, aliquots of these cell suspensions were taken and diluted 1 in 10 to prepare $1x10^4$ cells/ml solutions. These diluted culture suspensions were then seeded in triplicate onto 96 well plates, with additional complete media added to each well to yield a total volume of 200µl per well and blank wells containing only media set up for comparison. These plates were incubated over the course of several days – at 24h time points, a plate was taken and 20µl of Alamar Blue (Thermo Scientific) added to each well, with further incubation of the plate for 20h to allow greater reduction of the resazurin reagent and thus sufficient sensitivity in the final assay. The absorbance of each well at 570 and 600 nm was then assessed using a Versamax microplate reader and used to calculate optical density, with the average optical densities of each culture and blank control subsequently used to determine the adjusted optical density of each culture at the given timepoints.

Fibroblast freezing, storage and defrosting

To freeze and store fibroblast cultures for further culture at a later date, freeze media was initially prepared from 90% FCS (HyClone) and 10% DMSO (Sigma) and chilled on ice, and the cells washed and lifted from the plates with Trypsin as described above for passage. After centrifuging the cell suspensions for 5 minutes at 1200 rpm and removing the resulting supernatant, the cells were resuspended in 1ml of chilled freeze media and transferred to a 2ml cryovial. The cryovials were then wrapped in layers of paper for insulation and stored at -80^{0}C overnight to gradually reduce the sample temperature, before transferring the samples to liquid nitrogen storage. In order to defrost these frozen samples, the cryovials were taken directly from liquid nitrogen storage and transferred to a 37^{0}C water bath to defrost, following which the cell suspensions were immediately transferred to 6ml of complete fibroblast media to dilute the DMSO and thus limit its toxic effects on the cells. After centrifuging at 1200 rpm for 5 minutes, the supernatant was then removed and the cells resuspended in 6ml of complete fibroblast media before being transferred to a 25ml culture flask (Corning).

RNA extraction, reverse transcription and qPCR expression analysis

Cell cultures were extracted by washing each culture once with 1ml PBS and lysing the cells with 700µl of Qiazol, with RNA subsequently extracted from these lysates using a miRNeasy Mini Kit (Qiagen, catalog no 217004) to produce 30µl of elutant. RNA concentrations and purity were then assessed using a Nanodrop spectrophotometer, with higher 260:280 and 260:230 ratios indicating lower levels of protein, phenol and carbohydrate contaminants in each 1µl sample. For our experiments, 260:280 and 260:230 ratios over 2 were considered to represent pure RNA samples, with lower results indicating some protein or carbohydrate contamination of the extracted RNA. Additionally, if any RNA extraction had a substantially higher or lower concentration than the other samples, appropriate dilutions were then prepared from either this sample or the others respectively to ensure approximately equal total RNA concentrations across all extractions, generally around 50 – 100 ng/µl.

Reverse transcription was performed with 1µl of each sample, using 1µl random hexamer primers (Invitrogen, 100µM), 1µl dNTPs (ThermoFisher, 10mM), 0.5µl Superscript III (Invitrogen, 200U/µl), 1µl RNAse OUT (Invitrogen, 40U/µl), 2µl DTT (Invitrogen, 100mM), and 4µl 5x first-strand buffer (Invitrogen; 250 mM Tris-HCl, 375 mM KCl, 15 mM MgCl2), with 9.5µl nuclease free H_2O added to give a final 20µl cDNA solution. These reagents were used with the First-Strand cDNA synthesis reaction specified for SuperScript^{TM} III Reverse Transcriptase (ThermoFisher, catalog no 18080044). qPCR plates were set up from this cDNA, with each well containing 10µl 2X Fast SybrGreen master mix (ThermoFisher), 0.6µl each of 10µM primers, 7.8µl nuclease free H_2O and 1µl of the cDNA sample, with each cDNA sample run in duplicate and averaged to reduce any technical error from the plate

readings. For genes tested using Quantitect primer assays (Qiagen), these master mixes were adjusted so that each well contained the same volume of Fast SybrGreen master mix and cDNA, but 2µl of the 10X primer mixture and 7µl of nuclease free H_2O instead. These plates were run on a StepOnePlus qPCR platform using the protocol specified for Fast SYBR™ Green Master Mix (ThermoFisher, catalog no 4385612). The threshold cycle (Ct) at which the fluorescence of each sample rose above a threshold level relative to the baseline was recorded, with relative expression levels from each gene and cell type determined from the difference in Ct values (ΔCt) between the target gene and a housekeeping control as described by Levak and Schmittgen (2001), i.e:

$\Delta Ct = Target\ Ct - Control\ Ct$ $\quad\quad$ $Relative\ expression\ level = 2^{\Delta Ct}$

These were then normalised by dividing each result by the relevant control value.

Statistical testing

To assess for significant differences between the Wt and *Malat1* KO genotypes, relative expression levels were normalised to the average Wt results for each cell type and pooled from all experimental runs. Unpaired, two-tailed T tests were then used to test for significant differences in gene expression between Wt and KO, with Welch's correction used to compensate for differences in standard deviation between the genotypes. To compare expression of genes across different $CD4^+$ cell types, the samples were instead normalised to the naïve relative expression levels for each gene, and a one-way analysis of variance (ANOVA) used to test for significant expression differences between the pooled data sets for each $CD4^+$ cell type, with Tukeys multiple comparisons test used to identify differences between specific $CD4^+$ cell types. Expression levels of genes in $CD4^+$ and fibroblast cultures was compared based upon the non-normalised relative expression values in each cell type, with significant differences assessed using a one-way ANOVA and Tukeys multiple comparisons test as for expression between $CD4^+$ cell types.

Primer design, testing and optimisation

Expression levels of *Malat1*, *Neat1*, *U6*, *GAPDH*, *HPRT*, *IL-4*, *IL-10* and *IFNγ* were assessed using previously validated qPCR primer stocks, while qPCR primers for all other genes were designed using the online Roche primer design tool and ordered from Sigma. These were reconstituted with nuclease free water to 100µM stock solutions, and aliquots of these stocks taken and diluted 1 in 10 to produce the final 10µM concentration used in experiments. For some genes whose primer pairs were not suitable when generated through this system, due to either primer-dimer interactions or off target amplification producing multiple melting curve peaks, Quantitect primer assays (commercially

available SYBR Green qPCR primer pairs) were ordered and reconstituted in nuclease free H_2O, producing a 10X stock solution containing a mixture of forward and reverse primers.

To test the efficiency and purity of the new primer sets following resuspension, eight existing cDNA samples synthesised from liver and spleen tissues of Wt mice were pooled to create a stock cDNA solution for expression analysis, with 1 in 4 serial dilutions prepared from this stock solution to generate a series of cDNA concentrations ranging from 1 (the undiluted solution) to 1 in 256. qPCRs were carried out from these samples using the desired primers, with nuclease free H_2O as a negative control. The Ct values from each dilution were then plotted against the log10 values for each dilution, and the slope of the resulting graph used to assess the amplification efficiency of each primer pair via the ThermoFisher qPCR Efficiency calculator located here:

https://www.thermofisher.com/uk/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/qpcr-efficiency-calculator.html

In addition, the amplification and melt curves were examined to assess the quality of each primer pair, with multiple melt curve peaks indicating potential off target amplification or primer-dimer pair amplification and closely spaced amplification curve peaks indicating low amplification and possibly little to no expression of the gene of interest.

Results

Primer optimisation

Given that most of our qPCR primer pairs were newly designed for these experiments, each of these new primer sets were tested using stock cDNA solutions from liver and spleen samples to assess primer efficiency and check for any evidence of off target amplification or primer-dimer interactions. Amplification and melting curves of the *Malat1* primers were also assessed for the naïve $CD4^+$ samples in order to determine if the lncRNA was accurately amplified in these cultures. Based on the liver and spleen qPCR data, most of the primer pairs amplified a single product, with the sample melting curves forming single, close peaks and indicating little to no primer contamination or primer dimer pair formation while the amplification data showed the expected gradient when plotted against the different cDNA concentrations, and thus the desired level of amplification efficiency. Though the blank samples for several primer pairs, such as *Scyl1*, *Map3k11*, *Hist1h4a* and Las1l, showed some amplification of off target products (Figures 4B, 5B, 8B and 12B), this was only observed at high Ct values of around 34 – 40, and as such would not affect the comparison of the much lower Ct values observed in our $CD4^+$ and fibroblast cultures, typically around 15 – 30. Similarly, while some primer sets, such as *Scyl1*, *Hist1h4a*, *U2af1* and *Stau1* show some off-target amplification at around 81^0C (Figures 4B, 8B, 10B and 11B), this was only observed at the highest cDNA concentration, and thus likely indicates some slight contamination in the liver/spleen cDNA solution rather than in the primer pairs. While some primer sets showed good amplification efficiency, with *Hist1h4a*, *Ddx23* and *IFNγ* primers showing around 80-100% amplification with each PCR cycle, some were less efficient at around 75-80% efficiency, including the *Scyl1*, *Map3k11* and *U2af1* primer pairs (Table 2). While this efficiency is not ideal for expression analysis, it is still within an acceptable level of efficiency for further experiments, and as such these primers were used for further $CD4^+$ and fibroblast expression analysis.

Primer ID	Supplier	Sequence	Efficiency (%)
Malat1 forward (mouse)	Eurofins	TGCAGTGTGCCAATGTTTCG	Not assessed **
Malat1 reverse (mouse)	Eurofins	GGCCAGCTGCAAACATTCAA	Not assessed **
Malat1 forward (human)	Eurofins	GAATTGCGTCATTTAAAGCCTAGTT	Not assessed **
Malat1 reverse (human)	Eurofins	GTTTCATCCTACCACTCCCAATTAAT	Not assessed **
Neat1 forward (mouse)	Eurofins	CCTAGGTTCCGTGCTTCCTC	Not assessed **
Neat1 reverse (mouse)	Eurofins	CATCCTCCACAGGCTTACCG	Not assessed **
Neat1 forward (human)	Eurofins	ATGGGGAAGTAGTCTCGGGT	Not assessed **
Neat1 reverse (human)	Eurofins	TGAAGGCAATGTGATAGGGGTC	Not assessed **
Tigd3 forward	Qiagen	Not available (Qiagen catalog no QT00493500) *	204.92
Tigd3 reverse	Qiagen	Not available (Qiagen catalog no QT00493500) *	204.92
Scyl1 forward	Sigma	AACCGCTTTGTTGAGACCAA	79.14
Scyl1 reverse	Sigma	CTGCTTCTCAGCTGGCTCTT	79.14
Map3k11 forward	Sigma	CCCTTTGCACAACTCATGG	78.8
Map3k11 reverse	Sigma	CAGGATGGAGGCGAAGTC	78.8
U6 forward	Eurofins	TTCACGAATTTGCGTGTCAT	Not assessed **
U6 reverse	Eurofins	CGCTTCGGCAGCACATATAC	Not assessed **
Gapdh forward	Sigma	GGAGTCAACGGATTTGGTCGTA	Not assessed **
Gapdh reverse	Sigma	GGCAACAATATCCACTTTACCAGA	Not assessed **
Hprt forward	Sigma	GTTGGATACAGGCCAGACTTTGTTG	Not assessed **
Hprt reverse	Sigma	GATTCAACCTTGCGCTCATCTTAGGC	Not assessed **
IL-2 forward (mouse)	Sigma	GCTGTTGATGGACCTACAGGA	Not assessed **
IL-2 reverse (mouse)	Sigma	TTCAATTCTGTGGCCTGCTT	Not assessed **
IL-2 forward (human)	Sigma	GAATCCCAAACTCACCAGGATGCTC	Not assessed **
IL-2 reverse (human)	Sigma	TAGCACTTCCTCCAGAGGTTTGAGT	Not assessed **
IFNγ forward	Sigma	GGATGCATTCATGAGTATTGC	84.37
IFNγ reverse	Sigma	GCTTCCTGAGGCTGGATTC	84.37
IL-4 forward	Eurofins	CATCGGCATTTTGAACGAG	107.28
IL-4 reverse	Eurofins	CGAGCTCACTCTCTGTGGTG	107.28
IL-10 forward	Qiagen	Not available (Qiagen catalog number QT00106169) *	Not assessed **
IL-10 reverse	Qiagen	Not available (Qiagen catalog number QT00106169) *	Not assessed **
Hist1h4a forward	Sigma	CCCTGAAAAAGCGCTGTAAT	102.11

Hist1h4a reverse	Sigma	TCCAAAGGCACTCAAGGTTT	102.11
Ddx23 forward	Sigma	GAAAGATGGCGACTGTTCG	92.25
Ddx23 reverse	Sigma	ATGCGTCTCGGTCCTTTTT	92.25
U2af1 forward	Qiagen	Not available (Qiagen catalog number QT00175084) *	79.22
U2af1 reverse	Qiagen	Not available (Qiagen catalog number QT00175084) *	79.22
Stau1 forward	Qiagen	Not available (Qiagen catalog number QT00125069) *	77.92
Stau1 reverse	Qiagen	Not available (Qiagen catalog number QT00125069) *	77.92
Las1l forward	Sigma	GCATGGCAAGTAAGCTCTGA	77.33
Las1l reverse	Sigma	TTGGTCCAAAAGATACATGGTG	77.33
Cpsf1 forward	Sigma	GTCAGGCCTAAAGGGCTATGT	79.38
Cpsf1 reverse	Sigma	TCAGGCACCACTTCAATCAC	79.38

Table 2: Sources and sequences of all qPCR primer pairs used for expression analysis. N.B: * = sequences of Quantitect primer assays were proprietary, and thus unavailable for inclusion. Product numbers have been included instead. ** = primers were previously optimised in the Lagos lab from other experiments, and thus were not tested.

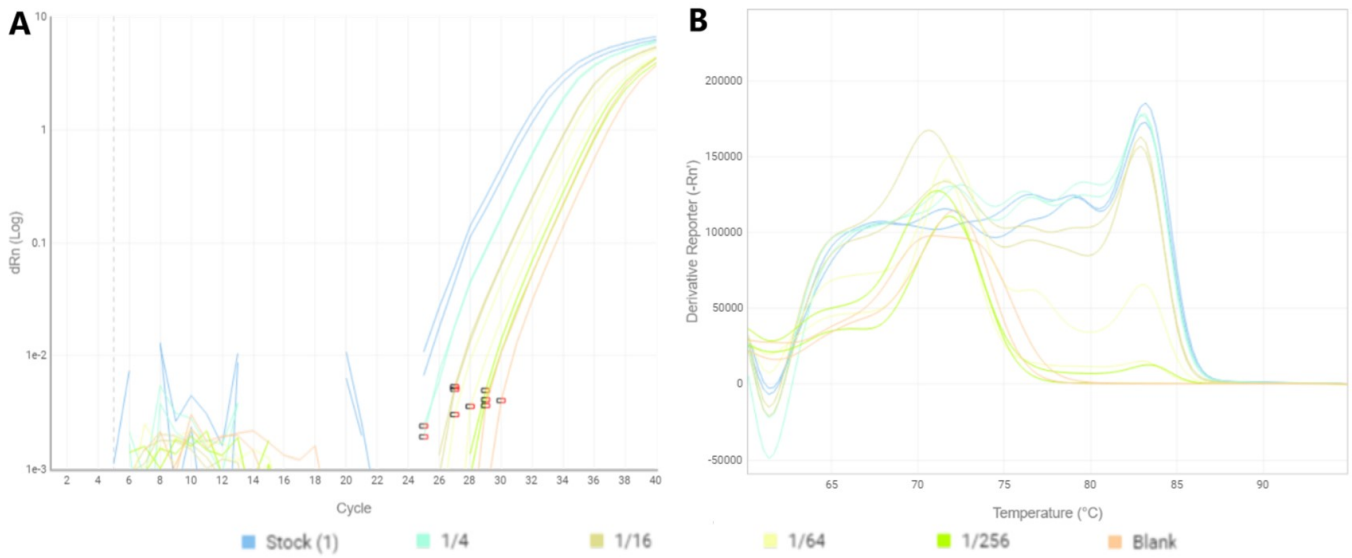

Figure 3: *Tigd3* primer test data, showing (A) amplification curve and (B) melting curve for liver/spleen samples

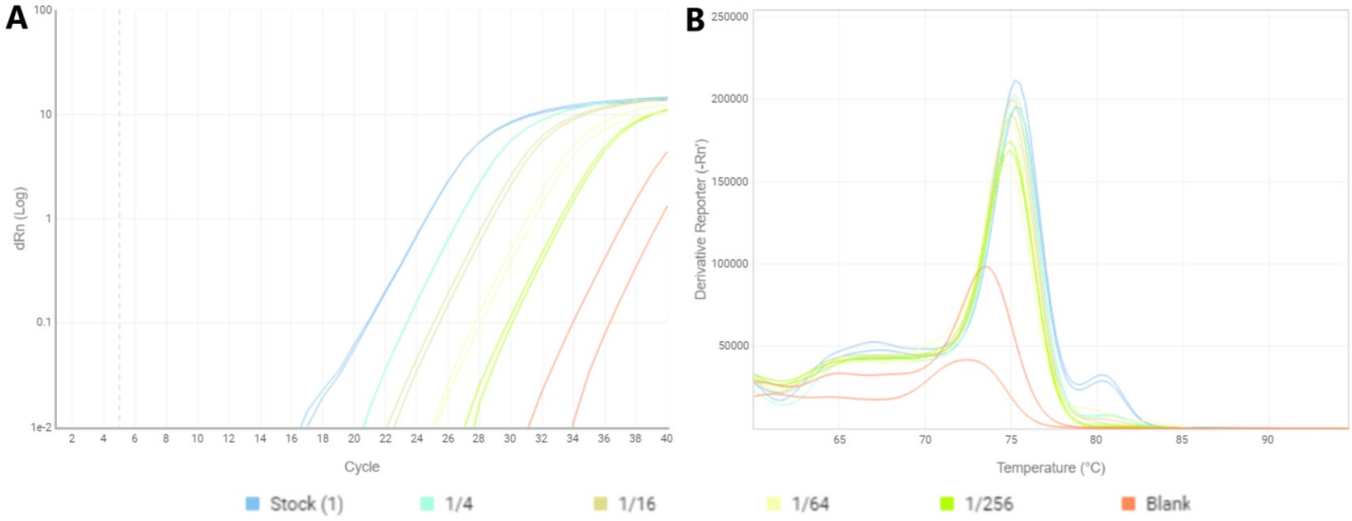

Figure 4: *Scyl1* primer test data, showing (A) Amplification curve and (B) melting curve for liver/spleen samples

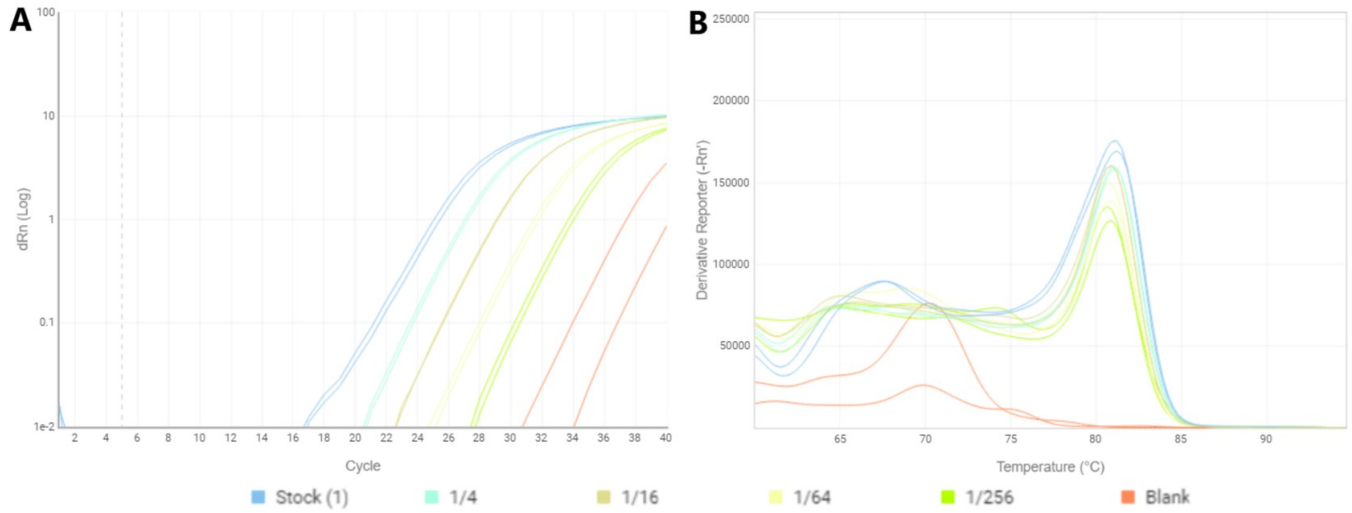

Figure 5: *Map3k11* primer test data, showing (A) amplification curve and (B) melting curve for liver/spleen samples

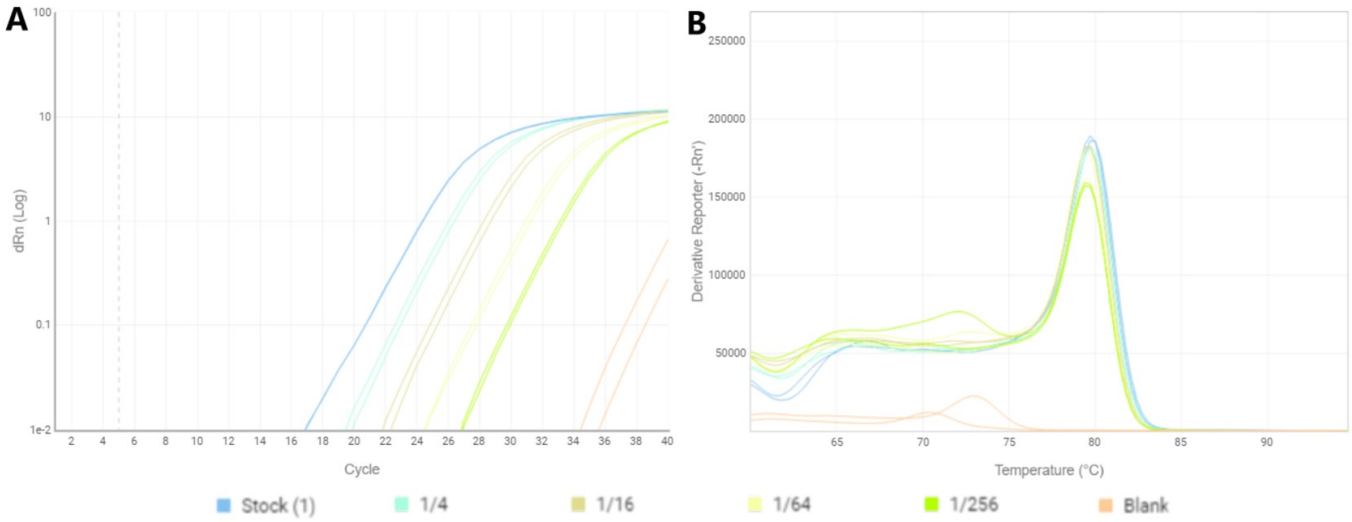

Figure 6: *IFNγ* primer test data, showing (A) amplification curve and (B) melting curve for liver/spleen samples

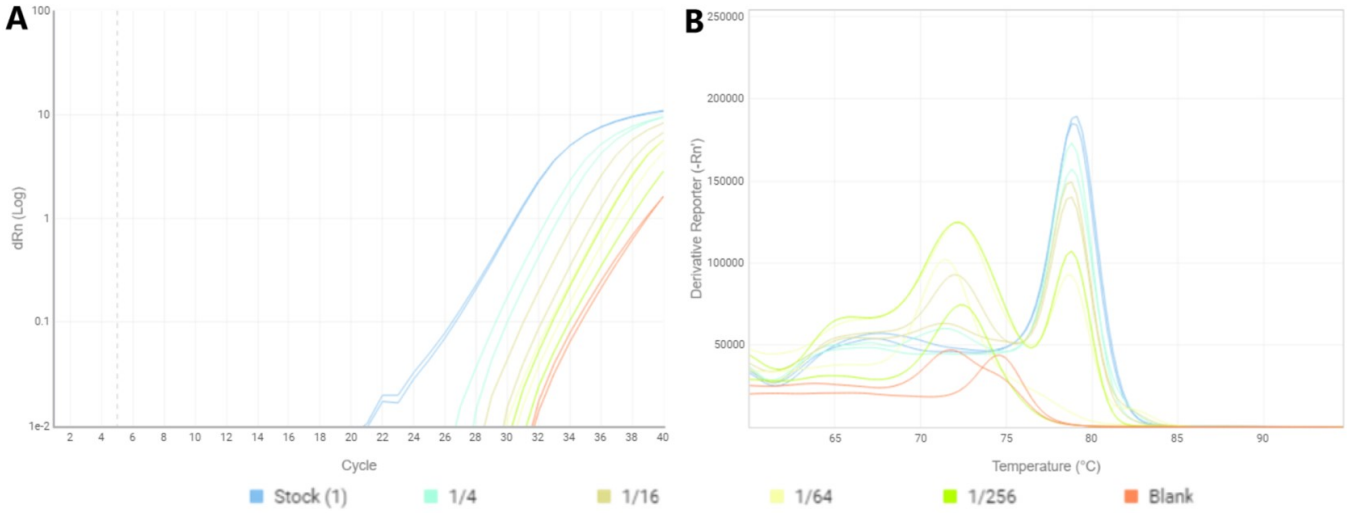

Figure 7: *IL-4* primer test data, showing (A) amplification curve and (B) melting curve for liver/spleen samples

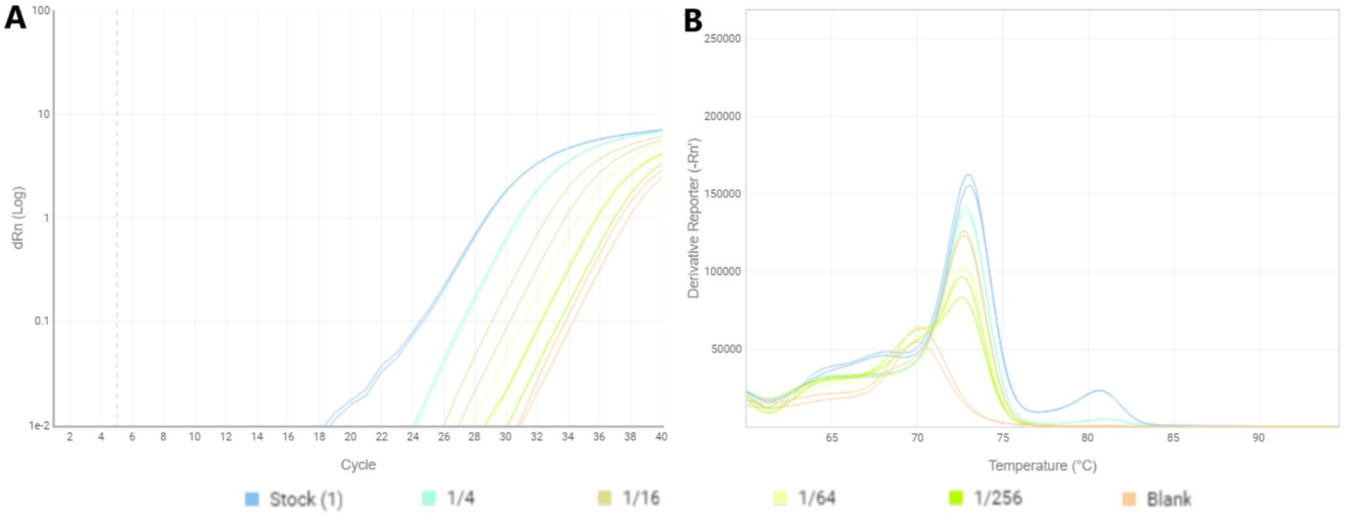

Figure 8: *Hist1h4a* primer test data, showing (A) amplification curve and (B) melting curve for liver/spleen samples

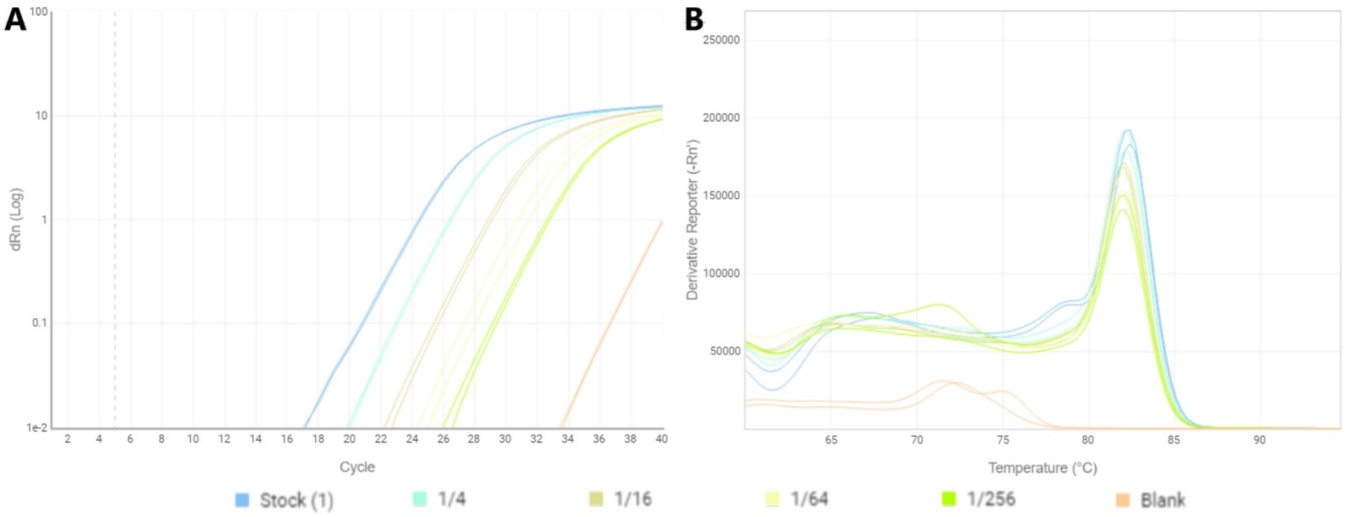

Figure 9: *Ddx23* primer test data, showing (A) amplification curve and (B) melting curve for liver/spleen samples

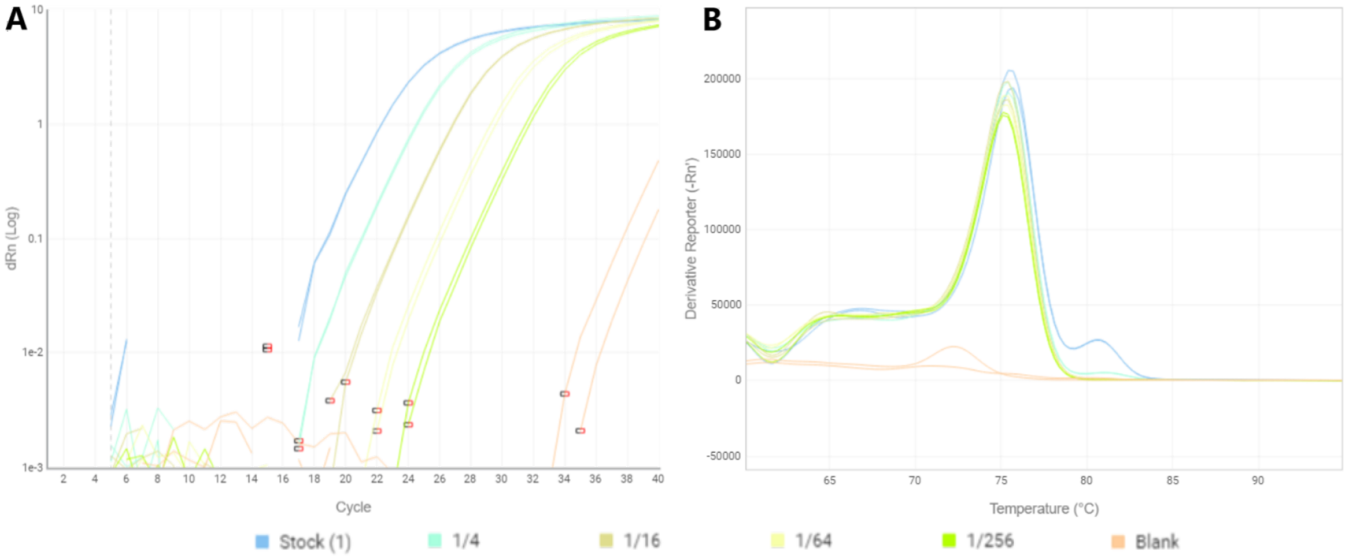

Figure 10: *U2af1* primer test data, showing (A) amplification curve and (B) melting curve for liver/spleen samples

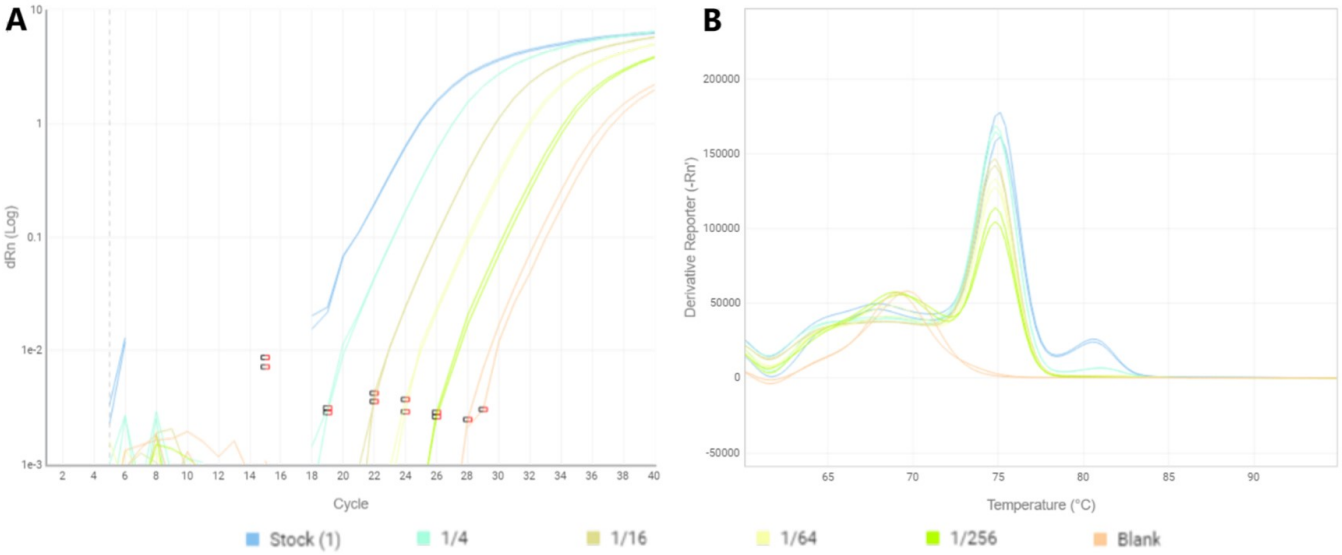

Figure 11: *Stau1* primer test data, showing (A) amplification curve and (B) melting curve for liver/spleen samples

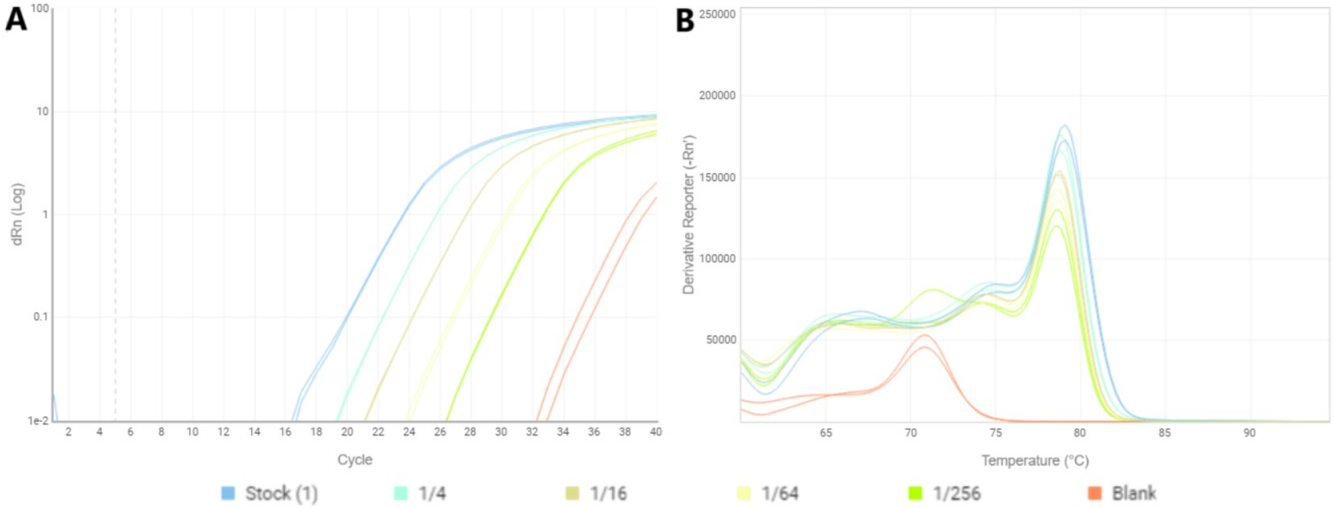

Figure 12: *Las1l* primer test data, showing (A) amplification curve and (B) melting curve for liver/spleen samples

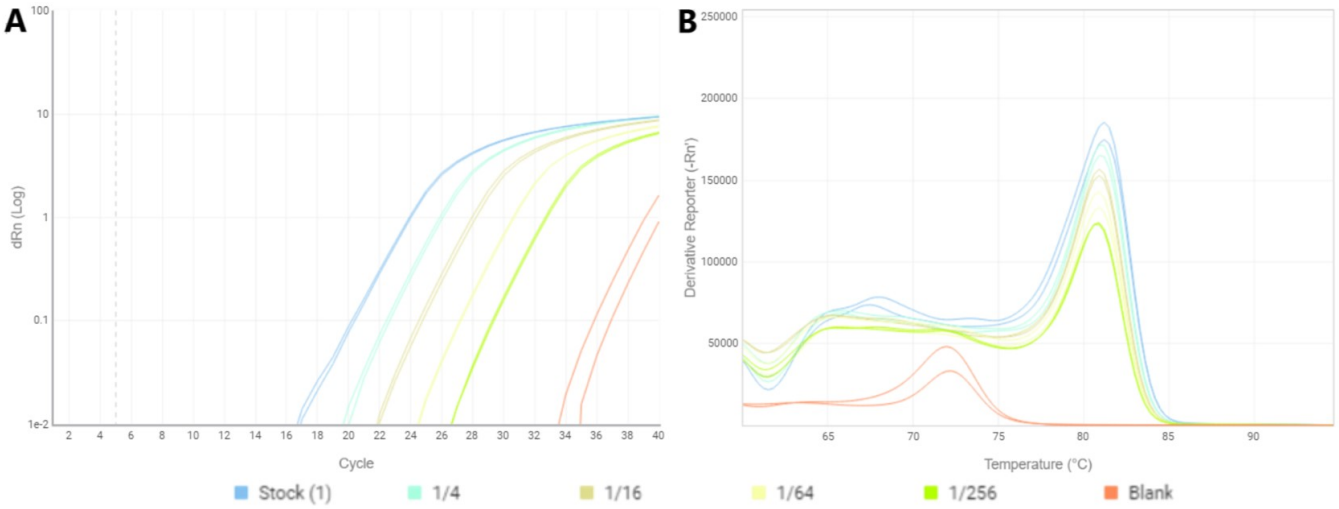

Figure 13: *Cpsf1* primer test data, showing (A) amplification curve and (B) melting curve for liver/spleen samples

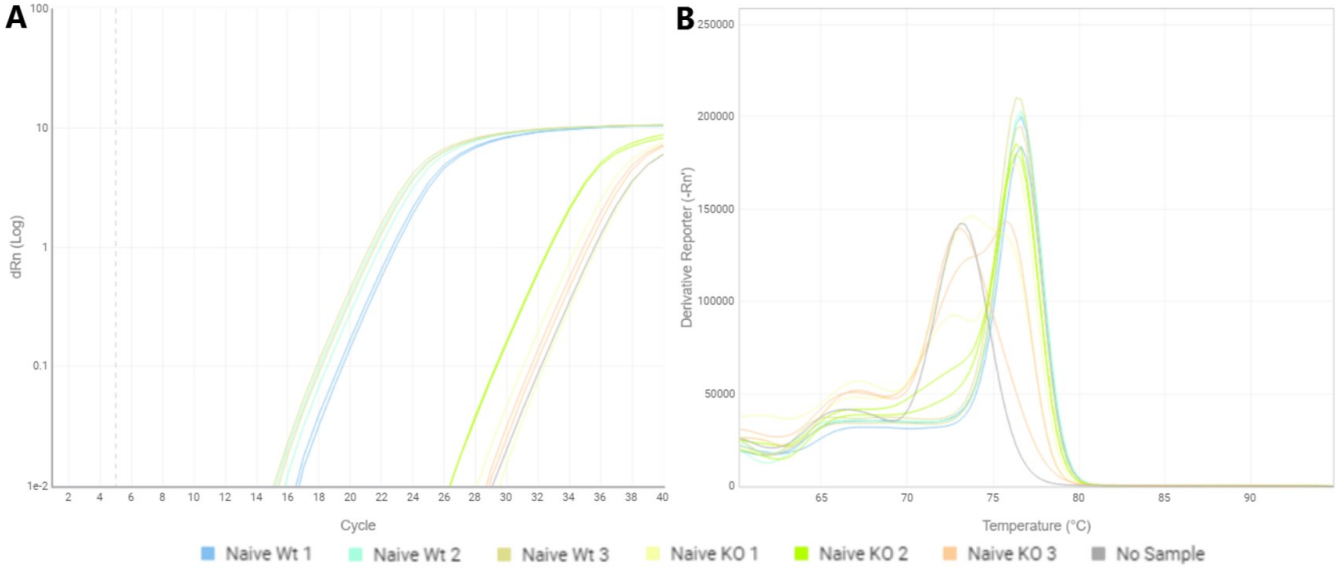

Figure 14: *Malat1* primer data, showing (A) amplification curve and (B) melting curve for naïve $CD4^+$ samples

The *Tigd3* primers produced very problematic primer test results, with numerous off target melting curve peaks and very little evidence of qPCR amplification observed in the liver/spleen qPCR data (Figures 3B and 3A). However, a possible solution may be found in the *IL-4* amplification data – while the *IL-4* amplification curves are similarly closely spaced with considerable off target amplification at around 72-73^{0}C (Figures 7A and 7B), this likely reflects the extremely low expression of this cytokine in liver and spleen cells (Yue *et al* 2014). Our *Tigd3* primers showed similarly close amplification curves – furthermore, previous transcriptomic data found very low expression of *Tigd3* in human liver, spleen and lymph node tissues, and similarly low expression in murine $CD4^+$ cell lines (Fagerberg *et al* 2014, Stubbington *et al* 2015). This suggests that the results observed for the *Tigd3* primers reflect little to no expression of the gene in our $CD4^+$ cell cultures, particularly given that three different primer sets, including a commercially available Quantitect primer assay**,** showed similar amplification and melt

curve results (data not shown). As such, the Quantitect primer pair was considered viable for further expression analysis of *Tigd3* – however, given the low expression observed in both T cell and fibroblast cultures, significant differences in *Tigd3* levels between these cultures should be considered with caution.

Similarly, while the amplification plots for the *Malat1* primers showed the expected difference between the two genotypes in naive $CD4^+$ cells, with relatively low Ct values for the Wt samples and very similar Ct values of around 31-32 for the *Malat1* KO samples and blank controls (Figure 14A), the melting curve results were unexpected, with strong off target peaks observed from both the *Malat1* KO and blank wells at around 73^0C (Figure 14B). This amplification likely represents some slight primer dimer interactions of the *Malat1* primers occurring in the absence of a fully complementary cDNA sequence, a possibility supported by both the similarity of the Ct values observed in the blank and *Malat1* KO samples and the smaller, wider peak observed for these melting curves, as these features are consistent with primer-dimer amplifications. However, the KO2 sample, which had a lower Ct value than the other KO's of 29, showed little to no off-target amplification, with the melting curve matching those of the Wt samples (Figure 14B) – this indicates that these primer dimer amplifications occur only at extremely low levels of *Malat1*. As such, although these results are not ideal, the *Malat1* primers appear reliable for assessing the genes expression in $CD4^+$ T cell cultures based on Ct values, provided these Ct values are sufficiently lower than that of the negative control.

Effects of stimulating immortalised cell lines on *Malat1* and *Neat1* expression

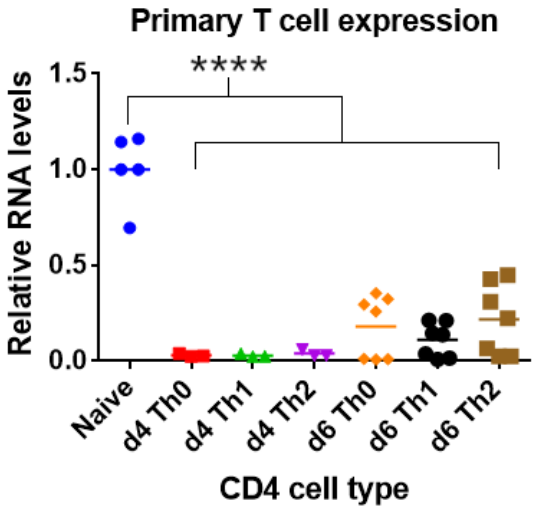

Figure 15: Relative *Malat1* RNA levels in different Wt primary CD4⁺ cell types, based upon ΔCt between *Malat1* and *U6* and normalised to the naïve average. Significance based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

Given that previous studies in the Lagos lab showed significant downregulation of *Malat1* upon activation of primary CD4⁺ T cells (Hewitson *et al* unpublished, Figure 15), this project initially aimed to determine if immortalised CD4⁺ cell lines, such as the human and murine cancer cell lines Jurkat and EL4, showed similar responses of *Malat1* to stimulation and thus might be a viable experimental model. These immortalised cell lines are far easier both to culture and maintain over long time periods and to manipulate *in vitro* than harvested primary cells, and thus would be preferable for further experiments. To assess this, cultures of Jurkat and EL4 cells were set up and stimulated with ConA or a mixture of PMA and PHA respectively, as these mitogens have previously been used in numerous studies for stimulation of T cell lines, including Jurkat and EL4 cells (Pang *et al* 2012, Lim *et al* 2016). These cultures were then incubated for either 24 or 48 hours and harvested, with subsequent RNA extraction, reverse transcription and qPCR expression analysis used to compare *Malat1* levels between stimulated and unstimulated cultures on both days (N.B. For the initial run of Jurkat cultures, RNA extraction, reverse transcription and *Malat1* and *IL-2* qPCR analysis were carried out by Dr Dimitris Lagos). EL4 cells were stimulated with different ConA concentrations to determine the ideal stimulation treatment for further experiments, using PMA stimulation as a positive control for CD4⁺ stimulation given its effective stimulation of the Jurkat cells, while *Neat1* expression was also tested to compare the response of another nuclear localised lncRNA potentially regulated by *Malat1* (Nakagawa *et al* 2012). Relative expression in Jurkat and EL4 cells was based upon ΔCt compared to *GAPDH* and *Hprt* respectively.

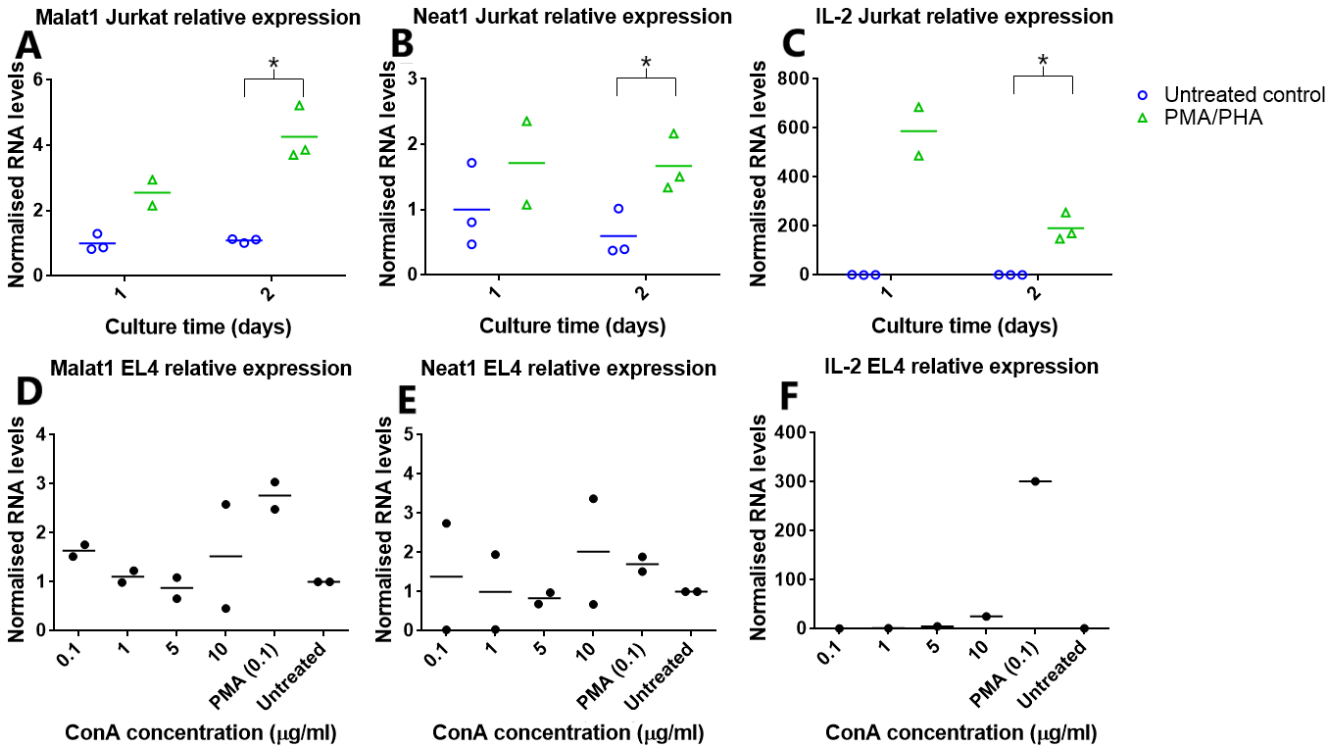

Figure 16: PCR expression data from immortalised CD4⁺ cell cultures. Relative RNA levels of (A) *Malat1*, (B) *Neat1* and (C) *IL-2* in Jurkat cultures are shown (N.B: RNA extraction for one day 1 PMA/PHA culture was unsuccessful and thus removed from analysis). Levels of (D) *Malat1*, (E) *Neat1* and (F) *IL-2* in EL4 cells, with expression in both cell lines normalised to the average of their respective Wt samples. Significance based on 2 tailed t-test; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001.

Jurkat culture	Gene			
	Malat1	***Neat1***	***IL-2***	***GAPDH***
d1 unstimulated	16.706	32.80155	31.337014	18.440049
d1 PMA/PHA	16.09186	32.67733	22.835433	19.191222
d2 unstimulated	15.12729	32.07893	29.467228	17.010846
d2 PMA/PHA	15.30054	32.58853	24.417203	19.138806

Table 3: Average Ct values of *Malat1*, *Neat1* and control genes from qPCR expression analysis of Jurkat cultures. Relative expression values for each culture calculated and normalised as detailed in the Methods section.

EL4 culture	Gene			
	Malat1	***Neat1***	***IL-2***	***Hprt***
0.1	20.87462	23.48581	24.796991	24.949281
1	20.531941	23.600416	26.008762	24.214249
5	21.340299	26.371088	29.656187	25.413683
10	20.47501	26.11797	28.009914	25.134259
PMA control	20.89315	24.661378	22.527192	26.30905
Untreated control	20.247962	23.341251	29.769643	24.242857

Table 4: Average Ct values of *Malat1*, *Neat1* and control genes from qPCR expression analysis of EL4 cultures. Relative expression values for each culture calculated and normalised as detailed in the Methods section.

While in earlier data (Hewitson *et al*, unpublished) *Malat1* was downregulated upon stimulation of primary mouse $CD4^+$ T cells, mitogenically stimulated EL4 cultures showed no significant change in *Malat1* levels. Although an approximately threefold increase was observed with PMA stimulation, the low sample size prevented statistical testing of this result (Figure 16D) – however, the highly similar Ct values for *Malat1* across the EL4 cultures indicates this difference is unlikely to be significant (Table 4). In Jurkat cells, on the other hand, significantly increased *Malat1* expression was observed in stimulated cells by the second day of culture – although some increase was observed in the day 1 cells, issues with one of the cultures produced too low a sample size for statistical testing (Figure 16A). However, the Ct values for *Malat1* showed little to no change between these cultures – as such, this significant difference in the day 2 cells likely occurred due to slight differences in the *GAPDH* Ct values between stimulated and unstimulated Jurkat cultures (Table 3).

Similarly, while *Neat1* also displayed increased RNA levels in the stimulated Jurkat cultures, reaching significance by day 2 similar to *Malat1* (Figure 16B), the highly similar Ct values in each culture (Table 3) suggest this is also due to the altered *GAPDH* Ct values. Neat1 likewise showed no clear expression change upon ConA stimulation of EL4 cells and a notable increase with PHA stimulation of unknown significance (Figure 16E) – as with *Malat1*, this increase might prove significant with larger sample sizes, but this is not reflected by the raw Ct values, which seem to indicate a decrease in Neat1 levels in this culture (Table 4). Taken together, this data indicates that *Malat1* expression patterns and function differ between immortalised and primary $CD4^+$ T cell lines, with significant downregulation of *Malat1* upon T cell receptor activation of primary $CD4^+$ cells (Figure 15) and either slight upregulation or no significant change in mitogenically stimulated T cell lines (Figures 16A and 16D). As such, these immortalised cell lines would not be a viable *in vitro* model for assessing regulation by *Malat1* in primary cells, and primary $CD4^+$ cell cultures were therefore used for all further expression analysis of *Malat1* and all other genes assessed in these experiments.

Selection of candidate genes for assessment of *Malat1 in cis* regulation

To identify candidate genes for further analysis from those surrounding *Malat1*, initially the genes identified by Zhang *et al* (2012) to show significant changes in *Malat1* knockouts were selected, as these had previously shown evidence of potential *in cis* regulation by *Malat1* in a similar KO model. Expression levels of these genes across different $CD4^+$ cell types were then compared using the online database Th Express, containing data from a previous transcriptomic study of mouse $CD4^+$ T-cells (Stubbington *et al* 2015). This data was assessed for any changes in expression corresponding to changes in *Malat1* levels, as such similarities might indicate regulation of the target gene by *Malat1*. Previous data on *Malat1* function and interactions was also considered, as well as the relative location

of each gene to the *Malat1* locus, to determine a final pool of candidate genes for CD4$^+$ expression analysis.

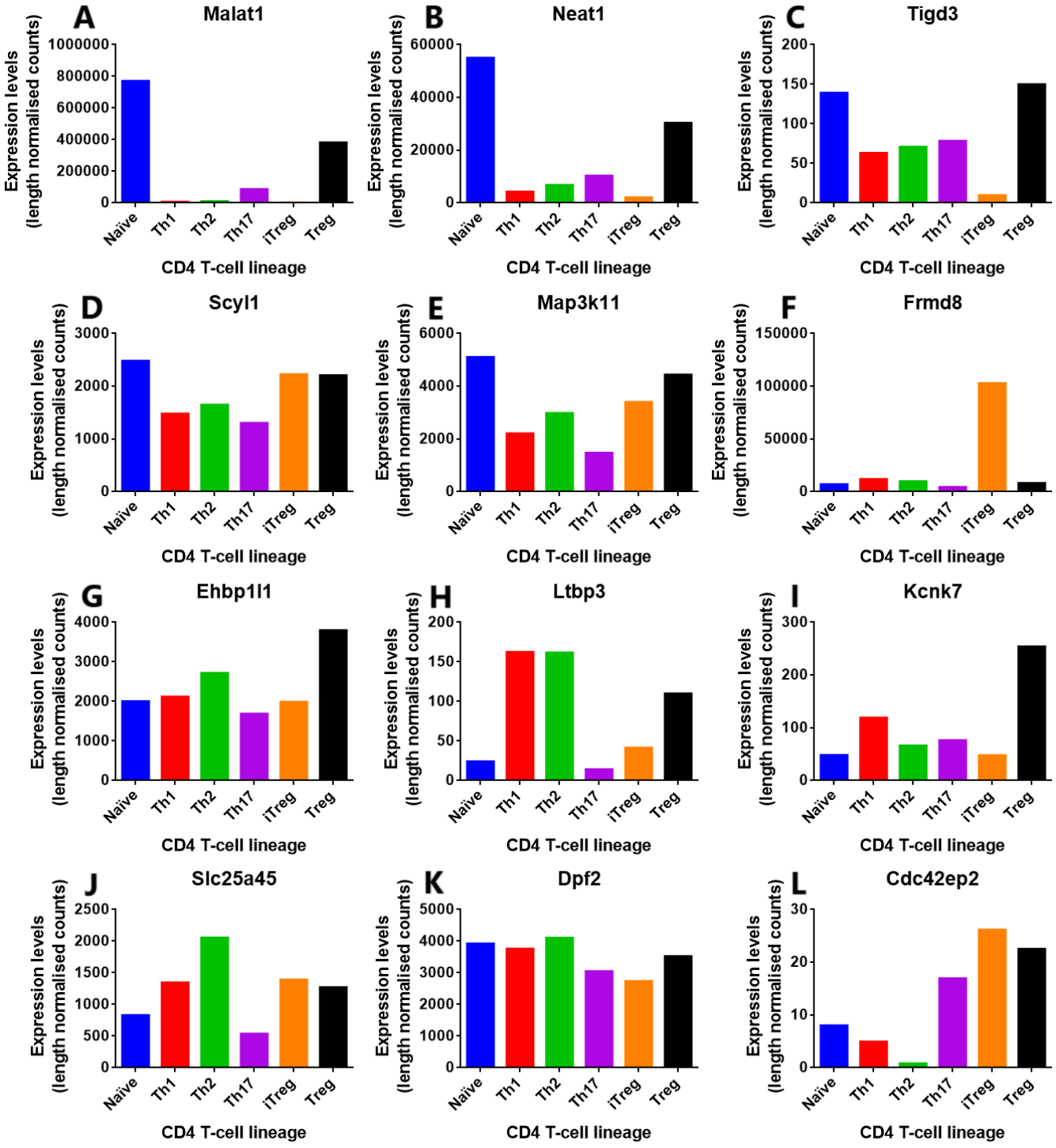

Figure 17: Relative CD4$^+$ expression levels of neighbouring genes of *Malat1*, based on previous transcriptomic data (Stubbington *et al* 2015). Expression levels of (A) *Malat1*, (B) *Neat1*, (C) *Tigd3*, (D) *Scyl1*, (E) *Map3k11*, (F) *Frmd8*, (G) *Ehbp1l1*, (H) *Ltbp3*, (I) *Kcnk7*, (J) *Slc25a45*, (K) *Dpf2* and (L) *Cdc42ep2* are plotted as length normalised counts for each CD4$^+$ subtype, including naïve, Th1, Th2 and Th17 lines, regulatory T cells (Treg) and induced regulatory T cells (iTreg). Expression was based on RNA sequencing counts, normalised by transcript length to allow more accurate comparison of mRNA levels between genes sequenced to different depths.

While most of the genes surrounding *Malat1* showed no clear pattern of $CD4^+$ expression**,** with very different patterns of up- and down-regulation across different cell types, several potential candidate genes were identified based on this transcriptomic database. Initially, Nuclear Enriched Abundant Transcript 1 (*Neat1*), another nuclear localised lncRNA with key functions in alternative splicing structures known as paraspeckles (Clemson *et al* 2009, Lin *et al* 2018), was assessed given both the similar patterns of downregulation to *Malat1* observed in stimulated $CD4^+$ cultures (Figure 17B) and the fact that two separate *Malat1* knockout papers found evidence of *Neat1* regulation by *Malat1* (Zhang *et al* 2012, Nakagawa *et al* 2012). Furthermore, other studies have shown matching functional roles for *Malat1* and *Neat1*, with the two lncRNAs targeting and binding jointly to numerous active genomic sites (West *et al* 2014). Additionally, both genes showed very similar patterns of expression across $CD4^+$ cell types (Stubbington *et al* 2015). Tigger Transposable Derived Element 3 (Tigd3), a nuclear localised paralogue of centromere binding proteins (Marshall and Choo 2012), also showed similar expression changes to both *Malat1* and *Neat1* across the $CD4^+$ transcriptomic data (Figure 17C) – this suggests that *Tigd3* may be regulated by or functionally linked to *Malat1* in mouse $CD4^+$ cells, and as such the gene was likewise selected for further analysis.

Scy Like 1 (Scyl1), a catalytically inactive pseudokinase with key regulatory functions in Golgi body homeostasis and neuronal survival (Burman *et al* 2010, Schmidt *et al* 2007), was also selected for further analysis**.** This choice was primarily based on the genes location immediately downstream of *Malat1* (Figure 1B) – given that numerous examples of *in cis* regulation by other lncRNAs have been found to regulate downstream or nearby gene expression by various mechanisms, such as transcriptional interference (Ard *et al* 2014, Ard *et al* 2016), *Scyl1* would be one of the most likely candidates for *in cis* regulation by *Malat1*. Finally, Mitogen Activated Protein Kinase Kinase Kinase 11 (Map3k11), a serine threonine kinase which functions as a positive regulator of JNK signalling and as a tumour suppressor in prostate cancers and B-lymphocytes (Gallo and Johnson 2002, Whitworth *et al* 2012, Knackmuss *et al* 2016), was found to show similar $CD4^+$ expression patterns to Scyl1 (Figures 17D and 17E). This suggested potential coregulation of *Scyl1* and *Map3k11* by a common factor which might be *Malat1* – furthermore, as *Map3k11* was the only one of these four genes transcribed from the opposite strand to *Malat1*, it allowed potential *in cis* regulation of this strand by *Malat1* to be assessed.

Effects of *Malat1* knockout on surrounding gene expression in CD4$^+$ T cells

To assess potential *in cis* regulation by *Malat1* of its surrounding genes, phenotypically naïve primary CD4$^+$ T cells (CD4+ CD62L+ CD44-) were sorted from naïve Wt and *Malat1* KO mice and differentiated *in vitro* to generate Th0, Th1 and Th2 primary cells. RNA was then isolated from three cultures of each cell type at 4 or 6 days of culture, and qPCR expression analysis used to assess expression levels of several candidate genes surrounding *Malat1*. The ubiquitous snRNA *U6* was used as a housekeeping control to determine relative expression, as *HPRT* and *GAPDH* expression were shown to differ between CD4$^+$ cell types in a previous CD4$^+$ transcriptomic study (Stubbington *et al* 2015), whereas our qPCR data showed very similar Ct values for *U6* across all CD4$^+$ subtypes. Normalised relative expression levels of each gene were then tested to identify any significant expression differences between the Wt and *Malat1* KO cultures. Expression levels were also compared between CD4$^+$ cell types to determine if any expression changes observed between cell types were consistent with existing transcriptomic data on mouse CD4$^+$ cells. The raw Ct values were used to screen for very low expression levels of any genes in the cultures, indicated by similar Ct values to the negative control, as any significant differences observed between such low levels should be treated with caution. Expression levels of *IFNγ*, *IL-4* and *IL-10* were tested as controls for CD4$^+$ polarisation, with *IFNγ* used as a positive control of Th1 activation and *IL-4* and *IL-10* as positive controls of Th2 activation (N.B. for *IFNγ*, day 4 *IL-10* data and the day 6 *IL-4* data, the qPCRs were performed by Laura Chaffey).

Gene	*IFNγ*		*IL-4*		*IL-10*		*U6*	
	Wt	**KO**	**Wt**	**KO**	**Wt**	**KO**	**Wt**	**KO**
Naïve	35.59	33.83	34.91	34.76	35.60	38.11	21.92	21.87
d4 Th0	31.48	30.71	30.70	32.29	36.05	36.05	21.72	21.59
d4 Th1	24.54	24.27	33.90	33.37	33.96	33.23	21.80	22.31
d4 Th2	33.14	32.42	27.28	26.66	28.59	29.64	21.86	21.66
d6 Th0	32.89	31.43	30.50	31.52	35.46	35.75	20.87	20.84
d6 Th1	23.18	26.42	31.50	34.48	33.18	36.23	20.84	21.35
d6 Th2	35.30	33.50	24.07	22.69	26.49	26.54	21.53	21.50
Fibroblasts	N/A *	N/A *	N/A *	N/A *	N/A *	N/A *	20.64	20.70
Negative control	Not tested **		34.60		38.66		35.76	

Table 5: Average Ct values of CD4$^+$ and relative expression control genes from qPCR expression analysis across CD4$^+$ and fibroblast cell types. Relative expression values for each culture calculated and normalised as detailed in the Methods section. N.B: * = cytokine controls were used to assess CD4$^+$ polarisation in response to stimulation, and thus were not tested in fibroblasts. ** = No negative control data could be found for the *IFNγ* qPCRs – as such, the high Ct values for the naïve cells and some stimulated CD4$^+$ cultures were taken to indicate minimal expression. However, as significant differences were only observed between these cultures and much lower Ct values, this was not deemed to affect expression analysis.

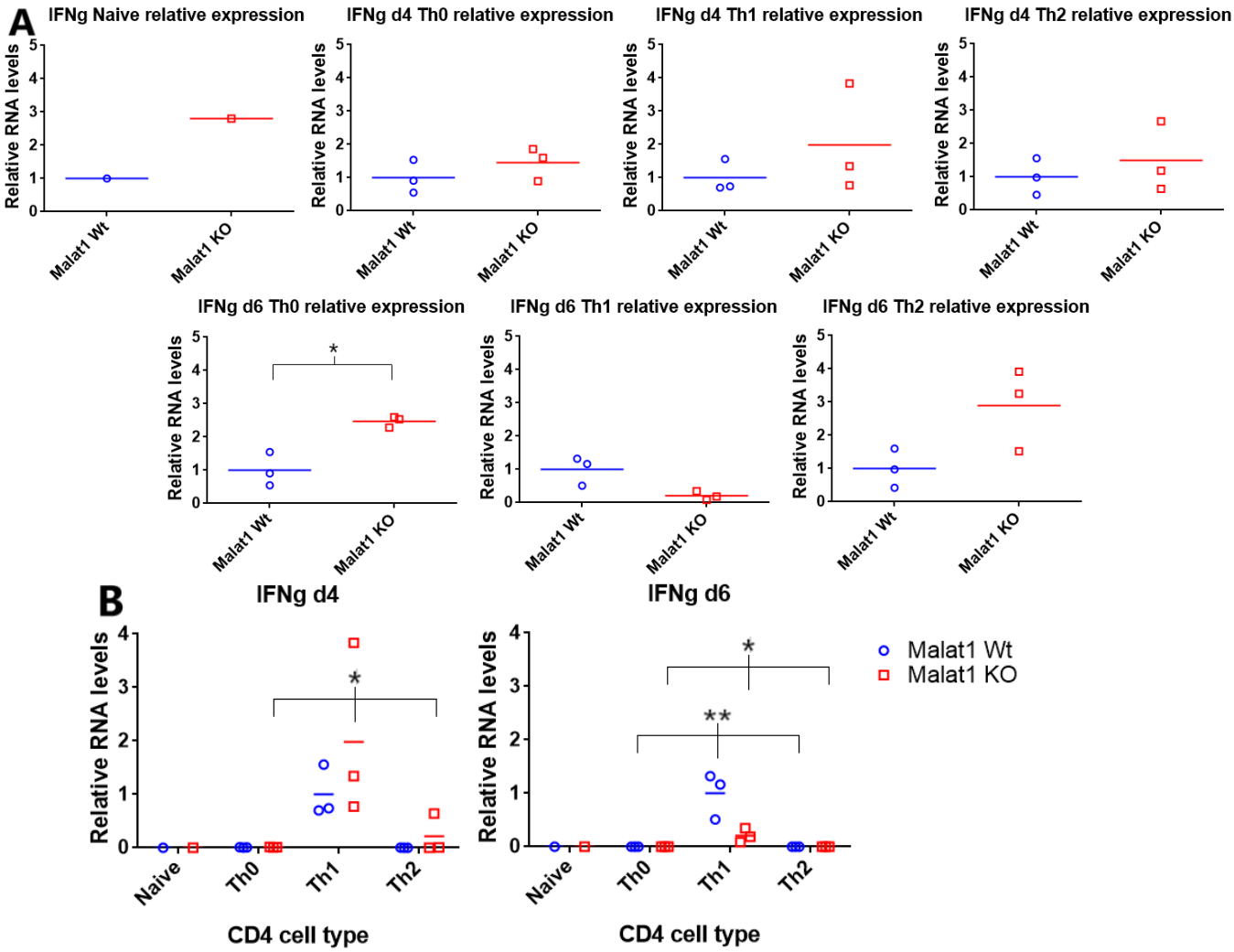

Figure 18: *IFNγ* qPCR expression data from CD4$^+$ cell cultures, showing relative RNA levels based upon differences in Ct values (ΔCt) between *IFNγ* and *U6*. (A) Relative *IFNγ* levels in Wt and KO cultures, normalised to the average of the Wt samples for each cell type. Significance based on 2 tailed t-test. (B) Relative *IFNγ* levels in different CD4$^+$ cell types, normalised to the Wt Th1 average for each culture time. Significance based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

Expression of cytokines associated with Th1 and Th2 differentiation differed as expected between the CD4$^+$ cell types – while significance could not be tested between the stimulated cultures and the single naïve cell sample, *IFNγ* expression was significantly increased specifically in Th1 cells compared to Th0 and Th2 for both the Wt cultures and the day 6 *Malat1* KO culture (Figure 18B). Similarly, Th2 cells showed significantly greater *IL-4* levels than Th0 or Th1 cells across all cultures (Figure 19B) and significantly greater *IL-10* levels in all *Malat1* KO cultures and the day 4 Wt cultures (Figure 20B). Both *IFNγ* and *IL-4* showed some significant expression changes between the Wt and *Malat1* KO cells, with *IFNγ* expression significantly increased in the day 6 Th0 knockouts and *IL-4* expression significantly decreased in the day 4 Th0 knockouts (Figures 18A and 19A). However, given that the Ct values of each cytokine in these cultures indicate very low expression in both genotypes (Table 5), the functional importance of such results is debatable. Overall, this indicates that stimulation of the different CD4$^+$

cell types was successful, with significant upregulation of CD4⁺ cell type specific cytokines in the expected cell types.

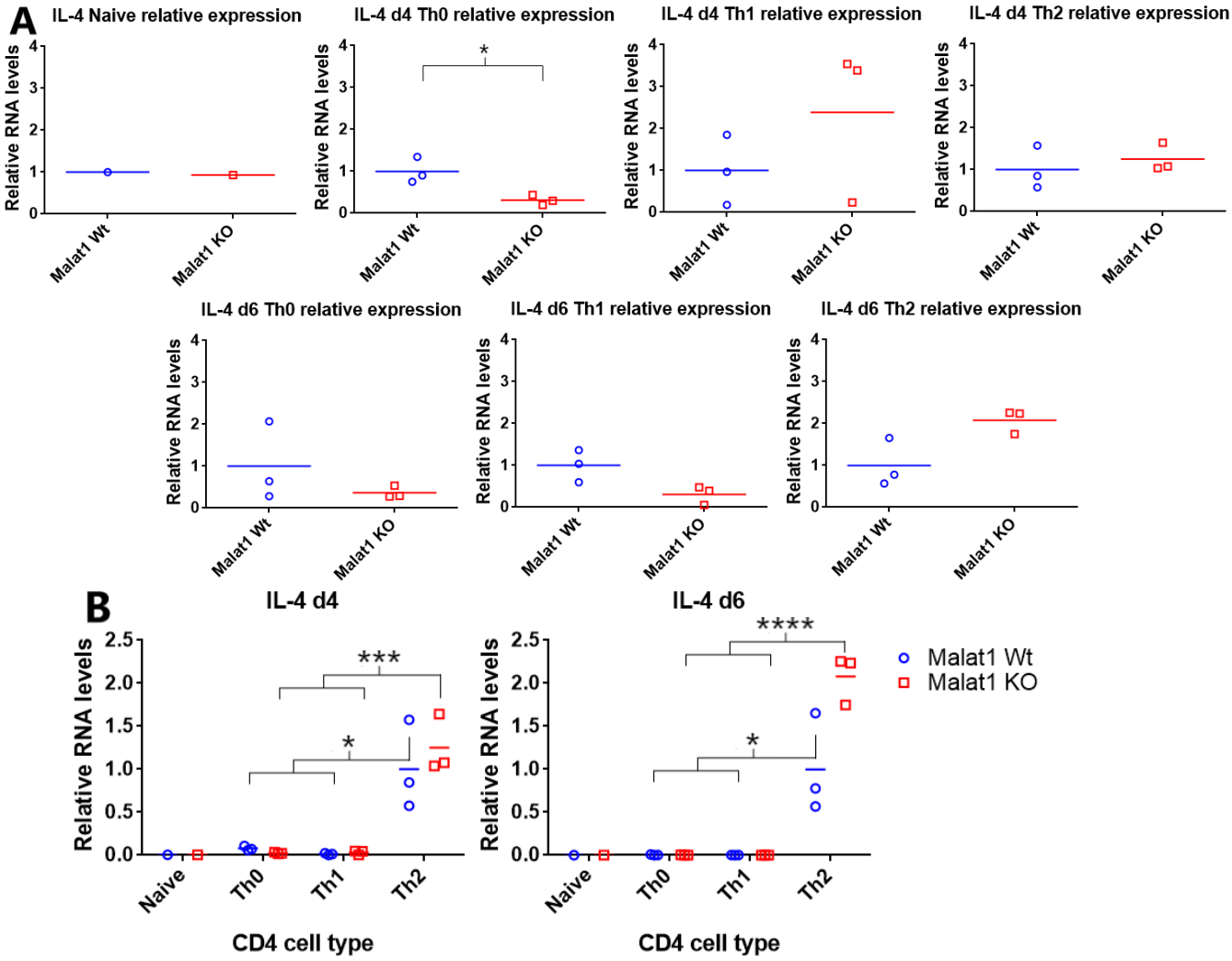

Figure 19: *IL-4* qPCR expression data from CD4⁺ cell cultures, showing relative RNA levels based upon differences in Ct values (ΔCt) between *IL-4* and *U6*. (A) Relative *IL-4* levels in Wt and KO cultures, normalised to the average of the Wt samples for each cell type. Significance based on 2 tailed t-test. (B) Relative *IL-4* levels in different CD4⁺ cell types, normalised to the Wt Th2 average for each culture time. Significance based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

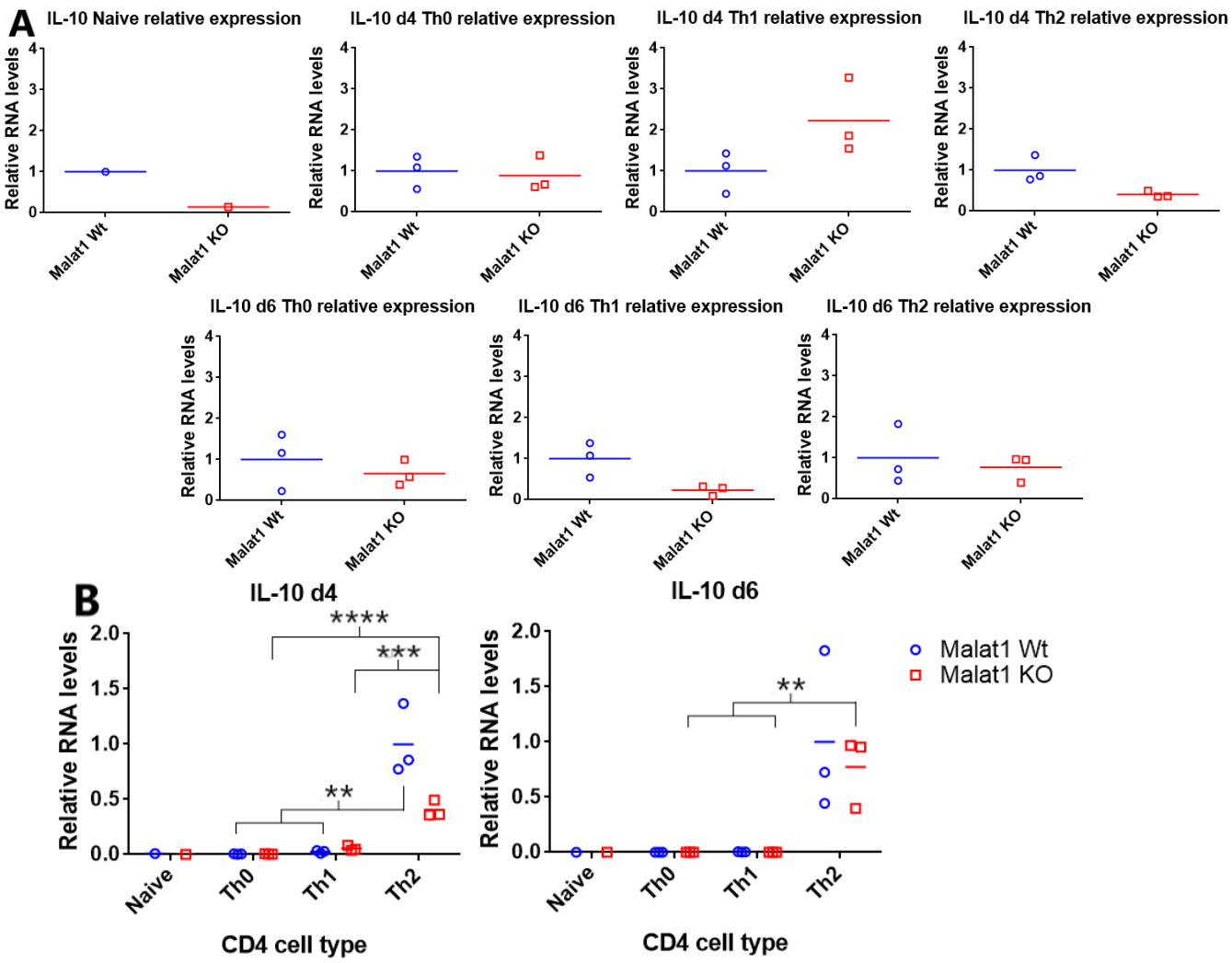

Figure 20: *IL-10* qPCR expression data from CD4⁺ cell cultures, showing relative RNA levels based upon differences in Ct values (ΔCt) between *IL-10* and *U6*. (A) Relative *IL-10* levels in Wt and KO cultures, normalised to the average of the Wt samples for each cell type. Significance based on 2 tailed t-test. (B) Relative *IL-10* levels in different CD4⁺ cell types, normalised to the Wt Th2 average for each culture time. Significance based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

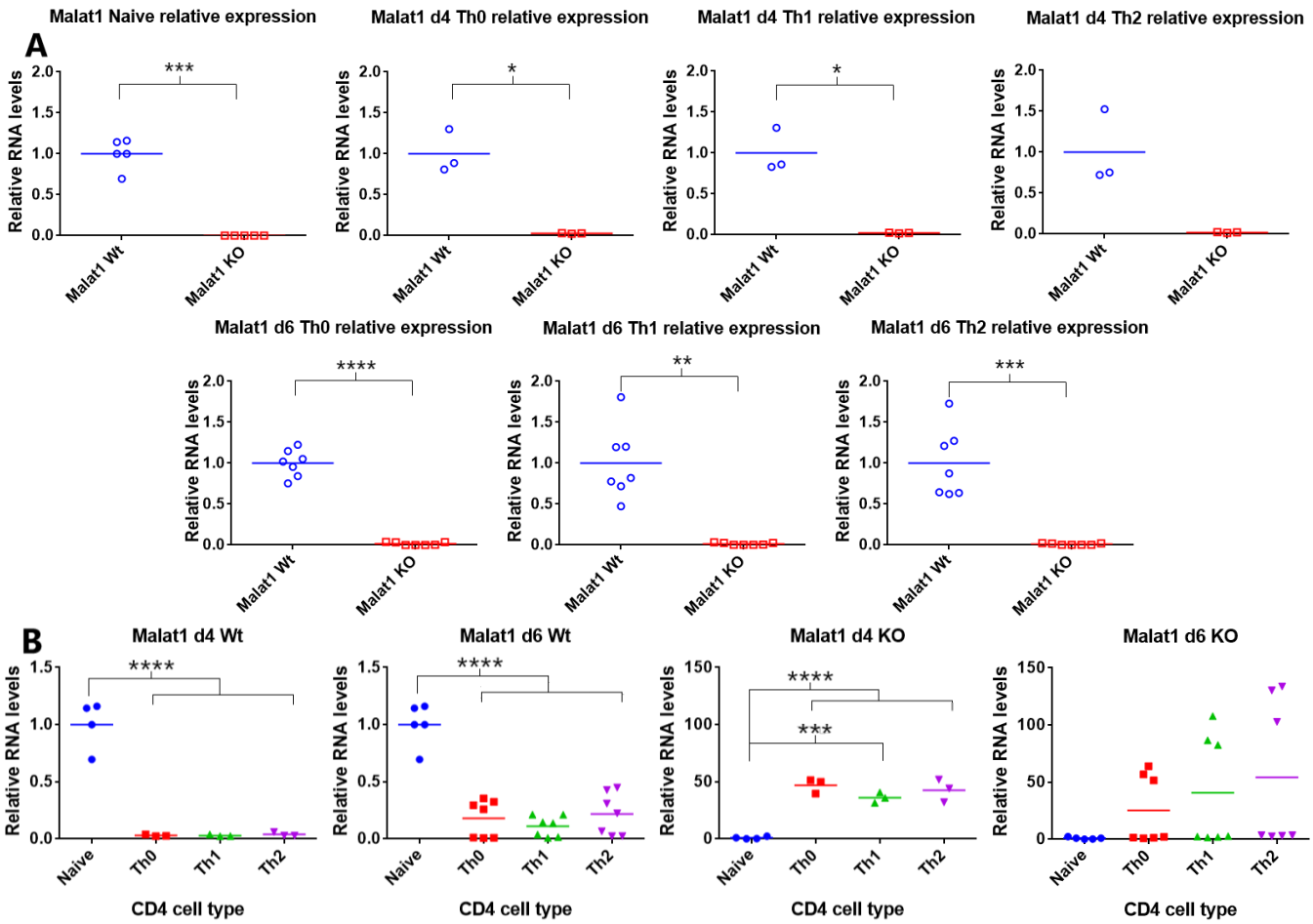

Figure 21: *Malat1* qPCR expression data from CD4⁺ cell cultures, showing relative RNA levels based upon differences in Ct values (ΔCt) between *Malat1* and *U6*. (A) Relative *Malat1* levels in Wt and KO cultures, normalised to the average of the Wt samples for each cell type. Significance based on 2 tailed t-test. (B) Relative *Malat1* levels in different CD4⁺ cell types, normalised to the naïve average for each genotype and culture time. Significance based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

In our CD4⁺ cultures, *Malat1* levels were significantly reduced in the *Malat1* KO cells compared to Wt for both the naïve CD4⁺ cells and most of the stimulated cultures after both 4 and 6 days of culture (Figure 21A). This was in agreement with unpublished data from the Lagos group (Hewitson *et al*, unpublished). Furthermore, the stimulated Wt cultures all showed significantly reduced expression compared to the Wt naïve results (Figure 21B), which is consistent with *Malat1* CD4⁺ expression data both from previous experiments in this lab and from prior transcriptomic studies (Hewitson, unpublished, Stubbington *et al* 2015). However, while the naïve *Malat1* KO cells showed little to no *Malat1* expression as expected (Table 6), all the stimulated KO cultures showed some level of residual *Malat1* expression (Figure 21B), to the extent that the day 4 Th2 Wt and KO samples did not significantly differ (Figure 21A, p = 0.065) due to a combination of the increased *Malat1* levels in the KO cells, reduced *Malat1* expression in Wt cells and the low sample number for this experiment. As such, our data indicates that CD4⁺ stimulation produces residual *Malat1* expression in the knockout

CD4⁺ cultures, though this may reflect increased formation of *Malat1* primer dimers at these relatively low levels of expression.

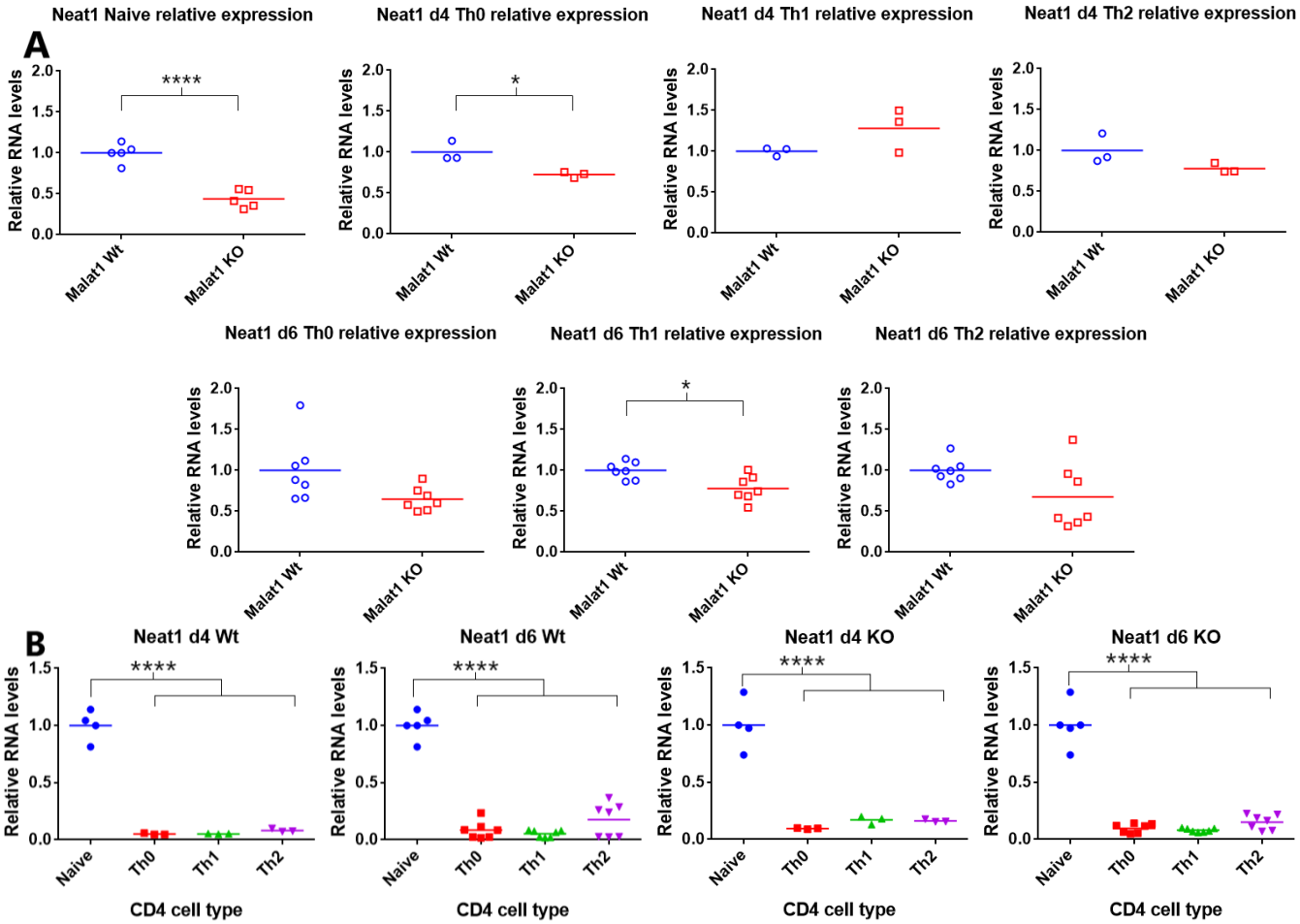

Figure 22: *Neat1* qPCR expression data from CD4⁺ cell cultures, showing relative RNA levels based upon differences in Ct values (ΔCt) between *Neat1* and *U6*. (A) Relative *Neat1* levels in Wt and KO cultures, normalised to the average of the Wt samples for each cell type. Significance based on 2 tailed t-test. (B) Relative *Neat1* levels in different CD4⁺ cell types, normalised to the naïve average for each genotype and culture time. Significance based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

Similar to *Malat1*, *Neat1* expression was significantly reduced in the stimulated CD4⁺ cultures compared to the naïve cells, though this reduction in stimulated CD4⁺ expression was consistent for both the Wt and *Malat1* KO cultures (Figure 22B). Furthermore, the KO cells showed significantly reduced *Neat1* expression compared to Wt in the naïve, day 4 Th0 and day 6 Th1 cultures (Figure 22A), indicating some potential regulation by *Malat1* under cell type specific conditions.

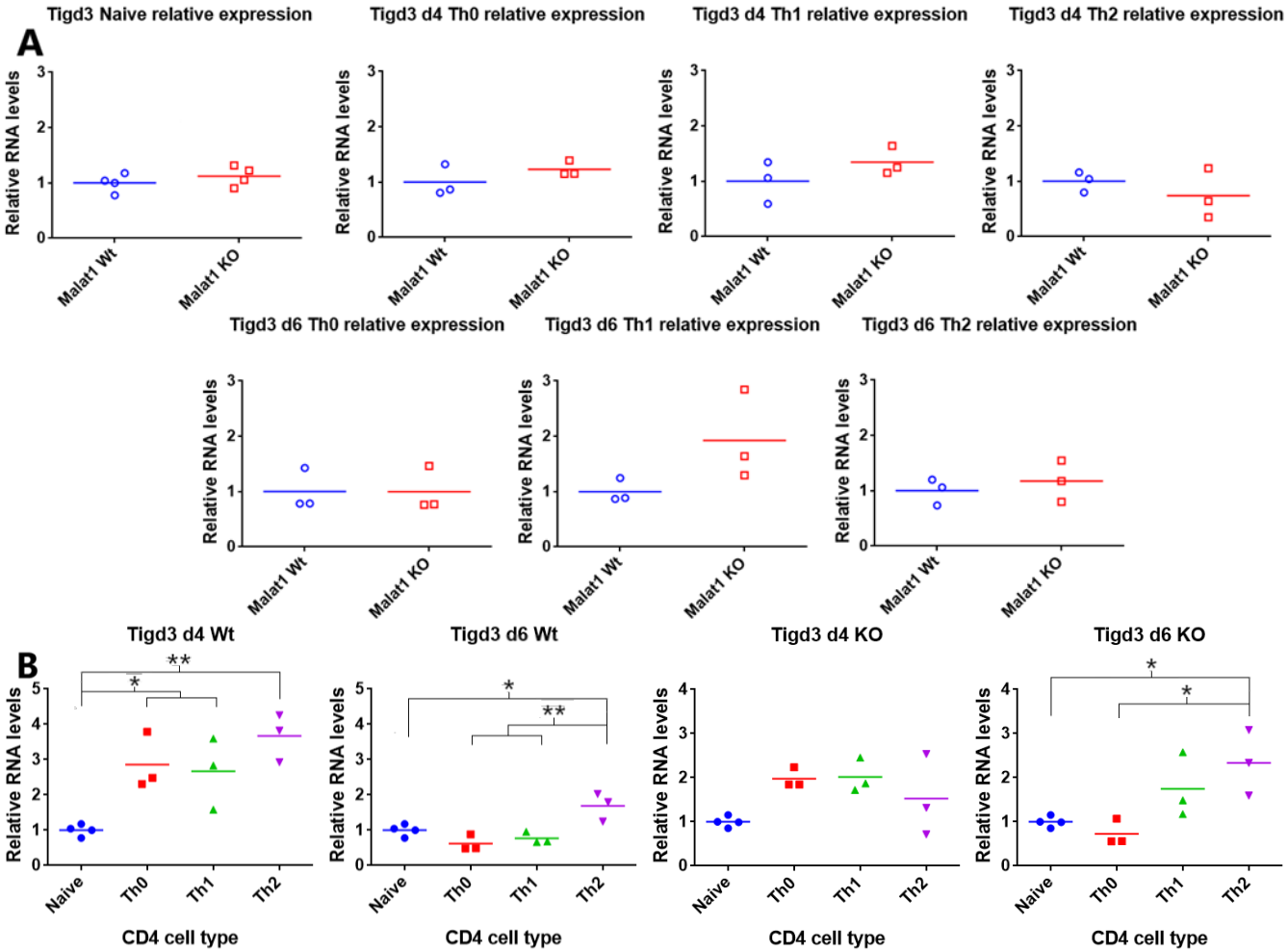

Figure 23: *Tigd3* qPCR expression data from CD4⁺ cell cultures, showing relative RNA levels based upon differences in Ct values (ΔCt) between *Tigd3* and *U6*. (A) Relative *Tigd3* levels in Wt and KO cultures, normalised to the average of the Wt samples for each cell type. Significance based on 2 tailed t-test. (B) Relative *Tigd3* levels in different CD4⁺ cell types, normalised to the naïve average for each genotype and culture time. Significance based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

Tigd3 levels showed some differences between naïve and stimulated Wt CD4⁺ T cells, with significantly increased expression in all the day 4 stimulated Wt cultures and in the day 6 Th2 Wt culture compared to both naïve cells and the other CD4⁺ cell types (Figure 23B). However, the *Malat1* KO cells showed very different expression patterns across the CD4⁺ cell types, with no significant differences between any of the day 4 cultures and significantly increased expression in the day 6 Th2 cultures compared to the naïve and Th0 CD4⁺ cells (Figure 23B). As previously detailed, the very low levels of *Tigd3* in these CD4⁺ cultures, as indicated by their Ct values (Table 6), makes it challenging to judge the relevance of any such differences in expression. Furthermore, none of these CD4⁺ cultures showed a significant difference in *Tigd3* levels between the Wt and *Malat1* KO cells (Figure 23A), indicating an apparent lack of regulation of the gene by *Malat1* in CD4⁺ T cells.

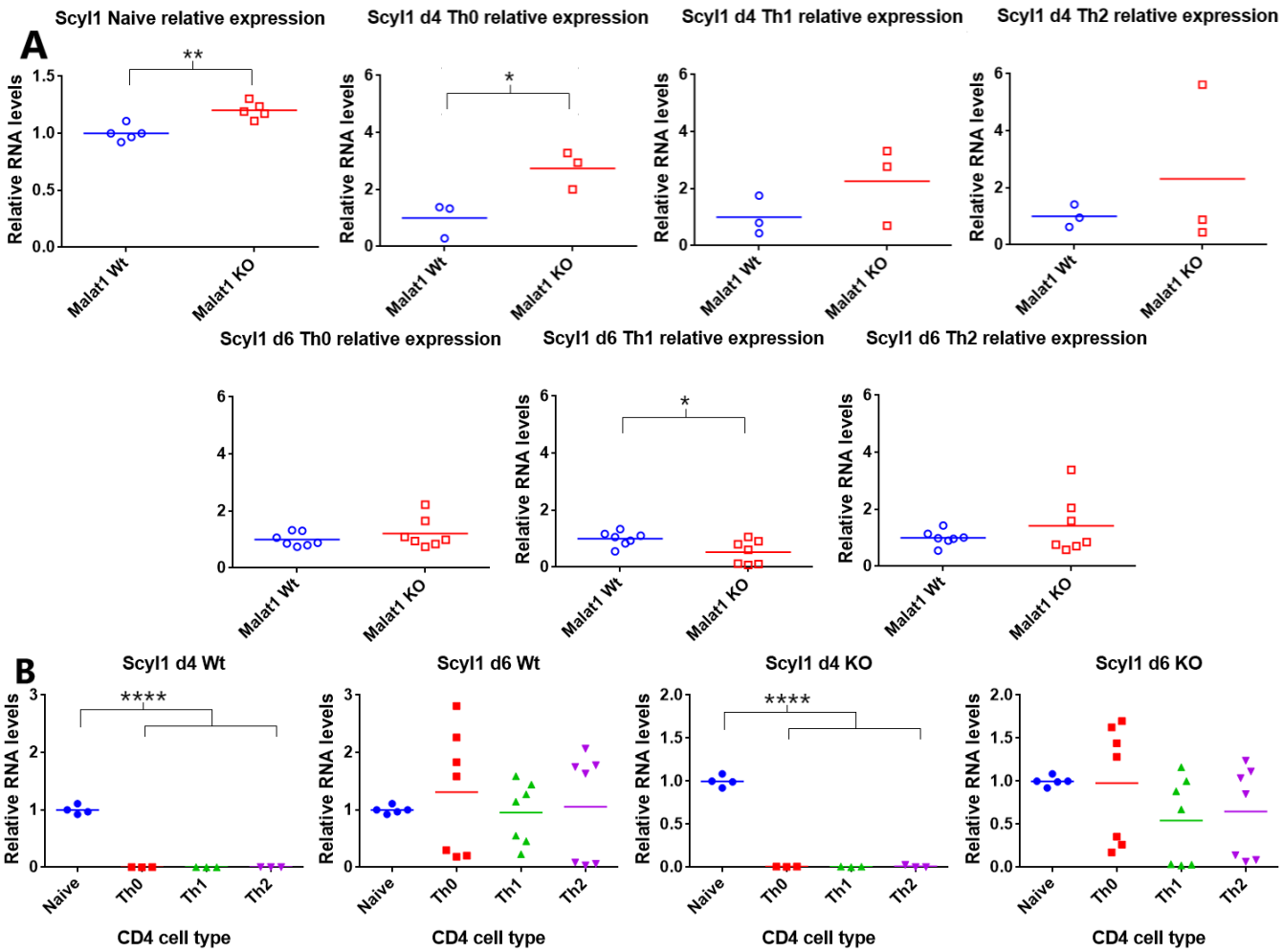

Figure 24: *Scyl1* qPCR expression data from CD4⁺ cell cultures, showing relative RNA levels based upon differences in Ct values (ΔCt) between *Scyl1* and *U6*. (A) Relative *Scyl1* levels in Wt and KO cultures, normalised to the average of the Wt samples for each cell type. Significance based on 2 tailed t-test. (B) Relative *Scyl1* levels in different CD4⁺ cell types, normalised to the naïve average for each genotype and culture time. Significance based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

Although *Scyl1* levels were significantly reduced in the day 4 stimulated CD4⁺ cultures, with the Ct values indicating negligible expression of the gene at this timepoint (Table 6), this significant difference was not observed in the day 6 cultures (Figure 24B). While the initial run of CD4⁺ cultures showed substantially reduced *Scyl1* expression in the stimulated cultures at day 6, the second sample set showed far less change in *Scyl1* levels between the stimulated and naïve cells (Figure 24B), indicating some variability in the responses of the two CD4⁺ culture sets to stimulation. Interestingly, *Scyl1* levels were both significantly increased in the naïve and d4 Th0 *Malat1* KOs and significantly decreased in the d6 Th0 *Malat1* KOs (Figure 24A). Although the negligible day 4 Th0 expression of *Scyl1* makes the relevance of the associated significant change difficult to assess (Table 6), these differing significant expression changes in the naïve and day 6 Th1 cultures both indicate that *Malat1* has some role in regulating *Scyl1* in CD4⁺ T cells and further highlighting the apparent cell type specificity of such regulation.

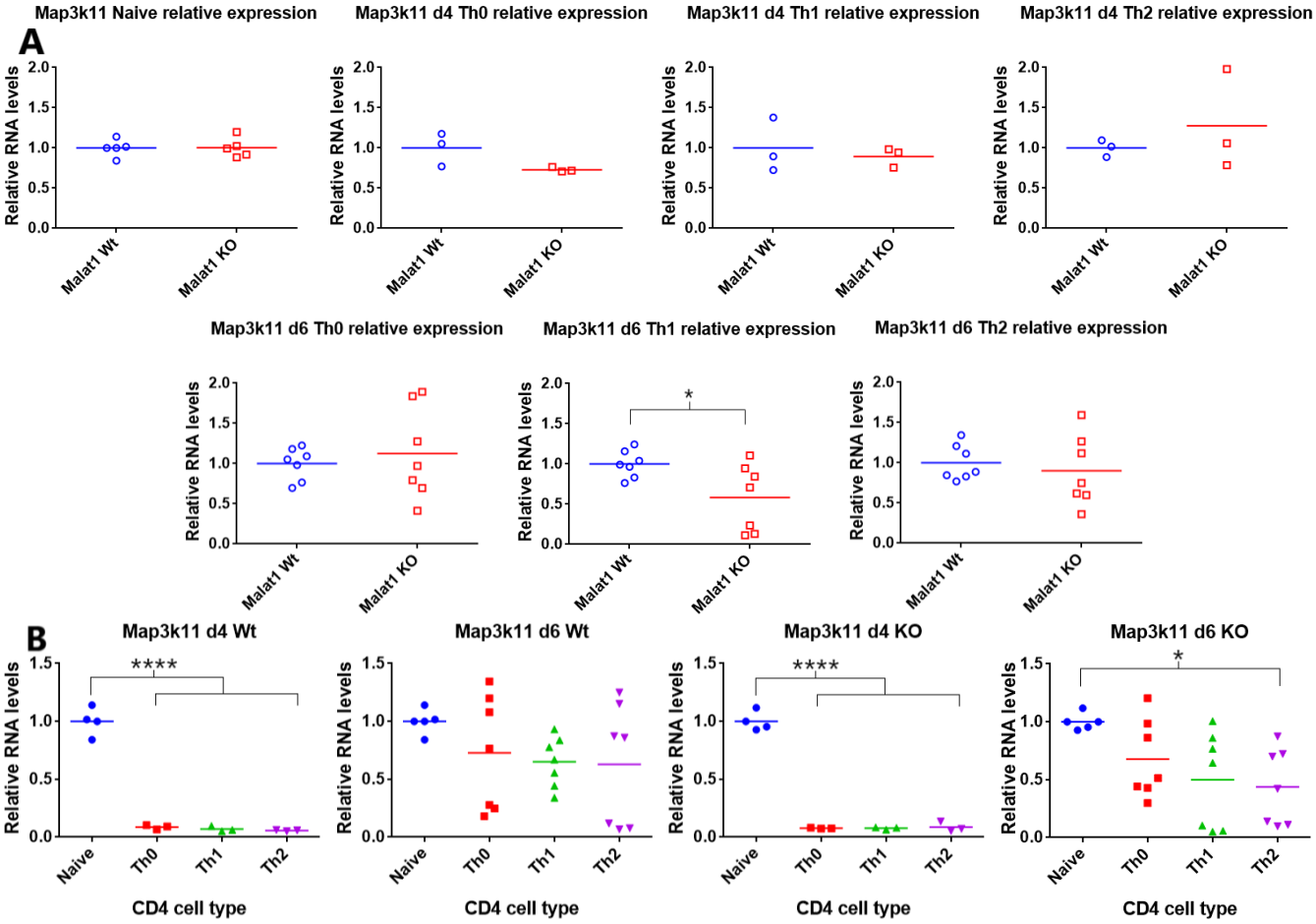

Figure 25: *Map3k11* qPCR expression data from CD4⁺ cell cultures, showing relative RNA levels based upon differences in Ct values (ΔCt) between *Map3k11* and *U6*. (A) Relative *Map3k11* levels in Wt and KO cultures, normalised to the average of the Wt samples for each cell type. Significance based on 2 tailed t-test. (B) Relative *Map3k11* levels in different CD4⁺ cell types, normalised to the naïve average for each genotype and culture time. Significance based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

Map3k11 showed similar expression changes to *Scyl1* across the Wt CD4⁺ cell types, with significantly reduced levels in the stimulated day 4 cultures compared to naïve cells and no significant expression changes in the day 6 cultures (Figure 25B). The *Malat1* KO cultures showed similar results, although here the day 6 Th2 KO cells showed significantly decreased *Map3k11* levels compared to the naïve cells (Figure 25B). Furthermore, the two day 6 sample sets showed a similar difference in *Map3k11* expression to *Scyl1*, with the initial CD4⁺ cultures showing substantially reduced expression and the second set of cultures showing little to no change in expression (Figure 25A). Only one CD4⁺ cell type showed significantly altered *Map3k11* levels in the *Malat1* KO cultures, with significantly decreased expression in the day 6 Th1 cells. Altogether, this CD4⁺ expression data suggests that *Malat1* does appear to regulate some of its surrounding genes, but this regulation is highly specific to certain cell types and conditions.

Effects of *Malat1* knockout on surrounding gene expression in fibroblasts

To determine if these effects were specific to $CD4^+$ T cell lineages, mouse tail fibroblast (MTF) cultures were established from tail samples of Wt and *Malat1* KO mice and harvested after 3 passages, with subsequent RNA extraction, reverse transcription and qPCR analysis carried out to compare expression of these genes in Wt and *Malat1* KO cultures. *Malat1* expression was, as expected, significantly lower in *Malat1* KO cultures than in Wt (Figure 26A), though of the genes surrounding *Malat1* only *Scyl1* showed a significant difference in RNA levels, with significantly greater expression in *Malat1* KO cultures than in Wt (Figure 26D). Therefore, regulation of these genes by *Malat1* differs between immune and non-immune cells, further indicating that such regulation is apparently cell type specific.

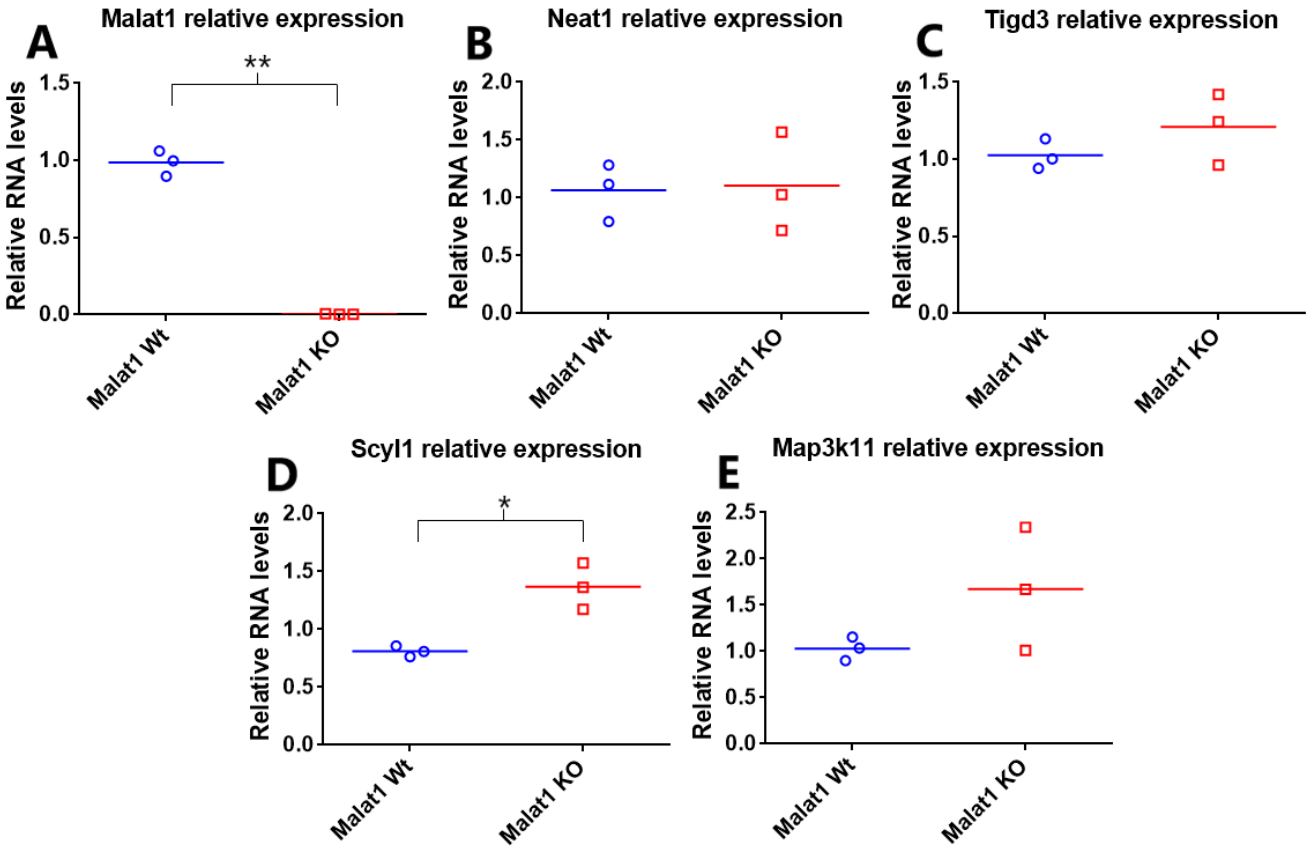

Figure 26: qPCR expression data from MTF cell cultures, showing relative RNA levels of (A) *Malat1*, (B) *Neat1*, (C) *Tigd3*, (D) *Scyl1* and (E) *Map3k11* based upon differences in Ct values (ΔCt) between the gene and *U6* and normalised to the average of the Wt samples. Significance based on 2 tailed t-test; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

Relative expression values for each gene were also compared between the $CD4^+$ and MTF cultures to compare expression levels of these genes in both cell types and determine if any differences observed corresponded to the variations in *Malat1* regulation observed. Relative expression levels of both *Malat1* and *Neat1* were significantly greater in fibroblasts compared to all $CD4^+$ cultures, with similarly significant increases across both Wt and *Malat1* KO cells (Figures 27 and 28). While *Scyl1* also showed significantly increased fibroblast expression compared to both Wt and *Malat1* KO Th cell cultures, the degree of significance varied considerably between different Wt cultures (Figures 30A and 30B). Although *Tigd3* showed numerous expression changes between the fibroblast and $CD4^+$ cell types, such as significantly reduced levels in fibroblasts compared to day 4 Th0 and Th1 cells of both genotypes (Figures 29A and 29B), the importance of these results should be considered carefully given the very low relative expression of *Tigd3* observed across all cell types (Table 4) and the off-target amplification observed in the primer test data (Figure 3B). Finally, relative RNA levels of *Map3k11* in fibroblasts were significantly lower than naïve $CD4^+$ cells in the Wt cultures (Figure 31A) and significantly greater than the day 4 stimulated cultures of *Malat1* KO cells (Figure 31B). Overall, while *Malat1* and these surrounding genes all show some significant differences in expression levels between the $CD4^+$ and fibroblast cultures, these changes do not correspond to the differing evidence of *Malat1* regulation between the two cell types, indicating that these distinct patterns of surrounding gene regulation by *Malat1* arise from cell type specificity rather than differing gene expression.

Gene	***Malat1***		**Neat1**		**Tigd3**		**Scyl1**		**Map3k11**		***U6***	
	Wt	**KO**	**Wt**	**KO**	**Wt**	**KO**	**Wt**	**KO**	**Wt**	**KO**	**Wt**	**KO**
Naïve	18.34	31.39	23.37	24.54	31.83	31.52	26.16	25.86	26.41	26.36	21.92	21.87
d4 Th0	24.79	29.79	28.99	29.32	31.43	30.97	34.73	32.87	29.30	29.60	21.72	21.59
d4 Th1	24.91	30.89	29.04	29.22	31.66	31.68	34.60	33.98	29.72	30.35	21.80	22.31
d4 Th2	24.61	30.01	28.47	28.62	31.20	31.60	33.71	33.06	30.05	29.60	21.86	21.66
d6 Th0	21.08	28.43	26.58	27.12	31.07	31.06	26.15	25.91	26.73	26.67	20.87	20.84
d6 Th1	21.16	28.29	26.93	27.82	31.22	31.19	26.13	28.16	26.55	28.31	20.84	21.35
d6 Th2	20.89	27.95	26.43	27.15	31.18	30.79	27.79	27.48	28.05	28.29	21.53	21.50
Fibroblasts	16.12	24.45	20.14	20.19	30.76	30.60	23.60	22.91	26.42	25.85	20.64	20.70
Negative control	34.88		33.29		32.24		34.50		36.59		35.76	

Table 6: Average Ct values of *Malat1* and surrounding genes from qPCR expression analysis across $CD4^+$ and fibroblast cell types, along with the U6 control. Relative expression values for each culture calculated and normalised as detailed in the Methods section.

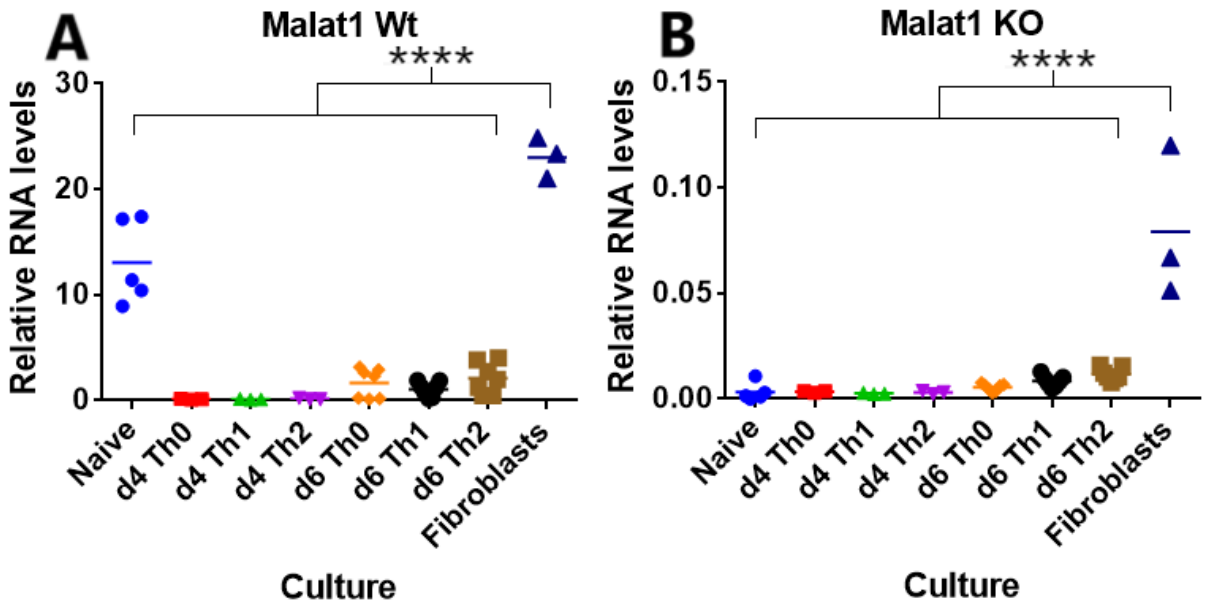

Figure 27: Comparison of relative *Malat1* RNA levels in CD4⁺ cell types and fibroblasts from (A) Wt and (B) *Malat1* KO cultures, based on ΔCt between the gene and *U6*. Significance based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

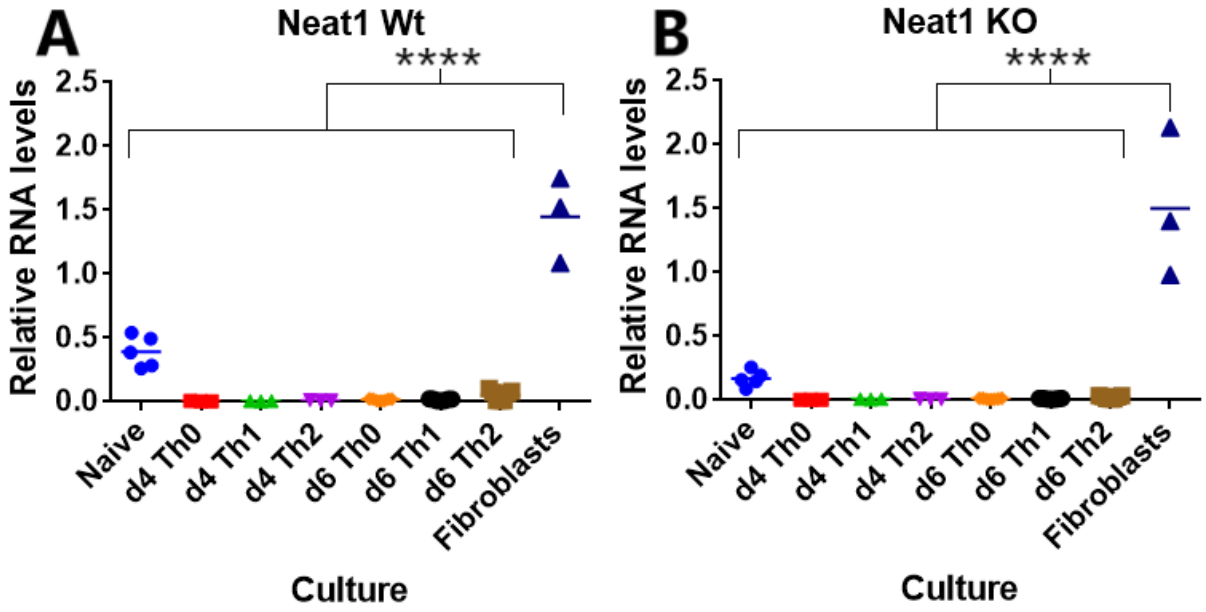

Figure 28: Comparison of relative *Neat1* RNA levels in CD4⁺ cell types and fibroblasts from (A) Wt and (B) *Malat1* KO cultures, based on ΔCt between the gene and *U6*. Significance based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

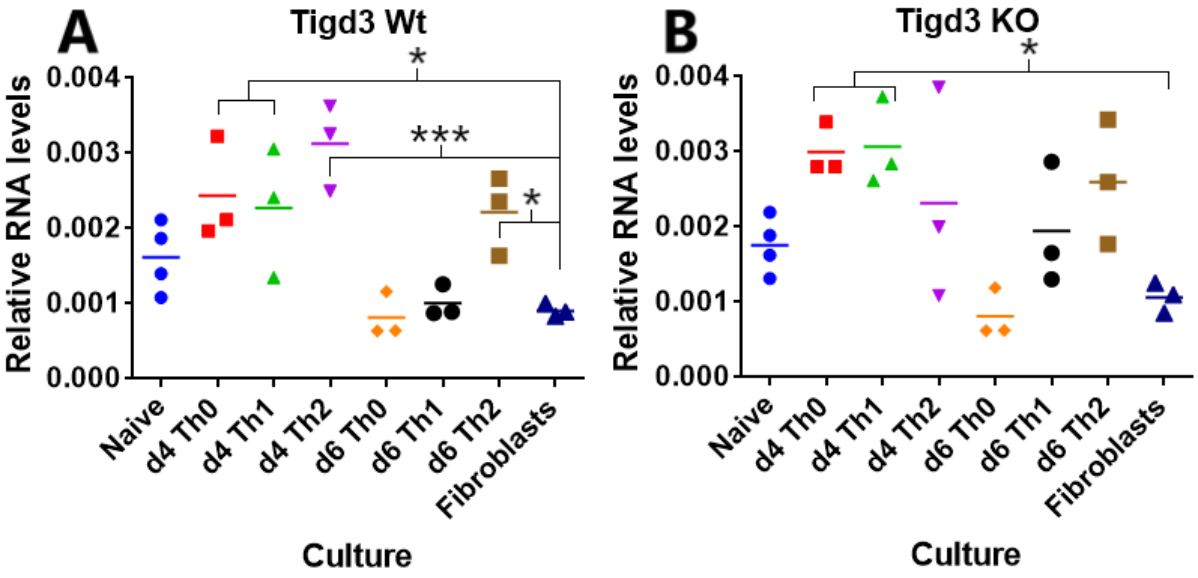

Figure 29: Comparison of relative *Tigd3* RNA levels in CD4$^+$ cell types and fibroblasts from (A) Wt and (B) *Malat1* KO cultures, based on ΔCt between the gene and *U6*. Significance based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

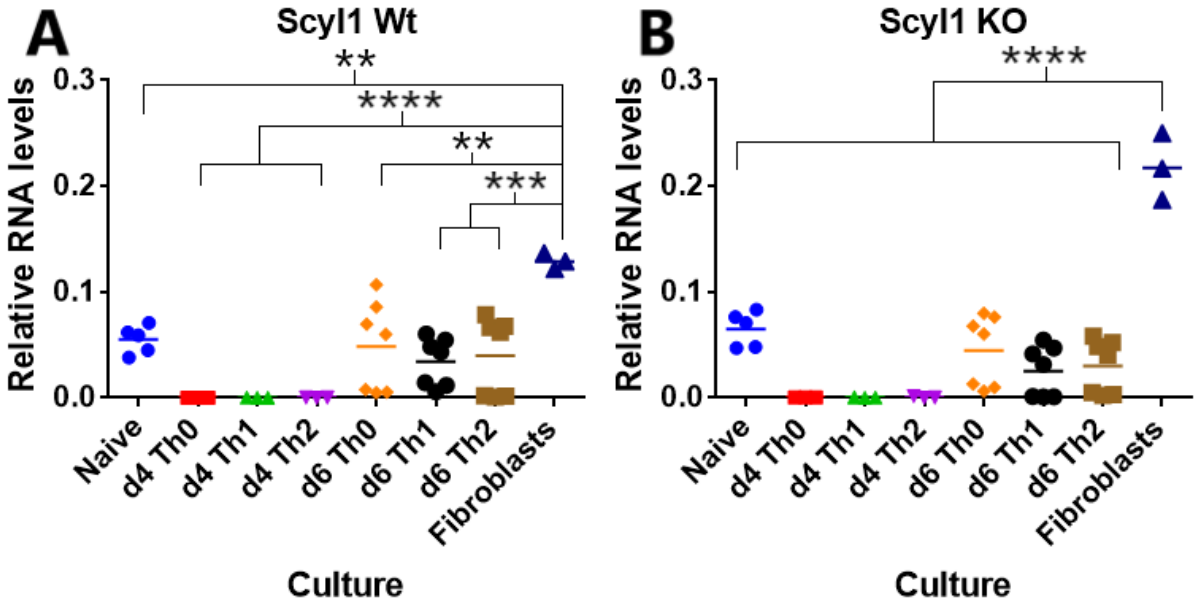

Figure 30: Comparison of relative *Scyl1* RNA levels in CD4$^+$ cell types and fibroblasts from (A) Wt and (B) *Malat1* KO cultures, based on ΔCt between the gene and *U6*. Significance based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

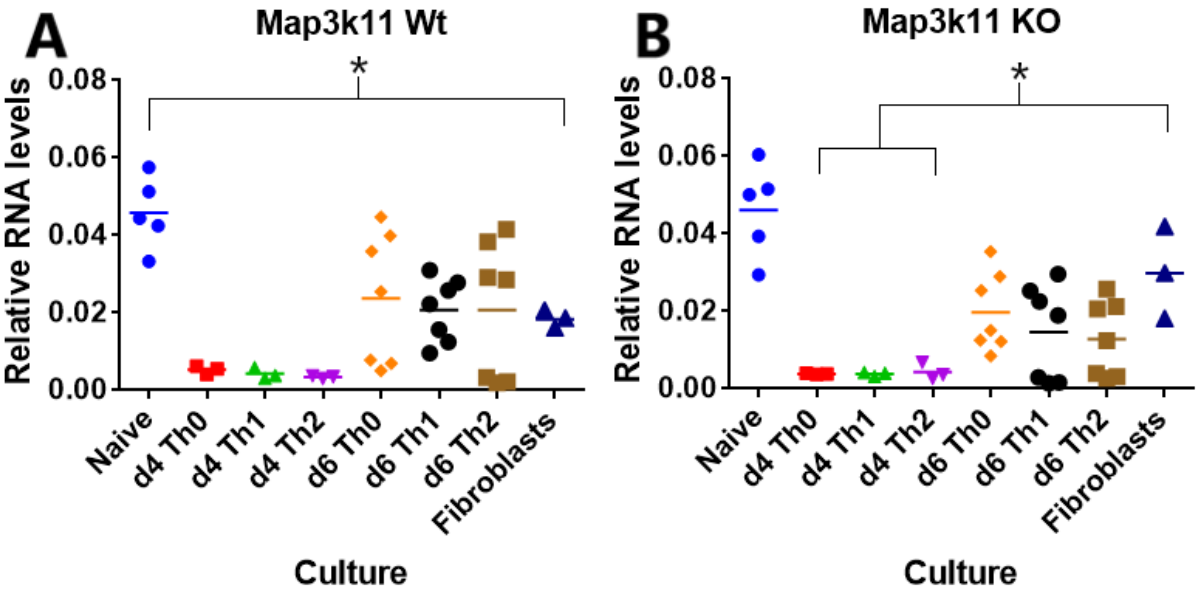

Figure 31: Comparison of relative *Map3k11* RNA levels in $CD4^+$ cell types and fibroblasts from (A) Wt and (B) *Malat1* KO cultures, based on ΔCt between the gene and *U6*. Significance based on one-way ANOVA; * = $p<0.05$, ** = $p<0.01$, *** = $p<0.001$, **** = $p<0.0001$

Effects of *Malat1* knockout on interacting RNA binding proteins

Given this apparent cell type specific regulation of the genes surrounding *Malat1*, the $CD4^+$ cultures were then used to assess if other genes interacting with *Malat1* showed evidence of similar regulation *in trans*. To this end, 6 genes were identified with products previously shown to interact with *Malat1* at either the protein or mRNA stage (Chen *et al* 2017, Engreitz *et al* 2014) and with various functions in RNA binding, processing and degradation. Specifically, the majority of these candidate genes have known functions in regulation and processing of the pre-mRNA spliceosome, a cellular process with which *Malat1* has been found to be very strongly associated (Chen *et al* 2017, Engreitz *et al* 2014) and thus is likely to have some role in regulating. Based upon this association, the genes selected for analysis were the histone subunit protein Hist1h4a (Marzluff *et al* 2002), the pre-miRNA processing DEAD-box helicase Ddx23 (Chu *et al* 2016), the pre-mRNA splicing complex component U2af1 (Okeyo-Owuor *et al* 2015), the double stranded RNA binding and mRNA decay protein Stau1 (LeGendre *et al* 2013), the ribosomal RNA processing and biogenesis protein Las1l (Castle *et al* 2010) and the alternative splicing specificity complex subunit Cpsf1 (Evsyukova *et al* 2012). While these genes were not the most strongly associated with *Malat1* in this previous interactome data (Engreitz *et al* 2014), they were all found to show similar patterns of expression to *Malat1* across $CD4^+$ cell types, with strong downregulation in stimulated cultures from earlier transcriptomic data (Stubbington *et al* 2015). Together with the *Malat1* interactome data, these genes show further evidence of possible regulation by *Malat1 in trans* – as such, expression levels of these genes in the $CD4^+$ cultures were therefore assessed by further qPCR expression analysis and compared between both the Wt and *Malat1* genotypes and the different $CD4^+$ cell types as with the surrounding genes. However, as only a single Naïve sample was available for the $CD4^+$ sample set tested here, significant expression differences between the naïve and stimulated cultures could not be fully assessed.

Gene	*Hist1h4a*		*Ddx23*		*U2af1*		*Stau1*		*Las1l*		*Cpsf1*		*U6*	
	Wt	**KO**	**Wt**	**KO**	**Wt**	**KO**	**Wt**	**KO**	**Wt**	**KO**	**Wt**	**KO**	**Wt**	**KO**
Naïve	34.18	36.52	34.39	38.61	32.53	31.78	31.88	31.68	33.40	34.09	34.70	35.66	21.93	21.60
d4 Th0	27.38	27.24	37.11	35.01	25.67	25.26	32.02	31.96	32.57	30.61	35.56	35.51	21.72	21.59
d4 Th1	27.03	28.09	34.66	36.02	25.73	26.35	31.60	32.02	31.54	31.42	35.13	34.97	21.80	22.31
d4 Th2	27.60	28.09	36.22	33.73	26.26	25.71	31.79	31.72	31.44	29.87	34.91	34.43	21.86	21.66
d6 Th0	28.47	29.09	27.21	26.29	25.64	26.19	31.40	31.30	35.41	34.31	38.67	38.45	20.75	20.73
d6 Th1	29.97	29.49	26.45	31.11	26.02	27.29	31.22	31.10	33.85	33.67	36.63	39.12	21.25	22.11
d6 Th2	31.47	31.30	32.79	29.28	26.65	26.08	31.22	31.10	33.74	34.01	37.35	39.72	22.33	22.15
Negative control	38.25		37.94		38.53		31.58		35.14		35.01		33.87	

Table 7: Average Ct values of *Malat1* interacting, RNA binding genes from qPCR expression analysis across $CD4^+$ cell types, along with the U6 control. Relative expression values for each culture calculated and normalised as detailed in the Methods section.

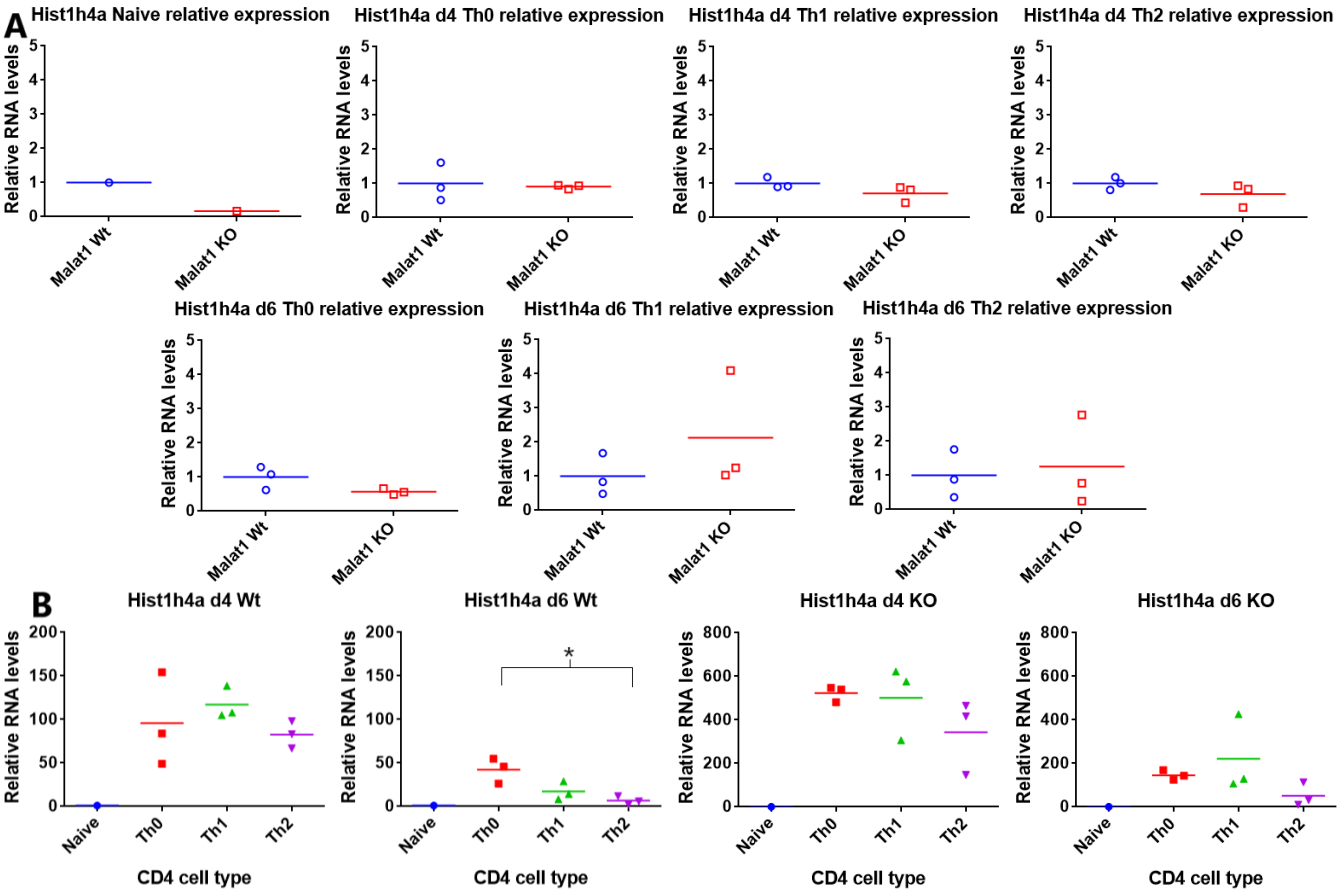

Figure 32: *Hist1h4a* qPCR expression data from CD4⁺ cell cultures, showing relative RNA levels based upon differences in Ct values (ΔCt) between *Hist1h4a* and *U6*. (A) Relative *Hist1h4a* levels in Wt and KO cultures, normalised to the average of the Wt samples for each cell type. Significance based on 2 tailed t-test. (B) Relative *Hist1h4a* levels in different CD4⁺ cell types, normalised to the naïve average for each genotype and culture time. Significance based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

Both the day 4 and day 6 stimulated Wt CD4⁺ cultures showed substantial increases in *Hist1h4a* levels compared to the minimal expression observed in Naïve cells (Table 7) – however, this increase was far less in the day 6 cultures than the day 4 cultures, with significantly lower *Hist1h4a* expression in the Th2 cells than in Th0 at this timepoint (Figure 32B). More ample increases in expression were observed in the stimulated *Malat1* KO cultures at both time points, although these showed similar patterns of less substantial increases on day 6 of culture than on day 4 (Figure 32B). While some differences in *Hist1h4a* expression levels were observed between the Wt and *Malat1* KO cultures, these differences were not significant in any of the CD4⁺ cell types (Figure 32A).

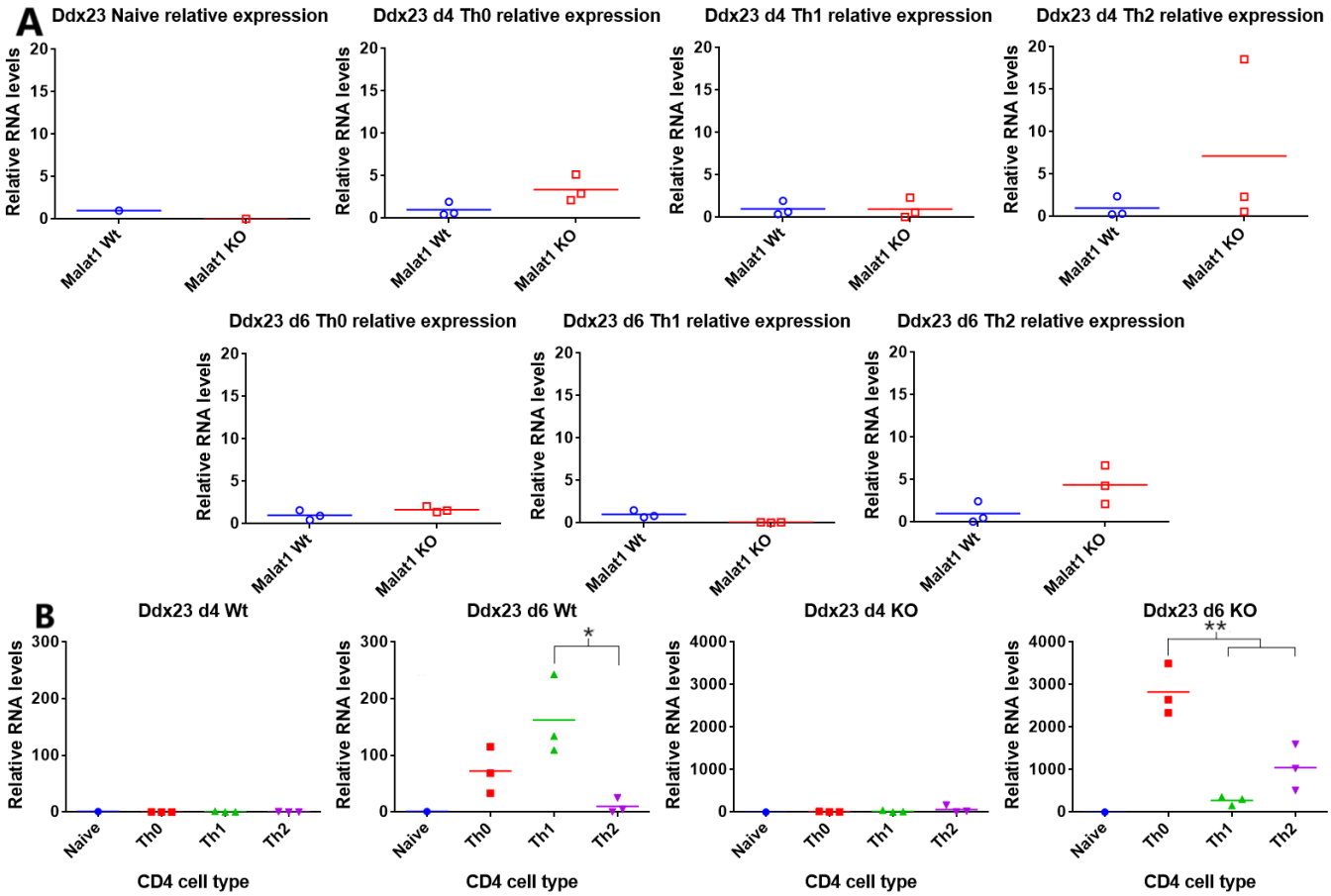

Figure 33: *Ddx23* qPCR expression data from CD4⁺ cell cultures, showing relative RNA levels based upon differences in Ct values (ΔCt) between *Ddx23* and *U6*. (A) Relative *Ddx23* levels in Wt and KO cultures, normalised to the average of the Wt samples for each cell type. Significance based on 2 tailed t-test. (B) Relative *Ddx23* levels in different CD4⁺ cell types, normalised to the naïve average for each genotype and culture time. Significance based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

While the Wt stimulated CD4⁺ cell cultures showed some slight reduction in *Ddx23* expression at day 4, with the Ct values indicating very low *Ddx23* levels in both the naïve and day 4 cultures (Table 7), this expression is substantially increased by day 6, with significantly higher expression levels in the Th1 cells compared to Th2 (Figure 33B). The *Malat1* KO CD4⁺ cultures showing increased expression compared to naïve cells at both timepoints, with significantly greater *Ddx23* expression in the d6 Th0 culture compared to both Th1 and Th2 cells (Figure 33B). Furthermore, while *Ddx23* expression levels did differ somewhat between the Wt and *Malat1* KO cells across the CD4⁺ cultures, none of these differences were significant (Figure 33A).

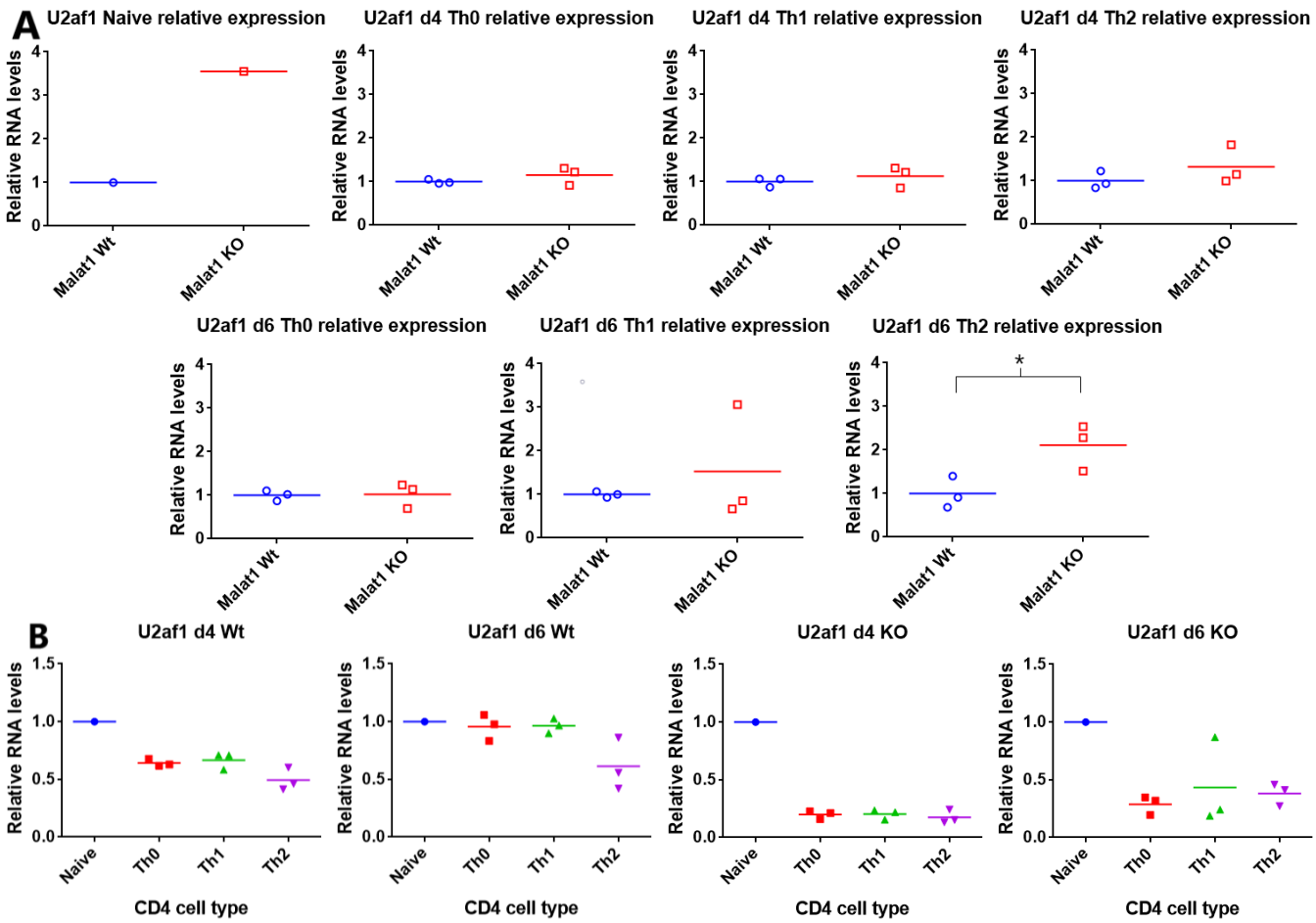

Figure 34: *U2af1* qPCR expression data from CD4⁺ cell cultures, showing relative RNA levels based upon differences in Ct values (ΔCt) between *U2af1* and *U6*. (A) Relative *U2af1* levels in Wt and KO cultures, normalised to the average of the Wt samples for each cell type. Significance based on 2 tailed t-test. (B) Relative *U2af1* levels in different CD4⁺ cell types, normalised to the naïve average for each genotype and culture time. Significance based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

The day 4 Wt CD4⁺ cultures showed a slight reduction in *U2af1* levels compared to naïve cells, with little to no reduction observed in the day 6 Wt cultures and more substantial decreases in expression in the *Malat1* KO cultures at both time points (Figure 34B). However, *U2af1* expression across the stimulated CD4⁺ cell types did not significantly differ in any of our cultures (Figure 34B). Notably, *U2af1* levels were significantly altered only in a single set of *Malat1* KO cultures, with increased expression in the day 6 Th2 culture compared to Wt cells (Figure 34A). This indicates some *in trans* regulation of the gene by *Malat1*, with similar cell type specificity to that observed in the genes surrounding *Malat1*.

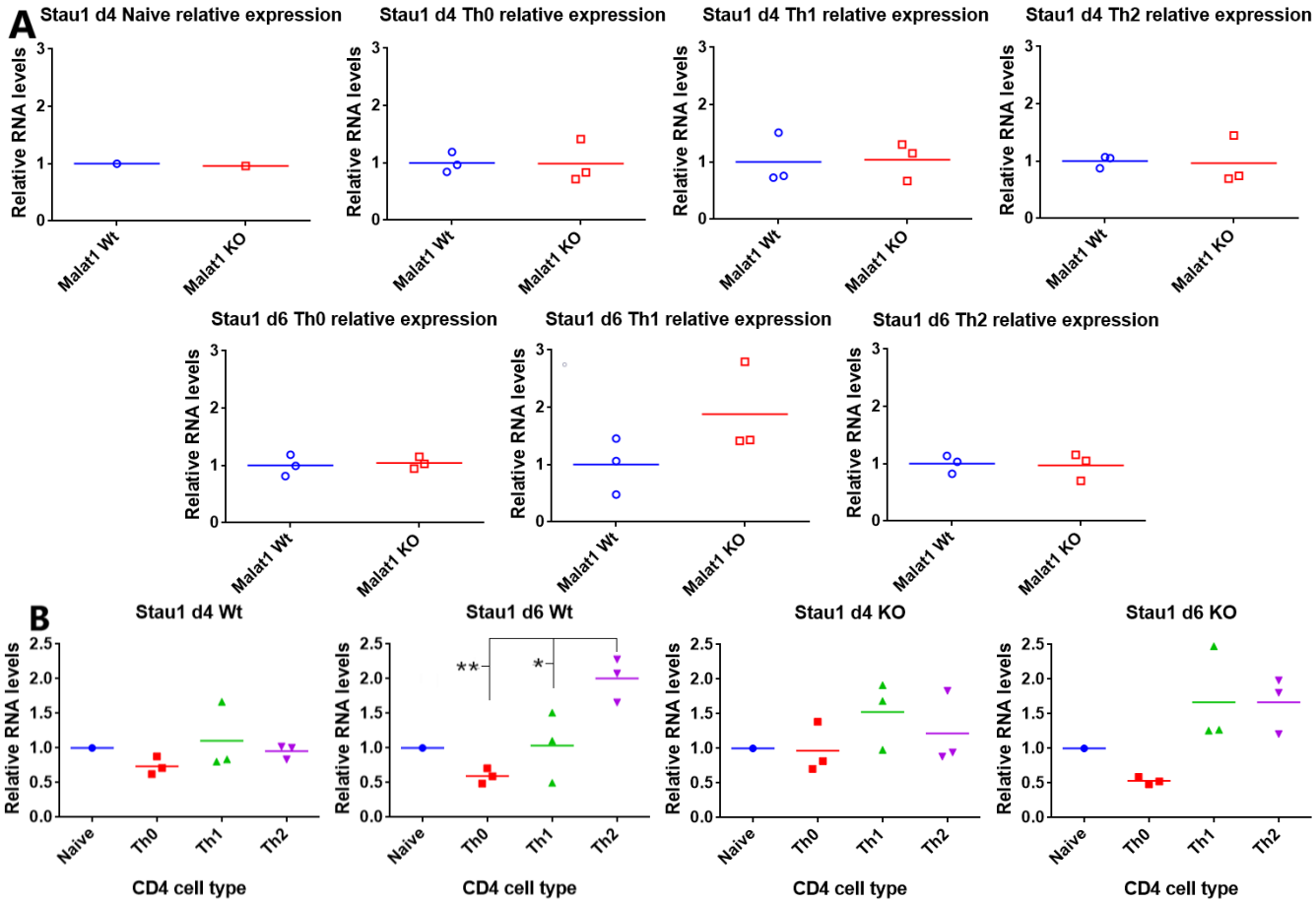

Figure 35: *Stau1* qPCR expression data from CD4⁺ cell cultures, showing relative RNA levels based upon differences in Ct values (ΔCt) between *Stau1* and *U6*. (A) Relative *Stau1* levels in Wt and KO cultures, normalised to the average of the Wt samples for each cell type. Significance based on 2 tailed t-test. (B) Relative *Stau1* levels in different CD4⁺ cell types, normalised to the naïve average for each genotype and culture time. Significance based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

Both the day 4 and day 6 Wt CD4⁺ cultures showed little to no difference in *Stau1* expression across the stimulated and naïve cells, save for a significant increase in the d6 Th2 culture compared to the Th0 and Th1 cells (Figure 35B). *Malat1* KO cultures showed similarly little change in *Stau1* levels at day 4, with a slight increase in the day 6 Th1 and Th2 culture (Figure 35B) – however, the Ct values for each culture indicated little to no expression of *Stau1* across all CD4⁺ cell types (Table 7), and as such the relevance of such differences is difficult to judge. Similarly, *Stau1* levels did not significantly differ between Wt and *Malat1* KO cultures across any of the CD4⁺ cell types (Figure 35A).

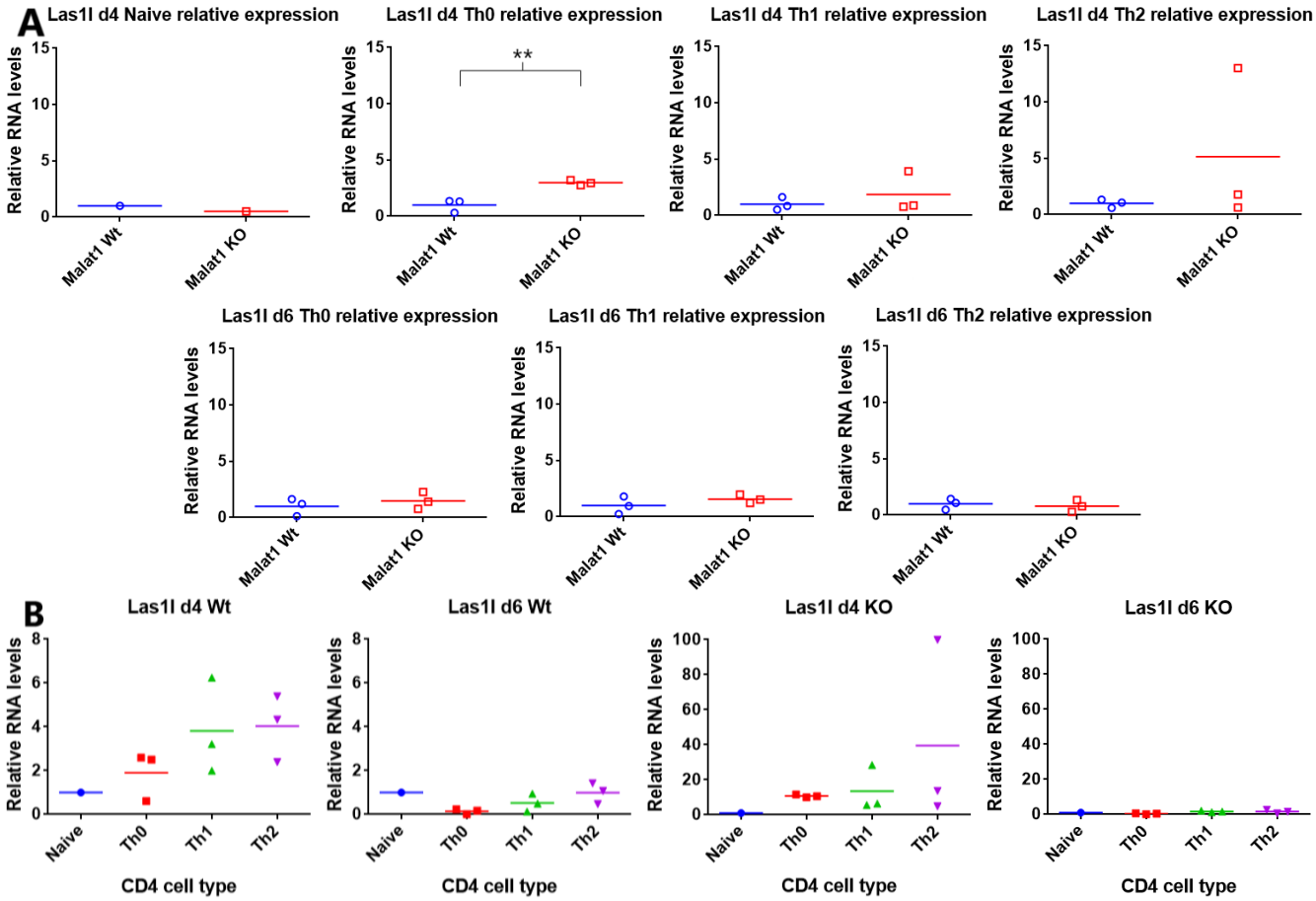

Figure 36: *Las1l* qPCR expression data from CD4⁺ cell cultures, showing relative RNA levels based upon differences in Ct values (ΔCt) between *Las1l* and *U6*. (A) Relative *Las1l* levels in Wt and KO cultures, normalised to the average of the Wt samples for each cell type. Significance based on 2 tailed t-test. (B) Relative *Las1l* levels in different CD4⁺ cell types, normalised to the naïve average for each genotype and culture time. Significance based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

While *Las1l* levels were increased in the stimulated day 4 Wt cultures compared to naïve cells (Figure 36B), the day 6 Wt cultures showed little change from the naïve, with the Ct values indicating negligible expression in these cell types (Table 7). Similarly, the KO cultures displayed considerably increased *Las1l* levels at day 4, but little to no change in *Las1l* expression at day 6 (Figure 36B). In spite of this, no significant difference in *Las1l* expression was observed between the stimulated CD4⁺ cell types across either genotype. Interestingly, *Las1l* showed some response to *Malat1* in specific CD4⁺ cell types, being significantly upregulated in the day 4 Th0 knockout cultures specifically (Figure 36A). As with *U2af1*, this suggests cell type specific *in trans* regulation of the gene by *Malat1*.

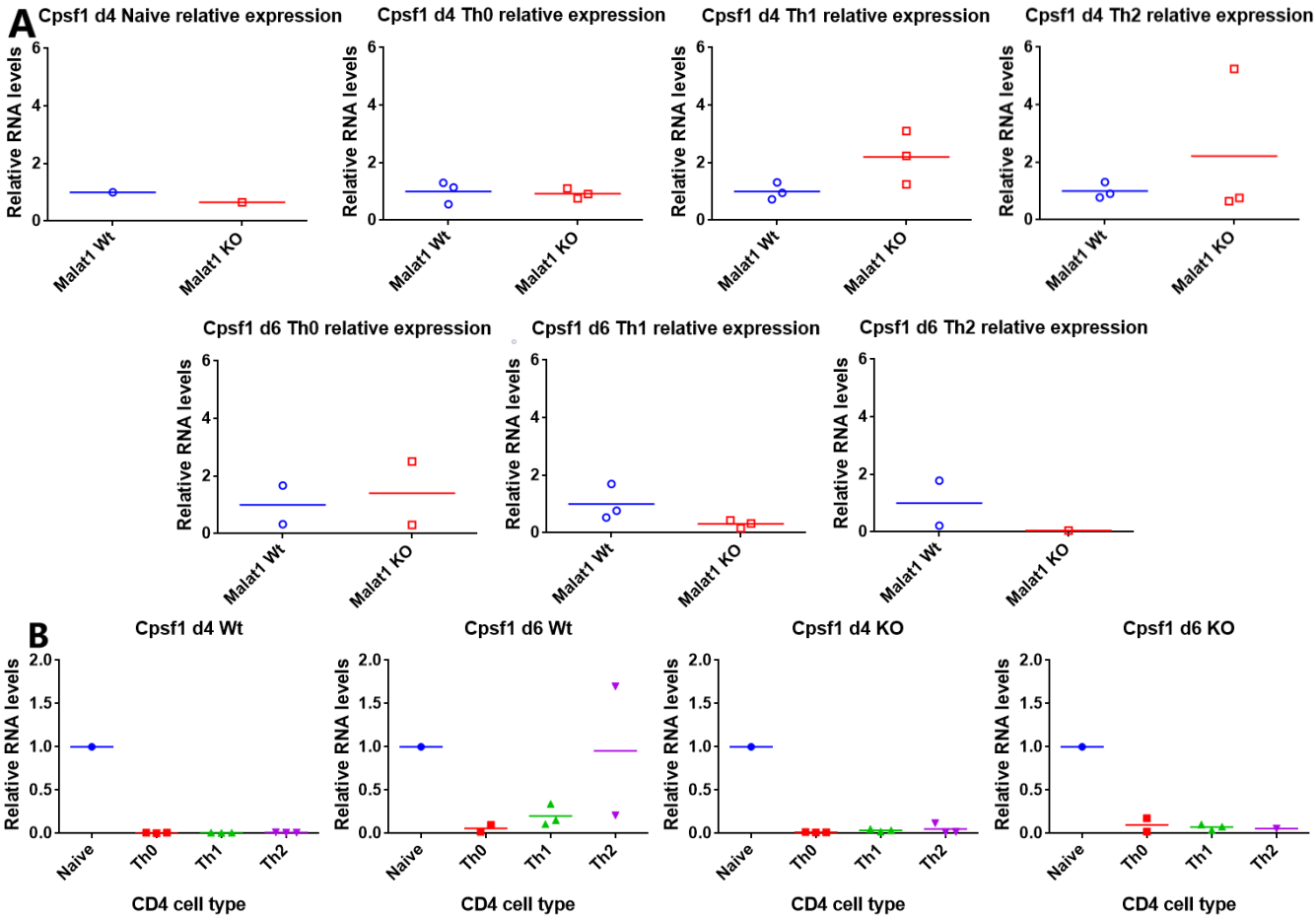

Figure 37: *Cpsf1* qPCR expression data from CD4⁺ cell cultures, showing relative RNA levels based upon differences in Ct values (ΔCt) between *Cpsf1* and *U6*. (A) Relative *Cpsf1* levels in Wt and KO cultures, normalised to the average of the Wt samples for each cell type. Significance based on 2 tailed t-test. (B) Relative *Cpsf1* levels in different CD4⁺ cell types, normalised to the naïve average for each genotype and culture time. Significance based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

Both the Wt and *Malat1* KO stimulated CD4⁺ cultures showed substantially reduced expression of *Cpsf1* compared to naïve cells at day 4, with far less of a decrease in *Cpsf1* levels observed in the Wt cultures at day 6 and little to no change in the day 6 Th2 cultures (Figure 37B). However, *Cpsf1* expression did not significantly differ between the simulated Th cell types in any culture, despite the varying expression observed in the day 6 Wt culture (Figure 37B), with the Ct values across the cultures indicating negligible expression of *Cpsf1* in all CD4⁺ cell types (Table 5). Similarly, none of these cell types showed significant differences in *Cpsf1* expression between the Wt and *Malat1* KO genotypes at either timepoint (Figure 37A).

Fibroblast culture analysis

As previous studies of *Malat1* have shown knockdown of the gene to significantly impair growth and proliferation of various cell types, including fibroblasts (Tripathi *et al* 2013), the effects of *Malat1* KO on growth of our MTF cultures were then tested for comparison. To this end, culture growth assays were set up from the MTF cultures at either 3 or 4 passages and run for up to 96h, with proliferation of each culture assessed by comparing the adjusted optical densities between the Wt and *Malat1* KO genotypes, between different initial cell densities and between different time points for each MTF culture. The plates were each visually examined on each day as well to assess any changes in confluence and the general health of the culture. The initial growth assay was seeded with either 500 or 1000 cells/well, while the subsequent final assay was set up with 5000, 10000 or 15000 cells in each well.

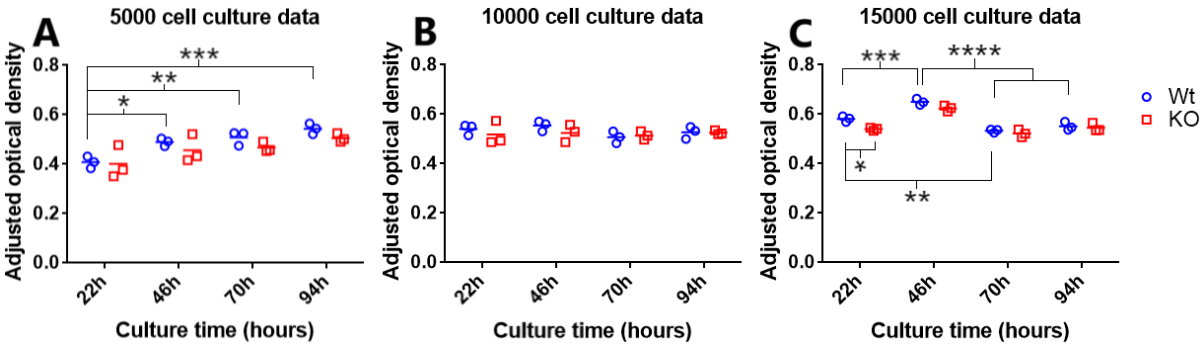

Figure 38: Fibroblast growth data, showing average optical density of plates seeded with (A) 5000, (B) 10000 or (C) 15000 cells taken after 20h of incubation with Alamar Blue. Optical density is adjusted relative to the average of 3 media controls. Significance between genotypes based on 2 tailed t-test, significance between time points based on one-way ANOVA; * = p<0.05, ** = p<0.01, *** = p<0.001, **** = p<0.0001

In the initial assay, while the optical density did increase with greater culture time, the difference between these time points was not significant for either initial cell density. This was unexpected, as longer culture times might be expected to produce greater changes in cell density. Given that very little change in confluence was observed for both culture densities at each time point, this suggested that the MTFs grow more slowly at lower densities. To further test this, the growth assay was repeated with greater initial cell densities as described above. Here, although significant increases in optical density were observed between some time points for the 5000 cell cultures, the later time points showed no significant increases in optical density (Figure 38A), while the 10000 cell cultures showed no significant changes in optical density whatsoever over 4 days (Figure 38B). Furthermore, the 15000 cell cultures showed an initial significant increase in optical density followed by a significant decrease (Figure 38C). Together with the observations of these cultures, this suggests that the 10000 and 15000 cell cultures rapidly reached confluence within the wells, while the 5000 well cultures showed strong initial growth which had begun to plateau by the end of the assay. This is supported by the visual

inspections of each well on each day (data not shown). However, save for a single significant reduction in the optical densities of the 15000 cell KO cultures at 22h, no significant difference was observed between the Wt and KO genotypes at any cell density tested (Figures 38A, 38B and 38C), indicating that *Malat1* KO has no effect on fibroblast growth rates over this period of time.

In order to further assess any effects of *Malat1* KO on fibroblast growth over a longer period of time, one Wt and one *Malat1* KO MTF culture were maintained for as long as possible, with each passaged at around 70-80% confluence and visually inspected daily to determine any differences in cell growth and proliferation. Although the two cultures showed similar rates of growth initially, by passage 6 (P6) the *Malat1* KO cells showed a noticeably reduced growth rate compared to the Wt cells, showing a larger, more spread out morphology associated with senescence (data not shown). Furthermore, while the Wt cultures were reliably maintained up to P10 before showing substantial cell death and the fore mentioned 'fried egg' senescent morphology at around P12, the KO cultures initially grew only to P8 before displaying arrested culture growth and senescent morphology (data not shown)**.** However, inspection of this culture later showed patches of strong cell growth, more similar to the Wt cultures at this timepoint, and on further passage these cells were maintained up to P12 by the end of the experiment. This indicates that *Malat1* has some effect on cell growth, senescence and cell cycle progression in MTFs over longer culture times.

Discussion

In this study we found some evidence of *Malat1* regulating some of its surrounding genes, as well as other RNA binding genes, in a cell type specific manner in $CD4^+$ T cells, with some regulation of *Malat1* neighbouring genes also observed in mouse tail fibroblast (MTF) cultures. Additionally, we observed some residual expression of *Malat1* in stimulated $CD4^+$ cultures, while our MTF cultures showed more rapid senescence and reduced growth and proliferation in the *Malat1* KO cells after 5 passages.

Stimulation of primary and immortalised $CD4^+$ T cells produces different significant changes in *Malat1* expression

The differing expression changes in *Malat1* observed upon stimulation of primary and immortalised $CD4^+$ T-cell lines indicate that the two sets of cells show different functional responses of *Malat1* to T cell receptor activation, either by CD3/CD28 antibody stimulation or by T cell receptor crosslinking by PHA/PMA or ConA. This may be ascribed to differences in the signalling responses produced by each system of stimulation – while ConA and PMA/PHA are demonstrated to induce proliferation of T cell lines, stimulating protein kinase C (PKC) activation and an influx of Ca^{2+} into the cytoplasm, studies have shown these signals to be produced via different mechanisms for ConA and CD3/CD28 activation (Kay 1991, Pang *et al* 2012). Furthermore, while PHA has been demonstrated to bind to and crosslink the T cell receptor/CD3 complex (Kay 1991), triggering downstream T cell receptor signalling, Jurkat cells have been recorded to show several differences in downstream T cell receptor signalling from primary human cells. This includes exaggerated Ca^{2+} signalling and a more limited range of cytokine production due to hyperphosphorylation and increased activation of several downstream signalling kinases, including several involved in activation of PKC (Bartelt *et al* 2009). As such, although both CD3/CD28 and ConA stimulation promote proliferation of $CD4^+$ T cells, the distinct signalling pathways involved in each would likely have differing effects on *Malat1* levels, particularly as the significant expression changes observed in our stimulated primary cells indicated *Malat1* to be regulated by T cell receptor signalling (Figure 20B). Stimulating further cultures of Jurkat and EL4 cells with the CD3/CD28 protocol used for the primary cell cultures and assessing *Malat1* expression with further qPCRs would allow this possibility to be further assessed.

As mentioned above, in both our primary $CD4^+$ experiments and in previous transcriptomic studies (Stubbington *et al* 2015), stimulation of $CD4^+$ cultures via T cell receptor activation significantly reduced *Malat1* expression, indicating that T cell receptor activation and subsequent downstream signalling regulate expression of *Malat1*. Such regulation of *Malat1* could be further evaluated by chromatin or RNA immunoprecipitation to test for association of the *Malat1* genomic locus with

downstream T cell signalling factors, such as T bet, GATA-3 or members of the Jnk signalling pathway (Hu *et al* 2013, Yang *et al* 1998), as would testing for any effects of knockdown of these T cell signalling factors on *Malat1* expression with and without T cell stimulation. Although both *IFNγ* and *IL-4* showed significant changes in expression in the KO cells from *Malat1* Th0 cultures (Figures 18A and 19A), suggesting a possible function for *Malat1* in regulating expression of cell type specific cytokines in $CD4^+$ Th0 cells, the minimal expression of both cytokines observed in this $CD4^+$ cell type make the functional relevance of these results difficult to judge. Furthermore, while the Th1 and Th2 cultures showed differing expression levels of ***I****FNγ*, *IL-4* and *IL-10* respectively in *Malat1* KO cells compared to Wt on a more substantial scale (Figures 18A, 19A and 20A), these differences were not significant on either day 4 ($p = 0.41$, $p = 0.52$, $p = 0.08$) or day 6 ($p = 0.07$, $p = 0.06$, $p = 0.65$), possibly due to the low sample numbers used for expression analysis of these cytokines. As such, further qPCR or ELISA analysis of *IL-4*, *IL-10* and *IFNγ* expression in *Malat1* KO $CD4^+$ cells would be advisable to better evaluate if the lncRNA has some function in regulating these cytokines.

Malat1* KO cells show some residual expression of *Malat1

The reduced expression of *Malat1* in the stimulated Wt $CD4^+$ cultures compared to naïve cells was consistent both with previous data from this research group (Hewitson, unpublished) and with transcriptomic data from mouse $CD4^+$ cells, as was the similarly reduced expression of *Neat1* in these cultures (Stubbington *et al* 2015). Furthermore, both *Scyl1* and *Map3k11* showed similar decreases in expression in stimulated $CD4^+$ T cells in the data from Stubbington *et al* (2015) to that observed in our day 4 cultures (Figures 24B and 25B). While the day 6 cultures did not show the same significantly lower expression for these genes, our data indicates this to be due to differences in the response of the two day 6 culture sets to stimulation, rather than a definite change in regulation of these genes in the day 6 cultures. Notably, although the day 4 Th2 KO cultures showed similarly reduced Malat1 expression compared to the other day 4 cultures (Figure 21A), our statistical analyses showed this difference to be non-significant, highlighting both the need for further replicates for greater reliability and that the current analysis of our results should be considered carefully.

While the residual expression observed in the *Malat1* KO cells from stimulated $CD4^+$ cultures might reflect an increase in primer dimer amplification with T cell activation and differentiation, this is unlikely given both the similarly low Ct values observed in the *Malat1* KO fibroblasts (Table 6) and the greatly reduced primer dimer peaks observed with lower Ct values in the naïve *Malat1* KO cells (Figure 14). Such residual expression, though unexpected, is consistent with data from the study that generated the *Malat1* KO system used in these experiments – in this study, some residual expression of a 3.2kb transcript was observed in brain tissues from *Malat1* KO mice, which was thought to reflect

a truncated transcript from the 3' end of the gene expressed through activation of an internal promoter sequence in the gene (Figure 38). This suggestion was also proposed by Zhang *et al* (2012), who observed some slight residual expression of *mascRNA* in *Malat1* KO brain samples and suggested that this reflected activity of a brain specific promoter upstream of *mascRNA*, i.e. within the 3' end of the *Malat1* sequence. This hypothesis is also consistent with a separate study of overexpression of *Malat1* fragments, which indicated *Malat1* to have a 3' functional motif with key biological functions in cell migration and proliferation (Xu *et al* 2010).

As the *Malat1* qPCR primers used for the qPCR expression analysis amplify within the region thought to be expressed in the truncated transcript observed by Nakagawa *et al* (2012) (Figure 39), this truncated expression may be occurring in the stimulated $CD4^+$ and MTF knockout cultures, with T cell receptor signalling stimulating expression from the internal *Malat1* promoter sequence. However, primer dimer amplification remains a possibility, and as such would need to be tested in the fibroblasts and stimulated $CD4^+$ cultures by using either further melting curve analysis or gel electrophoresis to assess the size of the product observed in these KO cells. Additionally, Nakagawa *et al* (2012) found no residual expression in mouse embryonic fibroblast (MEF) samples, which contradicts the results from the MTF samples. While this may reflect differences in fibroblast phenotype, as some studies have shown differences in gene expression between fetal and adult fibroblasts in response to stimuli such as wounding and tissue engineering protocols (Tang *et al* 2014), further analysis of the *Malat1* KOs would be required to confirm this, possibly by further qPCRs using primers at either the 5' or 3' end of the gene.

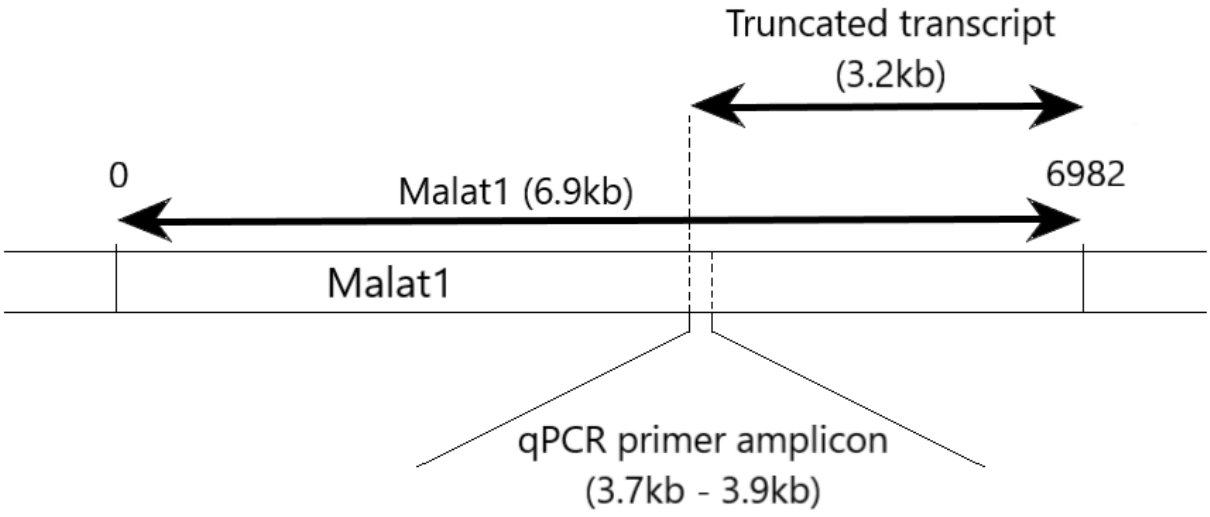

Figure 39: Simplified diagram of the *Malat1* locus, showing the relative positions of the truncated transcript identified by Nakagawa *et al* (2012) and the region amplified by the qPCR primers used in this study.

***Neat1*, *Scyl1* and *Map3k11* show regulation by *Malat1* in specific cell types**

Based on the expression analysis from $CD4^+$ and fibroblast cultures of both Wt and *Malat1* KO genotypes, *Malat1* does appear to regulate expression of *Neat1*, *Scyl1* and *Map3k11*, but only in specific cell types or conditions. Although our results differ between the two culture time points, with significantly altered surrounding gene expression in Th0 and Th1 cultures only on days 4 and 6 respectively (Figures 40B and 40F), this is likely due to the effects of expansion of the $CD4^+$ cultures in IL-2 after day 4. While application of IL-2 has been strongly associated with increased T cell growth and survival *in vitro*, and thus is commonly used for prolonged *in vitro* $CD4^+$ culture, it has also been found to have a key role in T helper cell differentiation, activating expression of downstream cytokine receptors and signalling proteins to modulate production of Th1 and Th2 effector cytokines (Hedfors and Brinchmann 2003, Liao *et al* 2011). Specifically, activation of STAT5 by IL-2 induces expression of IFNγ and IL-12Rβ2 during Th1 differentiation (Shi *et al* 2008, Liao *et al* 2011) and expression of IL-4 and IL-4Rα in Th2 cells (Cote-Sierra *et al* 2004, Liao *et al* 2008). As such, given the apparent effects of T cell receptor signalling on *Malat1* expression observed in these cultures, it is likely that the addition of IL-2 at day 4 promotes further Th1 and Th2 differentiation of the *in vitro* $CD4^+$ cultures, and this heightened Th1 and Th2 signalling in turn alters $CD4^+$ cell type specific expression and activity of *Malat1*. Given the differing directions of expression change observed at days 4 and 6 for both *Scyl1* and *Neat1*, this effect of IL-2 stimulation is more likely to account for the differences in significance than any statistical error arising from the differing sample sizes at each culture time – however, further sampling of primary $CD4^+$ cultures at the day 4 time point would be required to confirm this.

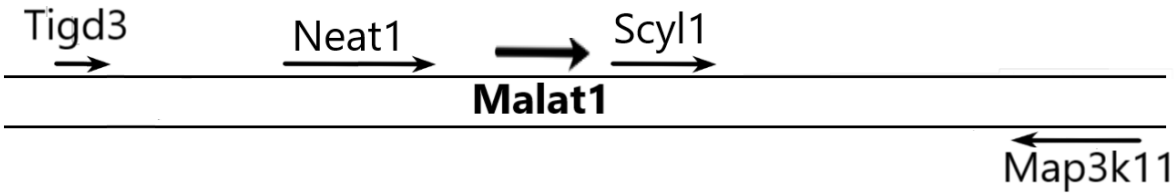

Figure 40: Simplified diagram of the relative locations of *Malat1* and the surrounding candidate genes

	Malat1	Neat1	Scyl1	Map3k11	Tigd3
Naïve	**0.00025*** **	**0.431**** **	**1.184** **	1.008	1.114
d4 Th0	**0.0281***	**0.725***	**2.738***	0.729	1.231
d4 Th1	**0.0222***	1.281	2.263	0.895	1.348
d4 Th2	0.0197	0.778	2.312	1.276	0.741
d6 Th0	**0.00329**** **	0.695	0.922	0.833	0.998
d6 Th1	**0.00797** **	**0.729***	**0.745***	**0.708***	1.928
d6 Th2	**0.00554*** **	0.442	0.753	0.620	1.173
Fibroblasts	**0.00344** **	1.037	**1.690***	1.624	1.178

Table 8: Summary of relative expression of Malat1 and surrounding genes in Malat1 KO cultures, expressed as a proportion of Wt expression averaged for each $CD4^+$ cell type and fibroblast cultures. Significant differences from Wt are highlighted in bold based on 2 tailed T-tests: * = $p<0.05$, ** = $p<0.01$, *** = $p<0.001$, **** = $p<0.0001$

For *Neat1*, these $CD4^+$ expression results are consistent with some previous data on its interactions with *Malat1*, as two independent *Malat1* KO studies showed regulation of *Neat1* by *Malat1* (Zhang *et al* 2012, Nakagawa *et al* 2012), with Nakagawa *et al* (2012) demonstrating that such regulation occurred only in specific cell types such as MEFs or intestinal cells. However, Eiβmann *et al* (2012) tested *Neat1* expression levels across a variety of tissues, including the liver, brain, intestine and colon tissues assessed by the other two studies, and observed no significant differences in *Neat1* expression across any of them. While none of these studies looked specifically at lymphocytes or $CD4^+$ T cells, these differing results highlight the current contrasts in the data concerning regulation by and function of *Malat1*. This discrepancy may be attributed to differences in knockout techniques – while Zhang *et al* (2012) and Nakagawa *et al* (2012) both utilised deletion or inactivation (respectively) of the *Malat1* transcriptional start site and recorded significant local effects of *Malat1* KO, Eiβmann *et al* (2012) deleted the entire gene and observed no significant effects. This suggests that the effects of *Malat1* knockout on *Neat1* levels observed by Nakagawa *et al* (2012) and Zhang *et al* (2012) are dependent on the *Malat1* locus itself, rather than the loss of *Malat1* transcription. Given the truncated *Malat1* transcript identified by Nakagawa *et al* (2012) in their knockout model, this might suggest aberrant binding of transcription factors to this truncated transcript, which in turn would affect transcription of *Neat1*. However, this could equally be due to differences in the gene knockout and expression analysis methods used in each study, and thus experiments comparing transcription factor binding at

the *Malat1* locus in these 3 knockout models and Wt cells, possibly via chromatin immunoprecipitation, would be required to further assess this possibility.

Notably, of our candidate genes, the two located closest to *Malat1*, *Neat1* and *Scyl1* (Figure 40), were most affected by its knockout across multiple $CD4^+$ cell types. Though this might indicate that insertion of the knockout cassette disrupts expression of these nearby genes, Nakagawa *et al* (2012) tested for such effects on *Neat1* expression in the original knockout study by comparing expression with knockout and antisense knockdown (KD) of *Malat1*. As both means of gene inactivation produced significantly reduced *Neat1* levels, they concluded that the loss of *Malat1* expression was responsible for the changes in *Neat1* expression, rather than disruption by the LacZ knockout cassette. While this study did not look specifically at Scyl1 expression in the *Malat1* KD cells, the gene might also be expected to show disrupted expression in the *Malat1* KOs, particularly given similar disruption observed downstream from other knockout cassettes (Pham *et al* 1996). However, such disruption of expression by *Malat1* knockout would be likely to produce more universal effects on Scyl1 levels across all $CD4^+$ cultures – furthermore, the differing effects of *Malat1* KO on *Scyl1* expression in different cell types indicate a more specific regulatory effect to be responsible, rather than general disruption of expression near the *Malat1* locus. As such, this increased effect of *Malat1* KO on closer genes may indicate the presence of one or more *cis* regulatory elements near *Malat1*, with greater effects on genes in the immediate proximity than on more distal transcripts. This possibility could be further tested with qPCR analysis of more of the genes surrounding *Malat1*, particularly those located between the candidate genes tested here such as *Ltbp3*, *Ehbp1l1* or *Frmd8*, and comparing the extent of *Malat1* KO effects on genes at different distances from *Malat1*.

Several mechanisms have been identified by which intergenic lncRNAs regulate their surrounding genes *in cis*, such as promoting formation of chromatin loops to allow interaction of enhancer and promoter sequences or increasing chromatin accessibility and Pol2 occupancy of nearby genes (Vance and Ponting 2014). While one of these mechanisms of *in cis* regulation may occur at the *Malat1* locus, these mechanisms have been found to positively regulate local transcription rather than repress it. As such, this type of regulation would be inconsistent with our observed *Scyl1* expression results, both with the significant upregulation observed with *Malat1* KO in both fibroblast and $CD4^+$ cultures and the diverging expression changes recorded in different $CD4^+$ KO cultures. This varying effect of *Malat1* KO on *Scyl1* expression in different $CD4^+$ cell types is unlikely to be due to changes in expression between the various cell types – while *Scyl1* expression did vary across $CD4^+$ cell types both here and in previous data (Stubbington *et al* 2015), these differences did not correlate with the cell types which observed significant differences in *Scyl1* levels between Wt and KO cultures. In addition, the interactome data of Engreitz *et al* (2014) found only a single direct interaction of *Malat1* with the

ncRNA *U1*, with their results suggesting the rest to be indirect interactions of *Malat1* with pre-mRNAs and active chromatin sites via association with splicing factors. As such, our *Scyl1* expression data indicates an indirect, highly cell type specific form of regulation by *Malat1*, likely through the activity of cell type specific transcription factors.

A more cell type specific form of *in cis* regulation by *Malat1* would be targeting of cell type specific transcription factors to the *Malat1* locus, with these factors subsequently interacting with nearby enhancer or promoter sequences to regulate surrounding gene expression. Such regulation would be consistent with the observed $CD4^+$ cell type specificity, as a previous study identified distinct genomic profiles of active enhancer regions in Th1 and Th2 cells, with binding of acetyltransferase p300 to enhancers of lineage specific genes in both cell types and enrichment of binding motifs for lineage specific transcription factors at these enhancers (Vahedi *et al* 2012). This possibility is also consistent with earlier work on the *Malat1* interactome, which observed only indirect interactions of *Malat1* with other loci or mRNAs and, in particular, a strong indirect association of *Malat1* with *Scyl1* and a less strong association with *Map3k11* (Engreitz *et al* 2014). Furthermore, such transcription factor binding would be consistent with the results of the initial *Malat1* KO studies – as detailed earlier, the differing effects of various *Malat1* KO systems on *Neat1* expression indicate the *Malat1* locus itself to have a greater effect on neighbouring gene expression than *Malat1* transcription, and therefore suggest binding of transcription factors to the *Malat1* gene influencing surrounding gene levels (Zhang *et al* 2012, Nakagawa *et al* 2012, Eiβmann *et al* 2012). Such regulation could occur via the formation of chromatin loops, allowing interaction of the *Malat1* bound factors with more distal genes such as *Map3k11*, and could be further assessed by using chromatin immunoprecipitation to compare binding of $CD4^+$ T cell signals at both the *Malat1* and surrounding gene loci between Wt and *Malat1* KO cultures.

Interestingly, our qPCR data showed evidence of neighbouring gene regulation by *Malat1* in only 3 of the $CD4^+$ cell types; naïve cells, the day 4 Th0 culture, and the day 6 Th1 culture (Table 8). Together with the apparent effects of T cell receptor signalling on *Malat1* expression and the possibility of transcription factors at the *Malat1* locus regulating the surrounding genes, this suggests regulation of these genes via binding of lineage specific downstream signalling proteins to the *Malat1* locus within these $CD4^+$ cell types. Given that the Th1 specific regulators STAT4 and T-bet have been observed to bind to and regulate expression of numerous lncRNAs, with STAT4 and STAT1 also identified as key regulators of enhancer region activation in Th1 cells (Hu *et al* 2013, Vahedi *et al* 2012), these transcription factors would be likely candidates for such cell type specific regulation of the *Malat1* locus. Additionally, transcription factors found to be specific to naïve and Th0 $CD4^+$ cells, such as those identified by Kanduri *et al* (2015) or Spurlock *et al* (2017) via transcriptional profiling, could also be

worth assessing as candidates for *Malat1* regulation. Downstream effectors of STAT5A, STAT5B and other members of the Jnk signalling cascade would also be worth investigating further, as these factors were found to be activated by IL-2 stimulation of $CD4^+$ cells and have roles in regulating Th1 differentiation (Gesbert *et al* 2005, Liao *et al* 2011), and thus would be consistent with the day 6 specific expression changes observed here. Possible interactions between *Malat1* and these $CD4^+$ regulatory proteins could be tested via chromatin immunoprecipitation (ChIP) analysis to test for binding of the transcription factors to the *Malat1* locus or surrounding enhancer sequences and gene loci as described above, and to compare any such binding between the different $CD4^+$ cell types.

Given that previous *Malat1* interactome studies indicated active gene expression to be required for *Malat1* association at the genomic locus (Engreitz *et al* 2014), this apparent regulation of *Neat1*, *Scyl1* and *Map3k11* suggests these genes to have potential functions in naïve, Th0 and Th1 $CD4^+$ subtypes. For *Scyl1*, this is consistent with previous data on its role in Golgi body morphology (Burman *et al* 2010) – IL-2, a characteristic Th1 cytokine, is thought to be secreted via the Golgi apparatus (Duitman *et al* 2008), and as such active expression changes in components of this cellular apparatus might be expected in the Th1 $CD4^+$ cultures. Similarly, other Map kinases and members of the Jnk signalling pathway have been found to have key functions in the function and differentiation of Th1 cells, correlating with the significant effects of *Malat1* KO on *Map3k11* in the day 6 Th1 cultures specifically (Yang *et al* 1998, Dong *et al* 2000). While no such $CD4^+$ cell specific functions have been identified for *Neat1*, human $CD4^+$ knockout models of the gene have shown enhanced HIV infection due to reduced splicing of viral RNAs by the paraspeckles, and thus indicate a role for *Neat1* in antiviral immunity of T helper cells (Liu *et al* 2018). However, these potential functions are contrasted somewhat by our qPCR expression data, as, aside from the significant downregulation upon T cell activation, no significant expression changes were observed in all 3 genes between the Th0, Th1 and Th2 cultures at any timepoint recorded (Figures 22B, 24B and 25B). Given these conflicting possibilities, assessing the potential functions of these surrounding genes in the naïve, Th0 and Th1 $CD4^+$ cell types might be prudent, possibly by using either KO or antisense knockdown models of each gene to compare $CD4^+$ culture growth and survival while using further qPCR analysis to assess expression of characteristic Th1 and Th2 cytokines in these cells.

Las1l **and** ***U2af1*** **are regulated** ***in trans*** **by** ***Malat1*** **in specific CD4$^+$ cell types**

The lack of significant effect of *Malat1* KO on expression of the majority of the RNA binding proteins tested here is consistent with existing data – while *Malat1* has been suggested to be essential to alternative splicing occurring in the nuclear speckles (Tripathi *et al* 2010), knockout studies showed no effect of *Malat1* KO on pre-mRNA splicing and nuclear speckle structure (Zhang *et al* 2012). Furthermore, while Engreitz *et al* (2014) found these RNA binding proteins to bind indirectly to *Malat1*, this does not necessarily indicate regulation of these genes by the lncRNA. However, while some substantial expression differences were observed between the naïve Wt and Malat1 KO cultures for these RNA binding protein genes, such as increased *U2af1* expression and reduced *Hist1h4a* and *Ddx23* expression in the naïve KO cells, none of these differences could be assessed for significance, given that only 1 Wt and 1 *Malat1* KO naïve cell sample was tested for each of these RNA binding proteins. Similarly, while substantial expression differences were observed across these RNA binding proteins between the naïve and stimulated CD4$^+$ cultures, the significance of these results could not be confirmed with only 1 naïve cell sample. As such, further qPCRs would be required to better assess the significance of any observed expression differences across these cell types.

Although further qPCR analysis would be required to test the significance of the observed CD4$^+$ expression differences, the direction of these expression changes was unexpected. While the CD4$^+$ transcriptomic data from Stubbington *et al* (2015) showed substantially reduced expression of all these RNA binding proteins in Th1 and Th2 cells compared to naïves, only *U2af1* and *Cpsf1* showed reduced expression in stimulated CD4$^+$ cultures, while the others showed either little expression change or substantially increased expression in these cultures. This contrast in expression data may be due to differences in technique – while both experiments used the same protocol for CD4$^+$ stimulation of Th1 and Th2 cell cultures, Stubbington *et al* (2015) determined expression values using microarray analysis, as opposed to the qPCR analysis used here. As such, given that *Ddx23* and *Las1l* both show evidence of alternative splicing (Rengasamy *et al* 2017, Joo *et al* 2013), it may be that the primer sets from both studies bind preferentially to different splice variants of each gene, with the reduced expression of one splice variant during T cell activation leading to concurrently increased expression of another. However, not all genes showing different expression changes from Stubbington *et al* (2015) have alternative splicing variants – as such, while this splice variant explanation is possible, particularly given the intron spanning qPCR primer pairs used in this study, it would not account for all differences in CD4$^+$ expression patterns between this study and previous data. As such, further qPCRs and statistical analysis would be crucial to better assess these conflicting data sets, while any differences in splicing variant expression and primer amplification between the two studies could be further tested via qPCRs with primers specific to both spliceforms of each gene.

Some evidence of cell type specific *in trans* regulation of these RNA binding proteins by *Malat1*, specifically *U2af1* and *Las1l*, was observed our $CD4^+$ cultures. As previous studies have indicated *Malat1* localisation to genes to be dependent upon active transcription of those genes (Engreitz *et al* 2014), this indicates some cell type specific role for U2af1 and Las1l in $CD4^+$ cultures. This is consistent with a study showing a key role of U2af2, regulated by U2af1, in coordinating T-cell activation, as well as transcriptomic studies showing strong association of *Malat1* with numerous pre-mRNA splicing factors (Whisenant *et al* 2015, Chen *et al* 2017). Furthermore, although previous *Malat1* KO studies found no significant changes in splicing factor expression or localisation or in global pre-mRNA splicing patterns (Zhang *et al* 2012, Nakagawa *et al* 2012), Zhang *et al* (2012) did observe significant changes in a handful of exons in both brain and liver tissues**,** highlighting the possibility of *Malat1* having specific effects on alternative splicing factors in our $CD4^+$ T cell cultures. This could be further investigated by looking at the effects of *U2af1* knockout on expression of characteristic Th2 cytokines, such as IL-4 and IL-10, as well as comparing alternative splicing of mRNAs known to interact with U2af1.

While no specific function has been observed for Las1l in $CD4^+$ T cells, as a subunit of the nucleolar rRNA processing complex, it has a ubiquitous function in ribosome biogenesis and cell proliferation, specifically in G1 to S phase progression (Castle *et al* 2010), and has been found to interact with *Malat1* at both the protein and pre-mRNA stages (Chen *et al* 2017, Engreitz *et al* 2014). In addition, earlier research showed T cell stimulation to increase ribosomal protein and RNA synthesis, with other rRNA processing factors such as RRS1 upregulated by the ERK MAPK pathway (Asmal *et al* 2003). Given this data, it is possible that *Las1l* is similarly upregulated in our stimulated $CD4^+$ cultures, with the apparent *in trans* regulation of *Malat1* specifically in day 4 Th0 cells potentially indicating either cell type specific targeting of the lncRNA to *Las1l* or regulation by *Malat1* during T cell activation being suppressed by subsequent Th1 or Th2 specific signalling. However, given the very low expression of *Las1l* observed across all $CD4^+$ cell types (Table 7), this possibility should be considered carefully, as such low levels of *Las1l* would be unlikely to have functional relevance. As such, using KO or knockdown models of *Las1l* to further assess the genes role in $CD4^+$ T cells, as suggested for the genes surrounding *Malat1*, might be prudent to follow up on these initial results. Similarly, any effects of increased *Las1l* expression in these $CD4^+$ cultures on rRNA processing could be further investigated by comparing levels of the *12S* and *32S* rRNAs in our Wt and *Malat1* $CD4^+$ cultures, as siRNA knockdown of *Las1l* has been found to substantially alter levels of these two ribosomal subunit components (Castle *et al* 2010).

Malat1 **KO affects cell proliferation of long term fibroblast cultures**

Observation of the MTF cultures over longer time periods indicated that *Malat1* affected the growth and proliferation of these cells after 5 passages (P5), with *Malat1* KO cultures beyond this point showing less frequent divisions and with many cells showing a larger, more spread out 'fried egg' morphology characteristic of senescent cells. Such effects on proliferation are consistent with *Malat1* data from numerous cell lines, where knockdown or silencing of the lncRNA in gastric, cervical or breast cancers led to significant G0/G1 cell cycle arrest and reduced cell growth and proliferation (Wang *et al* 2014, Guo *et al* 2010, Zhao *et al* 2014). Furthermore, a previous paper assessing *Malat1* depletion in human dermal fibroblasts (HDFs) observed similar G1 cell cycle arrest and senescent morphology in their cultures upon antisense knockdown of the lncRNA, with significant downregulation of numerous proteins involved in G1/S phase transition (Tripathi *et al* 2013). As such, the senescent morphology and reduced proliferation of the *Malat1* KO MTF cultures could occur through the mechanism proposed in this earlier study, with *Malat1* depletion leading to altered expression and alternative splicing of the transcription factor B-MYB and subsequent expression changes in the cell cycle genes regulated by this factor inhibiting mitotic progression (Tripathi *et al* 2013). However, this data only represents a preliminary set of results, particularly given the lack of significant difference observed in shorter term Alamar Blue assays (Figure 38). As such, these findings would need to be further investigated before such interpretations are made, possibly by using qPCRs to compare the expression of B-MYB and other associated cell cycle genes in Wt and *Malat1* KO MTF cultures, or the presence of senescence markers such as oxidative stress and shortened telomeres. Alternatively, supplementing *Malat1* KO cultures with cell cycle factors associated with G1/S or G2/M phase transition, such as FOXM1, would further assess these findings by testing for a growth phenotype more similar to the Wt MTF cultures.

An alternative explanation for the MTF culture growth was also suggested by the data of Tripathi *et al* (2013), which observed increased activity of p53 upon *Malat1* depletion in various cell lines. Given both the known function of p53 as a DNA damage response regulator (Williams and Schumacher 2016) and the importance of cell cycle checkpoint pathways, such as those disrupted by *Malat1* depletion, in regulating DNA repair mechanisms (Branzei and Foiani 2008), *Malat1* has been suggested to have a role in regulating DNA damage repair. As such, it may be that this process is impaired in the *Malat1* KO MTF cells, leading to the more rapid accumulation of DNA damage and subsequently more rapid senescence observed by P5. This might also be consistent with the increased proliferation of *Malat1* KO cells observed after P9 – as this growth was initially observed in patches among no clear growth in the other cells, this suggests individual MTFs to have acquired random mutations which restored similar proliferation to Wt cultures, possibly by restoring *Malat1* expression or knocking out

expression of another gene such as p53 which ablates the *Malat1* KO phenotype. As with the earlier explanation, however, this is based on initial findings, and would require further experimental verification to be fully assessed.

Conclusion

In summary, our data indicates a regulatory function for *Malat1* in naïve, Th0 and Th1 $CD4^+$ T cells, possibly *in cis* via transcription factor binding at the *Malat1* locus acting on nearby enhancer or promoter sequences. However, given both the variance observed between sample sets and the relatively low sample size, particularly for the studies of RNA binding proteins, further analysis would be required both to confirm the significant expression changes observed here and to further assess some substantial changes whose significance could not be determined. Assessing expression of these genes in Wt and *Malat1* KO cultures of other $CD4^+$ cell types, such as Th17 and Treg cells, would also be advisable to more clearly determine the effects of $CD4^+$ T cell receptor signalling and activation pathways on *Malat1*, as well as the expression and regulation of our candidate genes in response to such T cell signalling. Finally, our growth data from mouse tail fibroblast cultures correlates with known functions of *Malat1* in cell growth and proliferation, and may indicate regulation either of cell cycle gene activity or of DNA repair pathway components by the lncRNA.

References

Aune, T.M. *et al* (2016) Long noncoding RNAs in T lymphocytes. *Journal of Leukocyte Biology* **99**: 31 – 44

Ard, R. *et al* (2014) Long non-coding RNA-mediated transcriptional interference of a permease gene confers drug tolerance in fission yeast. *Nature Communications* **5**: 5576

Ard, R. *et al* (2016) Transcription-coupled changes to chromatin underpin gene silencing by transcriptional interference. *Nucleic Acids Research* **44** (22): 10619 – 10630

Arun, G. et al (2016) Differentiation of mammary tumors and reduction in metastasis upon Malat1 lncRNA loss. *Genes and Development* **30** (1): 34 – 51

Asmal, M. *et al* (2003) Production of Ribosome Components in Effector CD4+ T Cells Is Accelerated by TCR Stimulation and Coordinated by ERK-MAPK. *Immunity* **19** (4): 535 – 548

Bartelt, R.R. *et al* (2009) Comparison of T Cell Receptor-Induced Proximal Signaling and Downstream Functions in Immortalized and Primary T Cells. *PLoS One* **4** (5): e5430

Branzei, D. and Foiani, M. (2008) Regulation of DNA repair throughout the cell cycle. *Nature Reviews Molecular Cell Biology* **9**: 297 – 308

Brockdorff, N. (2013) Noncoding RNA and Polycomb recruitment. *RNA* **19** (4): 429 – 442

Brown, J.A. *et al* (2014) Structural insights into the stabilization of MALAT1 noncoding RNA by a bipartite triple helix. *Nature Structural & Molecular Biology* **21** (7): 633 – 640

Burman, J.L. *et al* (2010) Scyl1 Regulates Golgi Morphology. *PLoS One* **5** (3): e9537

Carninci, P. (2009) Molecular biology: The long and short of RNAs. *Nature* **457**: 974 – 975

Castle, C.D. *et al* (2010) Las1L Is a Nucleolar Protein Required for Cell Proliferation and Ribosome Biogenesis. *Molecular and Cellular Biology* **30** (18): 4404 – 4414

Chu, Y-D. *et al* (2016) A novel function for the DEAD-box RNA helicase DDX-23 in primary microRNA processing in *Caenorhabditis elegans*. *Developmental Biology* **409** (2): 459 – 472

Chen, R. *et al* (2017) Quantitative proteomics reveals that long non-coding RNA MALAT1 interacts with DBC1 to regulate p53 acetylation. *Nucleic Acids Research* **41** (17): 9947 – 9959

Clemson, C.M. *et al* (2009) An Architectural Role for a Nuclear Noncoding RNA: NEAT1 RNA Is Essential for the Structure of Paraspeckles. *Molecular Cell* **33** (6): 717 – 726

Cote-Sierra, J *et al* (2004) Interleukin 2 plays a central role in Th2 differentiation. *PNAS* **101** (11): 3880 – 3885

Dong, C. *et al* (2000) JNK is required for effector T-cell function but not for T-cell activation. *Nature* **405**: 91 – 94

Duitman, E.H. *et al* (2008) How a Cytokine Is Chaperoned through the Secretory Pathway by Complexing with Its Own Receptor: Lessons from Interleukin-15 (IL-15)/IL-15 Receptor α ▿ . *Molecular Cell Biology* **28** (15): 4851 – 4861

Eißmann, M. *et al* (2012) Loss of the abundant nuclear non-coding RNA *MALAT1* is compatible with life and development. *RNA Biology* **9** (8):1076 – 1087

Ernst, C. and Morton, C.C. (2013) Identification and function of long non-coding RNA. *Frontiers in Cellular Neuroscience* **7**: 168

Engreitz, J.M. *et al* (2014) RNA-RNA Interactions Enable Specific Targeting of Noncoding RNAs to Nascent Pre-mRNAs and Chromatin Sites. *Cell* **159** (1): 188 – 199

Evsyukova, I. *et al* (2012) Cleavage and polyadenylation specificity factor 1 (CPSF1) regulates alternative splicing of interleukin 7 receptor (IL7R) exon 6. *RNA* **19**: 103 – 115

Fagerberg, L. *et al* (2014) Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Molecular Cell Proteomics* **13** (2): 397 – 406

Gallo, K.A. and Johnson, G.L. (2002) Signalling: Mixed-lineage kinase control of JNK and p38 MAPK pathways. *Nature Reviews Molecular Cell Biology* **3**: 663 – 672

Gesbert, F. *et al* (2005) IL-2 Responsiveness of CD4 and CD8 lymphocytes: further investigations with human IL-2Rβ transgenic mice. *International Immunology* **17** (8): 1093 – 1102

Guo, F. *et al* (2010) Inhibition of metastasis-associated lung adenocarcinoma transcript 1 in CaSki human cervical cancer cells suppresses cell proliferation and invasion. *Acta Biochimica et Biophysica Sinica* **42**: 224 – 229

Gutschner, T. *et al* (2013) The non-coding RNA *MALAT1* is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Research* **73** (3): 1180 – 1189

Heo, J.B. and Sung, S. (2011) Vernalization-Mediated Epigenetic Silencing by a Long Intronic Noncoding RNA. *Science* **331** (6013): 76 – 79

Hedfors, I.A. and Brinchmann, J.E. (2003) Long-Term Proliferation and Survival of *In Vitro*-Activated T Cells is Dependent on Interleukin-2 Receptor Signalling but not on the High-Affinity IL-2R. *Scandinavian Journal of Immunology* **58**: 522 – 532

Hu, G. *et al* (2013) Expression and regulation of lincRNAs during T cell development and differentiation. *Nature Immunology* **14** (11): 1190 – 1198

Hutchinson, J.N. *et al* (2007) A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics* **8**: 39

Hrdlickova, B. *et al* (2014) Expression profiles of long non-coding RNAs located in autoimmune disease-associated regions reveal immune cell-type specificity. *Genome Medicine* **6** (10): 88

Ji, P. *et al* (2003) MALAT-1, a novel noncoding RNA, and thymosin bold italic beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene* **22**: 8031 – 8041

Joo, J-H. *et al* (2013) Transcriptomic Analysis of PNN- and ESRP1-Regulated Alternative Pre-mRNA Splicing in Human Corneal Epithelial Cells. *Investigative Ophthalmology and Optical Science* **54** (1): 697 – 707

Kanduri, K. *et al* (2015) Identification of global regulators of T-helper cell lineage specification. *Genome Medicine* **7**: 122

Kay, J.E. *et al* (1991) Mechanisms of T lymphocyte activation. *Immunology Letters* **29**: 51 – 54

Khan, M. and Gasser, S. (2016) Generating Primary Fibroblast Cultures from Mouse Ear and Tail Tissues. *Journal of Visualise Experiments* **107**: 53565

Knackmuss, U. *et al* (2016) MAP3K11 is a tumor suppressor targeted by the oncomiR miR-125b in early B cells. *Cell Death and Differentiation* **23**: 242 – 252

Kung, J.T.Y. *et al* (2013) Long Noncoding RNAs: Past, Present, and Future. *Genetics* **193** (3): 651 – 669

Kurts, C. *et al* (2007) Th17 cells: a third subset of CD4 + T effector cells involved in organ-specific autoimmunity. *Nephrology Dialysis Transplantation* **23** (3): 816 – 819

Lai, M-C. *et al* (2010) Long non-coding RNA MALAT-1 overexpression predicts tumor recurrence of hepatocellular carcinoma after liver transplantation. *Medical Oncology* **29** (3): 1810 – 1816

LeGendre, J.B. *et al* (2013) RNA Targets and Specificity of Staufen, a Double-stranded RNA-binding Protein in *Caenorhabditis elegans*. *Journal of Biological Chemistry* **288**: 2532 – 2545

Li, R. *et al* (2017) Functional dissection of NEAT1 using genome editing reveals substantial localization of the NEAT1_1 isoform outside paraspeckles. *RNA* **23** (6): 872 – 881

Liao, W. *et al* (2008) Priming for T helper type 2 differentiation by interleukin 2-mediated induction of IL-4 receptor α chain expression. *Nature Immunology* **9** (11): 1288 – 1296

Liao, W. *et al* (2011) Cytokine receptor modulation by interleukin-2 broadly regulates T helper cell lineage differentiation. *Nature Immunology* **12** (6): 551 – 559

Lim, P.S. *et al* (2016) Transcriptomic analysis of mouse EL4 T cells upon T cell activation and in response to protein synthesis inhibition via cycloheximide treatment. *Genomics Data* **7**: 148 – 151

Lin, Q. *et al* (2018) MALAT1 affects ovarian cancer cell behavior and patient survival. *Oncology Reports* **39** (6): 2644 – 2652

Lin, Y. *et al* (2018) Structural analyses of NEAT1 lncRNAs suggest long-range RNA interactions that may contribute to paraspeckle architecture. *Nucleic Acids Research* **46** (7): 3742 – 3752

Liu, H. *et al* (2018) HIV-1 replication in CD4+ T cells exploits the down-regulation of antiviral NEAT1 long non-coding RNAs following T cell activation. *Virology* **522**: 193 – 198

Livak, K.J. and Schmittgen, T.D. (2001) Analysis of Relative Gene Expression Data Using RealTime Quantitative PCR and the $2^{-\Delta\Delta Ct}$ Method. *Methods* **25**: 402 – 408

Luckheeram, R.V. *et al* (2012) $CD4^{+}$T Cells: Differentiation and Functions. *Clinical and Developmental Immunology*, Article ID 925135

Luo, S. *et al* (2016) Divergent lncRNAs Regulate Gene Expression and Lineage Differentiation in Pluripotent Cells. *Cell Stem Cell* **18** (5): 637 – 652

Ma, X.Y. *et al* (2015) Malat1 as an evolutionarily conserved lncRNA, plays a positive role in regulating proliferation and maintaining undifferentiated status of early-stage hematopoietic cells. *BMC Genomics* **16** (1): 676

Marshall, O.J. and Choo, K.H. (2012) Putative CENP-B paralogues are not present at mammalian centromeres. *Chromosoma* **121** (2): 169 – 179

Marzluff, W.F. *et al* (2002) The Human and Mouse Replication-Dependent Histone Genes. *Genomics* **80** (5): 487 – 498

Nakagawa, S. *et al* (2012) Malat1 is not an essential component of nuclear speckles in mice. *RNA* **18** (8): 1487 – 1499

Okeyo-Owuor, T. *et al* (2015) U2AF1 mutations alter sequence specificity of pre-mRNA binding and splicing. *Leukemia* **29**: 909 – 917

Pang, B. *et al* (2012) Differential pathways for calcium influx activated by concanavalin A and CD3 stimulation in Jurkat T cells. *European Journal of Physiology* **463** (2): 309 – 318

Peterlin, B.M. *et al* (2012) 7SK snRNA: a noncoding RNA that plays a major role in regulating eukaryotic transcription. *Wiley Interdisciplinary Reviews RNA* **3** (1): 92 – 103

Pham, C.T.N. *et al* (1996) Long-range disruption of gene expression by a selectable marker cassette. *PNAS* **93** (23): 13090 – 13095

Rengasamy, M. *et al* (2017) The PRMT5/WDR77 complex regulates alternative splicing through ZNF326 in breast cancer. *Nucleic Acids Research* **45** (19): 11106 – 11120

Rin, J.L. *et al* (2007) Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Non-Coding RNAs. *Cell* **129** (7): 1311 – 1323

Russ, B.E. *et al* (2013) T cell immunity as a tool for studying epigenetic regulation of cellular differentiation. *Frontiers in Genetics* **4**: 218

Sado, T. *et al* (2006) *Tsix* defective in splicing is competent to establish *Xist* silencing. *Development* **133**: 4925 – 4931

Schmidt, W. *et al* (2007) Mutation in the *Scyl1* gene encoding amino-terminal kinase-like protein causes a recessive form of spinocerebellar neurodegeneration. *EMBO Reports* **8** (7): 691 – 697

Spurlock, C.F. *et al* (2017) Profiles of long noncoding RNAs in human naïve and memory T cells. *Journal of Immunology* **199** (2): 547 – 558

Shi, M. *et al* (2008) Jak3-dependent signals induce chromatin remodeling at the *Ifng* locus during Th1 differentiation. *Immunity* **28** (6): 763 – 773

Stubbington, M.J.T. *et al* (2015) An atlas of mouse $CD4^+$ T cell transcriptomes. *Biology Direct* **10** (14): s13062-015-0045-x

Tang, Q-M. *et al* (2014) Fetal and adult fibroblasts display intrinsic differences in tendon tissue engineering and regeneration. *Scientific Reports* **4**: 5515

Tripathi, V. *et al* (2010) The Nuclear-Retained Noncoding RNA MALAT1 Regulates Alternative Splicing by Modulating SR Splicing Factor Phosphorylation. *Molecular Cell* **39** (6): 925 – 938

Tripathi, V. *et al* (2013) Long Noncoding RNA MALAT1 Controls Cell Cycle Progression by Regulating the Expression of Oncogenic Transcription Factor B-MYB. *PLoS Genetics* **9** (3): e1003368

Vahedi, G. *et al* (2012) STATs Shape the Active Enhancer Landscape of T Cell Populations. *Cell* **151** (5): 981 – 993

Vance, K.W. and Ponting, C.P. (2014) Transcriptional regulatory functions of nuclear long noncoding RNAs. *Trends in Genetics* **30** (8): 348 – 355

Wang, J. *et al* (2014) MALAT1 promotes cell proliferation in gastric cancer by recruiting SF2/ASF. *Biomedicine and Pharmacotherapy* **68** (5): 557 – 564

West, J.A. *et al* (2014) The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Molecular Cell* **55** (5): 791 – 802

Whitworth, H. *et al* (2012) Identification of Kinases Regulating Prostate Cancer Cell Growth Using an RNAi Phenotypic Screen. *PLoS One* **7** (6): e38950

Whisenant, T.C. *et al* (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

Williams, A.B. and Schumacher, B. (2016) p53 in the DNA-Damage-Repair Process. *Cold Spring Harbour Perspectives in Medicine* **6** (5): a026070

Wilusz, J.E. *et al* (2008) 3’ end processing of a long nuclear-retained non-coding RNA yields a tRNA-like cytoplasmic RNA. *Cell* **135** (5): 919 – 932

Xia, F. *et al* (2014) Dynamic Transcription of Long Non-Coding RNA Genes during CD4+ T Cell Development and Activation. *PLOS ONE* **9** (7): e101588

Xu, C. *et al* (2010) MALAT-1: A long non-coding RNA and its important 3' end functional motif in colorectal cancer metastasis. *International Journal of Oncology* **39**: 169 – 175

Yan, P. *et al* (2017) Cis- and trans- acting lncRNAs in pluripotency and reprogramming. *Current Opinion in Genetics and Development* **46**: 170 – 178

Yang, D.D. *et al* (1998) Differentiation of CD4+ T Cells to Th1 Cells Requires MAP Kinase JNK2. *Immunity* **9** (4): 575 – 585

Yang, M-H. *et al* (2015) MALAT1 promotes colorectal cancer cell proliferation/migration/invasion via PRKA kinase anchor protein 9. *Biochimica et Biophysica Acta* **1852**: 166 – 174

Yue, F. *et al* (2014) A comparative encyclopedia of DNA elements in the mouse genome. *Nature* **515** (7527): 355 – 364

Zhang, B. *et al* (2012) The lncRNA *Malat1* is dispensable for mouse development but its transcription plays a *cis*-regulatory role in the adult. *Cell Reports* **2** (1): 111 – 123

Zhao, Z. *et al* (2014) 17β-Estradiol treatment inhibits breast cell proliferation, migration and invasion by decreasing MALAT-1 RNA level. *Biochemical and Biophysical Research Communications* **445** (2): 388 – 393

Zhou, X. *et al* (2015) Long Non Coding RNA MALAT1 Promotes Tumor Growth and Metastasis by inducing Epithelial-Mesenchymal Transition in Oral Squamous Cell Carcinoma. *Scientific Reports* **5**: 15972

Zhu, J. *et al* (2010) Differentiation of Effector CD4 T Cell Populations. *Annual Review of Immunology* **28**: 445 – 489

Zong, X. *et al* (2016) Natural antisense RNA promotes 3' end processing and maturation of MALAT1 lncRNA. *Nucleic Acids Research* **44** (6): 2898 – 2908