



The  
University  
Of  
Sheffield.

# **Proteogenomic Identification of Membrane Trafficking Components affecting Antibody Secretion in Plasma Cells**

**Nabila Farah Rahman**

A thesis submitted in partial fulfilment of the requirements for the degree of  
Doctor of Philosophy in Biomedical Science

The University of Sheffield  
Department of Biomedical Science

November 2019

## ACKNOWLEDGEMENT

Firstly, I would like to thank Dr Andrew Peden for giving me the opportunity to work on this project. These past 4 years have been a difficult time with deteriorating health interrupting a large part of my Doctoral research. Therefore, I would like to convey my deepest gratitude to him for his understanding and patience regarding this unavoidable issue and for being a fantastic supervisor in his help, guidance and support.

I wish to thank my PhD advisors, Dr Kai Erdmann and Dr Marta Milo, for their advice and critiques regarding the direction of my project. I also thank Dr Mark Collins for helping me get a better grasp of the proteomics analysis related to this project. I'd like to acknowledge Dr Paul R. Heath for his help with various transcriptomics related queries.

I wish to thank my colleagues, Dr Elanchezian Rajan and Dr Asral Wirda Binti Ahmad Aswani, for performing wet lab validations, that were vital for this project and teaching me various wet lab techniques. To all the members of Peden lab, especially Wirda, Amber SM Shun Shion and Jessica A. Willmott, I thank you for being such fun people and making my Doctoral research an enjoyable experience. I would also like to thank the members of Kai Lab for their friendship and company.

I am grateful to the Department of Biomedical Science for funding my studentship, and Medimmune for providing the equipment I needed to carry out my research.

I wish to thank all my loved ones, my parents and my brother for their support throughout my PhD. In particular, I would like to convey my thanks to my amazing mother for rushing to England for my care. Lastly, I would like to express my gratitude to my wonderful partner for going above and beyond in caring for me through very tough times.

# CONTENTS

ACKNOWLEDGEMENT.....	i
LIST OF TABLES.....	iv
LIST OF FIGURES.....	v
LIST OF ABBREVIATIONS .....	viii
ABSTRACT.....	1
1 INTRODUCTION.....	2
1.1 The Endomembrane System and Vesicular Trafficking .....	2
1.1.1 Protein processing in the endoplasmic reticulum .....	2
1.1.2 Delivery of newly synthesized proteins to the Golgi apparatus.....	5
1.1.3 Retrieval of ER resident proteins .....	5
1.1.4 Post-Golgi transport of post translationally modified proteins .....	6
1.1.5 Vesicle tethering .....	7
1.1.6 SNAREs .....	10
1.2 Plasma cells – A Model for Studying Secretion.....	14
1.2.1 Plasma Cell Differentiation .....	14
1.2.2 Structural and Physiological changes during Plasma Cell Differentiation .....	18
1.3 HYPOTHESIS.....	18
1.4 Research Relevance.....	19
1.4.1 A Proteogenomic Atlas of Antibody Secreting Cells .....	19
1.4.2 Protein Secretion in the Plasma Cell Physiology.....	23
1.4.3 Cell Markers for Antibody Secreting Cells .....	23
1.5 Medical Relevance .....	24
1.6 Industrial Relevance .....	24
1.6.1 Production of Recombinant Proteins .....	24
2 PRELIMINARY TRANSCRIPTOMIC ANALYSIS .....	28
2.1 BACKGROUND .....	28
2.1.1 HYPOTHESIS .....	28
2.1.2 High Throughput Transcriptomics .....	28
2.1.3 Leveraging Publicly Available Transcriptomes of ASCs.....	32
2.1.4 AIMS.....	33
2.2 METHODS.....	34
2.2.1 Data Source.....	34
2.2.2 Differential Expression.....	36
2.2.3 Meta-Analysis of Cross-Platform Data.....	36
2.2.4 Ranking Differentially Expressed Hits .....	37
2.2.5 Functional Annotation .....	37
2.3 RESULTS.....	38
2.4 DISCUSSION .....	41
2.4.1 Isolating Known and Novel Targets .....	41
2.4.2 Evaluation of Methods.....	42
2.5 CONCLUSION .....	44
3 CROSS SPECIES MICROARRAY ANALYSIS.....	47
3.1 BACKGROUND .....	47
3.1.1 Of Mice and Men .....	47
3.1.2 Review of Microarray Normalization.....	48
3.1.3 Challenges of Meta-Analysis in Microarrays.....	53

3.1.4	Reannotation .....	55
3.1.5	Functional Analysis .....	56
3.1.6	AIMS & OBJECTIVES .....	59
3.2	METHODS .....	60
3.2.1	Workflow.....	60
3.2.2	Mouse B Cell Lineage .....	61
3.2.3	Human B cell Lineage.....	69
3.2.4	Cross Species Meta-Analysis .....	71
3.2.5	Functional Analysis .....	72
3.3	RESULTS.....	80
3.3.1	Transcript Classification .....	80
3.3.2	Differential Expression.....	83
3.3.3	Verification.....	87
3.3.4	Functional Analysis .....	89
3.3.5	Pathway Analysis .....	94
3.4	DISCUSSION .....	98
3.4.1	EDEM1-ERdj5 complex shows differential regulation in human.....	98
3.4.2	UGGT2 folding checkpoint enzyme is upregulated in human ASCs only .....	99
3.4.3	CREB3L2 may be a cargo selector for COPII vesicles .....	101
3.4.4	FNDC3B and TMEM184B may be potential biomarkers for ASCs .....	104
3.5	CONCLUSION .....	105
4	MULTI-OMICS ANALYSIS OF ASCs .....	107
4.1	BACKGROUND .....	107
4.1.1	RNA-Sequencing.....	107
4.1.2	RNA-Seq Data Analysis.....	111
4.1.3	Proteomics .....	115
4.1.4	Investigation of Membrane Trafficking Components.....	120
4.1.5	Cell Markers for Antibody Secreting Cells .....	120
4.1.6	Data visualisation.....	121
4.1.7	AIMS & OBJECTIVES .....	122
4.2	METHODs .....	123
4.2.1	Workflow.....	123
4.2.2	RNA-Seq .....	124
4.2.3	Proteomics .....	134
4.2.4	Low Throughput Validation .....	140
4.3	RESULTS.....	141
4.3.1	RNA-Sequencing.....	141
4.3.2	Proteomics .....	144
4.3.3	Proteogenomics .....	144
4.3.4	Verification.....	146
4.3.5	Known and Novel Signature Genes for ASCs .....	148
4.3.6	Functional Analysis .....	151
4.3.7	Pathway Analysis .....	152
4.4	DISCUSSION .....	154
4.4.1	EDEM1-ERdj5 complex does not show diverging regulation .....	154
4.4.2	UGGT2 may be specifically upregulated in human ASCs.....	155

4.4.3	Novel Markers of ASCs.....	156
4.4.4	Review of Genes identified in Preliminary study.....	163
4.4.5	RRBP1 – Ribosome Anchor .....	164
4.4.6	Regulation of Membrane Trafficking Components in ASCs.....	171
4.5	CONCLUSION .....	187
4.6	EVALUATION.....	189
4.7	FUTURE DIRECTIONS .....	192
5	REFERENCES.....	194
6	APPENDIX.....	208

## LIST OF TABLES

Table 2-1	Phenotype of Samples utilised for Cross-Study Meta-Analysis .....	35
Table 3-1	Phenotype of Mouse microarray profiles of PC cell Lineage from different studies .....	63
Table 3-2	Experimental Design - MG430.2A.....	67
Table 3-3	Phenotype of human microarray profiles of PC cell Lineage - HG133A (Jourdan <i>et al</i> ) .....	69
Table 3-4	Global summary of genes differentially expression in mouse and human ASCs..	87
Table 3-5	Cargo Proteins co-expressed with CREB3L2 .....	102
Table 4-1	Phenotype of mouse RNA-Seq profiles of PC cell Lineage – Illumina HiSeq 2500 (GSE60927).....	124
Table 4-2	Phenotype of human RNA-Seq profiles of PC cell Lineage – Illumina HiSeq 2500 (GSE81443).....	125
Table 6-1	Manually Curated Summary of Overlapping GO Biological Process Terms.....	208
Table 6-2	Manually Curated Summary of Overlapping GO Cellular Component Terms ....	209
Table 6-3	GO Biological Process enrichment of Upregulated Genes in ASCs across species (Microarray). .....	210
Table 6-4	GO Biological Process enrichment of Downregulated Genes in ASCs across species (Microarray) .....	211
Table 6-5	GO Cellular Component enrichment of Upregulated Genes in ASCs across species (Microarray).....	212
Table 6-6	GO Cellular Component enrichment of Downregulated Genes in ASCs across species (Microarray) .....	213
Table 6-7	Differentially upregulated genes that are known/ predicted to localise to ER, Golgi, proteasome and secretory granules. Includes genes enriched for ER stress response, anti apoptotic processes, membrane trafficking, glycosylation, and antigen presentation as indicated by GO enrichment analysis. ....	213
Table 6-8	Tables shows differentially upregulated genes in ASCs that are known to be upregulated as a result of TF overexpression/induction/gain of function mutation....	214
Table 6-9	TF co-expression results for genes differentially upregulated genes in ASCs (cont.).....	215
Table 6-10	Tables shows differentially upregulated genes in ASCs that are known to be upregulated as a result of TF inhibition/downregulation/loss of function mutation (cont.).....	218

Table 6-11   Novel components uniquely upregulated in ASCs identified by proteogenomic analysis. (Cont.).....	225
Table 6-12   Novel components uniquely downregulated in ASCs identified by proteogenomic analysis. (cont.) .....	226
Table 6-13   Upregulated Transcription factors in ASCs uniquely identified by our multi-omics analysis. Genes soft validated by proteomics are given in bold .....	227
Table 6-14   Downregulated Transcription factors in ASCs uniquely identified by our multi-omics analysis. Genes soft validated by proteomics are given in bold .....	227
Table 6-15   Contradictory regulation. Genes upregulated in mice ASCs but downregulated in humans ASCs across both RNA-Seq and microarray data .....	228
Table 6-16   Contradictory regulation. Genes downregulated in mice ASCs but upregulated in humans ASCs across both RNA-Seq and microarray data .....	228
Table 6-17   Cluster Differentiation markers upregulated in ASCs vs. NBCs across species in RNA-Seq and/or microarray data. Ordered by FDR adjusted <i>p</i> -value.....	228
Table 6-18   Upregulated proteins in the proteome of CD138+ plasmablast CD93+ plasmablast or both (cont.) .....	229
Table 6-19   Downregulated proteins in the proteome of CD138+ PB, CD93+ PB or both (cont.).....	231
Table 6-20   Genes/proteins uniquely upregulated in CD93+ isolated but not in CD138+ isolated cells.....	234
Table 6-21   Genes/proteins uniquely downregulated in CD93+ isolated but not in CD138+ isolated cells.....	234
Table 6-22   Genes/proteins consistently regulated in all three platforms studied amongst RNA-Sequencing, Microarray, Label Free Proteomics (cont.) .....	235
Table 6-23   Genes/proteins consistently regulated in two out of three platforms studied amongst RNA-Sequencing, Microarray, Label Free Proteomics (cont.) .....	236
Table 6-24   Genes upregulated in proteome data, but show inconsistencies in transcriptome. (cont.).....	238
Table 6-25   Genes downregulated in proteome data, but show inconsistencies in transcriptome. (cont.).....	239
Table 6-26   Genes predicted to be co-regulated with CREB3 and CREB3L2.....	241
Table 6-27   Top 500 Differentially regulated Genes/Proteins across species and platforms. Sorted by median adjusted <i>p</i> -value for multiple group comparison. (cont.).....	242

## LIST OF FIGURES

Fig 1-1   Coordination of Tethers and Coated vesicles in membrane trafficking. ....	7
Fig 1-2   Schematic diagram of structural and functional classifications of SNAREs. ....	10
Fig 1-3   Overview of SNARE complex localisation in the cell. (Adapted from Wang <i>et al</i> )....	11
Fig 1-4   Plasma cell Differentiation from Follicular B cells.....	14
Fig 1-5   Electron micrograph of A. Naïve B cell and B. Plasma Cell. ....	17
Fig 1-6   Microarray Signal Intensities in various stages of plasma cell differentiation. ....	19
Fig 1-7   Over 25 year of genetic integration of advantageous genes in CHO cells .....	25
Fig 2-1   Generic RNA detection method by microarrays. ....	29
Fig 2-2   A. Workflow of Metanalysis. B. Workflow of Merging. ....	32
Fig 2-3   Workflow of Preliminary Merging and Meta-analysis of microarray data. ....	34

Fig 2-4   Principal Component Analysis of raw and batch adjusted cross-study mouse microarray data.....	36
Fig 2-5   Venn diagram and Pie charts showing functional grouping of DEGs isolated in the preliminary study. ....	39
Fig 2-6   Heatmap of potential genes of interest found in the preliminary study.....	40
Fig 2-7   Boxplot of pre-normalised data and batch adjusted mouse microarray data.....	43
Fig 3-1   Density distribution curve of Log2 transformed probe intensities.....	50
Fig 3-2   Principle of quantile normalisation. ....	51
Fig 3-3   Principle of median polish. ....	52
Fig 3-4   Workflow of Methodology used for Cross-Species Meta-Analysis of Microarray Data.....	60
Fig 3-5   Box plot probe intensities in pooled vs separately normalised cross-study mouse arrays. ....	64
Fig 3-6   PCA plots indicating sample clustering as a result of MAS5, RMA and GCRMA normalisation in cross -study, mouse microarray data. ....	65
Fig 3-7   Boxplot of probe intensities in raw and GCRMA normalised human microarray data. ....	70
Fig 3-8   PCA Plot showing sample clustering in raw versus GCRMA normalised human microarray data.....	71
Fig 3-9   Example GO Ontology summarisation for GO Biological Processes generated from the results of EnrichR mining tool. ....	76
Fig 3-10   Example Visualisation of Pathway Enrichment results from EnrichR mining tool..	79
Fig 3-11   Proportion differentially expressed coding and non-coding transcripts in mice arrays, human arrays and in cross species analysis.....	82
Fig 3-12   Box plot of UTR lengths in A. Mice microarray and B. Human Microarray. Only UTRs differentially expressed in at least 3 out of 5 ASCs studied were considered. ....	83
Fig 3-13   Venn Diagram showing the overlap of differentially regulated genes in ASCs of mice and human microarray profile. ....	85
Fig 3-14   Visualisation of the number of DEG filtered out via cross-species microarray analysis.....	86
Fig 3-15   Differential regulation of transcription factors characteristic of naïve B cell in cross-species microarray analysis.....	88
Fig 3-16   Differential regulation of transcription factors characteristic of antibody secreting cells in cross-species microarray analysis. ....	89
Fig 3-17   Degree of consistency in the Differential Regulation of Functionally Enriched Genes and Visualisation of Summarised GO Biological Processes. ....	90
Fig 3-18   Enrichment Terms related to Unfolded Protein Response in cross-species microarray analysis. ....	91
Fig 3-19   Enrichment Terms overrepresented for Membrane Trafficking.....	92
Fig 3-20   Summary of GO cellular components enriched among A. Upregulated Genes, B. Downregulated Genes. ....	93
Fig 3-21   Bar charts showing potential positive regulators of genes upregulated in ASCs. ..	96
Fig 3-22   A. Bar charts showing potential negative regulators of genes upregulated in ASCs .....	97
Fig 3-23   Differential Expression profile of ER chaperones in the ASCs of mice and human compared to NBCs. ....	98
Fig 3-24   Expression profile of UGGT genes (microarray).....	100

Fig 3-25   Differential Expression of COPII vesicles components (microarray).....	101
Fig 3-26   Differential regulation of cargo proteins in ASCs vs NBCs (microarray). ....	102
Fig 3-27   Predicted Gene regulatory network for large protein transport in the ER.....	103
Fig 3-28   Differential Expression of RRBP1 and FOXO1 relative to CREB3L2 in ASCs vs NBCs. .....	104
Fig 4-1   Overview of Illumina Sequencing (Kircher & Kelso 2010).....	108
Fig 4-2   Schematic of RNA-Seq Analysis Workflow. ....	112
Fig 4-3   A Schematic Representation of Proteome Quantification.....	117
Fig 4-4   Workflow of multi-omics Analysis.....	123
Fig 4-5   Mean Variance Trend Plot depicting limma-voom normalisation and low count filtering efficacy. ....	127
Fig 4-6     Density Distribution Plot of RNA-Seq Counts Before and After Low-Count Filtering. .....	128
Fig 4-7   Box Plot of RNA-Seq Counts before and after TMM normalisation. ....	129
Fig 4-8   Schematic of our RNA-Seq Pipeline equipped with defensive programming and staged processes.....	132
Fig 4-9   Western Blot of resting B cells, CD138 purified plasmablasts and CD93 purified cells .....	135
Fig 4-10   Screenshots of input fields in the PlasmacytOMICs Web Application.....	137
Fig 4-11   Screenshot of an output from PlasmacytOMICs web application .....	139
Fig 4-12   Venn Diagram of RNA-Seq results showing overlap between different ASCs across species.....	141
Fig 4-13   Venn diagram of genes showing cross platform reproducibility in diverging gene regulation across species.....	144
Fig 4-14   Venn diagram of showing overlap of differentially expressed proteins in CD138 and CD93 purified cells. ....	144
Fig 4-15   Venn diagram showing consistency in differential expression across RNA-Seq, microarray and MS/MS.....	145
Fig 4-16   Differential Regulation of well characterised ASC and NBC markers according to multi-omics analysis.....	147
Fig 4-17   Venn diagram showing overlap of reproducible gene/proteins isolated by multi- omics versus those identified by Shi <i>et al.</i> .....	148
Fig 4-18   Transcription factors uniquely isolated by multi-omics analysis. ....	150
Fig 4-19   Summarised GO Enrichment Results for multi-omics analysis. ....	151
Fig 4-20   Transcription factors predicted to co-regulate with genes/proteins upregulated in ASCs according to multi-omics analysis. ....	153
Fig 4-21   Differential regulation of EDEM1-ERDJ5 complex according to multi-omics analysis.....	154
Fig 4-22   Differential Regulation of UGGT enzymes according to multi-omics analysis.....	155
Fig 4-23   Top 10 CD markers upregulated in ASCs according to multi-omics analysis. ....	156
Fig 4-24   A. Proteogenomic regulation of potential markers of ASCs. ....	158
Fig 4-25   A. Proteogenomic regulation of components related to regulated secretory pathway .....	163
Fig 4-26   A. Proteogenomic regulation of RRBP1.....	164
Fig 4-27   A-B. Proteogenomics regulation of ribosomal subunits in ASCs.....	167
Fig 4-28   A. CREB3L2 gene regulation across platforms and species.....	168

Fig 4-29   Differential Regulation of COPII vesicle components according to multi-omics analysis.....	169
Fig 4-30   Differential Regulation of CREB3-like isoforms (multi-omics). ....	171
Fig 4-31   Differential Regulation of COPI coat proteins (multi-omics). ....	172
Fig 4-32   Differential regulation of clathrin coated vesicles (multi-omics).....	174
Fig 4-33   Differential regulation of post-Golgi tethers (multi-omics). ....	177
Fig 4-34   Proteogenomic regulation of components of ER-Golgi tethers.....	178
Fig 4-35   Differential regulation of USO1 tethering factor and its interactors at the <i>cis</i> -Golgi. ....	179
Fig 4-36   A. Differential Regulation of NRZ subunits in ASCs (multi-omics). ....	180
Fig 4-37   Schematic diagram of TANGO1 interaction with NRZ complex, COPII coat proteins and the cargo, collagen.....	181
Fig 4-38   Differential Regulation of SNAREs (multi-omics). ....	183
Fig 4-39   Differential Regulation of SNARE complex components known to play a role in exocytosis in neurons. ....	185
Fig 4-40   Summary of regulation and coordination of coated vesicles, tethers and SNAREs in ASCs according to multi-omics analysis.....	186

## LIST OF ABBREVIATIONS

Ab	Antibody
ABC	Activated B cell
AP	Adaptor Protein
API	Application Programming Interface
ASC	Antibody Secreting Cell
BMPC	Bone Marrow Plasma Cell
CALR	Calreticulin
CANX	Calnexin
CD marker	Cluster Differentiation marker
CHO	Chinese Hamster Ovary
CPM	Counts per Million
DE	Differential Expression/Expressed
DEG	Differentially Expression Genes
dNTP	deoxyribose nucleotide triphosphates
ECM	Extracellular Matrix
EDEM	ER-degradation enhancing mannosidase-like protein
ER	Endoplasmic Reticulum
ERAD	ER associated degradation
ERGIC	ER-Golgi intermediate
ESI	Electrospray ioniser
FC	Fold Change
FDR	False Discovery Rate
GCRMA	guanine-cytosine Robust Microarray Average
GEO	Gene Expression Omnibus

GO	Gene Ontology
HG133A	Human Genome U133 Plus 2.0
LC	liquid chromatographer / chromatography
LFQ	Label Free Quantification
LPS	Lipopolysaccharide
mAb	monoclonal Antibody
MBC	Memory B cells
MG430.2A	Mouse Genome 430 2.0 Array
MM	Mismatch (probes)
mPB	mouse plasmablasts, typically isolated using CD138
mPB93	mouse plasmablast isolated using CD93
MS-MS	Tandem Mass Spectrometry
MV plot	Mean Variance trend plot
NBC	Naïve B Cell
NGS	Next Generation Sequencing
Non coding RNA	Non-Coding Ribonucleic Acid
PB	Plasma Blast
PC	Plasma Cell
PCA	Principle Component Analysis
PM	Perfect Match (probes)
prePB	pre plasmablasts
rER	Rough Endoplasmic Reticulum
RMA	Robust Microarray Average
RNA-Seq	RNA-Sequencing
snoRNA	small nucleolar Ribonucleic Acid
SpIPC	Splenic Plasma Cell
SRA	Sequence Read Archive
TF	Transcription Factor
TGN	<i>trans</i> -Golgi Network
TMM	Trimmed Mean of M value
TSC	<i>trans</i> -SNARE Complex
UGGT	UDP-glucose: glycoprotein glucosyltransferase
UPR	Unfolded Protein Response
UTR	Untranslated Region

## ABSTRACT

Plasma cells are professional antibody secreting cells unlike their naïve B cell (NBC) precursors. By exploiting the remarkable contrast in their secretory capacity, we set out to identify novel factors involved in antibody secretion and PC physiology using a multiplatform, cross-species proteogenomics approach. Using this methodology, we have reproducibly identified a large number of genes which were consistently upregulated in antibody secreting cells (ASCs). As expected, genes involved in protein folding and membrane trafficking are significantly enriched in our data set thus validating this method. As thousands of genes were differentially regulated in PCs, we generated a web based bioresource to aid in the visualisation of our data. Using this resource, we have investigated the regulation of several genes families during plasma cell differentiation including transcription factors, coat proteins, tethers and SNAREs. Our analysis suggests that ASCs specifically upregulate vesicle coats and tethers acting in the early but not the late secretory pathway. Interestingly, several genes implicated in collagen secretion are also significantly upregulated in ASCs suggesting that they may have a more general role in secretion than previously thought. For example, we have identified that NBAS a component of the ER localised tethering complex is upregulated in ASC which may explain why patients with mutations in this gene have defects in collagen secretion and hypogammaglobulinemia. Using the web-based resource, we have also identified a large number of poorly characterised genes which are significantly upregulated in ASCs so potentially having a role in antibody folding and/or trafficking (CRELD2, TMEM214 and HID1). Finally, this web-based resource will be useful for those aiming to identify novel biomarkers for plasma cells and factors which can be manipulated to enhance secretory capacity.

# 1 INTRODUCTION

## 1.1 THE ENDOMEMBRANE SYSTEM AND VESICULAR TRAFFICKING

Cell membranes are essential for the existence of living cells. Membranes are generally composed of lipid bilayers with a mosaic of embedded proteins that enclose cell contents in order to distinguish them from their external environment. In eukaryotic cells, membranes form a barrier between cytosol and intracellular organelles. This allows specialized compartments such as the nucleus, endoplasmic reticulum, Golgi apparatus, vesicles and so on to maintain their characteristic composition [1]. Eukaryotic cells employ vesicular trafficking to deliver newly synthesized lipids and proteins to their destination. Typically, proteins synthesized and folded in the Endoplasmic Reticulum (ER) are transported to the Golgi, where they undergo post-translational modifications (such as glycosylation, sulfation, etc) and packaged into transport carriers which bud from the *trans* Golgi network (TGN) and fuse with the destined target to deliver their protein contents [2].

### 1.1.1 Protein processing in the endoplasmic reticulum

Approximately 30% of newly synthesised proteins are incorrectly folded in mammalian cells [3]. Proper removal of these proteins is a necessity for cell survival as misfolded proteins can disrupt ER function by aggregating via exposed hydrophobic residues or compete with substrate binding of correctly folded proteins [4]. The unfolded protein response (UPR) acts to remove these proteins from the ER and restore ER homeostasis. PERK, ATF6 and IRE1 are ER transmembrane proteins that act as sensors of ER stress [5]. The PERK mediated pathway inhibits protein synthesis and reduces the generation of reactive oxygen species to relieve ER stress. ATF6 operates by upregulating the transcription of chaperones and protein folding enzymes and IRE1 performs the dual function of upregulating chaperones and protein folding

enzymes while also mediating degradation of terminally misfolded proteins [5]. If homeostasis cannot be achieved these pathways can commit to ER stress induced cell death [6]. In the following section we discuss key players in protein processing in the ER.

### 1.1.1.1 ER Chaperones mediating protein folding and degradation

#### 1.1.1.1.1 **CANX/CALR lectins**

The folding of nascent proteins in the ER is governed by lectin chaperones, calnexin (CANX) and calreticulin (CALR), and the protein disulphide isomerase, ERP57 (PDIA3) [7–9]. Misfolded proteins are recognised by the BiP chaperone (HSPA5), which facilitates refolding or entry into the proteasomal degradation pathway.

#### 1.1.1.1.2 **BiP**

The BiP chaperone (HSPA5) is a member of 70 KDa heat shock protein (Hsp70) family and acts as a master regulator of protein synthesis. When proteins are misfolded, BiP is able to bind to their characteristic exposed hydrophobic residues and thus prevent misfolded protein aggregation in the ER. This in turn allows these proteins to re-enter the CANX/CALR cycle and undergo further folding. If re-entry is delayed due to pressure from high levels of protein synthesis, or if the proteins are terminally misfolded, BiP promotes entry of these proteins into the ER associated degradation (ERAD) machinery through a complex mechanism [10, 11]. Degradation of soluble substrates and transmembrane proteins with luminal defects is dependent on BiP, whereas those with cytosolic defects enter the ERAD pathway in a BiP independent manner [12, 13].

#### 1.1.1.1.3 **EDEM**

EDEM (ER-degradation enhancing mannosidase-like protein) chaperones recognises terminally misfolded substrates from the CANX/CALR cycle and target them to the ERAD machinery [14–16]. Repeated demannosylation governs the degradation of misfolded glycoproteins. EDEM induces misfolded protein degradation by inhibiting the proteolysis of

the mannosidase, ERman1, and reglycosylation of its substrate, i.e. misfolded proteins [17, 18]. EDEM then targets these demannosylated proteins to the ER associated ubiquitin ligase complex via interaction with lectin chaperones, OS9 and XTP-3B (ERLEC1), and subsequent retrotranslocation via interaction with the translocon candidate, Derlin1 [19, 20].

#### 1.1.1.1.4 ERdj5

Aberrant proteins have to be unfolded via reduction of disulphide bonds before they can be retrotranslocated to the proteasome for degradation. The 40 kDa heat shock protein, ERdj5 (DNAJC10), is a luminal protein that acts as a BiP co-chaperone. It reportedly forms a complex with EDEM1 and reduces the disulphide bond of EDEM1 substrates to allow retrotranslocation of misfolded proteins via BiP [21].

#### 1.1.1.2 Protein Glycosylation in the ER

The majority of newly synthesized polypeptides are N-glycosylated with a high mannose containing core glycan group (Glc<sub>3</sub>GlcNAc<sub>2</sub>Man<sub>9</sub>) on entry into the ER [22]. Subsequent deglycosylation of 2 glucose residues allows nascent proteins to enter the CANX-CALR folding cycle [7–9]. Complete deglycosylation by  $\alpha$ -glucosidase II (GANAB/PRKCSH), allows proteins to dissociate from the lectin chaperones and is typically done after folding process is complete. If resultant proteins are misfolded, the BiP chaperones bind to exposed hydrophobic residues and prevent their aggregation. Subsequently, re-glycosylation of these proteins by UDP-glucose: glycoprotein glucosyltransferase (UGGT1 / UGGT2) allows them to re-bind to the lectin chaperones and undergo further rounds of folding[22].

If proteins have been correctly folded, ERMan1 mannosidases (MAN1A1, MAN1A1, MAN1B1, MAN1C1) demannosylate their N-glycan to produce GlcNAc<sub>2</sub>Man<sub>8</sub> isomers, which prevents any further re-glycosylation and re-entry into the folding machinery so that the proteins can then exit the ER. On the other hand, EDEM bound terminally misfolded proteins

undergo further rounds of demannosylation by ErMan1 complex and enter the ERAD pathway, where proteins are shunted to the proteasome for proteolytic degradation [22].

### 1.1.2 Delivery of newly synthesized proteins to the Golgi apparatus

Newly synthesized proteins that pass the folding quality check at the ER are typically packaged into vesicles and delivered to the ER-Golgi intermediate complex (ERGIC) and/or the *cis*-Golgi. The well characterised COPII vesicle is thought to mediate this process [1]. This vesicle consists of a cytosolic small GTPase, called SAR1. In its GTP bound form SAR1 anchors to ER exit sites and recruits the heterodimeric inner coat proteins, SEC23-SEC24 via specific interaction with SEC23. SEC23 can later promote GTP hydrolysis, to promote SAR1, and thus COPII complex, disassociation from the donor membrane after it has been fully assembled [23].

The SEC24 subunit is thought to play a key role in cargo sorting by specifically binding to export signals on cargo proteins and drawing them to the ER exit site. Some cargo are known to be loaded into COPII independently of SEC24, i.e. via transport adaptors or simple diffusion [24]. SAR1 and SEC23 are thought to interact with the outer COPII heterotetrameric coat proteins, SEC13-SEC31 and complete the COPII complex. The outer COPII coats of fully assembled COPII complexes are able to polymerise and give shape to the newly forming COPII vesicle. The resultant extruded portion of the ER membrane is then able to bud off and travel to the ER-Golgi intermediate complex (ERGIC) or the *cis*-Golgi [23].

### 1.1.3 Retrieval of ER resident proteins

After delivering cargo to Golgi apparatus, ER resident and escaped misfolded proteins have to be trafficked back to maintain the volume and integrity of the ER and prevent cell surface expression of non-natively folded proteins. COPI-coated vesicles are implicated in this process [1]. However, this vesicle not only functions in the replenishment of ER membrane components but also operates in intra-Golgi membrane trafficking for the correct localisation of Golgi

resident glycosyltransferases. COPI has also been proposed to operate in endosomal trafficking [25].

COPI is made up of 7 coat proteins,  $\alpha$ -COP (COPA),  $\beta$ -COP (COPB1),  $\beta'$ -COP (COPB2),  $\gamma$ -COP (COPG),  $\delta$ -COP (ARCN1),  $\epsilon$ -COP (COPE) and  $\zeta$ -COP (COPZ) [25]. At the Golgi-apparatus, cytosolic small GTPase, ARF1, can anchor to the Golgi membrane in its GTP bound form that then likely co-recruits  $\gamma$ -COP along with the ARF activating protein, ARFGAP2 or ARFGAP3. These ARFGAPS can sense membrane curvature and function to promote ARF1-GTP hydrolysis and subsequent disassembly of the COPI coat complex from the membrane of a fully formed vesicle [25].  $\alpha$  and  $\beta'$  COP subunit have been implicated in cargo sorting by specifically binding to export signals on cargo proteins and concentrating them at the Golgi membrane. Adaptor proteins such as KDEL and ERGIC2-3 complex also mediate sorting by binding to COPI subunits and simultaneously recognising specific signals on cargo proteins. As six out of seven COPI subunit are adjacent to the donor membrane, these proteins could potentially participate in cargo sorting, however, the role of specific COPI subunits other than  $\alpha$  and  $\beta'$  COP are yet to be explored [25].

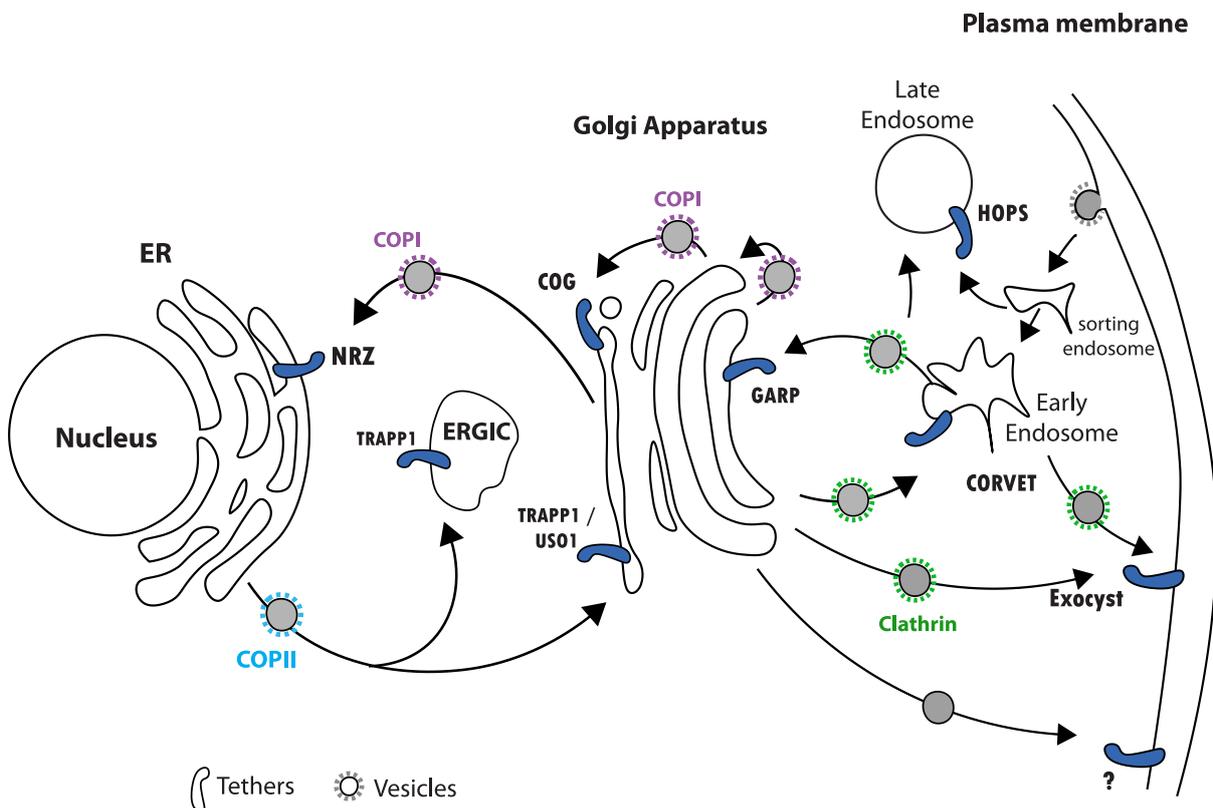
#### 1.1.4 Post-Golgi transport of post translationally modified proteins

At the Golgi apparatus, correctly folded proteins undergo post-translation modification and are then shunted to their target destination. As in the early secretory pathway, Golgi and plasma membrane volume and function is maintained by receptor mediated endocytosis from the cell surface to the Golgi via endosomal compartments and subsequent endocytic recycling of cargo from the Golgi back to the plasma membrane. Clathrin coated vesicles have been implicated in the transport of cargo between the *trans*-Golgi, endosomal compartments and the plasma membrane [26].

Clathrin coat complex is composed of three light chains and three heavy chains that interact to form a triskelion shape [27]. There are two known types of clathrin light chains,

CLTA and CLTB, and two types of known clathrin heavy chains (CLTC, CLTCL1) [27]. The clathrin triskelion structure can polymerise to create a polyhedral lattice that envelops the budding vesicle. Unlike COPI and COPII, clathrin does not bind to or sort cargo. Instead, adaptins or adaptor proteins have been implicated to perform this function [26]. The adapter protein 1 complex (AP-1), mediates endocytic recycling, i.e. the forward transport of protein from the *trans*-Golgi and/or recycling endosome to the plasma membrane. In contrast, AP-2 mediates receptor mediated endocytosis, i.e. delivery of cell surface cargo to endosomal compartments [28].

### 1.1.5 Vesicle tethering



**Fig 1-1** | Coordination of Tethers and Coated vesicles in membrane trafficking.

We have illustrated an overview of tethering factors and coat proteins in Fig 1-1. Incoming vesicles from a donor membrane have to be recognised, docked and then fused to the acceptor membrane for the delivery of cargo. Large proteins or multi-protein complexes mediate

interaction between target membrane and the incoming vesicles and restrain the latter in close proximity to SNAREs at the target membrane [29]. This promotes the formation of a *trans*-SNARE complex between a vesicle embedded SNARE (v-SNARE) and its counterparts at the target membrane (t-SNARE). This promotes vesicle fusion at the target membrane [30].

#### 1.1.5.1 Tethering at the early secretory pathway

##### 1.1.5.1.1 TRAPP complex

Studies in yeast have demonstrated that the heptameric TRAPP I complex may act as a tethering factor for trapping COPII coated vesicles at ERGIC and *cis*-Golgi. Its homolog, TRAPP II, containing 3 additional protein subunits, is speculated to function in the tethering of COPI vesicles for intra-Golgi transport [29].

##### 1.1.5.1.2 COG

The Conserved Oligomeric Golgi (COG) complex is an octameric tethering factor that serves as a tether in intra-Golgi retrograde transport to “catch” COPI vesicles arriving at the *cis*-Golgi. COG has been implicated in the correct localisation of glycosylation enzymes and other Golgi resident components [31].

##### 1.1.5.1.3 USO1 / p115

USO1 was one of the first tethering factors implicated in ER-Golgi transport. It is a large coiled-coil protein that is believed to primarily operate at the *cis* Golgi. The mechanism of action of this protein is a topic of debate. One model suggests that USO1 acts in intra-Golgi transport by binding to the GOLGAB1/giantin localised to incoming COPI vesicles, and then binding to GOLGA2/ GM130 at the *cis*-Golgi to mediate tethering at this membrane [29].

Another model suggests that two of the four coiled-coil domains of USO1, i.e. CC1 and CC4, may simultaneously bind a COPII v-SNARE and *cis*-Golgi t-SNAREs to promote the formation of the *trans*-SNARE complex for anterograde transport of cargo from the ER to Golgi apparatus [32].

#### 1.1.5.1.4 **NRZ Complex**

In mammalian ER, retrograde transport from Golgi to ER is thought to be mediated by the NRZ complex. This is a multi-subunit tethering factor made up of the proteins, NBAS, RINT1 and ZW10. NRZ is thought to promote trans-SNARE assembly and vesicle fusion by binding to the v-SNARE, SEC22B, on incoming COPI vesicles as well as the ER bound t-SNARE complex composed of STX18, USE1, and BNIP1 [33].

#### 1.1.5.2 **Post-Golgi tethering**

##### 1.1.5.2.1 **CORVET/HOPS**

The CORVET and HOPS are heterohexameric tethering factors that share 4 subunits. HOPS localises to lysosomes and late endosomes and is implicated in the tethering of incoming vesicles from the early endosomes, TGN and autophagosomes containing cargo destined for lysosomal degradation [34]. The less understood CORVET complex is speculated to work as a tether for homotypic fusion, wherein sequential endosomal fusion promotes the sorting of membrane proteins in late endosomes and endocytic recycling to the plasma membrane. Thus, this tether is believed to localise to early/sorting endosomes [35].

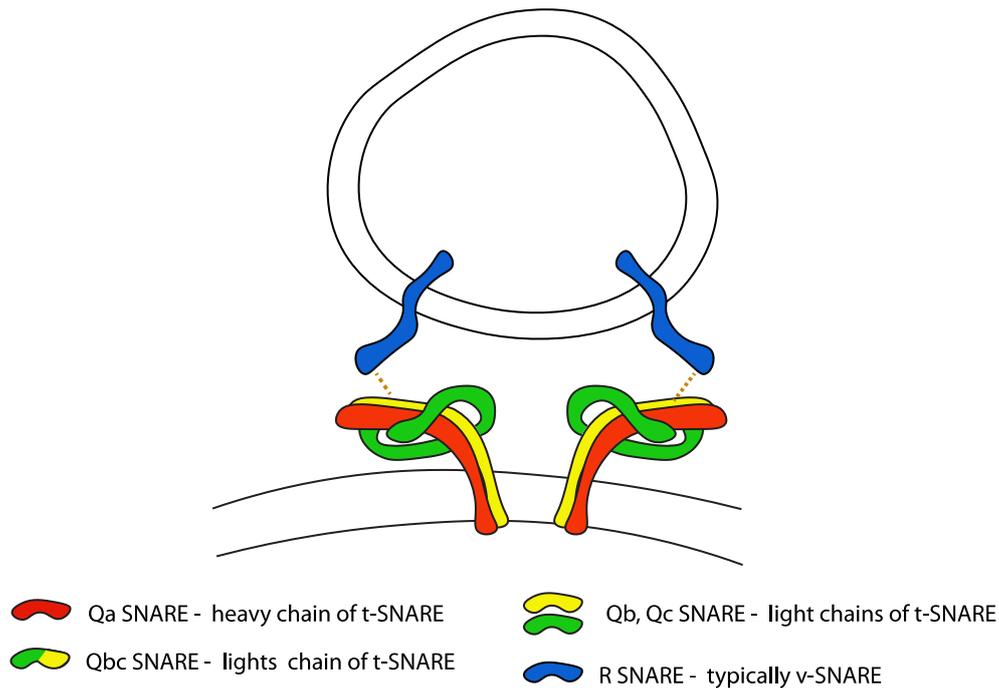
##### 1.1.5.2.2 **GARP Complex**

The GARP complex is a heterotetrameric complex proposed to perform as the tether for retrograde transport of vesicles arriving from the endosomes to the TGN [29].

##### 1.1.5.2.3 **Exocyst**

The Exocyst complex is an octameric complex that serves at the plasma membrane as a tether for vesicles arriving from TGN and recycling endosomes. Exocyst complex is implicated in polarised exocytosis in budding yeast, synaptogenesis as well as neurite outgrowth during nervous system development [29]. However, alternative tethering complexes may act at the plasma membrane as synaptic vesicles are found to undergo membrane fusion independent of the Exocyst complex [29]

## 1.1.6 SNAREs



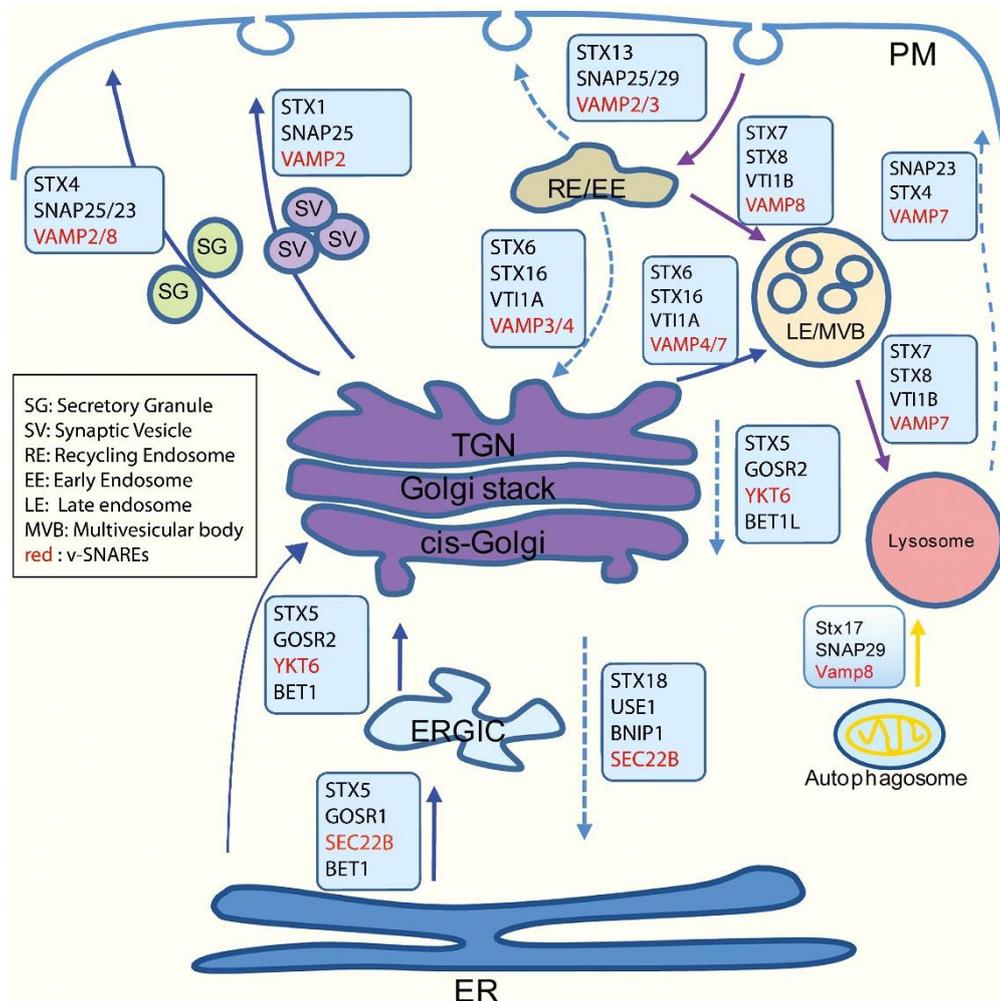
**Fig 1-2** | Schematic diagram of structural and functional classifications of SNAREs. v-SNAREs are always R-SNAREs. t-SNAREs are made up of 2-3 Q-SNAREs (2 light and 1 heavy coiled-coil motif). The heavy chain is typically a Qa SNARE. The light chains can be a single Qbc SNARE or a combination of Qb and Qc SNARE.

SNAREs mediate the last step of vesicle docking and fusion to the target membrane. SNARE proteins on the vesicle and recipient membrane combine to form a 4- $\alpha$ -helical coiled-coil called the *trans*-SNARE complex (TSC) that underpins the docking and fusion of incoming vesicles [36]. Functionally SNARE proteins are divided based on their association to vesicles (v-SNAREs) or the target membrane, (t-SNAREs) [36].

t-SNAREs contribute three coiled-coils to the TSC, i.e. 1 heavy and 2 light chain motifs [37]. t-SNAREs can also be classified as Q-SNAREs and divided into 4 subcategories. Qa SNAREs typically contribute one heavy chain motif to the TSC. Qb and Qc SNAREs each contribute a light chain motif to heterotrimeric t-SNAREs, whereas Qbc SNAREs contributes 2 coiled-coil light chains to a heterodimeric t-SNARE [37].

v-SNAREs are tail anchored proteins that always contribute a single coiled-coil motif to the TSC. Structurally, these proteins are classified as R-SNAREs. As shown in Fig 1-2, one

R-SNARE and 3 Q-SNAREs are needed to form a fusogenic *trans*-SNARE complex. An overview of SNARE localisation is given in Fig 1-3 [38].



**Fig 1-3** | Overview of SNARE complex localisation in the cell. (Adapted from Wang *et al*).

### 1.1.6.1 Early Secretory Pathway

#### 1.1.6.1.1 ER-Golgi/ERGIC Anterograde transport

YKT6 or SEC22B typically act as the v-SNAREs for COPII coated vesicles as illustrated in Fig 1-3 [36]. Studies have shown that YKT6 is able to fully substitute the role of SEC22B in COPII coated vesicles [36]. At the Golgi apparatus, these proteins can interact with the Q-SNAREs: STX5, BET1 and GOSR2 [33]. A similar Q-SNARE complex operates at the

ERGIC for the fusion of COPII, but in place of GOSR2, its isomer, GOSR1, lends a light chain coiled-coil motif to the ERGIC t-SNARE.

#### 1.1.6.1.2 Golgi-ER retrograde transport

At the ER, the Q-SNAREs, syntaxin 18 (STX18), BNIP1 and USE1 make up a heterotrimeric t-SNARE. The NRZ complex, discussed in Section 1.1.5.1.4, traps incoming COPI vesicles by binding to the R-SNARE, SEC22B, and subsequently promotes the formation of the TSC [33].

#### 1.1.6.1.3 Intra-Golgi Retrograde transport

Retrograde transport of cargo from *trans*-Golgi stacks to the *cis*-Golgi is thought to be mediated by the vesicular R-SNARE, YKT6, and the heterotrimeric t-SNARE complex consisting of the Q-SNAREs: syntaxin 5 (STX5), GOSR2 and BET1L. This SNARE complex is also believed to also operate in the tethering of vesicles returning from recycling endosomes to the TGN [36].

### 1.1.6.2 Post-Golgi Pathway

#### 1.1.6.2.1 Retrograde transport from plasma membrane to TGN via endosomes

Another heterotrimeric t-SNARE complex composed of STX16, VTL1A and STX6 mediates vesicle fusion in retrograde transport from the plasma membrane to the *trans*-Golgi via recycling endosomes through interaction with the vesicular R-SNARE, VAMP4 [39].

#### 1.1.6.2.2 Anterograde transport from the TGN to late endosome and/or lysosomes

Fusion of vesicles destined for lysosomal degradation is mediated by the t-SNARE complex composed of the Q-SNAREs: STX7, VTL1B and STX8. At late endosomes this complex typically form a *trans*-SNARE complex with vesicles containing VAMP8, whereas lysosomal t-SNAREs interact with VAMP7 [40].

#### 1.1.6.2.3 Exocytosis of lysosomal cargo

Heterodimeric SNAREs operate at the plasma membrane. Secretory granules budding from lysosomes are thought to be tethered by the cell surface Q-SNAREs, STX4-SNAP23, that

forms a complex with the R-SNARE, VAMP7 [41]. Fusion of lysosome and the plasma membrane typically occurs in order to quickly patch leaks at the plasma cell membrane [42].

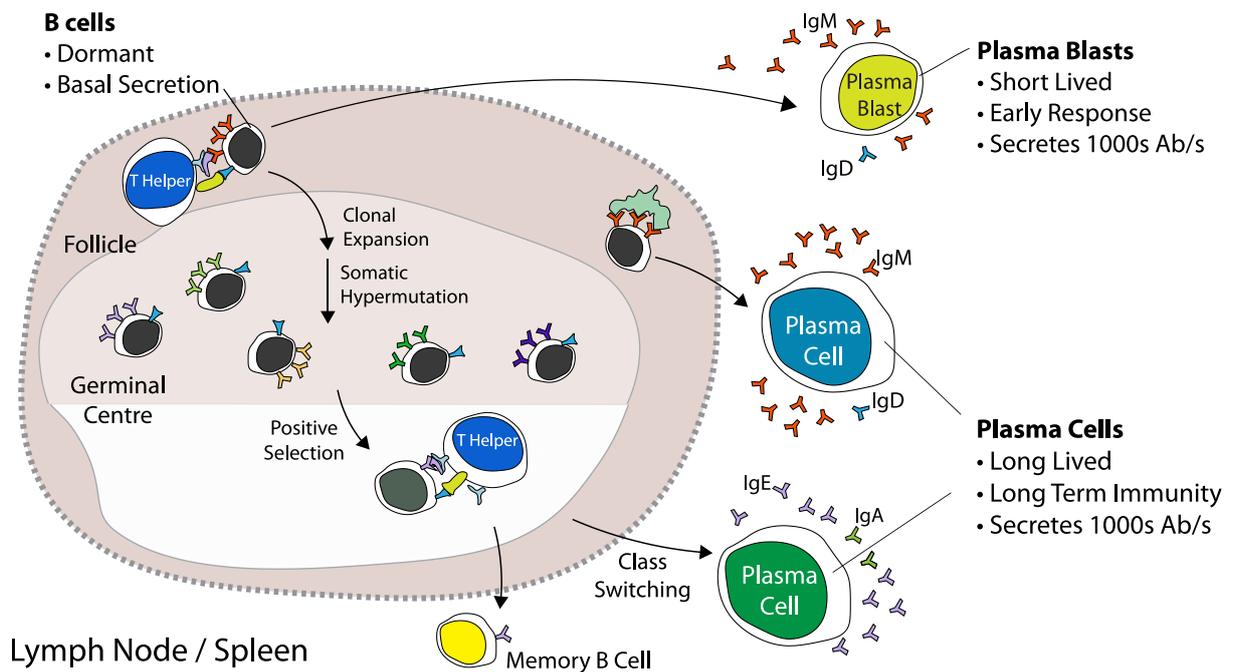
#### 1.1.6.2.4 **Regulated secretion of synaptic vesicle**

In the regulated secretory pathway, cargo destined for exocytosis are held within transport carriers until a stimulus promotes their release. In response to stimuli, synaptic vesicles release their cargo via fusion with the plasma membrane. This is thought to be mediated by the heterodimeric Q-SNARE complex consisting of STX1, SNAP-25 and the R-SNARE, VAMP2 [43].

#### 1.1.6.2.5 **Exocytosis of secretory granules**

Plasma membrane fusion of pre-docked secretory granules carrying cargo, such as insulin, is thought to utilise STX4-SNAP23 t-SNARE complex and the VAMP2 v-SNARE [44]. Fusion of newly arriving secretory granules, on the other hand, reportedly involves the Q-SNARE, STX3 and the R-SNARE, VAMP8, both of which are located in the secretory granule itself [45]. Recently, the t-SNARE, STX1A, has been reported to interact with VAMP8 to replenish insulin granules at the plasma membrane of pancreatic  $\beta$ -cells [44].

## 1.2 PLASMA CELLS – A MODEL FOR STUDYING SECRETION



**Fig 1-4** | Plasma cell Differentiation from Follicular B cells

Unlike regulated secretion, constitutive secretion does not require a stimulus to deliver cargo to the cell surface. This pathway is a fundamental process by which cells transport the majority of newly synthesized lipids and proteins to the plasma membrane. In this process, transport carriers are spontaneously secreted from the cell within minutes of exiting the Golgi apparatus and this process is independent of external stimuli [46]. Examples of constitutive cargo include collagen, fibronectin, and albumin [1]. The constitutive pathway operates in all cells at a basal level. However, professional secretory cells, such as plasma cells (PC), enhance this constitutive secretory pathway to expressly secrete thousands of proteins, in this case antibodies, per second [47].

### 1.2.1 Plasma Cell Differentiation

When resting B cells encounter an antigen they rapidly proliferate and differentiate into antibody secreting plasma cells. These are key effectors of the adaptive immune system that

protect our body from invading pathogen. Plasma cells can arise from 3 different types of B cells: B1 B cells, Marginal zone B cells and follicular B cells.

#### *B1 B cells*

B1 B cells typically reside in the peritoneal cavity, pleural cavity and mucosal sites. They function in an innate-like manner and rapidly respond to invading pathogen. They produce poly-specific antibodies which means that they can recognise many different antigens with low affinity. These cells are responsible for maintaining humoral level of IgM and are also implicated in a protective role during infancy. B1 B cells are activated in a T-independent manner and, therefore, most abundant during vaccination or infection [48].

#### *Marginal Zone B cells*

Marginal Zone B cells are typically located in the marginal zone of lymphoid organs such as spleen, tonsils and lymph nodes. They are polyreactive and mount a rapid, T-independent response to blood borne pathogens that become trapped in these organs. Furthermore, they are also able to present antigen to follicular B cells in T-dependant responses [48].

#### *Follicular B cells*

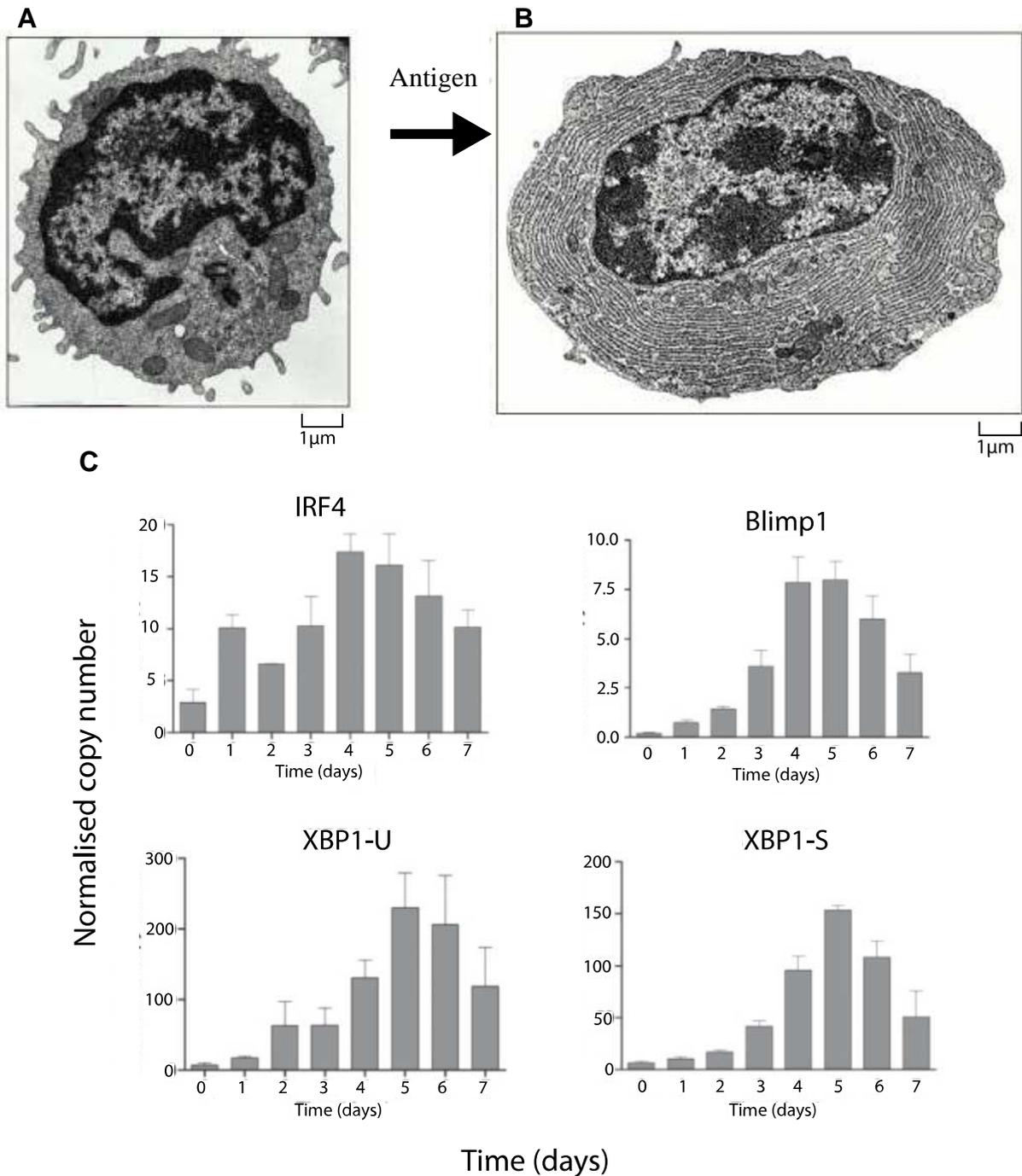
Follicular B cells are located in the follicles of lymphoid organs. These cells typically recognise antigen in a T-dependant manner with higher specificity than both B1 and marginal zone B cells. Upon T-dependant activation, a portion of these cells rapidly proliferate and differentiate into short-lived plasmablasts for an early response to infection (Fig 1-4A). These plasmablasts primarily secrete IgM and a small amount of IgD that bind antigen with moderate specificity [48].

Another portion of activated B cells will form a germinal centre (GC) in the lymphoid follicle through rapid proliferation followed by hypermutation (Fig 1-4B). Hypermutation generates B cell populations with varying affinity for the antigen. Those B cells with highest

antigen affinity will be positively selected via CD4+ T-Helper cell while the rest will undergo apoptosis. This then leads to the generation of high affinity memory B cells and also antibody secreting cells that can class switch to secrete IgE, IgA, IgG etc. While most of these proliferating early plasma cells / plasma blasts are relatively short lived, a small portion of GC derived plasma cells are able to migrate to bone marrow to become long-lived terminally differentiated plasma cells [49].

When naïve B cells (NBCs) encounter antigens (e.g. lipopolysaccharide) that cause cross linking of B cell receptors, the resultant antigen-receptor bond is strong enough to induce activation independent of T-helper cells (Fig 1-4C). In these cases, follicular B cells can form long lived plasma cells independent of affinity maturation at the GC [50].

In 1986, Hibi *et al* used ELISA sandwich assays to estimate the rate of antibody secretion *in vitro* and found that plasma cells secreted between 13-27 thousand antibodies per second. [51]. In 1993, Werner-Favre showed that human plasma cells are capable of secreting ~5000-6000 antibodies per second using limiting dilution [52]. A more recent study in 2009 utilised a flourosport ELISA assay to estimate antibody secretion per cell and found that plasma cells in the blood secreted ~150 thousand antibodies per second while splenic plasma cells secreted 30 thousand antibodies per second [53]. As such it is apparent that plasma cells and plasmablasts are very highly specialised for secreting antibodies per second unlike NBCs [51]. In this project we exploit this remarkable difference in the secretory capacity of antibody secreting cells (ASC) and their follicular B cell precursor to understand antibody secretion.



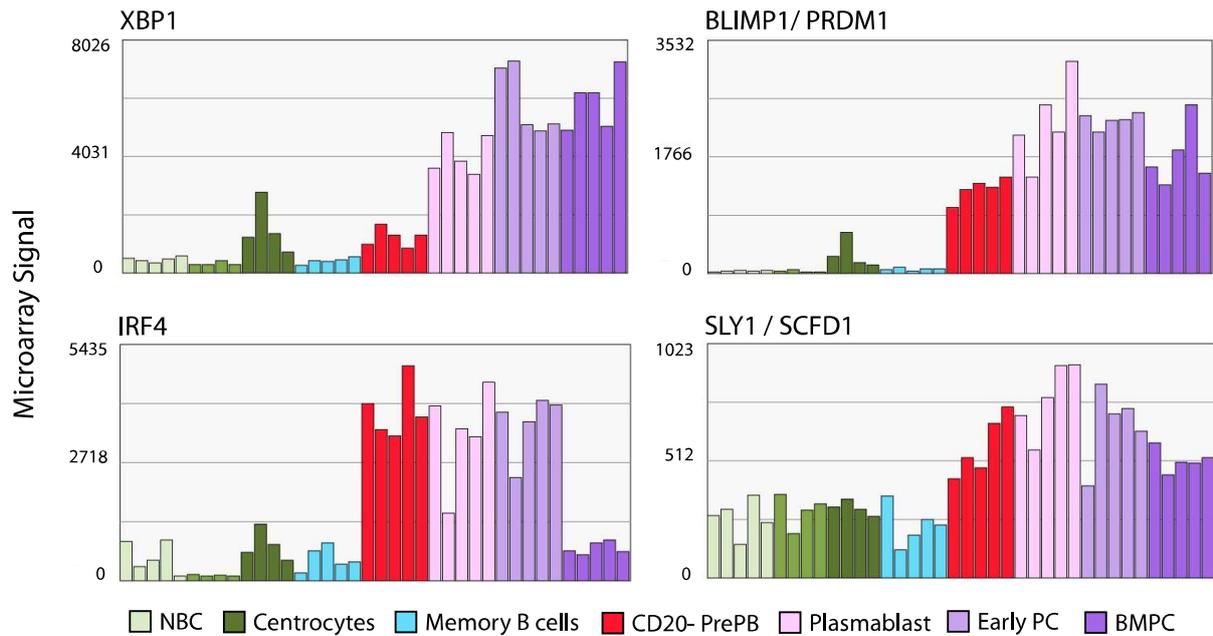
**Fig 1-5** | Electron micrograph of **A**. Naïve B cell and **B**. Plasma Cell. Unlike NBCs, an extensive rough endoplasmic reticulum is apparent in PC and the rER appears swollen due to pressure from newly synthesized antibodies [1]. **C**. qRT-PCR results showing the upregulation of positive regulators of protein secretion in NBCs over 7 days after antigen activation. Purified B cell were *in vitro* activated with CpG, PWN and SAC at day 0.

### 1.2.2 Structural and Physiological changes during Plasma Cell Differentiation

Unlike NBCs, which sport minimal endomembrane (Fig 1-5A), PCs exhibit an extensive rough endoplasmic reticulum which appears swollen, due to high levels of newly synthesized antibodies (Fig 1-5AB) [1]. Kirk *et al* quantified this organellar expansion and found that the rough endoplasmic reticulum (rER) of PCs was up to 4 fold larger and Golgi apparatus was up to 6.5 fold larger than NBCs to accommodate for the increased demand for protein biosynthesis and secretion [54]. The study also measured the mRNA expression of *in vitro* activated NBCs over time and found that known positive regulators of secretion, such as XBP1 and BLIMP1, are markedly upregulated as NBCs transition to ASCs (Fig 1-5C) [54]. In each case, NBCs show little or no expression while PCs show a marked upregulation in positive regulators of protein secretion. Using Amazonia! Gene Expression Atlas we see that these expression patterns are mirrored in high-throughput microarray profile of NBCs and PCs (Fig 1-6).

### 1.3 HYPOTHESIS

Based on the remarkable structural and physiological changes during the transition of NBCs to ASCs, we hypothesize that the transcriptome of NBCs can be used as a non-secreting control and contrasted against their antibody secreting progeny to isolate components required for antibody secretion. This may give valuable insights into the coordination and regulation of known membrane trafficking components and determine which specialised trafficking machinery (if any) contribute to enhanced antibody secretion. Furthermore, this analysis may help identify novel membrane trafficking components and regulatory mechanisms that play a role in this process.



**Fig 1-6** | Microarray Signal Intensities in various stages of plasma cell differentiation. Derived from Amazonia! Gene Expression Atlas [55, 72].

## 1.4 RESEARCH RELEVANCE

### 1.4.1 A Proteogenomic Atlas of Antibody Secreting Cells

#### 1.4.1.1 Transcriptomics

A number of research groups have generated high throughput transcriptome profiles for comparison of NBCs, PCs and their intermediates in mice and human model [55–58]. These studies primarily focus on identifying key transcription factors regulating the stages of PC differentiation. Thus the potential role of downstream membrane trafficking components have not been explored [55–58]. We note that some of these studies were carried out using RNA-Seq, while others with microarrays.

##### 1.4.1.1.1 Microarrays vs RNA-Seq

###### *Microarrays*

Arrays utilised in these studied were primarily probe based and are ideal for testing the expression of known sequences. Detection of genes using the same microarray platform has been shown to have 70-85% reproducibility across studies [59]. The median correlation

between differentially expressed genes (DEG) across different microarray platforms is reported to be 87% on average. We utilise Affymetrix arrays in this project and MAQC benchmarking show that intra-platform differential expression calls for Affymetrix arrays exhibit 91-98% concordance [59].

#### *RNA-Seq*

RNA-Sequencing is ideal for finding both known and novel genes as they do not rely on pre-designed probes. SEQC benchmarking reports show that different RNA-Seq pipelines can reproducibly detect the expression of 90% of all known genes and 95% of differentially expressed genes [60]. As such RNA-Seq results are considerably more reproducible than microarray and has also been shown to have a higher dynamic range [61]. Nevertheless, SEQC studies comparing differential expression values from microarray and RNA-Seq show high correlation (89-92%) [60].

#### 1.4.1.1.2 **qPCR**

Real time polymerase chain reaction (RT-PCR) is a low throughput transcriptomic method that utilises fluorescently labelled probes to quantify the absolute or relative copy number of a transcript during PCR amplification. It is thought to be the “gold standard” of transcript quantification. Differential expression values for both microarray and RNA-Seq show good correlation (over 90%) with qPCR studies [60]. Nevertheless, qPCR is used as downstream validation for microarray results. This is because of the limited and somewhat fixed dynamic range of microarrays as opposed to qPCR [62]. A higher dynamic range means that qPCR can accurately detect transcripts with very high or very low copy numbers more accurately. In contrast, qPCR is not recommended for RNA-Seq cross-validation unless the experimental design involves few or no replicates. This is because RNA-Seq has a broader dynamic range, which is primarily governed by sequencing depth [59, 60].

#### 1.4.1.1.3 High Throughput Transcriptomics

High-throughput transcriptomics is predictive in nature not only due to technical and biological limitations, but also because it does not substantiate protein expression. This makes genes beyond the “top hits” unattractive for cell biologists to carry forward for downstream analysis. Sufficient number of replicates are required to overcome these limitations as discussed in Section 2.1.2.3. Unfortunately, generating enough replicates has been difficult due to previously high cost of RNA-Sequencing and in some cases sample scarcity. In recent years, combining data from comparable transcriptomes generated by independent studies, otherwise known as “meta-analysis”, has proven to be a powerful tool for improving the reproducibility of transcriptomic data and thus candidate selection [63, 64]. For example, Shi *et al* compared a number of antibody secreting cell types in the mouse model to find signature genes for these cell types as opposed to plasma cells [56]. We note that meta-analysis of the transcriptomes of the plasma cell lineage across species and across platforms have not yet been performed. Therefore, we aim to leverage existing transcriptomes of ASCs accumulating on public repositories to build a meta-analysis bioresource that improves the identification of reproducibly changing genes by combining data from across studies, species and transcriptomic platforms. The resultant database may enable us to isolate membrane trafficking machinery in the PC physiology with more accuracy and serve as a useful resource for biologists who are interested in the physiology of ASCs in general.

#### 1.4.1.2 Proteomics

During the era of microarrays, large scale studies showed that less than 27% of the changes occurring in human transcript levels related to consequent shifts in protein abundance [65]. In the recent years, considerably improved correlation ( 60- 90%) depending on the cell line was observed [66–68]. This is mainly due to recent advances in RNA-Sequencing and high-

resolution mass spectrometry that alleviated imperfections in the previous systems used to perform these analyses.

The proteome of the plasma cell lineage has previously profiled using I.29 $\mu$ + lymphoma cells, which does not accurately reflect the wild type physiology [69, 70]. Moreover, the total number of detected proteins were small (~2000 proteins) compared to the yield achievable today by high resolution mass spectrometry (~5000-10000 protein) [69, 70]. In order to study the physiological proteome of the PC lineage, our colleagues E. Rajan and A.W Asral have isolated B cells and *in vitro* generated ASCs, whose proteome were externally profiled using high resolution mass spectrometry. Using this data, we aim to analyse the correlation between mRNA and their corresponding protein product in the physiological model of the PC lineage. This may not only strengthen the validity of utilising transcriptomic analyses for candidate selection in individual members of the PC lineage but also serve as soft validation for differential transcriptomics.

#### 1.4.1.3 Visualization of Proteogenomic Data

High throughput differential expression analyses typically output large excel sheets of hundreds to thousands of rows of differentially regulated genes. This makes it difficult for cell biologists to discern important expression patterns hidden beyond the top hits. A number of web-based tools such as Amazonia! and Genomicscape allow the mining and visualization of the microarray profiles of the PC lineage [71, 72]. However, these studies do not allow visualization of comparable experiments from different labs nor do they incorporate recently available RNA-Seq data or any proteomic analyses. Therefore, we aim to create a web application that summarises our proteogenomic bioresource to overcome these limitations. By visualising changes in mRNA expression alongside their corresponding protein product, this web application may allow biologists to better judge candidates for downstream analysis and potentially improve the cost effectiveness of downstream validation.

### 1.4.2 Protein Secretion in the Plasma Cell Physiology

To date, functional genomics approaches to identify components of protein transport have relied on artificially inducing the secretion of labelled fusion proteins and then performing genome wide RNAi screening to determine whether the loss of a gene perturbs this secretion [73–75]. However, these studies were typically performed using embryonic *Drosophila* cells, HeLa and HEK cells, none of which are naturally optimized for high levels of protein secretion [73–75]. As such, bulk protein secretion has not been studied in a physiological model. To address this issue, we aim to utilise the proteogenomic bioresource of plasma cells, discussed in 1.4.1, as they are specialized for secreting thousands of antibodies per second [51]. This may help us identify patterns of expression that would otherwise be absent in an artificial model, isolate novel trafficking components and gain valuable insights into how known membrane trafficking components are coordinated and regulated in a physiological manner.

### 1.4.3 Cell Markers for Antibody Secreting Cells

To generate antibody secreting cells *in vitro*, NBCs have to be purified from the spleen, activated by a selected antigen and resultant ASCs have to be isolated from a mixed population of non-secreting B cells and ASCs. Typically ASCs are isolated by positive selection using the CD138/syndecan-1 [76]. However, CD138 expression has been found to be heterogeneous in plasma cells and inconsistent in the various stages of differentiation. For example, short lived ASCs express lower levels of CD138 compared to long-lived plasma cells [77]. This means that using CD138 as the sole marker ASCs purification only allows partial yield of antibody secreting cell population. Therefore, we aim to utilise proteogenomic analysis of ASCs to identify novel biomarkers common to antibody secreting cells regardless of their maturation stage and half-life. These biomarkers may potentially play a role in improving the yield of ASCs in a laboratory setting as well as aid in the targeted destruction of plasma cells in cancer therapy.

## 1.5 MEDICAL RELEVANCE

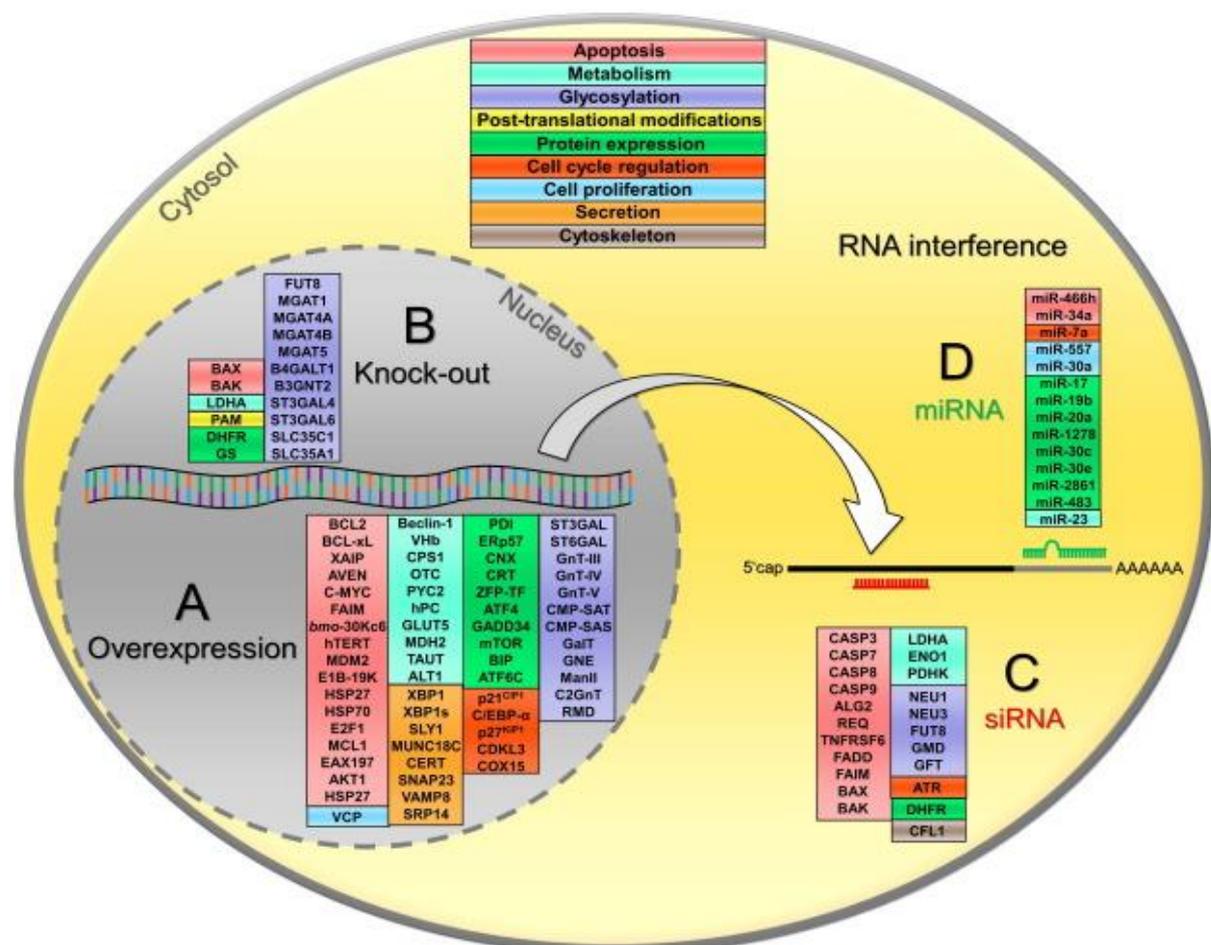
Plasma cell related autoimmune disorders have been implicated in diseases such as lupus erythematosus, rheumatoid arthritis, and multiple sclerosis. Emerging treatments for multiple sclerosis, for example, involves aggressive chemotherapy to destroy circulating lymphocytes with subsequent stem cell transplantation [78]. Unfortunately, aggressive chemotherapy is not suitable for all patients and a safer alternative would improve the success and applicability of such treatment. A possible solution is the usage of cell surface markers to allow selective targeting and destruction of circulating lymphocytes. Rituximab, a monoclonal antibody (mAb) against the CD20 marker, is one of the most successful candidates for treating plasma cell malignancies [79]. Unfortunately, this treatment is not successful in multiple myeloma, where plasma cells downregulate CD20 expression [80]. As discussed in the previous section, the commonly used CD138 marker does not give full coverage of existing antibody secreting cells. This means that if this marker is solely used for targeted destruction, malignant ASCs may be overlooked thus increasing the chances of relapse. Therefore, we utilize proteogenomic meta-analysis to isolate novel biomarkers of ASC. These may potentially be used in combination with existing markers to achieve better coverage of ASCs targeted for destruction and improve the prognosis of stem cell therapy.

## 1.6 INDUSTRIAL RELEVANCE

### 1.6.1 Production of Recombinant Proteins

The global market for biologics was worth over 221 billion dollars in 2017, 43% of which is accounted for by the sale of mAbs such as Rituximab and erythropoietin [81]. Industrial production of biologics typically involves the culturing of transfected cell lines in batch culture bioreactor tanks where the cells produce and secrete gram quantities of therapeutic proteins that are then harvested from the media. This means that yield is directly related to the ability

of host cell lines to efficiently exocytose the recombinant protein. Chinese Hamster Ovary (CHO) cells are a widely used cell line for the industrial production of recombinant proteins including antibodies due to their ability to often correctly glycosylate human proteins [82, 83]. However, this system has relatively low yield, stress resistances, is more expensive than its bacterial or yeast counterparts. Most of these problems arise from the fact that the CHO cell line is not naturally optimized for large levels of protein secretion as well differences in species-specific glycosylation.



**Fig 1-7** | Over 25 year of genetic integration of advantageous genes in CHO cells (Fischer, 2015).

Genetic modification of host CHO cells has been proven to improve the performance of protein production [82, 84]. This strategy typically involves the overexpression of beneficial genes or the repression of unfavourable ones through RNAi silencing or gene knockout. In Fig 1-7, the orange block highlights a list of genes that have been overexpressed to increase the level of

secretion in CHO cells. These include XBP1 and BLIMP1, which are known to be highly upregulated in plasma cells compared to B cells [85]. While genetic engineering techniques, media and vector optimization over the years have considerably improved the CHO cell performance[86], there is still a significant difference in the titre of plasma cell antibodies (over 1000 mg/L) compared to the best optimized CHO cell biologics (~863 mg/L) [87, 88]. Therefore, if we establish how known and novel components of membrane trafficking are coordinated and regulated in ASCs, the same behaviour can potentially be replicated in CHO cells to optimize the yield of recombinant proteins. Doing so may improve the cost effectiveness of biologics and minimize the growing cost associated with the increasing demand for therapeutic proteins in the healthcare industry.

# Chapter 2

## Preliminary Transcriptomic Reanalysis

## 2 PRELIMINARY TRANSCRIPTOMIC ANALYSIS

### 2.1 BACKGROUND

When naïve B cells (NBC) encounter an antigen, they become specialised antibody secreting cells (ASCs) that release thousands of antibodies per second to fight infection [51]. ASCs are key components of adaptive immunity and make up 1% of circulating cells in the human body [47]. Proliferating ASCs, known as plasmablasts (PB), mount a rapid, early response to pathogens, while the terminally differentiated plasma cells (PC) provide long term humoral immunity [89, 90]. As NBCs differentiate into ASCs they undergo a 4 fold expansion in ER, 6.5 fold expansion in Golgi volume and exhibit a sharp upregulation in the mRNA and protein level of known positive regulators of secretion such as XBP1 and BLIMP1 [54].

#### 2.1.1 HYPOTHESIS

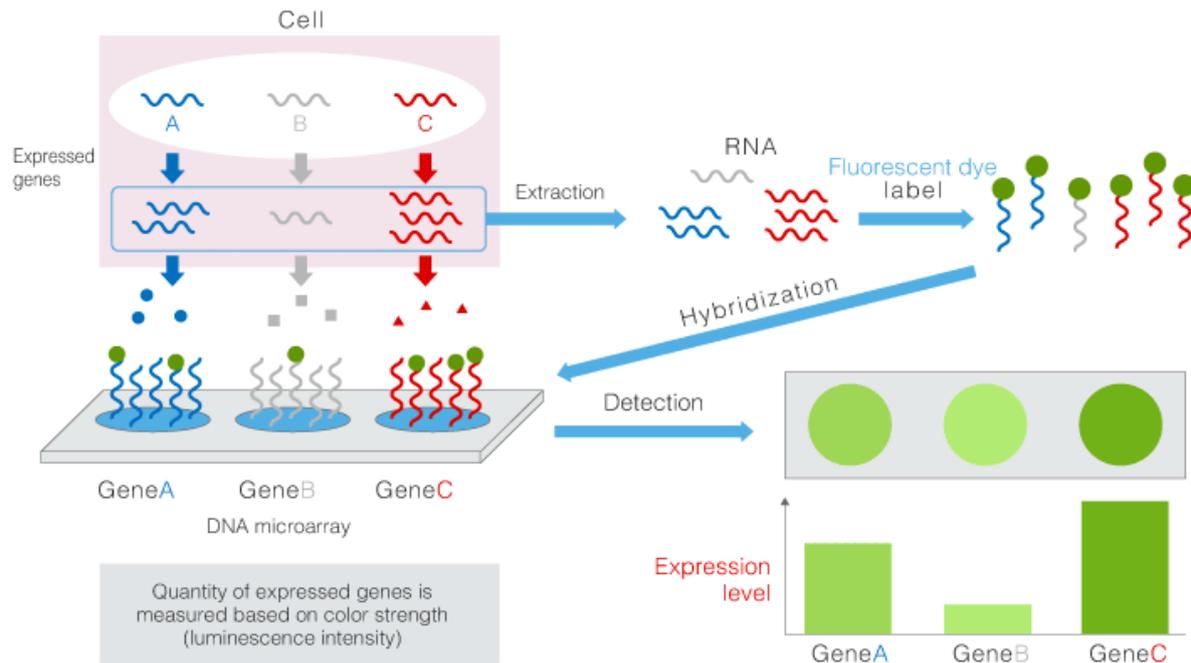
Based on these extraordinary structural and physiological changes occurring in NBCs as they transition to factories of antibody production and secretion, we hypothesize that if the whole cell transcriptome of antibody secreting cells are compared to non-secreting naïve B cell precursor we are likely to find enrichment in known and potentially novel components of protein processing, protein trafficking as well as antibody secreting cell markers amongst genes upregulated in ASCs.

#### 2.1.2 High Throughput Transcriptomics

High throughput transcriptomics such as RNA-Sequencing (RNA-Seq) and gene-chip based microarrays allow the simultaneous study of thousands of genes within a cell in a quick and cost-effective manner.

##### 2.1.2.1 DNA Microarrays

While RNA-Seq is a more robust and accurate means for transcriptome analyses, microarrays remain the fastest, cheapest and the simplest way to measure total gene expression within a



**Fig 2-1** | Generic RNA detection method by microarrays. From TORAY, 2013

cell. DNA microarrays typically consist of clusters of immobilized probes, where each cluster contains hundreds of short (25-50bp) oligonucleotide strands that are pre-designed to hybridise to a specific genomic sequence. Sample RNA is reverse transcribed *in vitro* to cDNA, biotinylated or fluorescently labelled and added onto these gene chips. These cDNA bind to their complementary probe while unreacted reagents are washed away. Thus, probe clusters fluoresce where labelled cDNA have hybridised. Fluorescent scan of the Genechip yields optical intensity values for probes that correlate to the cDNA fragments' abundance. Therefore, these fluorescent intensities can be exploited to estimate the relative expression of mRNA sequences and generate a snapshot of the transcriptome of a given cell (Fig 2-1).

### 2.1.2.2 Sources of Variation for Microarray Data

Systematic variations or noise exert consistent effects on multiple measurements and typically arise during RNA extraction, reverse transcription, labelling or photodetection methods [91]. As such the raw intensity values from microarrays cannot directly be classified as mRNA

expression. Fortunately, systematic variations can be estimated and corrected through well characterised normalization methods.

Stochastic/random variations specifically affect a few genes or corresponding probe clusters. These can arise due to variations in cDNA quality, *in vitro* amplification, probe binding efficiency, probe cluster size, non-specific or cross hybridization, stray fluorescent signals as well as biological differences between subjects arising from genotypic differences such as single nucleotide polymorphisms (SNPs) [91]. Unfortunately, not all of these can be estimated. While RNA-Seq alleviates probe related limitations, other stochastic variations affect this technique the same way as microarrays. This is one of the core reasons individual high-throughput transcriptomics studies often generate non-reproducible results.

### 2.1.2.3 Replicates for normalisation and differential expression analysis

In order to correct for systematic noise arising from experimental procedures, technical replicates are required. A minimum of three technical replicates are recommended for estimating and removing noise in microarrays [92]. However, because correlation in gene expression between technical replicates tend to be high (86-93%), many experimental designs have foregone the use of technical replicates due to sample scarcity or run cost, in the case of RNA-Seq. Instead, more importance is given to biological replicates as correlation of gene expression across subjects is considerably lower (66-78%) [93]. Statistical power of differential analysis is directly related to the number of replicates which confers the degree of freedom for calculating  $p$ -values. For example, the probability of correctly detecting 2-fold change with 10 replicates is 100% in RNA-Seq, which is reduced to 87% when 3 biological replicates are used [94]. Therefore, best practise for both microarray and RNA-Seq is to use a minimum of 3 biological replicates. Sample scarcity and cost of analysis have caused some studies to generate samples with no replicates. The  $p$ -values for such experiments have little meaning and is thought to be exploratory at best. However, differential expression in stimulation studies, such

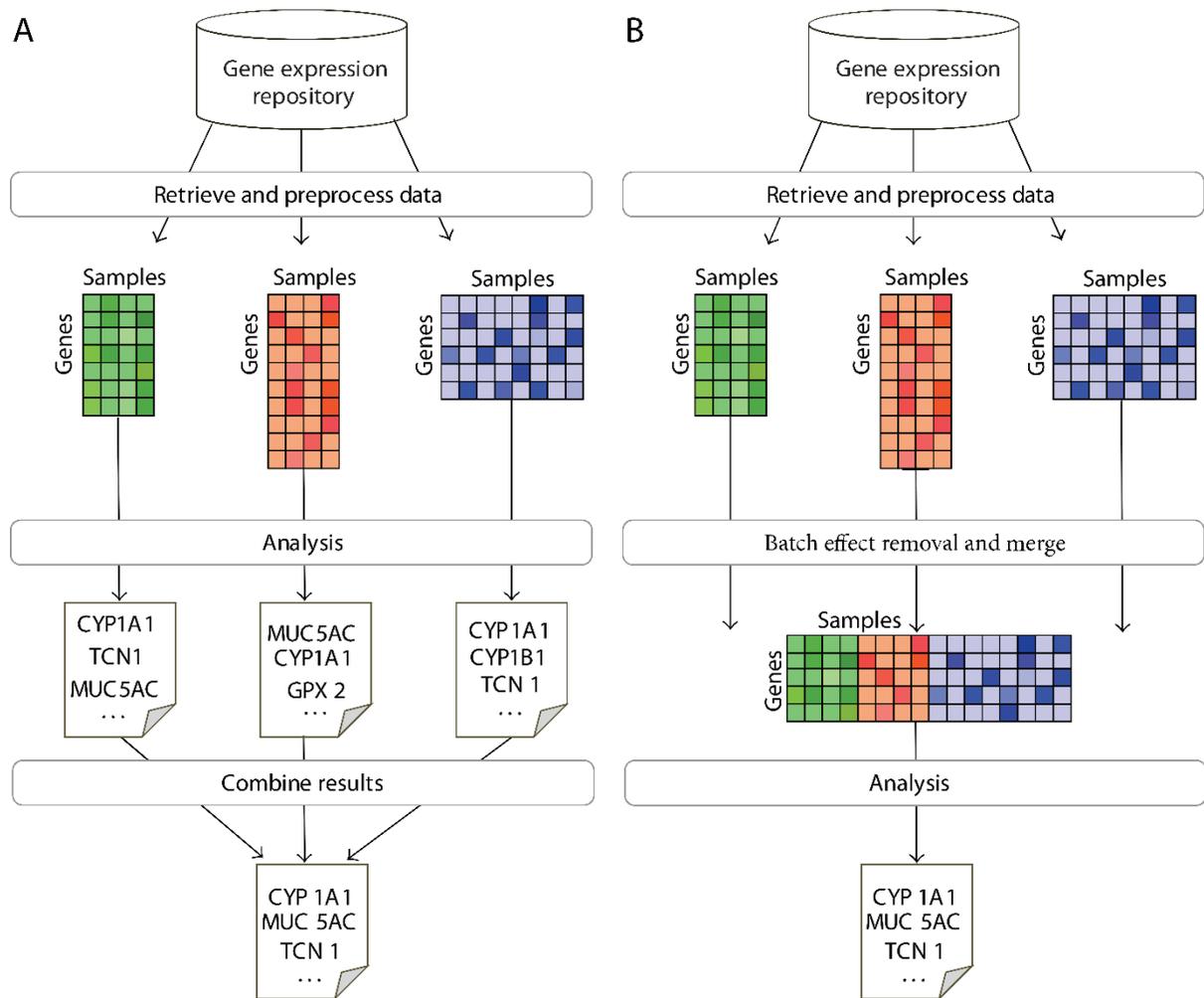
as antigen activation of immune cells, has been shown to outweigh both biological and technical variation and should not be ignored [94]. This is especially true when conducting meta-analysis where the evidence for differential expression can be cross validated from other studies having suitable number of replicates.

#### 2.1.2.4 Utilising comparable Omics data

As discussed, sufficient replicates to overcome these above-mentioned limitations can be cost-inhibitive. As such, the practise of combining data from comparable transcriptomes generated by independent studies has proven to be a powerful tool for improving the reproducibility of transcriptomic data and thus candidate selection [63, 64]. This can be done in two ways depending on the homogeneity of the data: “merging” or meta-analysis.

“Merging” is a method where comparable data from different labs is pooled together prior to performing statistical analyses [95]. This method requires adjustment to remove any unwanted variation arising from the data set being generated by different research groups or in different conditions (Fig 2-2A). Meta-analysis, on the other hand, involves pooling data together after requisite statistical analyses has been performed. As such, this method does not require any special adjustment and has been found to be more conservative than merging [95]. Although, “merging” results in more differentially expressed hits, the approach is difficult to apply in heterogeneous data generated from different species or different platforms due to disagreements in gene/transcript identifiers, orthologs and measurement types (RNA-Seq transcript count vs. microarray optical intensity).

## 2.1.3 Leveraging Publicly Available Transcriptomes of ASCs



**Fig 2-2 | A.** Workflow of Metanalysis. **B.** Workflow of Merging. From Taminau *et al* 2014

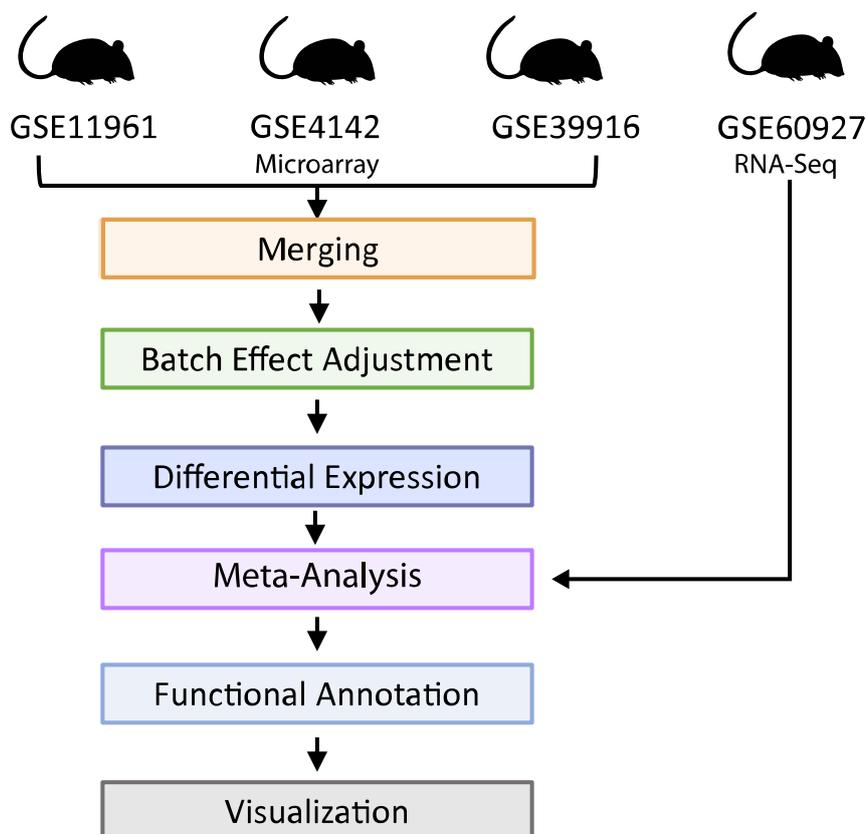
Over the years a number of the microarray studies of *in vitro* and *ex vivo* ASCs have been published. Recently a comparable RNA-Seq study was performed on the mouse models [55–58]. These studies primarily focus on identifying key transcription factors regulating the stages of PC differentiation. Thus, the coordination and regulation of downstream membrane trafficking components potentially contributing to enhanced secretory capacity in ASCs have been overlooked.

#### 2.1.4 AIMS

In this preliminary chapter, we “merge” three microarray datasets to improve reproducibility of differential expressed genes. Then add comparable RNA-Seq data (GSE60927) via meta-analysis. We do this to:

1. Evaluate our initial merging and meta-analysis protocol
2. Determine whether upregulated genes in ASCs compared to NBCs are overall enriched for (a) Protein processing components, (b) Membrane Trafficking Components and (c) Cargo proteins

## 2.2 METHODS



**Fig 2-3** | Workflow of Preliminary Merging and Meta-analysis of microarray data. Microarray-generated, pre-processed gene expression profiles from three individual studies were downloaded and merged. After batch effect adjustment comparable gene expression profiles generated by RNA-Seq was added via meta-analysis. Differentially upregulated genes were functionally annotated, and the number of genes found to be in relevant functional groups were visualized.

### 2.2.1 Data Source

The data used in this chapter is summarised in Table 2-1. We utilised microarray-generated gene expression profiles from three distinct studies. All three data sets originated from *Mus musculus* and were generated using the same microarray gene chip: Affymetrix GeneChip Mouse Genome 430 2.0. Therefore, we assume there is sufficient homogeneity in the data to perform “merging” between these three datasets. GSE11961 and GSE39916 profiled NBCs and transitional PCs from the spleen while GSE4142 profiled mature plasma cells from bone marrow (GSE4142). *In vivo* generated plasma cells were extracted 7 days post antigen

**Table 2-1 | Phenotype of Samples utilised for Cross-Study Meta-Analysis**

Study	Cell Type	Tissue	Strain	<i>Ex vivo</i> Stimulus	Minimal Marker	Replicate
GSE11961	NBC	Spleen	C57BL/6	-	CD45+ CD138-CD23+	3
GSE11961	PC	Spleen	C57BL/6	NP-CGG/alum	CD45+ CD138+ IgG1-low	3
GSE4142	NBC	Spleen	C57BL/Ka	-	CD45+ CD23+ IgM-low IgD+	3
GSE4142	PC	Spleen	C57BL/Ka	NP-CGG/alum	CD45-low CD138+ IgM-low	3
GSE39916	NBC	Spleen	C57BL/6	-	CD45+ CD23+	3
GSE39916	PC	Bone Marrow	C57BL/6	KLH	CD45- CD138+	3
GSE60927	NBC	Spleen	C57BL/6	-	CD45+ CD23+	2
GSE60927	PC	Bone Marrow	C57BL/6	-	CD138+ Blimp+	1

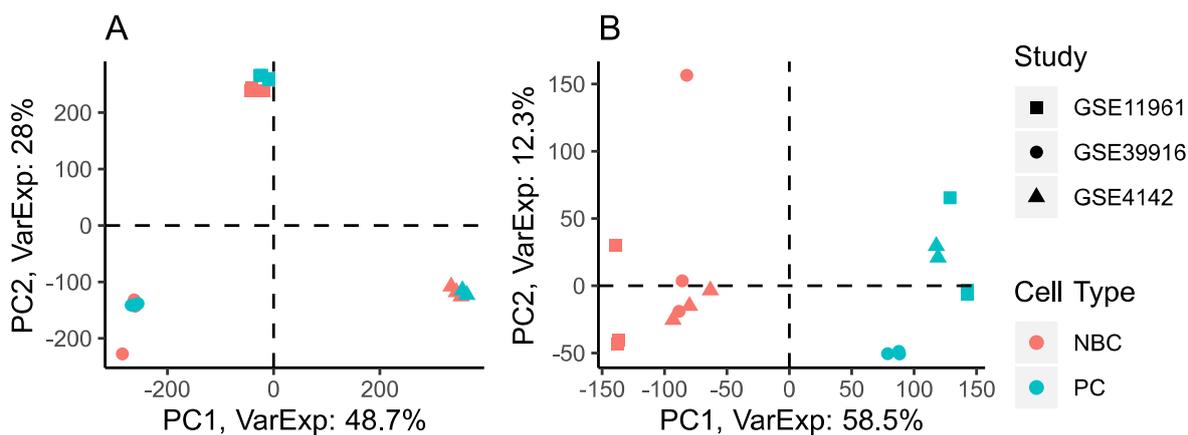
activation in GSE11961 and GSE4142, while ASCs in GSE60927 were extracted 28 days post immunisation. In addition to these data, we utilise the recently submitted RNA-Seq study (GSE60927), that profiled the transcriptome of splenic NBCs and *in vivo* bone marrow PCs. The PC sample for this profile was extracted from three different mice to account for the rarity of ASCs in the organs of unstimulated mice [56].

### 2.2.1.1 Merging

Pre-normalised profiles were downloaded from NCBI Gene Expression Omnibus (GEO). Replicates for follicular B cells and plasma cells from each study were merged together as shown in Fig 2-3. We assigned splenic and bone marrow PCs to the same group as we are interested in finding genes that show same direction of regulation in antibody secreting cells regardless of maturity

### 2.2.1.2 Batch Effect Removal

When merging data, unwanted sources of variation has to be adjusted before further analyses can be performed. While standard normalisation methods such as RMA caters for intra and inter array variations, unaccounted variables such as environmental, genetic or technical differences can have a strong effect on gene expression. These factors can be: the time when a sample was prepared, the person conducting the experiment, differences in reagents or transcriptomics platform used, as well as differences in the species or strain of the source



**Fig 2-4** | Principal Component Analysis of raw and batch adjusted cross-study mouse microarray data. **A.** Direction of variance before batch effect adjustment show all samples clustering by study. **B.** After batch effect adjustment samples cluster by cell type.

material. If these “batch” effects are not accounted for, gene expression will not be representative of the variable of interest. ComBat (Combatting batch effects when combining batches of microarray data) algorithm adjusts for known batch affects and is appropriate for small sample sizes.

As shown in Fig 2-4A, principal component analysis of our data showed that the direction of variation in the data was due to cross-laboratory differences. Therefore, we removed batch effect using the above mentioned ComBat approach [96]. Batch adjusted data showed expected variation to be greatest between B cells and plasma cells. (Fig 2-4B).

### 2.2.2 Differential Expression

Batch adjusted data was fitted to the linear regression model for microarrays (limma) and empirical Bayes moderated t-statistics was used to calculate differential expression and adjusted for global false discovery [97]. A conservative threshold of 2-Fold Change (FC) and Benjamini & Hochberg adjusted  $p$ -value less than 0.05 was enforced to filter out noisy hits.

### 2.2.3 Meta-Analysis of Cross-Platform Data

For cross-platform comparison of RNA-Seq and microarray data, we extracted pre-calculated differential expression data from GSE60927. This data was merged to our microarray analysis

using gene identifiers and contradictory results were filtered out. Genes that were upregulated in the microarray data and downregulated in RNA-Seq or vice versa were considered contradictory. Differentially expressed genes missing genes in either data set were retained for further analysis.

#### 2.2.4 Ranking Differentially Expressed Hits

We found a large number of genes with very low adjusted *p-values* and were unable to utilise this statistic for meaningful ranking. The alternative way to sort differentially expressed data is to use gene fold changes. However, as FC relies on ratios, small differences can produce inflated FC (e.g. going from 0.001 to 10 gives 10000 FC), while large changes (e.g. 1000 to 2000) will only give an understated 2 FC. To circumvent this problem, raw intensity values were sorted into 5 bins with increments of 50, with the 5<sup>th</sup> bin holding intensity values greater than or equal to 200. Using these bins differentially expressed genes were custom sorted by largest plasma cell intensity then smallest B cell intensity followed by the greatest magnitude of change.

#### 2.2.5 Functional Annotation

Differentially expressed genes were annotated with Entrez, Uniprot and Tocris Summary using Lifemap's GeneAlaCart tool [98]. We used the 2015 version of David tools to retrieve Gene Ontology (GO) and Swiss-Prot Protein Information Resource (SP\_PIR) terms[99, 100]. The knowledge based COMPARTMENTS database was used to group secreted and membrane proteins based on their localization [101].

*Drosophila* orthologs of differentially expressed genes were obtained using the DRSC Integrative Ortholog Prediction Tool (DIOPT) tool [102]. As invertebrates such as *Drosophila melanogaster* lack an adaptive immune system, matching orthologs should theoretically filter out genes related to adaptive immunity.

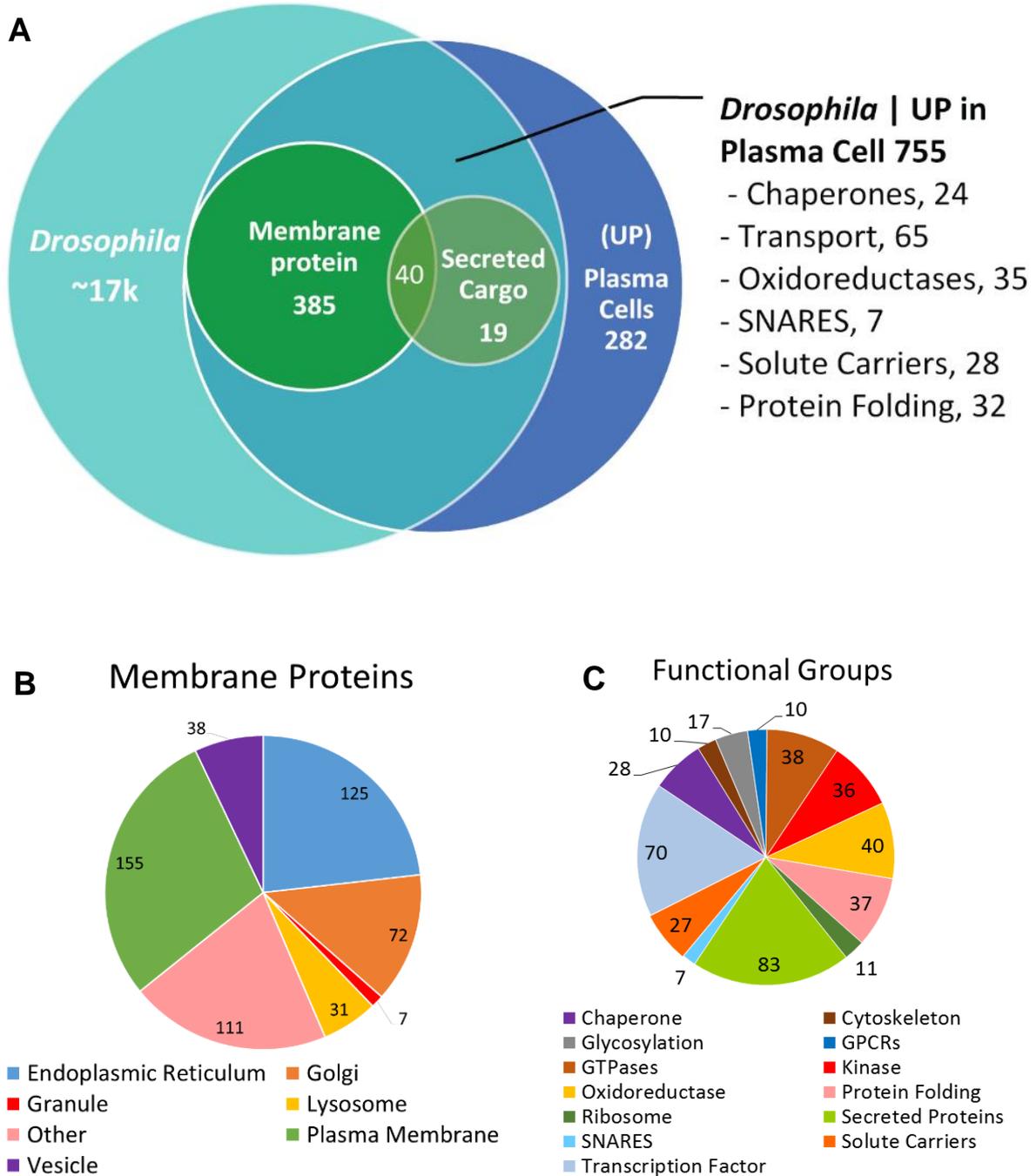
## 2.3 RESULTS

Differentially expressed genes were defined by adjusted  $p$ -value less than 0.05 and fold change greater than or equal to 2. Microarray based transcriptome profiles of mouse splenic plasma cells (SplPC) and B cells were reanalysed. From arrays of 45101 transcripts, 1873 (4.15%) were statistically significant ( $q$ -value<0.05) with at least 2-fold change. 72% of these results correlate with the equivalent RNA-Seq data [56]. By considering the respective missing genes in either platform, a total of 2042 differentially expressed genes with at least 2 FC were found. Of these, 1037 showed upregulation.

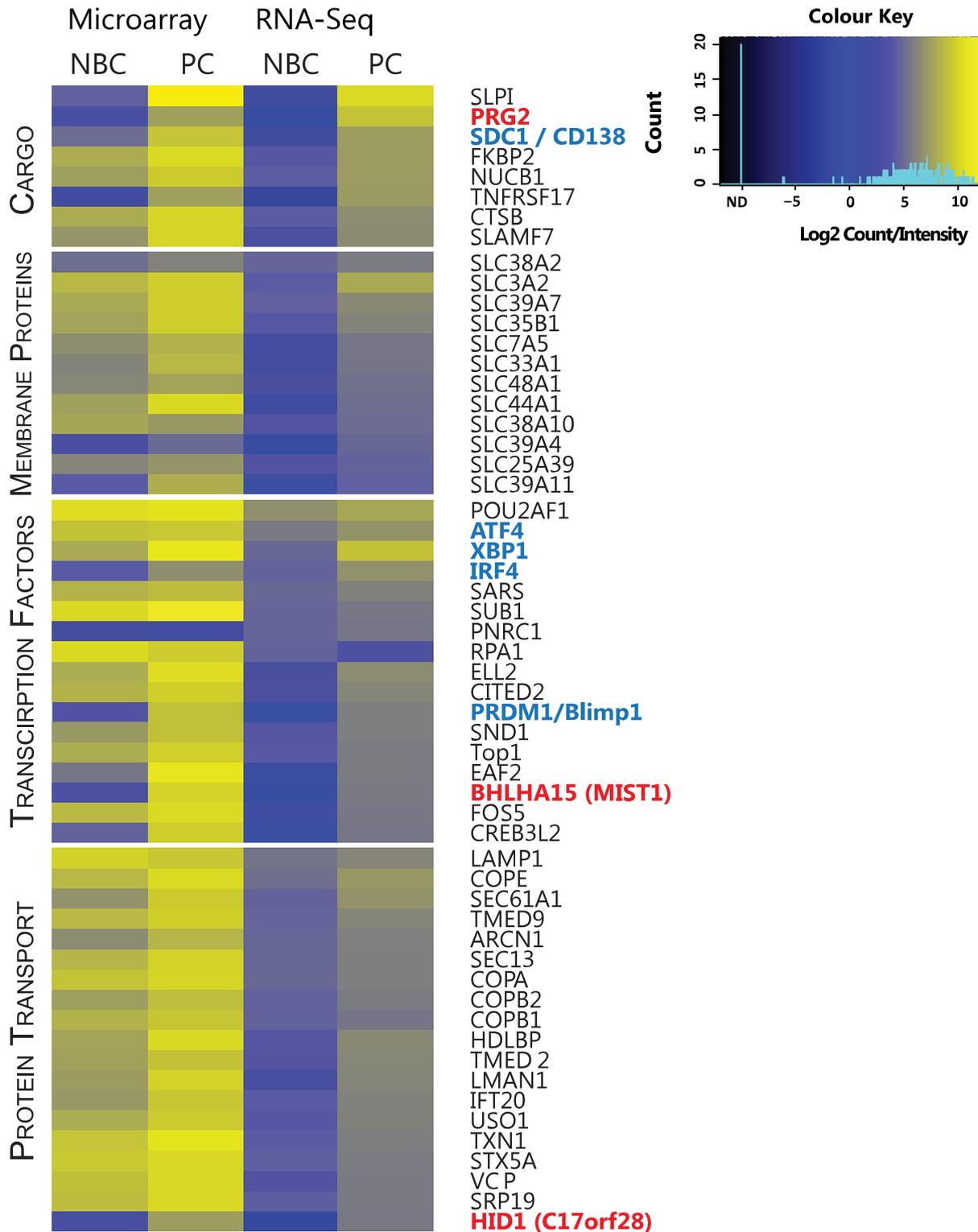
### 2.3.1.1 Functional Analysis

Counting number of genes with relevant functional annotations, showed 755 upregulated genes related to bulk transport and conserved among mice and drosophila (Fig 2-5A). At least 200 genes mapped to groups of interest, such as SNAREs, chaperones, protein transport, etc., that are known or predicted to be related to vesicular trafficking. Fig 2-5B illustrate the localisation of upregulated genes in mice ASCs. At least 190 genes were predicted to be localized to ER and Golgi and another 125 at the plasma membrane. Fig 2-5C shows the number of genes mapping to functional categories of interest.

Fig 2-6 illustrates the intensities/transcript counts of the most highly upregulated plasma cell cargo and transcription factors. Known markers of PC such as XBP1, PRDM1 and CD138 / SDC1 were upregulated and BCL6 a marker for B cells was downregulated. From functional analysis and the localization of the genes we identified the HID1 gene, which is implicated in the formation of large dense core vesicles and secretion of regulated secretory cargo, such as PRG2, which was also upregulated in ASCs [103, 104].



**Fig 2-5** | Venn diagram and Pie charts showing functional grouping of DEGs isolated in the preliminary study. **A.** Venn diagram of upregulated genes in plasma cells that are also present in *Drosophila* genome. Gene overlapping between groups are not unique, as one gene can belong to multiple groups. As *Drosophila* does not have an adaptive immune system, 282 genes that do not overlap are likely to be specific to adaptive immunity in plasma cells. **B.** Pie chart of membrane protein localization upregulated in plasma cells. Gene count is not distinct, as a single gene can be localized to multiple regions. 476 membrane proteins were upregulated in plasma cells. **C.** Functional groupings of upregulated genes. Gene count is not distinct, as a single gene can belong to multiple groups. 371 genes fall into these groups. Remaining 336 genes are ungrouped and require further analysis.



**Fig 2-6 |** Heatmap of potential genes of interest found in the preliminary study. Heatmap shows median raw intensities and transcript counts of microarray and RNA-Seq datasets respectively. Top upregulated hits among 4 categories (vesicular cargoes, solute carriers, transcription factors, vesicular transport related proteins) are given. Well characterised markers of plasma cells are indicated in blue (XBP1, Blimp1 and SDC1). In red are potentially novel genes related to antibody secretion.

## 2.4 DISCUSSION

### 2.4.1 Isolating Known and Novel Targets

Results from the transcriptomic analysis were used to screen for upregulated genes encoding known and predicted components that play a role in protein processing or transport. As we hypothesized, a considerable number of genes were shown to be related to protein synthesis, post-translational modification and transport. This shows that using NBCs as a control for secretion competent ASCs allows isolation of genes contributing to enhanced protein secretory capacity (Fig 2-5C). Among these genes we noted components of the regulated secretory pathway.

#### 2.4.1.1 Regulated Secretory Pathway in ASCs?

Plasma cells are known to utilise the constitutive pathway to secrete proteins as they are made. Unlike ASCs, specialised cells such as pancreatic exocrine cells and intestinal epithelial cells utilise another secretory pathway where cargo is secreted on demand in response to external stimuli. This is known as the regulated secretory pathway [46, 105, 106]. It has previously been noticed that plasma cells and gastric zymogenic cells are both specialised for high levels of protein secretion and arise from a non-secretory precursor; a study comparing their lineage has shown that these cells share 269 upregulated genes [107]. This included *Mist1* (bHLHa15), a transcription factor that may be involved in the formation of large secretory granules. *Mist1* is involved in the regulated secretion of digestive enzymes in gastric zymogenic cells [108]. However, Benjamin *et al* has demonstrated that *Mist1* double knockout has no effect on the secretion of IgM in plasma cells [107], indicating that MIST1 is not directly involved in constitutive secretion.

Nevertheless, the mRNA for one of the top upregulated cargo is PRG2, whose gene product is sorted into granule fractions in eosinophils and to a lesser extent in neutrophils [104, 109–111]. As a further indication of a possible regulatory secretory pathway in PCs, our

preliminary results also showed a significant upregulation in V-ATPases (ATP6V0A1 and ATP6V0A2), Zinc transporters (multiple members of the SLC39 family) and Golgi-localized calcium transporters (NUCB1, SDF4 and TMEM165). Consistent with existing literature, MIST1 was also found to be highly upregulated. Therefore, we hypothesized that these genes together may play a role in sorting regulated cargo such as PRG2 into granules and act in a regulatory secretory pathway, hitherto unheard of in plasma cells.

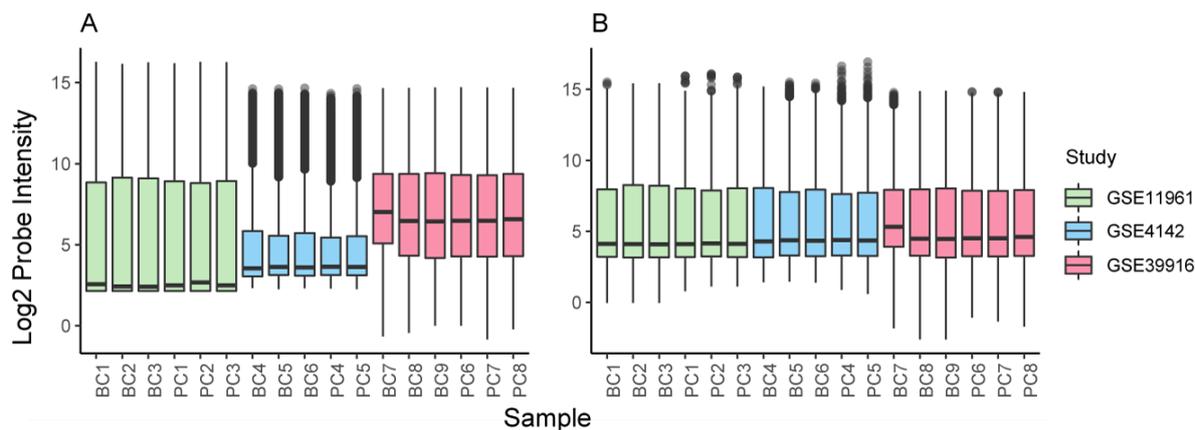
Unfortunately, when my colleagues E Rajan and AWA Aswani Western blotted PRG2 they saw no significant increase in this gene in mouse PBs compared to NBCs. To understand whether this discrepancy was due to errors in our transcriptome analysis I evaluate the preliminary methods in the following section.

## 2.4.2 Evaluation of Methods

### 2.4.2.1 Normalisation and Batch effect Adjustment

We used pre-processed data from three studies and each of them had used a different normalisation method. Arrays from GSE4142 were normalised using Robust Microarray Average (RMA), GSE11961 used GCRMA (guanine-cytosine RMA), while GSE39916 was normalised by proprietary Gene Pattern module from Illumina. Fig 2-7B shows that inter array bias was accounted for as the mean intensity within studies were fairly similar after batch effect adjustment. However, the batch affect variable was likely stronger than it should have been as cross-study differences were compounded with cross-normalisation differences.

In the process of batch effect removal, the distribution of the three studies are normalised as evident in Fig 2-7B. However, removing batch effect prior to fitting the data to the regression model introduces an extra degree of freedom in the limma *t*-test because the function does not know that the batch effect has already been removed. This is why our resultant *p*-values were too small meaning changes in gene expression had over exaggerated statistical power and thus the number of false positives were increased. Ultimately, this affected



**Fig 2-7** | Boxplot of pre-normalised data and batch adjusted mouse microarray data. **A.** Pre-normalized intensity values from 3 difference studied **B.** Pre-normalized intensity values from 3 studied after batch adjustment.

the ranking of genes whereby we had to resort to custom sorting rather than using adjusted  $p$ -values as standard.

#### 2.4.2.2 Lack of Replicates in RNA-Seq Data

Ideally, the issue with  $p$ -values in microarray will have been dealt with when performing cross platform meta-analysis using the RNA-Seq data. However, the BMPCs profiled in the RNA-Seq data had no replicates. Even though the input sample was mixture of 3 biological samples, having only 1 replicate prevents the use of normalisation techniques and  $p$ -values have little meaning as discussed in Section 2.1.2.3. Thus, differentially expressed genes in the RNA-Seq data would not have reliably filtered or ranked genes based on these  $p$ -values. If done correctly, the microarray analysis would have helped filter out noisy genes and solve the issue of ranking.

#### 2.4.2.3 Phenotype Differences

Other factors to take into consideration, is that samples from two of the studies used the C57BL/6 inbred mouse strain whereas one study used C57BL/Ka which can introduce some unwanted variations.

Furthermore, while the authors of the 3 studies labelled the test array as plasma cells, they all used somewhat different markers for classifying plasma cells. Meaning the three

different studies may have been studying plasma cells or plasma blasts at differing levels of maturity.

#### 2.4.2.4 Functional Annotation

The list of terms we retrieved from SP-PIR and GO, had both computationally predicted and experimentally validated hits. As such, some genes could have been assigned to the wrong category. Annotation and counting number of recurring functional terms do not take into account the importance or rank of a gene in a differential expression study. While this did not affect our hypothesis related to regulated secretion, we acknowledge that the functional analysis can be improved using some form of gene set enrichment analyses [112, 113].

## 2.5 CONCLUSION

Despite the errors and limitations mentioned above, known markers of plasma cells were differentially expressed in this study. This is likely due to meta-analysis of the microarray and RNA-Seq data, which will have removed a considerable number of false positive genes arising from the microarray analyses as these would show contradictory regulation in the two platforms. This strengthens the argument for cross-study meta-analysis as it prevents errors. Nevertheless, noise introduced due to batch effect adjustment of differently normalised data will have suppressed hits, which would otherwise be present if we used the processed RNA-Seq results on their own. To avoid such errors and improve accuracy of our analysis, in the following chapters, we forego the use of processed data and begin our analysis from scratch to streamline the data processing methodology. Furthermore, we calculate enriched functional groups based on gene ranking and also utilise our own curated database of functional protein complexes to avoid the inclusion of poor-quality annotation.

It is important to note that the errors in the microarray analysis did not explain the presence of PRG2 as this gene was strongly upregulated in the RNA-Seq analysis as well. The web tool Amazonia! allows users to view gene expression in human ASCs and NBCs [72].

Interestingly, PRG2 was not detected in the human transcriptome despite the microarray having probes hybridizing to this gene. Therefore, we theorised that the upregulation of PRG2 mRNA could potentially be some common form of contamination in the mouse RNA isolation and treatment protocol or a transient artefact in the mouse plasma cell transcriptome. Therefore, cross-species meta-analysis of the plasma cell lineage may potentially solve this issue.

# Chapter 3

## CROSS SPECIES MICROARRAY ANALYSIS

## 3 CROSS SPECIES MICROARRAY ANALYSIS

### 3.1 BACKGROUND

Post antigen activation, naïve B cells (NBC) of the immune system show dramatic increase in their ability to secrete antibodies. Preliminary comparison of the mice transcriptome of non-secreting NBCs to their antibody secreting progeny showed hundreds of genes upregulated in the differentiated antibody secreting cells (ASCs) that were related to protein processing and trafficking. However, we also identified a highly upregulated cargo known to be transported by the regulated secretory pathway whose protein product could not be detected in mice. Potential sources of errors in the merging procedure did not explain the presence of the confounding results as they were present across comparable RNA-Seq and microarray studies in the mice species. Upon further investigation, we noticed that these genes were absent in comparable human transcriptomes.

#### 3.1.1 Of Mice and Men

The mouse model is used extensively in biomedical research due to their phylogenetic closeness to humans and the ease of breeding and handling in laboratories [114]. Protein coding regions of mice and human share about 85% similarity with gene identity ranging from 60 to 99% while non-coding regions are much more dissimilar [115, 116]. Higher genetic similarity in coding regions between mammals suggest conservation of protein function across species. This property is particular useful in high throughput functional genomics for cross-referencing the reproducibility of candidate genes [63, 64]. Therefore, matching the regulation of genes across microarray profiles of mice and human ASCs may aid in filtering out erroneous results such as those identified in the previous chapter.

### 3.1.2 Review of Microarray Normalization

In the previous chapter, we analysed Affymetrix arrays from different studies and came to the conclusion that normalisation methods for each array must be the same to allow for accurate quantification. However, each study had used a different technique and we wanted to know which procedure would be optimal for our analysis. Therefore, in this section we discuss the design and functionality of Affymetrix arrays and review the three major normalisation methods utilised for them. These normalisation methods are: Microarray Analysis Suite 5.0 (MAS5), Robust Microarray Average (RMA) and its refined counterpart Guanine Cytosine Robust Microarray Average (GCRMA) [117–119].

#### 3.1.2.1 Affymetrix Gene chips

Affymetrix systems utilise probes that perfectly match genomic regions called perfect match (PM) probes. Probe intensities in Affymetrix arrays always yield positive results due to optical noise and non-specific hybridization when there is no gene fragments present. Mismatch (MM) probes are used to detect these background noise in Affymetrix arrays. These probes are identical to PM probes except for a single nucleotide in the middle of the sequence, which is replaced by its complement [117]. The purpose of these probes is to allow for calculation of the true signal by subtracting the MM probe signals from the PM signals. Affymetrix Genechips typically have 11 to 20 different probe clusters that hybridize to the same sequence evenly spread throughout the chip. These are collectively called “probe sets”. Probe intensities have to be summarized to retrieve a representative value for each unique mRNA species.

#### 3.1.2.2 MAS5

##### 3.1.2.2.1 Background Correction & Summarisation

MAS5 corrects for non-specific binding and summarises probe set level intensities in a single step. It uses the classical method of subtracting MM signals from PM intensities to calculate true signals and summarises these signals based on a robust average. Robust averages utilise

statistics based on median instead of the arithmetic mean to ensure that the probe set level intensities are insensitive to extreme outliers

#### 3.1.2.2.2 Criticisms

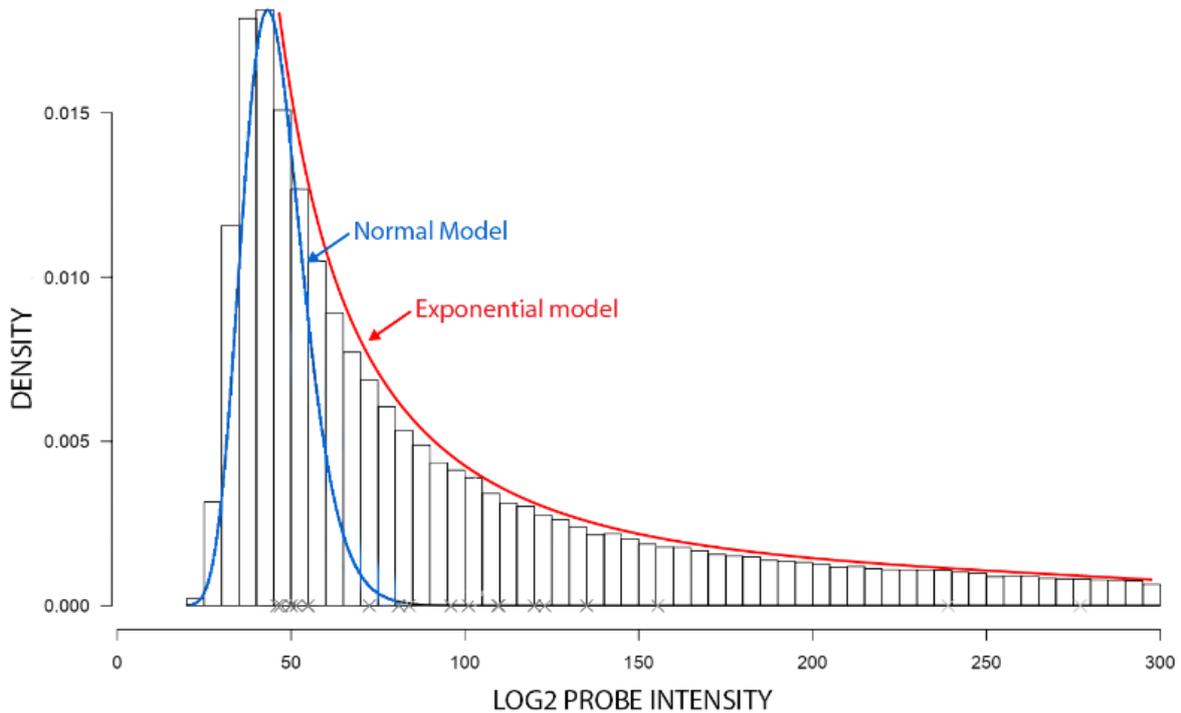
Optical signals can only be present (positive values) or absent (zero). Often MM probes bind to relevant transcripts which can lead to the MM signal being larger than the PM intensities. This results in nonsensical negative signals when MM is subtracted from PM. The MAS5 algorithm discards these PM and MM probes altogether to avoid this issue. Unfortunately, these signals can account for up to a third of all probes and lead to considerable loss in probe level information. Furthermore, subtraction of MM signals can introduce unwanted noise at low intensities [117].

Array bias arise from different samples having different overall intensity distributions. Unfortunately, MAS5 algorithm does not account for these differences as each array is normalized sequentially in an independent manner.

#### 3.1.2.3 RMA

##### 3.1.2.4 Background Correction

For these reasons, RMA is now widely recommended for Affymetrix array analysis. RMA assumes that the MM derived background noise within an individual array follows a normal distribution to the left of the log<sub>2</sub> transformed probe intensities, while PM probe signals follows an exponential model (Fig 3-1). Based on this assumption, subtracting an overall background estimate fitted to a normal model from the exponential model-fitted intensities of each probe yields a background corrected result that always remains positive [118]. Simply put, a background noise threshold is calculated and PM probes below this threshold are removed. The primary criticism of the background correction step of RMA is that it fully disregards MM signals and fails to utilise all available information on non-specific binding.



**Fig 3-1** | Density distribution curve of Log2 transformed probe intensities (Adapted from Irizarry et al). This plot illustrates the model fitting used for background correction in RMA. Typically, most probes remain unhybridized and show 0 or very low optical intensity. These can be classified as background noise represented by the normal model (in blue) and appears to the left of the curve in the low intensity region, while the exponential model (in red) represents real PM probe signals. Background correction in RMA, works by subtracting an overall background estimate from the exponential fitted PM values. In other words, RMA filters out PM probes that show optical intensity below a noise threshold.

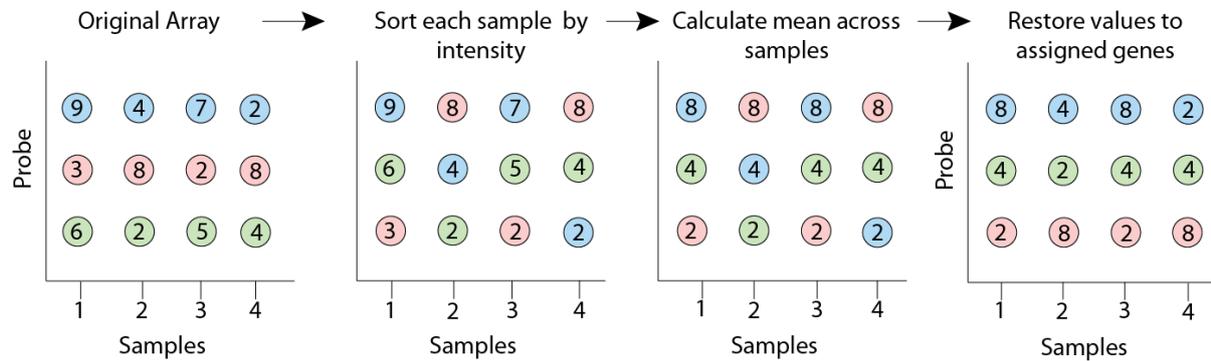
### 3.1.2.5 GCRMA

#### 3.1.2.5.1 Background Correction

A modified form of RMA, known as, GCRMA overcomes this issue. It utilises probe sequence information to calculate specific and non-specific binding affinity of MM signals, meaning, MM probes that uniquely bind to relevant transcripts are reassigned as PM. This prevents the generation of negative intensities and retains both PM and MM information, which would have otherwise been discarded in MAS5 and RMA. The inter array normalisation and summarisation step of both RMA and GCRMA are identical and is discussed below [119].

## 3.1.2.6 RMA/GCRMA

## 3.1.2.6.1 Normalisation across arrays – RMA/GCRMA

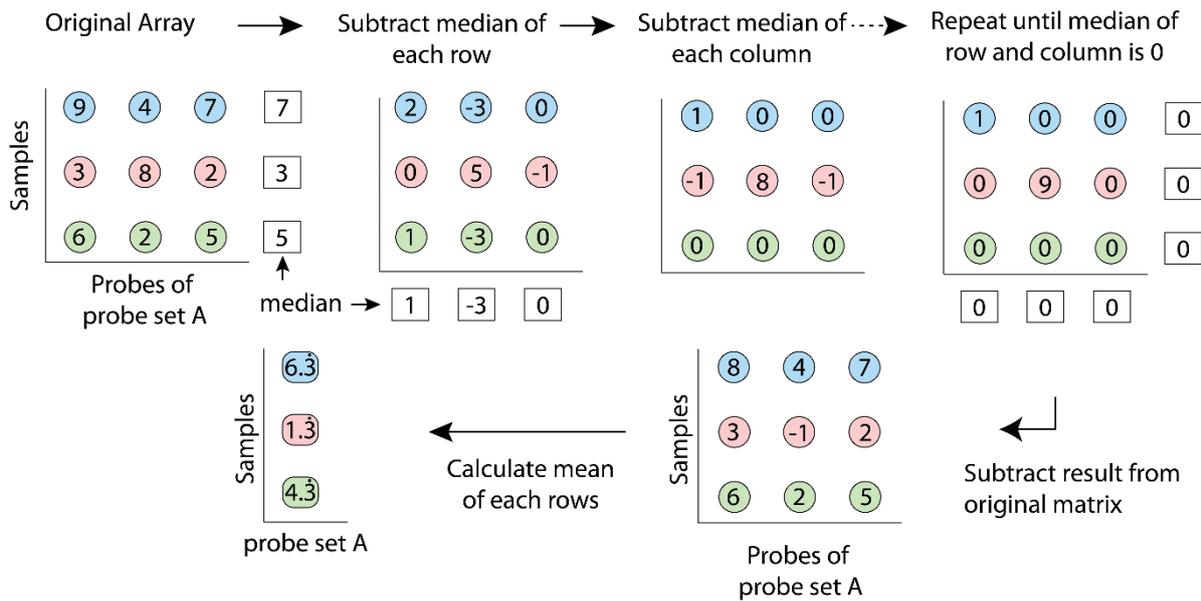


**Fig 3-2** | Principle of quantile normalisation. To make arrays comparable quantile normalisation assumes that the distribution of intensities across arrays must be the same. In this example rows are genes/probes and columns are samples. First each column is sorted by size (and their gene assignment is remembered by the algorithm). The sorting occurs such that the first row shows maximum values found in each sample while last row shows minimum values found in each sample. Then the mean of each row is calculated and replaces the original values. Then these mean intensity values are returned to their original position. This method now ensures that if you calculate the average of a column (sample or array) it will be the same as all the other columns. Thus, direct comparison of genes across arrays are possible.

When biological replicates are measured using microarrays, array bias arises due to affinity differences, faulty/dead probes, or minor batch differences. These types of systematic errors are compensated for in RMA/ GCRMA by quantile normalisation. As illustrated in Fig 3-2, this procedure ensures that the range of values within each sample is the same. Background corrected intensity values are arranged in a matrix where each row is a probe, and each column is a sample. Quantile normalisation first sorts each sample column by size while remembering which probe each of the intensity values belonged to. The values of each resultant rows are then replaced by their average. Finally, the corrected values are returned to their original gene rows to give the normalised matrix.

Simply put, if array A is naturally scaled by 2, and array B is scaled by 7.5, the arrays are not comparable. Without knowing these scaling factors, quantile normalisation will adjust the distribution of the two arrays such that the mean and interquartile range of the two arrays are fairly similar. This allows for direct intensity comparison.

## 3.1.2.6.2 Probe Set Summarisation – RMA/ GCRMA



**Fig 3-3** | Principle of median polish. This procedure adjusts the intensity of probes belonging to probe set for outliers to prevent a skewed arithmetic mean. This process determines how much a probe differs from the median of all probes within an array and the median of all probes within a probeset. This is done by repeatedly subtracting these medians from each probe intensity until most probes are equal or close to 0, while outliers remain considerably larger. By subtracting these differences from the original matrix, outlying probes are adjusted. This ensures the final mean of all probes is minimally affected by outliers.

As discussed in Chapter 1, Section 2.1.2.1, probes of the same gene / transcript are scattered throughout an Affymetrix arrays. The intensity values from these probes have to be aggregated and summarised. In the summarisation step, the log transformed, background corrected and normalised intensity of each probe is used to obtain a single intensity value for each probe set. The summarisation process used in RMA/GCRMA is called “median polish” [118].

Median polish is a method where each column (probe) and each row (sample) of a matrix are normalised to their median (Fig 3-3) [120]. Take for example a matrix of processed intensity values arranged in a matrix where each column is a probe belonging to probe set A, and each row is a sample. The median intensity of each probe across samples is subtracted from the original intensities. Then the median of each samples across a probeset is subtracted from the result. This is repeated a maximum of 5 times or until the median converges, i.e. the median of each row and column is 0. The resultant median polish residuals are subtracted from the original

matrix of probe intensities and finally the average of each row is calculated to yield the most representative intensity of a given probe set in each sample.

Simply put, this process calculates the median intensity of all probes within a probeset by first adjusting for outliers, judged based on their distance from the median probe intensity.

#### 3.1.2.6.3 Criticisms and Recommendations

RMA/GCRMA is preferred over MAS5 because it provides fewer false positives, reduces noise at relatively low intensities and produces more reproducible fold changes. However, RMA/GCRMA assumes that all arrays have equal distributions. This can confound hidden structures in data that can be essential for downstream quality control such as identifying batch effects [117, 119]. Therefore, it is recommended to utilise MAS5 for downstream quality control and then redo the processing step using RMA/GCRMA for differential expression analysis [121].

### 3.1.3 Challenges of Meta-Analysis in Microarrays

As mentioned previously, in order to improve the robustness of our results, we intend to carry out meta-analysis of mice and human ASCs. Pooling RNA-Seq experiments from different studies is relatively simpler as data can be combined as long as the same genome assembly and gene annotation or same gene repository for cross species comparison is used. Unfortunately, DNA microarray platforms have distinct sets of probe sequences identified by distinct probe IDs. This means pooling probes from two different microarray platforms based on probe ID alone, is impossible. Furthermore, redundant probes mapping to coding as well as non-coding regions of the same gene complicate cross species comparison in microarrays.

The most common and straightforward way to pool microarray data is to use gene ID, gene symbol or other gene level annotations to merge the datasets. However, this approach can lead to loss of relevant data and introduce biases into the analysis.

### 3.1.3.1 Batch Effect Estimation

#### *Microarrays*

A number of batch effect removal methods exist for microarrays as reviewed by Lazar *et al* [122]. However, these methods require over 25 samples for batch effect estimation, while we only have a total of 18 samples across 3 studies. The ComBat method used in the previous chapter requires a minimum of 5 samples and is a good tool for estimating batch effect when the source is not known [96]. However, this method has been reported to overexaggerate  $p$ -values when used with limma's moderated  $t$ -test [123]. Therefore, we utilise the blocking function available in the limma R package as it also caters to lower sample sizes.

This method allows our desired groups, ASC and NBC, to be compared in separate “blocks” [97]. It thus allows batch effect adjustment while preventing errant degrees of freedom that occurs with the ComBat method, as discussed in Chapter 2, Section 2.4.2.1. An alternative to using the blocking function in limma is to estimate batch effect using ComBat first and then specify covariates in limma for blocking. However, this has been shown to produce somewhat deflated  $p$ -values [123].

### 3.1.3.2 Inconsistency in probe design

The library of probes used in each platform depends entirely on the manufacturer, which is generally influenced by the reference genomes and transcript annotations available at the time. We cannot avoid the underlying limitation that microarrays cannot bind to all genes and gene isoforms. Thus, when different platforms are matched based on gene identifiers, lack of consistency in expressed genes can be purely due to absence of probes for those genes or gene isoforms. Manufacturer's discretion in designing probes also means that probes designed for the same gene will map to different regions in different platforms [61]. As a result, it is hard to determine whether difference in signal in a one-to-one mapping is purely due to platform

variation or due to the probes mapping to distinct differentially expressed transcripts / gene isoforms.

### 3.1.3.3 Probes without annotations

Probes with missing gene symbol or other identifiers may be intronic, pseudo genes or novel genes yet to be fully annotated. Therefore, merging data where probes have missing gene information, can lead to loss of information and potentially the removal of genes of interest.

### 3.1.3.4 Redundant Probes

Redundant probes exist to consolidate the expression pattern of a gene. However, the curse of redundant probes manifests itself when expression pattern for different probes for the same gene do not agree with each other. In one-to-one mapping, only one probe per gene is used, usually the first one that appears on a gene matrix. This leads to gene signal bias, as incorrect or inflated signals can be piped into downstream analysis. Contradictory regulation, commonly seen in microarrays, is found to be caused by probes that: (a) map to intergenic or intronic regions; (b) map to multiple transcripts; or (c) contain outdated or incorrect annotations [61].

### 3.1.4 Reannotation

Realigning and reannotating probes to the updated reference genome can overcome some of these issues. With microarrays become an aging technology, most tools developed for microarray probe reannotation have become out of date or unusable due to lack of updates [124–126]. Saka *et al* has recently proposed a protocol for Affymetrix probe reannotation but has not provided a tool to streamline the process [127]. The Re-Annotator pipeline, last updated in 2014, relies on stepwise alignment with Burrow-Wheeler Aligner (BWA) of probe sequences to a custom-made *in silico* mRNA database [125]. Successful alignments are trimmed for mismatches, filtered to retain coding strands and then updated with gene annotations. The tool also specifies whether probes are intergenic, exonic, intronic, etc. For this project, this information is particularly important as it allows us to give weight to exonic

regions, as we are mainly interested in protein components of membrane trafficking. Unfortunately, the sequence aligner used in this pipeline (BWA), is inherently unsuitable for short oligomers such as microarray probes so Re-Annotator allows mismatches to account for this. While this works for probes sequences of 50 or more base pairs, shorter Affymetrix probes of 25bp length show poor alignment. Since this pipeline was written, better aligners specific for short sequences have become available [128]. In this chapter we update this tool for the purpose of updating our microarray annotations.

### 3.1.5 Functional Analysis

As mentioned in the previous chapter, hundreds of genes are differentially regulated in ASCs compared NBCs. It is difficult to pinpoint which genes are important if we consider their function one at a time. Looking at sets of genes collectively instead of in isolation can improve the statistical power of the analyses. If a set of gene with a common biological function is consistently regulated in the same direction, then we can derive biological meaning from it. Various databases of functionally relevant gene sets exist such as KEGG and Gene Ontology. The EnrichR tool allows mining of these databases and provides access to a useful collection of perturbation experiments that lets users determine whether a set of genes has previously been shown to be expressed as a result of a perturbation.

EnrichR calculates the statistical validity of gene set enrichment for a functional term or gene perturbation using rank-based ranking [113]. This method utilises iterative Fisher Exact Test, which calculates the probability of a ranked set of genes overlapping with the members of a known functional group by random chance. In order to adjust for false discovery, the Fisher Exact test is run repeatedly for many random gene sets to calculate a mean rank as well as the standard deviation from the expected rank. The  $p$ -value is then adjusted based on this deviation [112].

#### 3.1.5.1.1 Limitations of Functional Enrichment analysis

The KEGG database is a concise database of well-known genes acting in largely experimentally characterised pathways. While this database is great for understanding the biology behind well characterised genes, it's less useful for exploratory analysis for candidate selection. Therefore, Gene Ontology (GO) databases, containing both predicted and experimentally validated hits, are typically used to analyse the functional characteristics of differentially expressed genes.

##### *Summarising GO Terms*

Unfortunately, one of the caveats of GO analysis is that it typically outputs many overlapping and redundant GO terms or functional categories. The REVIGO tool uses machine learning algorithms to find highly similar GO terms and thus reduces redundancy in GO analysis results. Although, these machine learning algorithms summarise GO terms considerably well, some discrepancies remained that require unavoidable manual correction.

Although the REVIGO tool tells us what GO terms fall under the same overall category, it does not solve the issue of differing number of genes, some overlapping, some not, appearing under GO terms falling under the same category. In order to fully summarise similar GO terms, genes belonging to these groups have to be combined, duplicates have to be removed, and unique genes have to be counted to allow for visualisation. Unfortunately, this data mining process, though simple in concept, is laborious, time consuming and prone to user errors when attempted via excel or similar user-friendly software. Therefore, in this chapter we create a data mining tool to refine the results of GO Ontology analysis.

##### *Refining Pathway Enrichment Results*

Coexpression or perturbation databases, such as those generated from transcription factor or kinase perturbations, allow for complex pathway analysis. However, to make sense of these databases it is important to (a) check that genes known to be affected by the perturbation were

regulated in the same direction as the regulation seen in the user's data (b) check that perturbed gene showed equivalent regulation in the users data, i.e. if knockdown of transcription factor X is known to result in the upregulation of gene set Y, we determine this perturbation to be potentially relevant if transcription factor X is among our differentially downregulated genes and a large portion of gene set Y is among our upregulated genes. Unfortunately, these checks cannot be carried out by EnrichR as the tool cannot distinguish between upregulated and downregulated genes.

In addition, like GO ontologies, perturbation terms sometimes overlap and affected genes require summarisation. For example, knockdown, knockout, inhibition of the same gene can be loosely categorised as “downregulation”. Therefore, in this project we create a tool to automate the mining and refinement of gene perturbation or co-expression experiments using specific EnrichR databases.

### 3.1.6 AIMS & OBJECTIVES

To date, a number of research groups have generated high throughput microarray profiles for of NBCs, PCs and their intermediates in mice and human model [55–58]. However, a comprehensive cross species or cross platform meta-analysis of the transcriptomes have not been performed. Therefore, we aim to leverage these transcriptomes accumulating on public repositories to:

1. Build a bioresource for the identification of reproducibly changing genes by combining microarray data from across studies, species and platforms.
2. Perform robust functional enrichment analysis to identify conserved genes related to protein processing and transport.

In the process of completing these objectives we create a set of bioinformatics tools that:

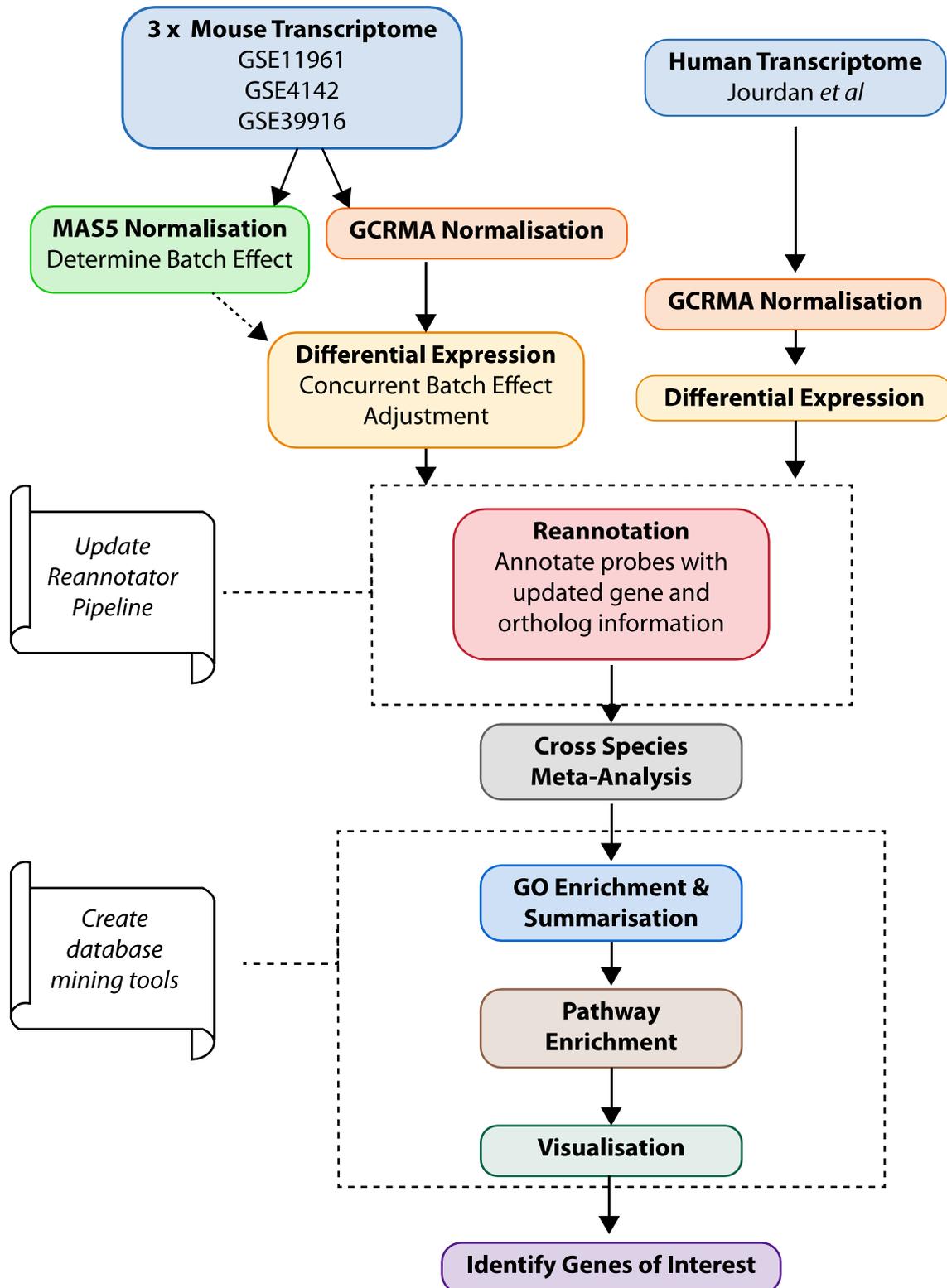
3. Refine and summarises results of GO Ontology analysis.
4. Automate the mining and refinement of gene perturbation or co-expression experiment databases and output a result suitable for visualisation.

Furthermore, we update an existing tool such that:

5. The Re-Annotator pipeline can accurately update the annotation of microarrays with relatively short probes.

## 3.2 METHODS

### 3.2.1 Workflow



**Fig 3-4** | Workflow of Methodology used for Cross-Species Meta-Analysis of Microarray Data.

## 3.2.2 Mouse B Cell Lineage

### 3.2.2.1 Data Source

#### 3.2.2.1.1 GSE11961

Kaji *et al* isolated naïve, follicular B cells from the spleen of unstimulated, female, C57BL/6 mice aged between 8-10 weeks using Magnetic-activated cell sorting (MACS) system. Cells were negatively selected for CD5, CD11b, CD95, CD138, CD43, Gr-1, TER119, F4/80, CD3 and CD90 using biotinylated antibodies (Ab) followed by sorting for B220+, AA4.1-, CD21int, CD23+ fractions [57].

The research group generated T-cell dependent ASCs by immunising equivalent mice models with 4-hydroxy-3-nitrophenylacetyl conjugated to chicken gamma globulin (NP-CGG) and aluminium adjuvant. 7 days post immunisation, splenic PCs (SplPC) were isolated by the MACS system [57]. This involved negative selection using biotinylated Abs against IgM, IgD, CD5, Gr1, NK1.1, DX5, TER119, F4/80, CD3 and CD90, followed by sorting for NIP-binding B220+, IgG1low, CD138+, Igλ+ fractions [57].

Kaji *et al* extracted whole cell RNA using triZOL (Invitrogen), amplified the RNA using MessageAmp aRNA kit (Ambion), and labelled them with biotin using Bioarray High Yield RNA Transcription Labelling kit (Enzo Life Sciences)[57]. This was followed by cRNA fragmentation and hybridization to GeneChip Mouse Genome 430 2.0 Array (MG430.2A) as instructed by Affymetrix [57].

#### 3.2.2.1.2 GSE4142

Luckey *et al* isolated naïve, follicular B cells from the spleen of unstimulated, male, C57BL/Ka mice aged between 4-8 weeks using double Fluorescence-activated cell sorting (FACS). This involved sorting for B220+, CD23+, IgM-low, IgD+, Igλ+, CD11b-, Ter119-, Gr-1-, CD3-, CD4-, CD8- and CD5- fractions [129].

In order to generate ASCs, Luckey *et al* immunised equivalent mice with NP-CGG/alum. SplPCs were harvested 7 days after immunisation. These ASCs were sorted for B220-low, IgM-low, Igλ+, CD138+, NP-PE+, CD11b-, Ter119-, Gr-1-, CD3-, CD4-, CD5-, CD8- and PNA-negative cells fractions [129].

Luckey *et al* extracted whole cell RNA using triZOL (Invitrogen), amplified the RNA using Arcturus RiboAmp kits and labelled them with biotin using Affymetrix IVT Labelling kits [129]. This was followed by cRNA fragmentation and hybridization to MG430.2A as instructed by Affymetrix [129].

#### 3.2.2.1.3 **GSE39916**

Benson *et al* isolated naïve B cells from C57BL6/J mouse strain. Age or gender of mice has not been declared by the author. NBCs were isolated from spleen using MACS system and sorted for B220+ and CD23+ fractions [130].

In order to generate mature PCs, Benson *et al* immunised mice on day 0 with keyhole limpet hemocyanin (KLH) and Freund's adjuvant, rested them for 21 days and re-challenged pre-immunised mice with KLH. At day 28, bone marrow plasma cells (BMPCs) were isolated using MACS system and sorted for B220- and CD138+ fractions [130].

Benson *et al* extracted RNA using Ambion RNAqueous Micro kit, amplified the RNA using NuGen OvationRNA Amplification V2. This was followed by cRNA fragmentation and hybridization to MG430.2A as instructed by Affymetrix [130].

#### 3.2.2.1.4 **Treatment Differences across Studies**

The studies discussed so far use different kits for RNA extraction, purification and amplification. Table 3-1, summarise the major phenotypic differences and similarities between the three data sets. As these experiments utilise a number of disparate markers, we focused on common markers shared across studies to define ASCs and NBCs. This is mainly

**Table 3-1** | Phenotype of Mouse microarray profiles of PC cell Lineage from different studies

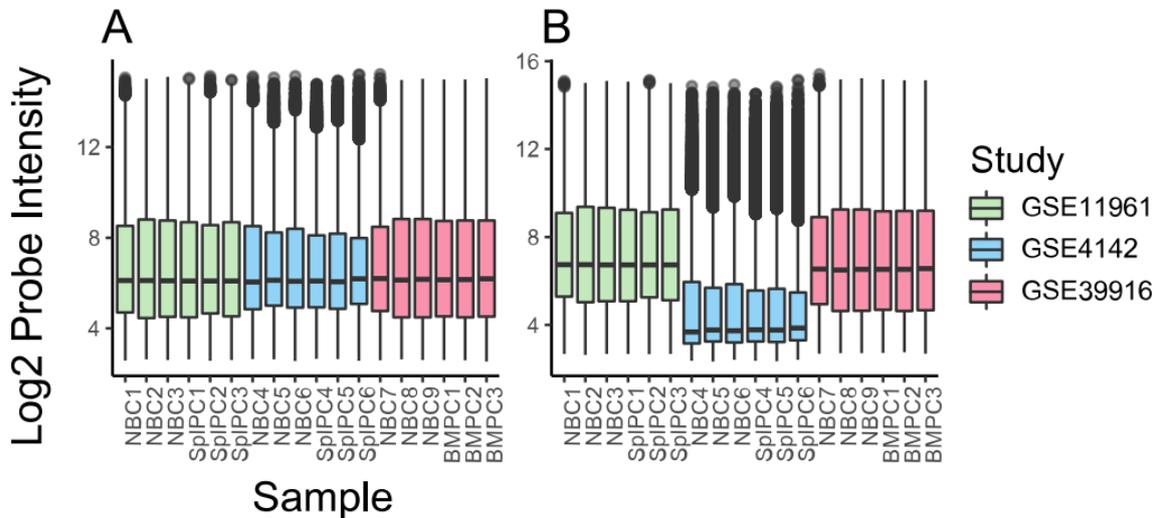
Study	Type	Tissue	Gender	Strain	Day	Stimulus	Minimal Marker
GSE11961	NBC	Spleen	Female	C57BL/6	0	-	B220+ CD138- CD23+
GSE11961	PC	Spleen	Female	C57BL/6	7	NP-CGG/alum	B220+ CD138+
GSE4142	NBC	Spleen	Male	C57BL/Ka	0	-	B220+ CD23+
GSE4142	PC	Spleen	Male	C57BL/Ka	7	NP-CGG/alum	B220-low CD138+
GSE39916	NBC	Spleen	Unknown	C57BL/6	0	-	B220+ CD23+
GSE39916	PC	Bone Marrow	Unknown	C57BL/6	28	KLH	B220- CD138+

to prevent confusion between PC and PB, which are used interchangeably by some authors.

We have previously discussed CD138, a common plasma cell marker. Contrary to CD138, B220 or the high molecular mass isoform of CD45 is a marker of unstimulated B cells. This marker is gradually replaced by the low molecular weight isoform, B200, as NBCs mature to antibody secreting plasma cells [131]. Therefore, B220 is a marker for identifying the maturity of plasma cells, with B220+ and CD138+ plasma cells likely to be less mature than B220- and CD138+ fractions. CD23+, on the other hand, differentiates between naïve and memory B cells [132].

### 3.2.2.2 Pre-Processing

As discussed in Section 3.1.2, RMA or GCRMA is the best normalisation method for differential expression analysis. Due to the nature of quantile normalisation, inherent to these algorithms, normalising all three studies together using RMA/GCRMA assumes that cross study distribution of data is the same (Fig 3-5A) and masks underlying inter study differences. Normalising each study individually highlighted similarities in the distribution of GSE11961 and GSE39916, which then differed from GSE4142 (Fig 3-5B). This difference is best explained by strain differences between C57BL/6 and C57BL/Ka (Table 3-1).

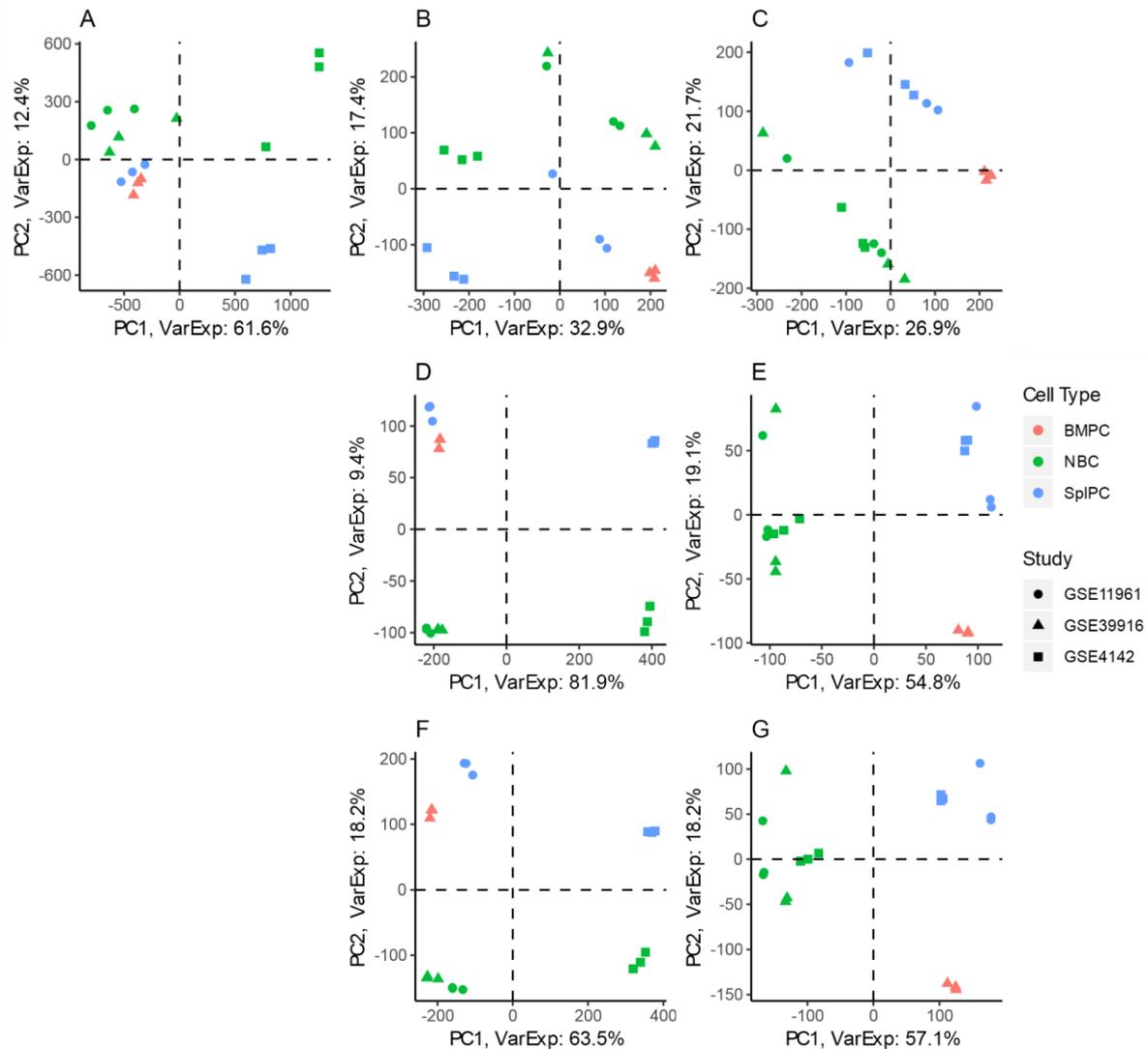


**Fig 3-5** | Box plot probe intensities in pooled vs separately normalised cross-study mouse arrays. **A.** RMA normalisation on “merged” studies. We pooled intensity data from all three experiments into a single matrix or table and carried out RMA normalisation. This method adjusts the data such that the mean of each array across studies is the same. This hides underlying cross study differences. **B.** RMA normalisation on individual studies. Arrays from each study were normalised separately. From the boxplot, it is apparent that the mean intensity of GSE11961 and GSE39916 groups away from the mean of GSE4142. As the former two studies are from a separate strain to the latter, from this data alone we could presume that the batch effect is primarily cross-strain.

### 3.2.2.2.1 Batch Groups

Our experimental design involves multiple cell types from different studies. Therefore, we estimated cross-study correlation, i.e. the estimated size of the cross-study batch effect, using the `duplicateCorrelation` function in `limma` R package. This was considerably large i.e. 94.4%, when each study is considered a separate batch. Based on the Fig 3-5B, it is logical to assume that the two experiments carried out using the C57BL/6 mouse strain should be considered a single batch. However, this assumption would ignore protocol differences between the two studies and other unwanted variables discussed in Section 3.2.2.1. Furthermore, the computed cross-strain correlation was 94.8%, which is not significantly different from the cross-study correlation.

In order to determine the correct batch grouping we visualised the multidimensional clustering of samples via principal component analysis (PCA). Raw intensities (Fig 3-6A) were normalised by MAS5 algorithm in order to determine the structure of the batch effect. Fig 3-6B



**Fig 3-6** | PCA plots indicating sample clustering as a result of MAS5, RMA and GCRMA normalisation in cross-study, mouse microarray data. **A.** Raw intensities show that SplPC and BMPC of GSE11961 and GSE39916 studies cluster very close to NBCs, while difference between SplPC and NBC of GSE4142 is immediately obvious. **B.** After MAS5 normalisation, we see similar distance between the ASCs and NBCs on the y-axis. However, each study cluster separately on the x-axis. This suggests the presence of cross-study batch effect. **C.** Cross-study batch correction allows ASCs to cluster away from NBCs; however, the variation is poor (only 26.9%). **D.** RMA normalisation, and **F.** GCRMA normalisation suggest cross-strain differences based on the clustering on the x-axis. **E, G** Cross-Study batch correction, based on MAS5 quality check, improve final variation from 26.9% to 54.8 – 57.1%. We opt for GCRMA as it is likely to give somewhat superior differential expression results based on higher ASC vs NBC variation.

shows that the primary variation across samples according to MAS5 normalised samples were between the three separate studies rather than strain as there was limited overlap between GSE39916 and GSE11961. Consistent with literature, RMA and GCRMA normalisation (Fig

3-6D, F) exaggerates strain differences by artificially altering the intensity distribution and is a poor choice for identifying batches. Therefore, we chose to adjust for three way cross-study differences as it takes into account strain as well as other possible confounding variables such as reagent, gender and age of mice.

#### 3.2.2.2.2 Normalisation & Batch Effect Adjustment

We modelled batch adjusted probe intensities using the `removeBatchEffect` function from `limma` R package. MAS5 is an ideal pre-processing method for quality control. However, after batch effect adjustment our factors of interest, i.e. NBCs versus ASCs, only accounted for 26.9% of the variation in the data using this normalisation method (Fig 3-6C). On the other hand, cell type differences accounted for 54.8% and 57.1% of the variation when RMA and GCRMA normalisation was used, respectively (Fig 3-6E, G). As GCRMA best represented cell type differences and is considered a superior method to RMA (as discussed in Section 3.1.2), we chose to utilise this method followed by cross-study batch effect adjustment for downstream differential expression analysis.

#### 3.2.2.2.3 Experimental Design

We utilise the blocking method provided by the R `limma` package to adjust for cross-study batch effects. This allows our desired variable (ASC vs. NBC) of each study to be calculated separately or in “blocks”, while allowing for multiple group comparison i.e. comparison of differential expression between different ASCs [97]. This method allows batch effect adjustment while preventing errant degrees of freedom that occurs with the `ComBat` method as discussed in Chapter 2, Section 2.4.2.1.

From Fig 3-6G, it is evident that SplPCs from two different studies clustered close together and away from BMPCs. This is a clear indication that the 7-day old SplPCs utilised in GSE4142 and GSE11961 are likely to share similar phenotype

**Table 3-2** | Experimental Design - MG430.2A

Study	Type	Block	Replicates
GSE11961	NBC	1	3
GSE11961	SplPC	1	3
GSE4142	NBC	2	3
GSE4142	SplPC	2	3
GSE39916	NBC	3	3
GSE39916	BMPC	3	3

in terms of PC maturity and should be grouped away from the 28-day old BMPC which differs in both tissue origin and age (Fig 3-6BC). Therefore, instead of treating ASCs from all three studies as a single group we assign two groups: SplPC and BMPC to be compared to the NBC control. The experimental design for calculating differential expression is shown in Table 3-2.

### 3.2.2.3 Differential Expression Analysis

The processed data was fitted to the limma linear regression model while taking into account the inter-study batch effect. F-statistic for multi-group comparison was calculated using empirical Bayes moderation provided in the limma R package. Computed *p*-values were adjusted using Benjamini & Hochberg method for global false discovery [133].

A threshold of 2-Fold Change (FC) was enforced to differentiate up/downregulated genes from those showing no change. However, in order to avoid removal of potentially important genes, we retained genes that met our differential expression (DE) criteria in at least one cell type, in this case genes upregulated or downregulated in BMPC, SplPC or both compared to NBCs. For intraspecies analysis we enforce an adjusted *p*-value cut-off of 0.05. For cross species analysis, we do not enforce *p*-value cut-offs until after data pooling.

### 3.2.2.4 Re-annotation

As discussed in 3.1.4, when reannotating probes, BWA utilised in Re-Annotator pipeline is unsuitable for aligning short reads compared to available tools. Therefore, we edited the existing Re-Annotator Pipeline to use the Bowtie 1.2.2 aligner instead of BWA. Replacing BWA alignment command as shown below, allows robust output of uniquely mapping probe

sequences with no mismatches. This ensures that shorter microarray probes, such as those found in Affymetrix arrays, are correctly aligned. For mouse annotation, we used NCBI's mm10 genome assembly submitted on Dec 2011, and Gencode gene annotation version 16 submitted in Dec 2017.

#### *#Original Code*

```
bwa aln -l 100 -t $numCPU -n $numMismatch $db $inReads > $outAln
```

#### *#Commits by Nabila Rahman*

```
$bowtie1 -v 0 -m 1 $db -f $inReads -S $samFile -p 8
```

Using the modified Re-Annotator Pipeline, we have updated the annotation Affymetrix Mouse Genome 430 2.0 Array. This annotation was merged with the manufacturer annotation, last updated in March 2016, to account for any probes not correctly identified by the Reannotator tool [134]. Mouse and human orthology information were obtained from Ensembl database version 91.

### 3.2.2.5 Removing Redundant Probes

By reannotating, we were able to identify which probes hybridised to coding and non-coding transcripts. Differential expression analysis contained redundant probes for most genes. We removed these redundancies based on the type of transcript region the probes hybridised to. Unique genes were retained in the following order (most relevant first): exons, 5' or 3' untranslated region (UTR), introns, splicing regions, sequence up or downstream of open reading frame, intergenic transcripts and probes with one or more sequence mismatches. If redundant probes mapped to the same type of transcript region, the probe with lowest FDR adjusted *p*-value across test groups were retained.

### 3.2.3 Human B cell Lineage

**Table 3-3** | Phenotype of human microarray profiles of PC cell Lineage - HG133A (Jourdan *et al*)

Sample	Type	Stimulus	Day	Markers	Replicates
E-MTAB-1771	NBC	-	0	CD19+ CD27-	5
E-MEXP-2360	MBC	-	0	CD19+ CD27+	5
E-MEXP-2360	ABC	CpG, CD40L	4	CD20+ CD38-	5
E-MEXP-2360	PB	CpG, CD40L	7	CD20- CD38+	5
E-MEXP-2360	PC	CpG, CD40L	10	CD20- CD138+	5
E-MEXP-2360	BMPC	-	0	CD138	5

#### 3.2.3.1 Data Source

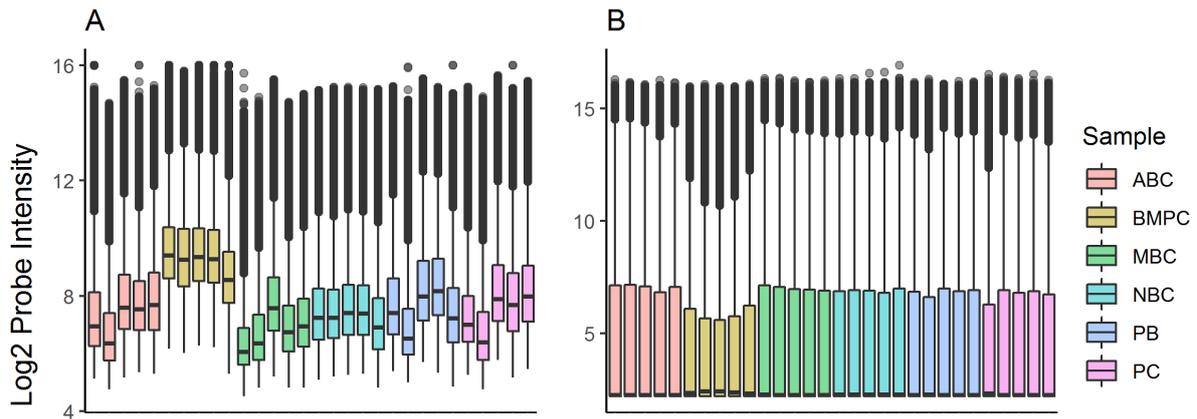
Jourdan and Kassambara *et al* isolated cells of the PC lineage from adult human peripheral blood using FACS [55, 71, 135, 136]. Unlike the mouse lineage previously discussed, this study initially used *in vitro* generated ASCs from memory B cells (MBCs) rather than NBCs.

MBCs were immunised at day 0 using CpG oligonucleotide, recombinant human CD40L in the presence of IL2, IL10 and IL15 [55, 71, 135, 136]. At day 4, non-secreting Activated B cells (ABC) fraction were harvested [55, 71, 135, 136]. From day 4, cells were cultured with IL2 and IL6 and harvested for plasmablasts (PB) at day 7 [55, 71, 135, 136]. Cells were cultured with IL6, IL5 and IFN- $\alpha$  until day 10 and harvested for PCs [55, 71, 135, 136]. Markers used for sorting and harvesting relevant cell types is summarised in Table 3-3.

Total RNA was extracted using Qiagen RNeasy Mini Kit. This was followed by P-AFFY-2 Affymetrix biotin labelling and hybridization to Human Genome U133 Plus 2.0 (HG133A) as instructed by Affymetrix [71, 135].

#### 3.2.3.2 Normalisation

All arrays were normalised using GCRMA in order to conform to the pre-processing methods applied to the mouse transcriptome profile described in Section 3.2.2.2. The resultant array distributions are visualised in Fig 3-7B.



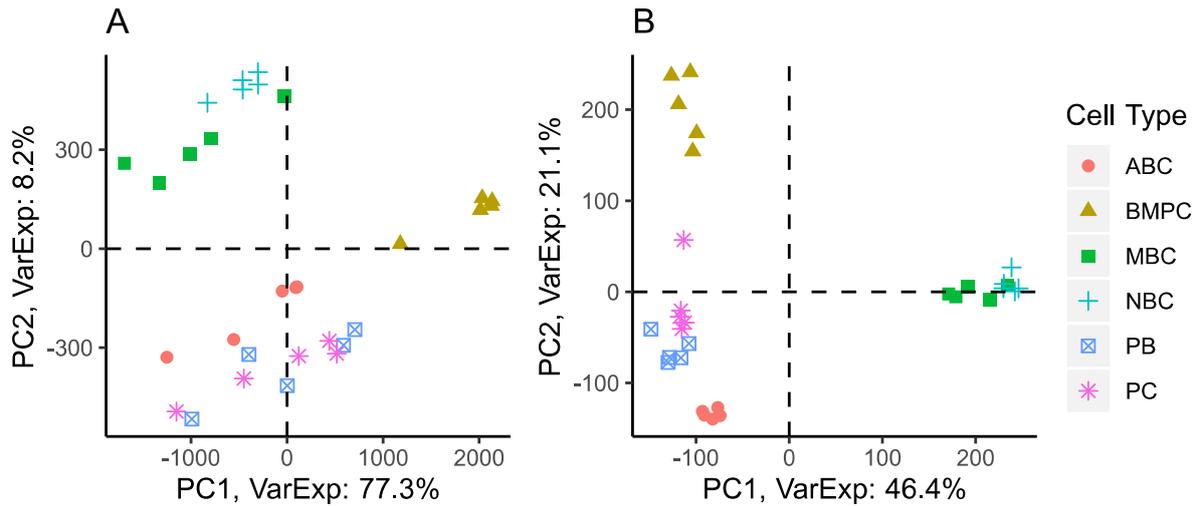
**Fig 3-7** | Boxplot of probe intensities in raw and GCRMA normalised human microarray data. **A.** Raw Data. The distribution of optical intensities differs for each array. **B.** GCRMA normalisation, establishes background noise cut-off and normalises the range of intensities present in each array. There does not seem to be any obvious batch effect as all arrays conform to a similar range of intensity values.

### 3.2.3.3 Experimental Design

After normalisation we carried out principal component analysis to determine correct experimental design. We wanted to know whether utilising NBCs as the non-secretory control would be suitable for ASCs generated from MBCs. As shown in Fig 3-8B, both MBCs and NBCs cluster very close together and are likely to show very similar differential expression when compared to ASCs. Therefore, to conform to previous methods, we utilise NBCs as the control group and disregard MBCs for further analysis. As we are interested in ASCs, we primarily focus on contrasting BMPC, *in vitro* generated PC and PB groups to NBCs for cross species meta-analysis. Study of early stages of B cell differentiation, ABC and PBs, are not in the scope of this project.

### 3.2.3.4 Differential Expression Analysis

Fold changes and statistical test for DE were nearly identical to those described in Section 3.2.2.3. The only difference was the exclusion of blocking methods as no cross-study merging was involved in this section.



**Fig 3-8** | PCA Plot showing sample clustering in raw versus GCRMA normalised human microarray data. **A.** Raw data. Before normalisation the MBCs and NBCs together cluster away from ASCs. However, this variance is very small, 8.2%. **B.** After GCRMA normalisation, the variation between ASCs and NBCs are as much as 46.4%.

### 3.2.3.5 Reannotation and Removing Redundancies

Probes were reannotated using NCBI's hg38 genome assembly submitted on Dec 2013, and Gencode gene annotation version 27 submitted in Jan 2017 through the modified Re-Annotator tool. Updated annotations were merged with latest manufacturer (March 2016) annotations for better probe coverage. Mouse orthologs were obtained from Ensembl mart release 91. Post annotation redundant probes were removed as described in Section 3.2.2.5.

## 3.2.4 Cross Species Meta-Analysis

In this section we combine processed DE results from Mouse and Human B cell lineage.

### 3.2.4.1 Known orthologs

Genes with known orthologs were merged using Ensembl ortholog annotations. As each of these genes now had two separate FDR adjusted  $p$ -values, one from mice and another from human studies, we combined this statistic by calculating the median of the adjusted  $p$ -values per gene. If a gene had multiple orthologs present in either species, the ortholog pair with

homology lower than 50% were removed and then ortholog pairs with lowest FDR adjusted  $p$ -value for multi group comparison was retained.

#### 3.2.4.2 Genes with No orthologs or Missing probes

A number of mouse genes had no known human orthologs and vice versa. In addition, certain genes with known orthologs had probes in the mouse microarray but were missing in the human microarray or vice versa. As these are due to lack of annotation and design issues respectively, we wish to retain these genes. We took into consideration that these genes will not have cross-species evidence to back their reproducibility and retained results from probes that only hybridised to exons and 5' or 3' UTR. We scaled the FDR adjusted  $p$ -values for these genes to conform to the median  $p$ -values calculated in Section 3.2.4.1 to prevent them from being over or underestimated.

#### 3.2.4.3 Cross Species Differential Expression

Once cross-species data was combined, we used a voting-based system to select for DE genes conserved among species. We compared 5 ASCs types to NBCs: (1) mouse BMPC, (2) mouse SplPC (3) human BMPC, (4) human generated *in vitro* PB and (5) *in vitro* differentiated human PCs. If a gene showed DE (2 FC) in the same direction in half the test groups and the FC in the remaining groups showed no significant change/ was missing, the gene was considered differentially expressed. Same rules were applied to genes with no orthologs with the number of test groups and requisite number of “votes” adjusted accordingly. We then ranked our genes by median adjusted  $p$ -value, lowest value (i.e. most likely to be true positive) first and finally enforce a 0.05 FDR adjusted  $p$ -value cut-off.

#### 3.2.5 Functional Analysis

After obtaining our differentially expressed gene set, we wanted to see what biological functions these genes related to. As previously discussed EnrichR allows users to query biological databases such as GO, KEGG, etc. In order to derive meaningful and statistically

valid results, we calculated the overlap of our differentially expressed genes (DEG) and biological databases using rank-based ranking via the EnrichR R package.

The EnrichR package can be used to query a comprehensive list of curated databases. For our purposes we utilised the following databases:

- GO Cellular Component 2017 – a database of gene localisation relative to cell structure.
- GO Biological Process 2017 – a database of biological functions relating to genes
- ARCHS4 Transcription Factor Coexpression – database derived from large scale RNA-Seq meta-analysis of thousands of datasets. It allows us to find transcription factors (TF) that are commonly co-regulated with sets of genes.
- Transcription Factor Perturbations Followed by Expression – allows mining of TF perturbation experiments to see how they relate to gene sets.

### 3.2.5.1 Summarising GO Terms

In order to summarise these GO terms, we used REVIGO API (Application Programming Interface) via R [137]. Overlapping terms were given a *simRel* score, and any terms with 0.5 similarity score were grouped under the same category. The “whole Uniprot database” was used for obtaining GO terms. The *simRel* score is a machine learning algorithm based on Resnik’s and Lin’s similarity measure [138, 139]. We manually curated discrepancies in overlapping GO terms assigned by this algorithm.

Once GO terms were summarised, we ran into the issue of having multiple GO terms with differing numbers of member genes, some overlapping, some not. In order to combine these gene members and only retain unique ones, and then count how many unique genes mapped to each summarised group, we created R functions to automate the data mining process to allow for robust and reproducible results. These functions are available at:

<https://github.com/NabilaRahman/EnrichR-mining-tools>.

In the next section we explain the usage of these tools.

## 3.2.5.1.1 GO Ontology Summary and Visualisation Protocol

*Step 1: Database Query*

The following code chunk queries EnrichR to obtain enriched GO terms and relevant statistics for genes upregulated in ASCs across species. The grey boxes represent our R script and the white boxes show the processing and result of the script.

```
## EnrichR -----
library(enrichR)
my_chosen_databases <- c(
  "GO_Biological_Process_2017" )

enriched.Up <- enrichr(Upregulated_genes, my_chosen_databases)
```

```
## Uploading data to Enrichr... Done.
## Querying GO_Cellular_Component_2017... Done.
## Querying GO_Biological_Process_2017... Done.
## Parsing results... Done.
```

*Step 2: Filter for True Positives*

Our `getEnrichmentResults` function filters out GO terms with adjusted  $p$ -value greater than 0.01 (this threshold can be adjusted by the user). Member genes of these GO terms are not shown due to space constraint.

```
## GO Biological Process --
goBP.UP <- getEnrichmentResults(
  data = enriched.Up
  , goTYPE="go.BP"
  , direction = "UP"
  , dbgo = "GO_Biological_Process_2017" ) [ ["Enrichment"] ]
head(goBP.UP) [, c(1,2,3,5)]
```

```
##          Term Overlap Adjusted.P.value
## 1 GO:0036498      32    1.866856e-14
## 2 GO:0006890      35    6.938853e-12
##
##                               Description
## 1                IRE1-mediated unfolded protein response
## 2  retrograde vesicle-mediated transport, Golgi to ER
```

*Step 3: Obtaining Revigo Summary*

The `getEnrichmentResults` function is also used to obtain a result with format suitable for inputting directly into the REVIGO tool manually or automatically using our `getRevigoTable` function.

```

revigo.BP.UP <- getEnrichmentResults(
  data = enriched.Up
  , goTYPE="go.BP"
  , direction = "UP"
  , dbgo = "GO_Biological_Process_2017" ) [ ["revigo"]]

revigo.BP.UP.input <- pasteDataFrame(revigo.BP.UP)

revigo.BP.UP.output <- getRevigoTable(goList = revigo.BP.UP.input
  , cutoff= "0.5"
  , isPValue= "yes"
  , measure = "SIMREL"
  , goSizes = "0"
  )

```

#### Step 4: Manually Curated Summary of GO terms

Once REVIGO results are obtained manual intervention is required to group overlapping gene sets into a common group. Example of the summary table is shown below.

```

go.BP.ref <- read.table("F:/4.Microarray-Function/GO.BP.Ref.manual.txt", header = T, sep="\t")
head(go.BP.ref, 3)

```

##	Term	Description	Summary
## 1	GO:0006260	DNA replication	mitotic cell cycle
## 2	GO:0007062	sister chromatid cohesion	mitotic cell cycle
## 3	GO:0055085	transmembrane transport	transport

#### Step 5: Summarise Go Terms for Visualisation

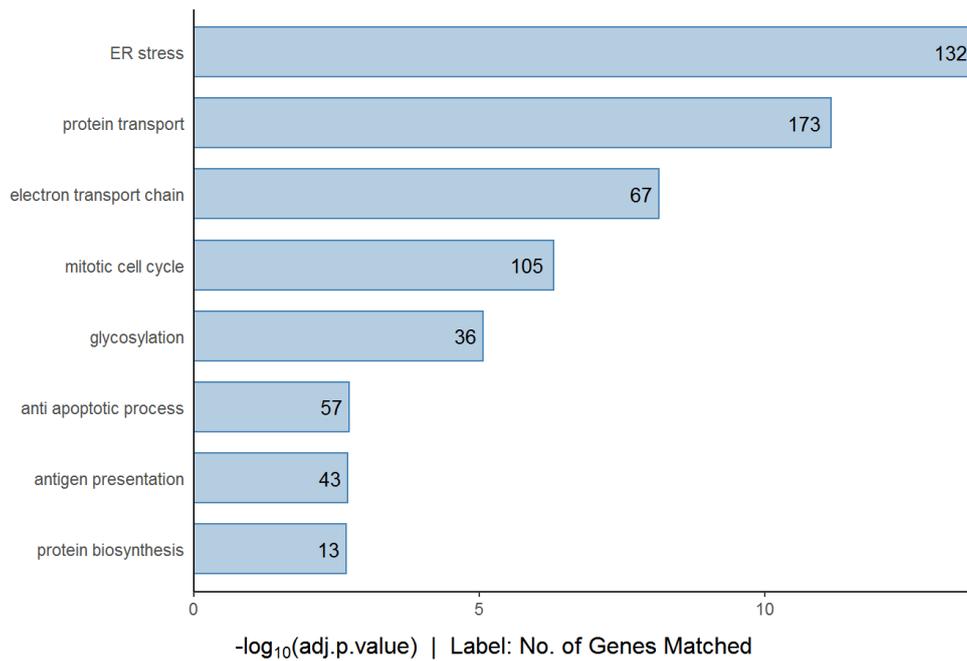
The summary table is then used to update the processed enrichR results from step 2. Finally, our `mergeGoSummary` function is used to aggregate all the unique genes belonging to the same overall summary group and counted for visualisation as shown in Fig 3-9.

```

goBP.UP <- merge(goBP.UP, go.BP.ref[,c(1,3)], by="Term", sort = F, all.x = T)

```

```
goBP.UP.summary <- mergeGoSummary(goBP.UP)
```



**Fig 3-9** | Example GO Ontology summarisation for GO Biological Processes generated from the results of EnrichR mining tool. Under “ER stress” a number of redundant terms were found: IRE1 mediated pathway, IRE1-mediated unfolded protein response, ER unfolded protein response, response to ER stress, ubiquitin-dependent ERAD pathway, retrograde protein transport, ER to cytosol. Without summarisation over 60 GO terms persist with too many genes overlapping in each group.

### 3.2.5.1.2 Pathway Analysis – protocol for robust pathway enrichment

Databases for transcription factor or kinase perturbation allow for complex pathway analysis. A kinase perturbation database allows us to query genes that were previously shown to be regulated in response to a kinase perturbation. To make sense of perturbation databases we retain hits where the perturbed genes and their affected genes match the direction of regulation found in the user’s data. As in the case of GO ontologies, we also summarise overlapping perturbation terms. Although we do not discuss kinase perturbations in this project, in the following example we utilise the kinase perturbation as an example of generic pathway analysis that can be performed using our protocol. Essentially, the same procedure applies to transcription factor perturbation.

*Step 1: Database Query and Filtering*

We use `EnrichR` package and `getEnrichmentResults` function to obtain enrichment terms with adjusted  $p$ -value less than 0.01. In this example, differentially upregulated genes are shown.

```
## EnrichR -----
library(enrichR)
my_chosen_databases <- c(
  "Kinase_Perturbations_from_GEO_up" )

enriched.Up <- enrichr(Upregulated_Genes, my_chosen_databases)

##Kinases - Gene UP
Kinase.result.UP <- getEnrichmentResults(
  data = enriched.Up
  , goTYPE="kinases"
  , direction = "UP"
  , dbgo = "Kinase_Perturbations_from_GEO_up"
  , revigo = F )["Enrichment"]
```

*Step 2: Output based on kinase regulation*

The `getGenePerturbation` function is used to query our DEG and only keep kinase terms whose perturbation matches the direction of regulation in our dataset.

```
Kinase.result.UP <- getGenePerturbation(data=Kinase.result.UP, direction="UP",
  genesUp = uniqueUp, genesDown = uniqueDown
  , type = "kinases")
```

*Step 3: Direct Correlation*

The direct correlation of kinases and our DEGs can be visualised with the code chunk below. This is typically meant to summarise how many genes were upregulated when a specific kinase experienced upregulation, activation or overall gain of function. The results here show that no genes were upregulated when kinases were upregulated.

```
kinase.Up.Gene.Up <- Kinase.result.UP["DirectSummary"]
head(kinase.Up.Gene.Up) [, c(1:3,5:7)]
```

```
## [1] Term          Count          Adjusted.P.value pertType
## [5] pertDir        geneDir
## <0 rows> (or 0-length row.names)
```

*Step 4: Inverse Correlation*

The inverse correlation ( `InverseSummary` ) of kinases and our differentially upregulated genes is shown below. For example, inhibition, downregulated or overall loss of function of TRIM33 led to the upregulation of 79 genes in our dataset. Furthermore, TRIM33, BRD4, ATM, etc are kinases that were differentially downregulated in our dataset. Using such output, we can visualise enriched kinase perturbation the same way as the Gene Ontology example in Fig 3-9.

```
kinase.Down.Gene.Up <- Kinase.result.UP[["InverseSummary"]]
head(kinase.Down.Gene.Up) [,c(1:3,5:7)]
```

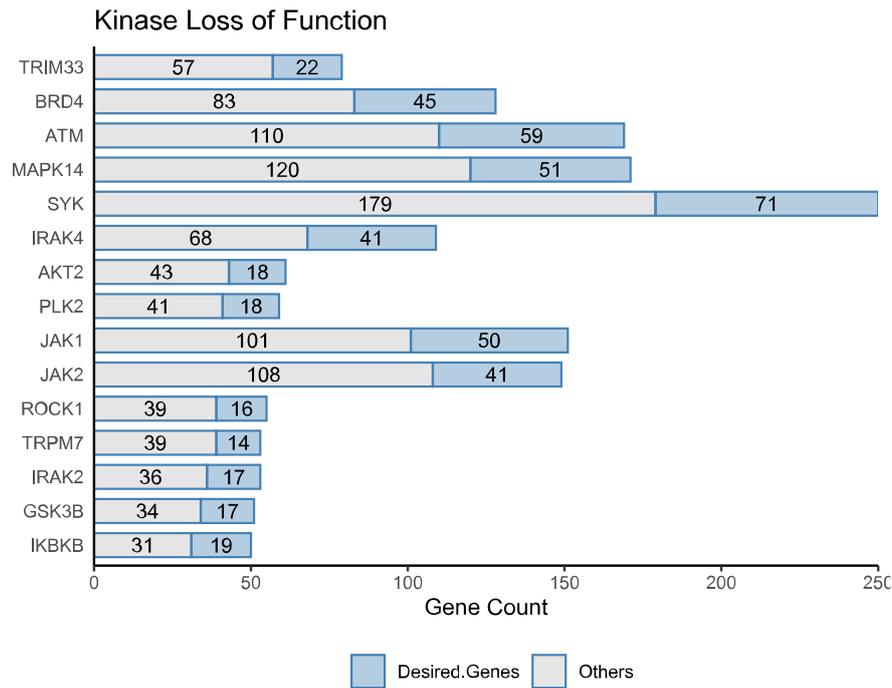
##	Term	Count	Adjusted.P.value	pertType	pertDir	geneDir
##	TRIM33	79	9.208714e-14	knockdown	DOWN	UP
##	BRD4	128	3.386253e-12	druginhibition;knockdown	DOWN	UP
##	ATM	169	4.696190e-09	knockout;knockdown	DOWN	UP
##	MAPK14	171	9.850537e-09	knockout;knockdown	DOWN	UP
##	SYK	250	1.149229e-08	druginhibition	DOWN	UP
##	IRAK4	109	5.453364e-08	knockout;defectivemutant	DOWN	UP

*Step 5: Extract Functionally Relevant Genes*

Using GO Enrichment, we were able to identify a number of genes that are likely to be relevant to membrane trafficking. If we want to know how many of these genes were affected by kinases, querying kinase databases with this filtered set of genes would underestimate the statistical power of the enrichment analysis due to smaller degree of freedom. Therefore, retain results of and highlight how many of the genes perturbed by kinases were among our genes of interest. This was done using our `getDesiredEnrichment` function.

```
kinase.Down.Gene.Up.Desired<- getDesiredEnrichment(enrichData = kinase.Down
.Gene.Up , desiredGenes = desiredGenes )
```

##	Term	Total	Adjusted.P.value	desiredGeneCount	Remainder
##	17 TRIM33	79	9.208714e-14	22	57
##	4 BRD4	128	3.386253e-12	45	83
##	2 ATM	169	4.696190e-09	59	110
##	11 MAPK14	171	9.850537e-09	51	120
##	15 SYK	250	1.149229e-08	71	179



**Fig 3-10** | Example Visualisation of Pathway Enrichment results from EnrichR mining tool. The figure shows how many genes are upregulated as a result of Kinase loss of function. Kinases are order by statistical significance, highest first. User can input a separate set of “Desired Genes” to determine how many of genes of interest are being perturbed by a kinase or a transcription factor. Our “desired genes” were those that had GO terms related to ER, Golgi, and membrane trafficking components. Note that we do not discuss kinase perturbation in this chapter, and this database is used as an example only.

Using the above result, we can visualise how our genes of interest may be affected by kinase perturbation as shown in Fig 3-10.

The above described protocol was utilised for mining other databases such as “ARCHS4 Transcription Factor Coexpression” and “Transcription Factor Perturbations Followed by Expression”.

### 3.3 RESULTS

#### 3.3.1 Transcript Classification

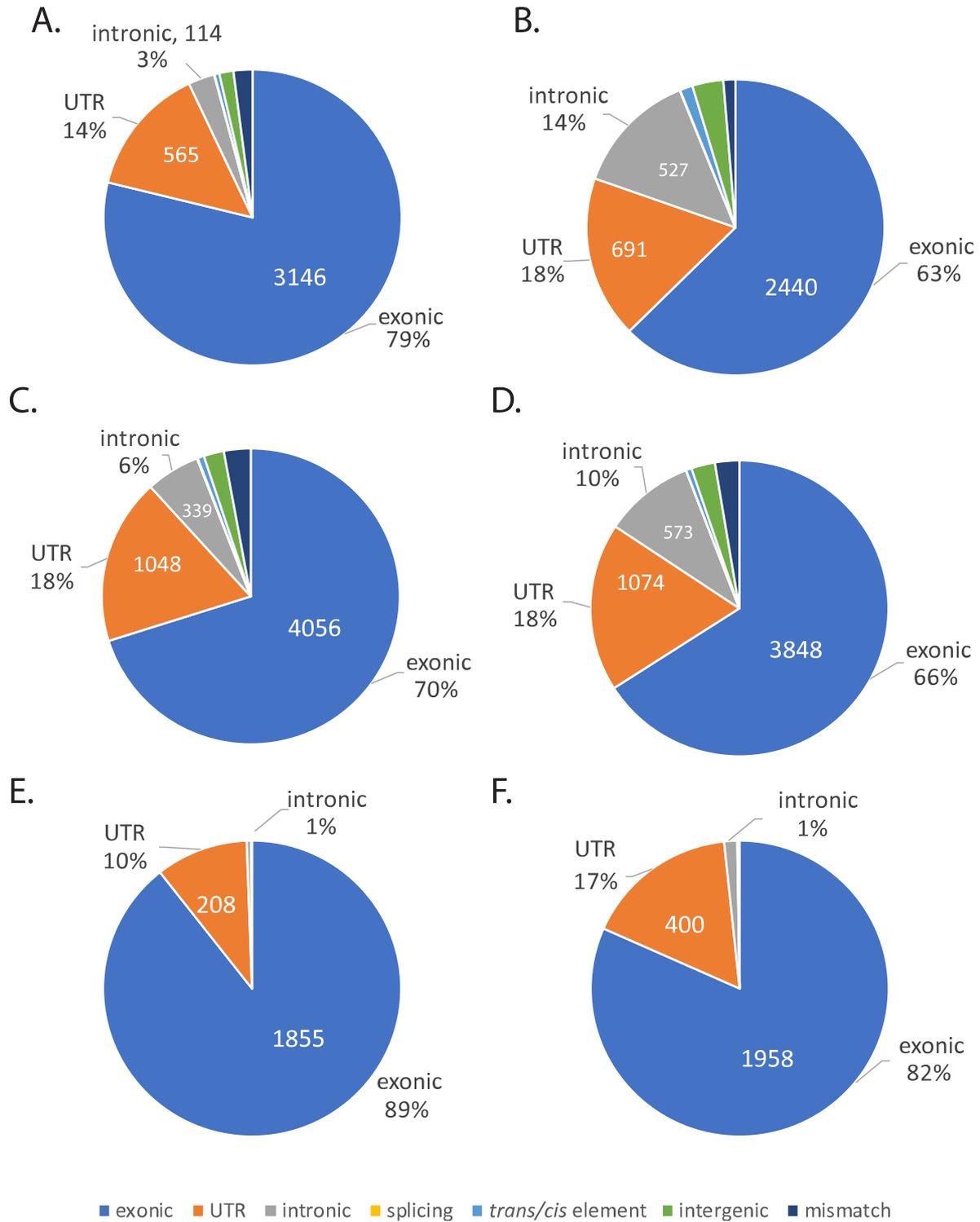
We updated the annotation on Affymetrix Genechips and found that the Mouse Genome 430 2.0 Array have a total of 25596 unique genes or gene isoforms and the Human Genome U133 Plus 2.0 Array had 30496 after re-annotation. Reannotation gave us valuable insight into the quality of the probesets. This information can be used to evaluate the how much of the differentially expressed transcriptome is protein coding.

##### 3.3.1.1 Cross Species Analysis verifies integrity of low-quality transcripts

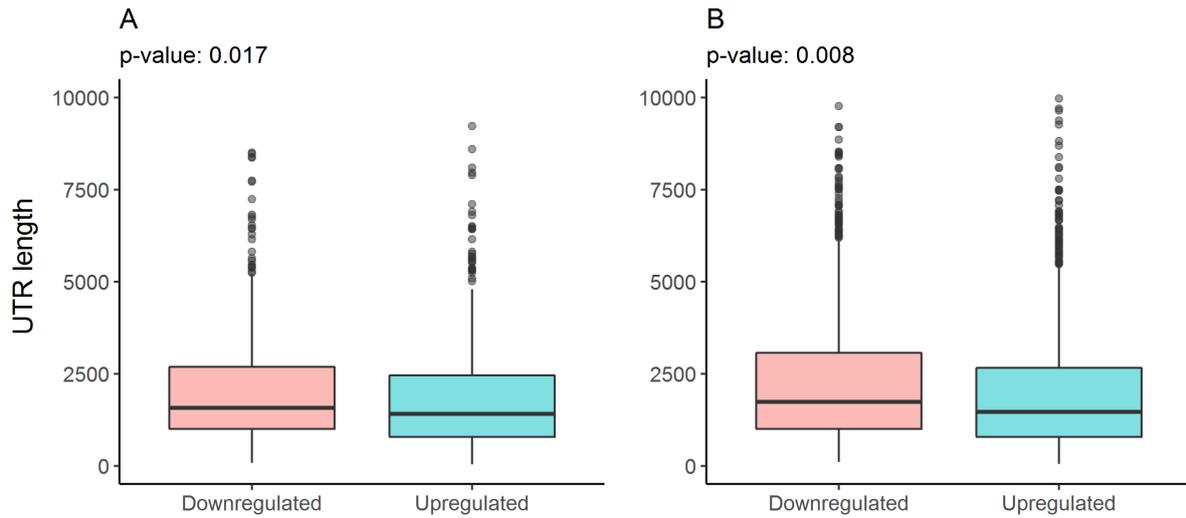
When we combine the differentially expressed genes (DEG) from mice and human, we are able to better ascertain whether genes showing differential expression (DE) via hybridization to poor quality probes are real. For example, if gene A in mouse showed DE via an intergenic probe, which maps to a non-coding region between gene A and gene C, it is unclear whether gene A or gene C is showing differential expression. But if at the same time the human dataset shows DE in gene A via a uniquely mapping exonic probe, we can say that the gene A is likely a valid hit. Likewise, if gene B in mouse showed DE via an intronic probe and the same gene in human dataset showed DE via a uniquely mapping exonic probe, we can say with higher confidence that gene B is likely to be protein coding. As shown in Fig 3-11A-D, before cross species analysis, intergenic as well as mismatching transcripts accounted for approximately 10-20% of DEG in both mice and human. After cross species meta-analysis, these low-quality transcripts accounted for less than 1% of the results. By removing low quality genes and prioritising exonic probes our methodology increased the representation of differentially expressed protein coding transcripts by up to 20% as shown in Fig 3-11E, F.

### 3.3.1.2 Length bias in UTR regions

While removal of lower quality transcripts improves data quality, the importance of non-coding RNA cannot be ignored. We found that distribution of intronic gene and UTR regions were proportionately higher in downregulated genes as opposed to the upregulated set even after cross-species data analysis (Fig 3-11E, F). Previous studies have shown that a sizable portion of mammalian introns and UTR regions encode for gene regulatory elements [140, 141]. We, therefore, speculated that the overrepresentation of non-coding regions in the downregulated set may suggest a role in the B cell lineage. However, upon closer inspection we found that the lengths of the UTR regions found in the downregulated set were ~18% longer in mouse and ~30% longer in human compared to the corresponding upregulated set ( $p$ -value<0.05). Therefore, this length bias may be the primary cause of differential number of UTR genes detected in the up- and downregulated sets



**Fig 3-11** | Proportion differentially expressed coding and non-coding transcripts in mice arrays, human arrays and in cross species analysis. **A-B.** Number of genes DE in at least 1 cell type after standalone analysis of mice microarrays. Non-coding regions make up **A.** 21% of upregulated and **B.** 37% of downregulated genes. **C-D.** Number of genes DE in at least 1 cell type for standalone analysis of human microarrays. Non-coding regions make up **C.** 30% of upregulated and **D.** 34% of downregulated genes. **E-F.** Number of genes DE in at least half the ASCs studied across species, this set also includes genes with missing probes or orthologs for either species. Non-coding regions make up **E.** 11% of upregulated and **F.** 18% of downregulated genes.



**Fig 3-12** | Box plot of UTR lengths in **A.** Mice microarray and **B.** Human Microarray. Only UTRs differentially expressed in at least 3 out of 5 ASCs studied were considered.

### 3.3.2 Differential Expression

#### 3.3.2.1 Mouse ASC lineage

In the mouse model, splenic plasma cells, bone marrow plasma cells or both showed differential upregulation in 3991 unique genes compared to naïve B cells (fold change  $\geq 2$ , false discovery adjusted  $p$ -value  $< 0.05$ ). A similar number of genes, 3896, were downregulated.

#### 3.3.2.2 Human ASC lineage

In the human model, bone marrow PCs, *in vitro* generated PBs and PCs showed differential upregulation in 5783 unique genes in at least one cell type compared to NBCs. A similar number of genes, 5836, were downregulated.

#### 3.3.2.3 Cross Species Analysis

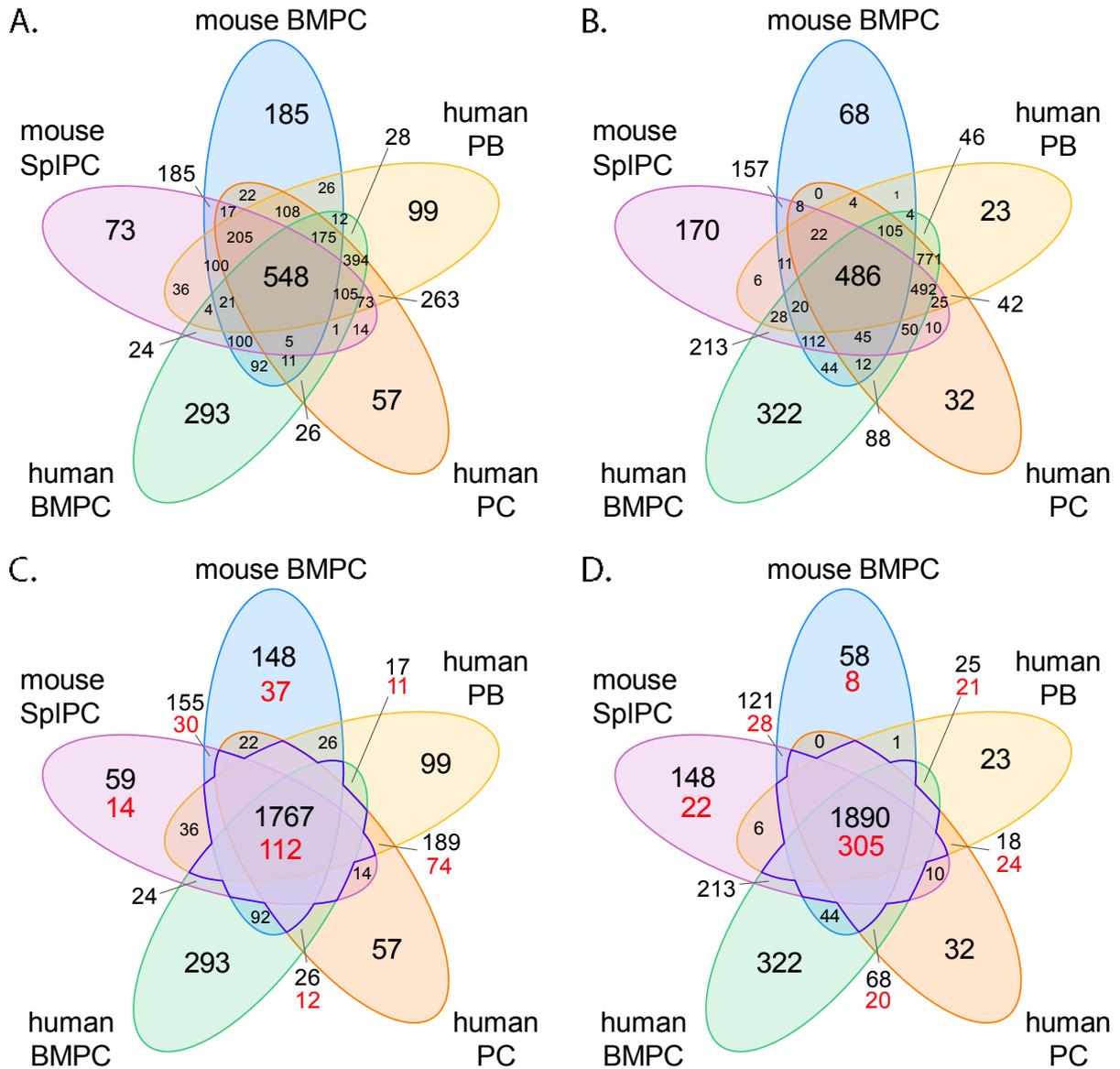
##### 3.3.2.3.1 Upregulated Genes

Fig 3-13A, C summarises the overlap of differentially upregulated genes in mice and human. 548 genes show conserved upregulation across species in all 5 ASC studied (Fig 3-13A). 511

genes were upregulated in 4 out of 5 ASCs and 820 were upregulated in 3 out of 5 ASCs. We carried forward these 3 groups of upregulated genes for further analysis (Fig 3-13C, middle, purple area). In addition to these subsets, differentially expressed genes (DEGs) that were present in one species but had no orthologs or array probes for another were included for downstream analysis (Fig 3-13C, in red).

#### 3.3.2.3.2 Downregulated Genes

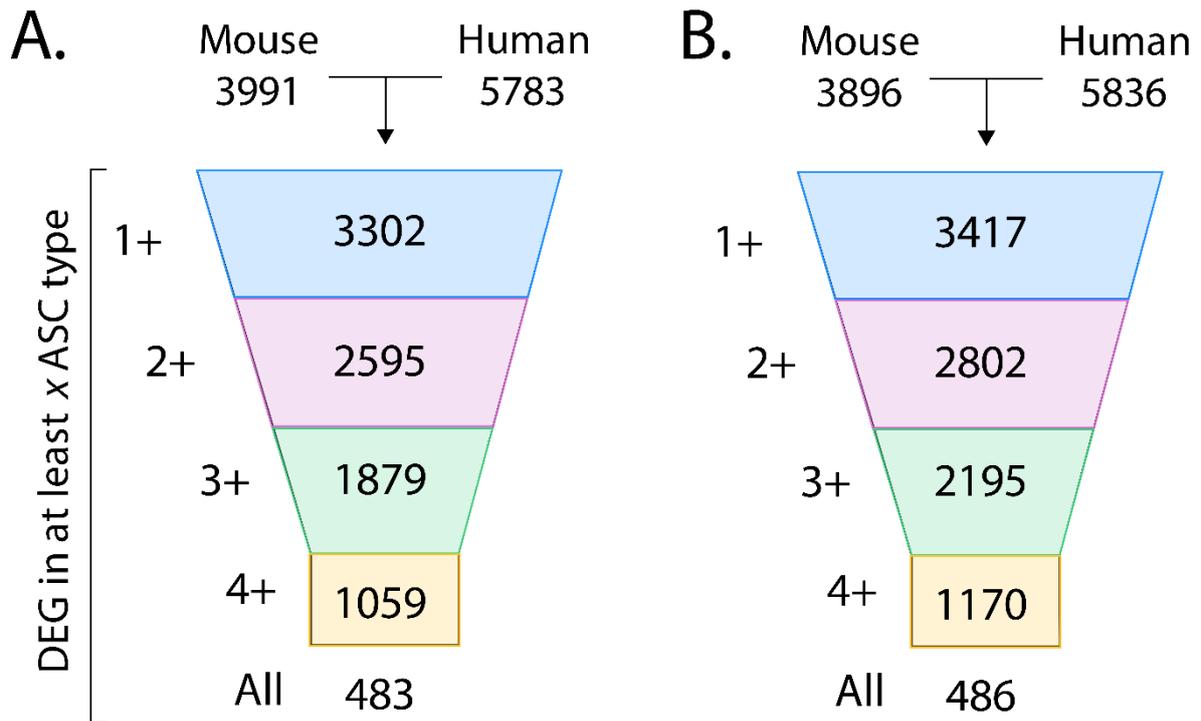
Fig 3-13BD summarises the overlap of differentially downregulated genes in mice and human. 486 genes are consistently downregulated in all 5 ASCs (Fig 3-13B). 684 genes were upregulated in 4 out of 5 ASCs and 1025 were downregulated in 3 out of 5 ASCs. We carried forward these 3 groups of genes for further analysis (Fig 3-13D, middle, purple area).



**Fig 3-13 |** Venn Diagram showing the overlap of differentially regulated genes in ASCs of mice and human microarray profile. **A.** Upregulated in mice and human ASCs. **B.** Downregulated in mouse and human ASCs. **C-D.** Genes differentially expressed in at least 3 ASCs are indicated by middle, purple region. An additional set of genes without known orthologs or having missing probes in either species (indicated in red), were included in the downstream analysis. **C.** Upregulated in mice and human. **D.** Downregulated in mice and human

**3.3.2.3.3 Cross species analysis filters out up to 86% of DE genes.**

Fig 3-14 summarises the consistency of differential expression across species and ASC types. Filtering genes by the reproducibility of transcript expression across ASC types allowed us to narrow down around 500 (14%) upregulated genes from approximately 3500 potential candidates (Fig 3-14A). A similar proportion of genes were isolated from the downregulated



**Fig 3-14** | Visualisation of the number of DEG filtered out via cross-species microarray analysis. **A.** Upregulated genes overlapping in at least x ASCs and no significant change in either direction in any remaining groups. **B.** Downregulated genes overlapping in at least x ASCs and no significant change in any remaining groups. Note that prior to multi group filtering, low quality transcripts (intergenic, mismatches, introns) without cross species evidence were removed. This removes a up to 2500 hits. As there were only 2 mouse ASCs studied, the gene counts do not represent genes that are DE in mouse platform but had no ortholog or probes in the human one.

set (Fig 3-14B). These genes show the most reproducible transcript expression and thus are supposedly the ideal candidates for downstream analysis.

However, a caveat of meta-analysis is that the fold change values across species or platforms are very rarely the same. Typically, genes exhibit some scatter in the fold change values across similar tested conditions. We acknowledge that sometimes the minimum value of this range falls below the absolute 2-fold threshold. In this particular study, these changes could be due to (a) minor optical measurement differences pushing result below 2 FC threshold; (b) natural scatter or batch differences; or (c) phenotypic differences not accounted for. This is why, considering only the 14% genes expressed in every ASC studied in this chapter would be overly stringent. Therefore, we implemented a voting system to allow for genes that are

differentially expressed in the majority of ASCs. This method retains 57% upregulated genes and 64.2% downregulated genes, respectively.

### 3.3.2.4 Global Gene Regulation

**Table 3-4 | Global summary of genes differentially expression in mouse and human ASCs**

Direction of Regulation		Total Genes
Mice ASC	Human ASC	
Up	Up	2074
Down	Down	2399
Up	Down	514
Down	Up	574

#### *Consistent Across Species*

Fig 3-14 only considers genes that were present in both microarrays and had orthologs across mouse and humans. After considering genes with unknown orthologs or lacking probes in one platform, we find a total of 2074 upregulated genes, and 2399 downregulated genes for further analysis.

#### *Differential Regulation between Species*

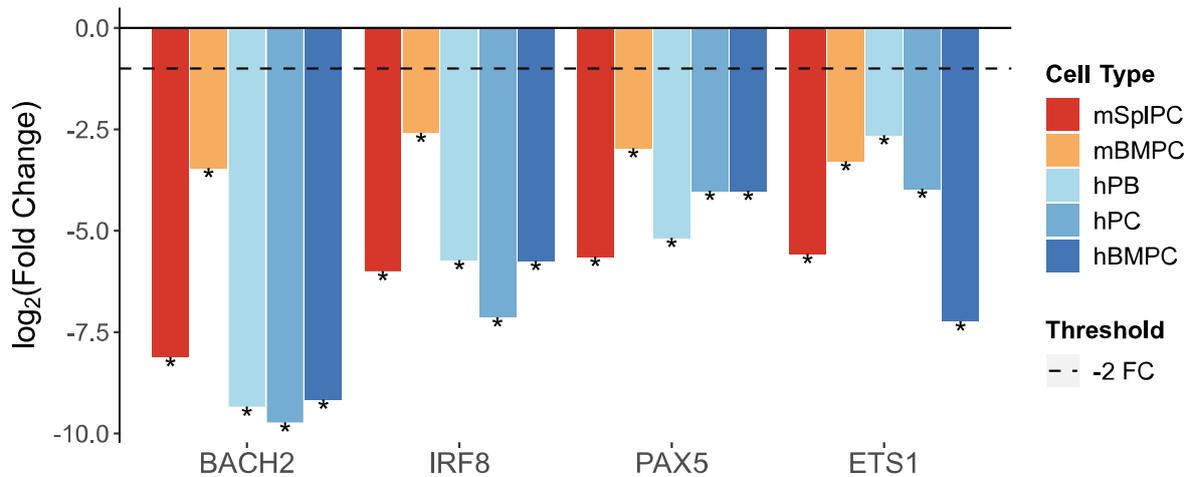
The premise of cross species analysis is that genes showing consistent regulation are likely to more reliable. Nevertheless, there may be some true positive contradictory results that highlight differences between mouse and human. We found 1085 genes showing differential regulation in mouse vs. human ASCs. 514 of these are upregulated in mouse but show no change or downregulation in humans and 574 genes are downregulated in mouse and upregulated in humans. However, we acknowledge that a considerable number of these genes are likely to be erroneous results brought on by systemic and/or non-systemic noise. Therefore, we mainly focus on genes showing species conserved regulation.

### 3.3.3 Verification

In order to determine whether our results have been correctly processed, we exploit well characterised markers that can be used as a reference to ensure the integrity of our data.

### 3.3.3.1 Known B cell biomarkers

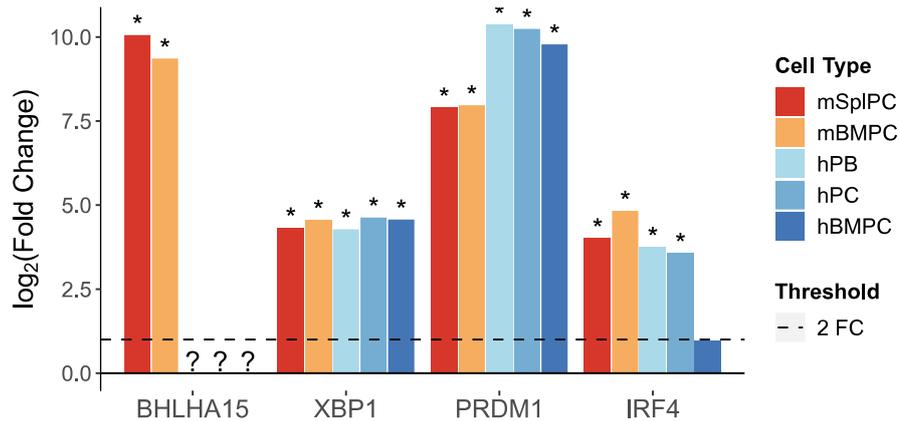
PAX5, BCL6, BACH2 and FOXO1 are well studied transcription factors known to maintain the B cell phenotype [142, 143]. The expression of these genes are known to decrease when B cells commit to plasma cell differentiation. As shown in Fig 3-15, these genes were consistently downregulated in both human and mice ASCs compared to resting B cells in accordance with existing literature.



**Fig 3-15** | Differential regulation of transcription factors characteristic of naïve B cell in cross-species microarray analysis. Literature shows these genes are characteristic of naïve B cells and decrease in expression when NBCs commit to antibody secretion. In line with previous studies, cross species analysis show that these genes are consistently downregulated in ASCs versus NBCs. (\*) indicate FDR adjusted  $p$ -value < 0.05

### 3.3.3.2 Known ASC biomarkers

IRF8, PRDM1/Blimp1, XBP1 as well as BHLHA15/MIST1 are known enhancers of the ASC phenotype [142, 144]. Our analysis indicates that IRF8, PRDM1/Blimp1, XBP1 genes are consistently upregulated across species. While BHLHA15/MIST1 also shows upregulation in mice, the human Affymetrix Genechip used to profile the human transcriptome did not have a probe for this gene. This justifies our experimental design in incorporating genes with missing probes and orthologs into the cross-species analysis.



**Fig 3-16** | Differential regulation of transcription factors characteristic of antibody secreting cells in cross-species microarray analysis. Literature shows that these genes are characteristic of ASCs and increase in expression when NBCs commit to antibody secretion. In line with previous studies, cross species analysis show XBP1, PRDM1 and IRF4 are consistently upregulated in ASCs versus NBCs. We also observe strong upregulation of BHLHA15/MIST1 in mice array, although, no data was available in human microarray due to missing probes for this gene. (\*) indicate FDR adjusted  $p$ -value  $< 0.05$ . (?) indicates missing ortholog or probes.

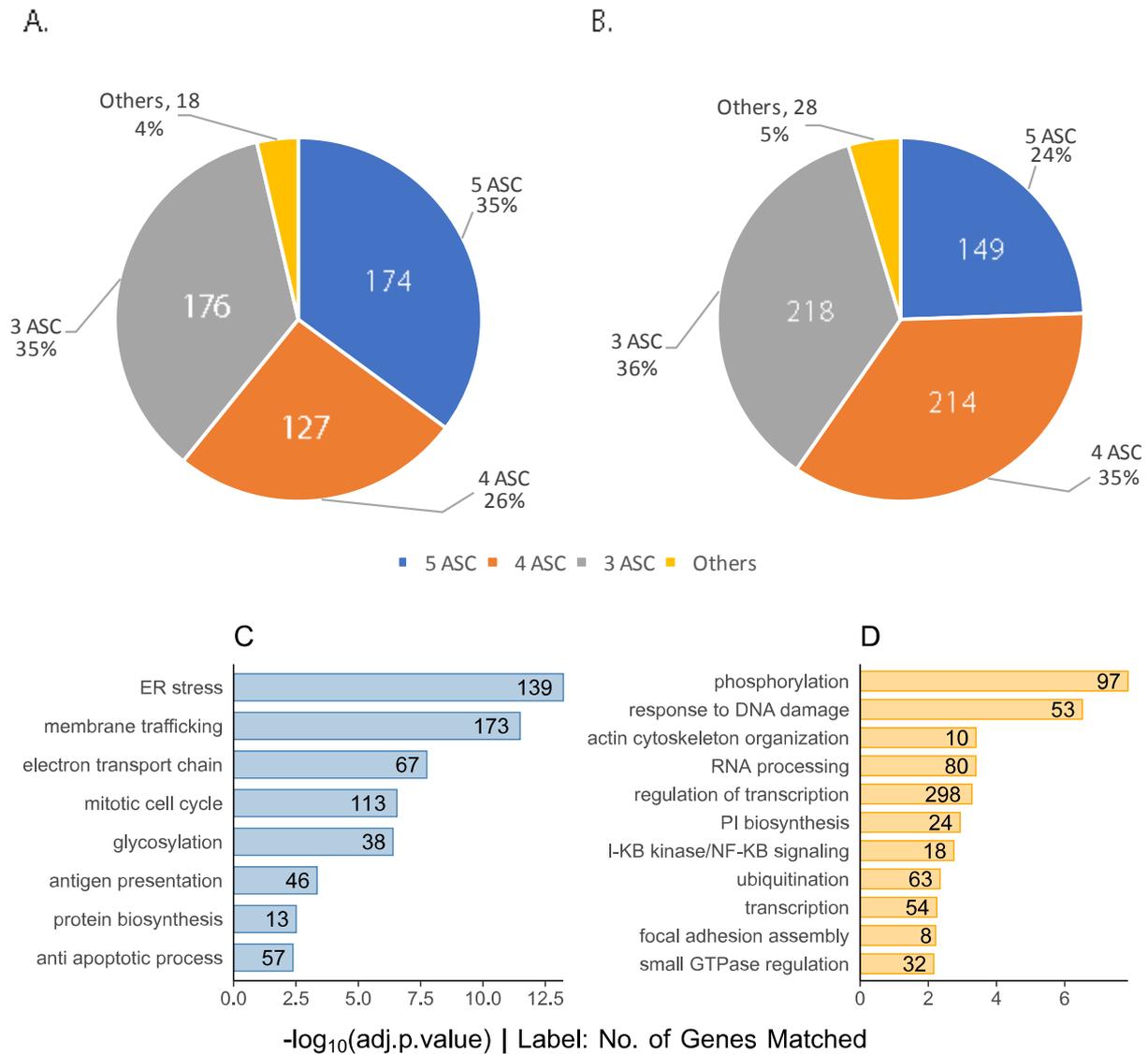
### 3.3.4 Functional Analysis

After narrowing down unique genes showing conserved differential regulation in both mouse and human antibody secreting cells, we studied what biological function and cellular localisation was enriched in this set.

#### 3.3.4.1 GO Biological Processes

We interrogated the “GO Biological Processes 2017” database using EnrichR and found 66 statistically significant GO terms (rank based ranking adjusted  $p$ -value  $< 0.01$ ). We summarised these GO terms based on overlapping genes by utilising the REVIGO API and manual curation. The summarised terms that we assigned to overlapping GO biological process in this chapter are given in Appendix Table 6-1, along with the number of genes enriched per GO term.

Out of 2074 genes showing conserved upregulation across mouse and human ASCs, 495 (23.8%) unique genes showed statistically significant enrichment for at least one GO biological process while downregulated genes showed enrichment in 609 (25.4%) genes out of 2399. While most of the remaining genes had associated GO terms, there were either not

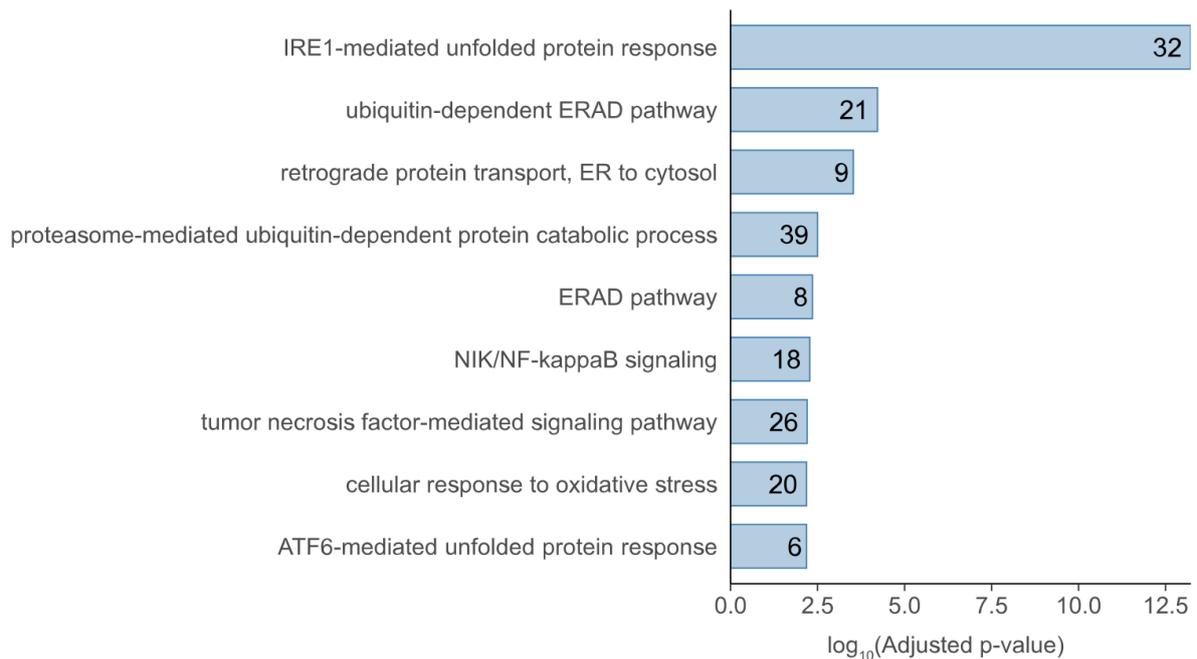


**Fig 3-17** | Degree of consistency in the Differential Regulation of Functionally Enriched Genes and Visualisation of Summarised GO Biological Processes. **A-B.** Pie chart showing the robustness of genes enriched for GO biological functions, as implied by conserved differential expression in up to 5 ASCs. Genes showing conserved regulation in some but not all ASCs help improve the statistical power of enrichment analysis and highlight the importance of avoiding overly stringent filtering. **A.** Upregulated Genes. **B.** Downregulated Genes. “Others” imply genes that were analysed separately due to lack of orthologs or probes across species. **C-D** Summary of GO Biological processes. Bar labels indicate number of genes enriched for each summarised GO term. The enrichment results are ranked by adjusted *p*-value denoting the chance of an enrichment occurring by random chance. ER stress had lowest adjusted *p*-value or highest -log<sub>10</sub>(adjusted *p*-value) meaning it is very likely to be a true positive enrichment. Note that these summarised categories are aggregated from overlapping GO terms as detailed in Appendix Table 6-1. **C.** Upregulated Genes, **D.** Downregulated Genes.

enough members for each term to consider them statistically significant or the member genes were too scattered (rank wise) in the inputted gene lists. It may be tempting to assume that

members of enriched GO terms mostly belong to the gene sets that show consistent upregulation in all 5 ASC types studied. However, Fig 3-17A, B shows that this is not the case, as enriched genes were evenly distributed across gene sets showing varying degree of overlap between different ASC type in mice and humans. The results of GO Biological process enrichment is summarised in Fig 3-17C,D. ER stress response and membrane trafficking were the top enriched functional categories for genes upregulated in ASCs compared to NBCs (Fig 3-17C). On the other hand, downregulated genes were mostly associated with transcriptional regulators and modulation of protein activity through phosphorylation (Fig 3-17D). The specific DEGs that map to these functional groups can be found in Appendix Table 6-3 and Table 6-4.

#### 3.3.4.1.1 Protein Processing in the ER



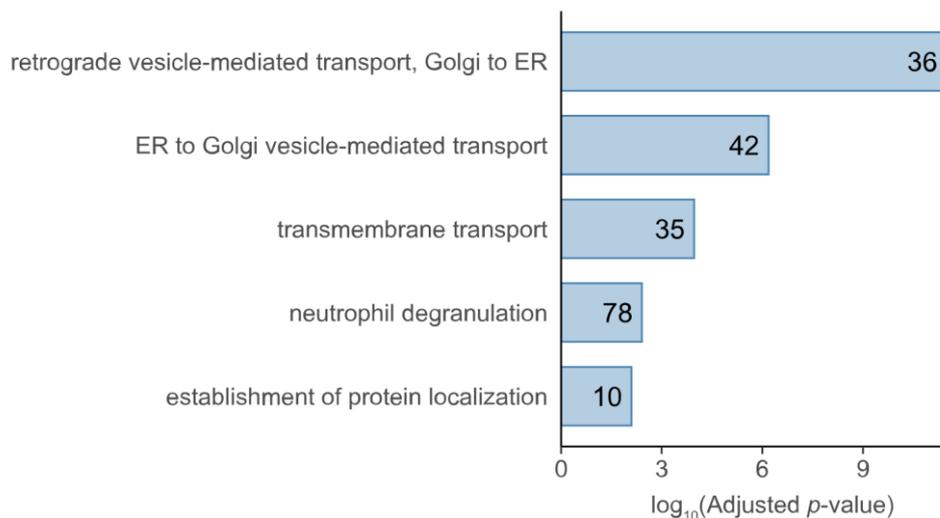
**Fig 3-18** | Enrichment Terms related to Unfolded Protein Response in cross-species microarray analysis. The enrichment results are ranked by adjusted  $p$ -value denoting the chance of an enrichment occurring by random chance. IRE1 mediated UPR had lowest adjusted  $p$ -value or highest  $-\log_{10}(\text{adjusted } p\text{-value})$  meaning it is very likely to be a true positive enrichment. All GO terms shown had adjusted  $p$ -value  $<0.01$ .

ASCs synthesise and secrete thousands of antibodies per second [51]. Under such forward pressure, it is expected that large number of unfolded and misfolded proteins would arise and

accumulate in the ER. This explains why the top biological process enriched for upregulated genes in ASCs compared to NBCs is ER Stress or the Unfolded Protein Response (UPR). We found a total of 139 unique genes upregulated in ASCs likely to be involved in maintaining ER homeostasis and Fig 3-18 shows the breakdown of GO terms under the UPR category.

IRE1 mediated UPR response was the top GO term enriched among biological processes. IRE1 (ERN1) and its downstream splicing target, XBP1, are well-known mediators of UPR and their role in ASCs is well characterised [1]. Splicing of XBP1 mRNA by IRE1 allows XBP1 to upregulate ERAD proteins, protein folding enzymes, chaperones and also lipid biogenesis [145]. As such we see enrichment for the ERAD pathway, proteasomal degradation and oxidative stress likely from disulphide bond reduction for the unfolding of misfolded proteins.

#### 3.3.4.1.2 Membrane Trafficking

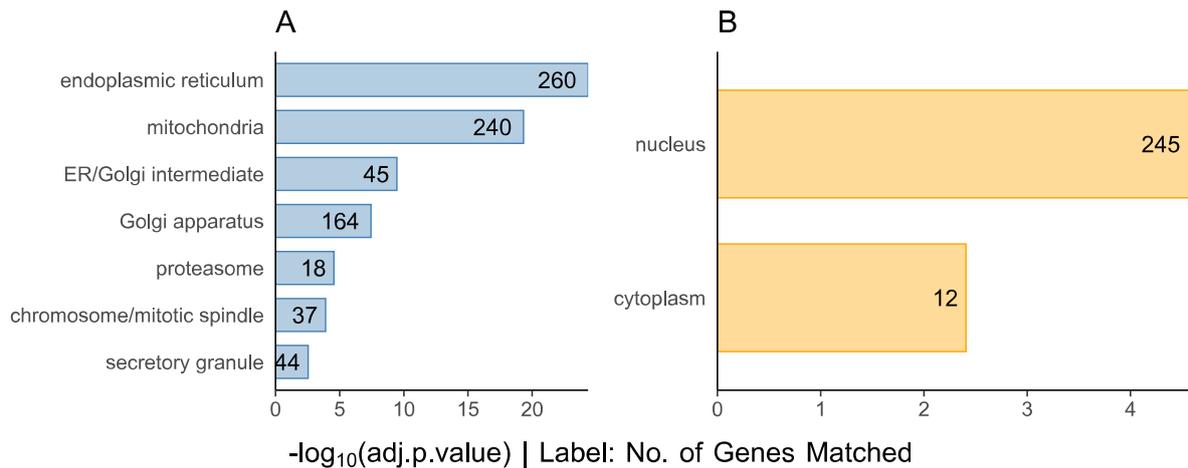


**Fig 3-19** | Enrichment Terms overrepresented for Membrane Trafficking. The enrichment results are ranked by adjusted  $p$ -value denoting the chance of an enrichment occurring by random chance. Retrograde vesicle mediated transport from Golgi to ER had lowest adjusted  $p$ -value or highest  $-\log_{10}(\text{adjusted } p\text{-value})$  meaning it is very likely to be a true positive enrichment. All GO terms shown had adjusted  $p$ -value  $<0.01$ .

When naïve B cells encounter an antigen, they become factories of protein (antibody) production and secretion. As such, we predicted that comparing ASCs to their non-secreting counterparts would allow us to isolate trafficking components responsible for this enhanced

secretory phenotype. In accordance with this hypothesis, we found that membrane trafficking was the second most enriched functional category among genes upregulated in ASCs compared to NBCs (Fig 3-17C). Within this category, enrichment for vesicular transport between ER and Golgi was most statistically significant (Fig 3-19).

### 3.3.4.2 GO Cellular Component



**Fig 3-20** | Summary of GO cellular components enriched among **A.** Upregulated Genes, **B.** Downregulated Genes. Note that these categories are aggregated from overlapping GO terms as detailed in Table 6-2. Genes localised to ER had lowest adjusted  $p$ -value or highest  $-\log_{10}(\text{adjusted } p\text{-value})$  meaning it is very likely to be a true positive enrichment. All GO terms shown had adjusted  $p$ -value  $<0.01$ .

We interrogated the “GO Cellular Component 2017” database using EnrichR and found 31 statistically significant GO terms (rank based ranking adjusted  $p$ -value  $<0.01$ ). The summarised terms that we assigned to overlapping GO biological process are given in Appendix Table 6-2.

Out of 2074 genes showing conserved upregulation across mouse and human ASCs, 676 (32.5%) unique genes showed statistically significant enrichment for specific compartments while downregulated genes showed enrichment in 256 (10.6%) genes out of 2399. The results of GO Cellular Component enrichment is summarised in Fig 3-20. In line with biological function enrichment, the most overrepresented localisation among upregulated genes was the ER followed by mitochondria and the Golgi apparatus. Downregulated genes,

on the other hand, mostly localised to the nucleus and cytoplasm. The specific DEGs that map to these cellular locations can be found in Appendix Table 6-5 and Table 6-6.

#### 3.3.4.2.1 Genes of Interest based on GO enrichment Analysis

We isolated gene sets that are enriched for GO terms of interest and performed further analysis to determine how they are being regulated in ASCs. We selected genes localised to the ER, Golgi, proteasome and secretory granules as well as any DEG associated with ER stress response, anti apoptotic processes, membrane trafficking, glycosylation, and antigen presentation as indicated by GO enrichment analysis. These genes account for a total of 534 upregulated genes detailed in the Appendix Table 6-7. As downregulated genes were mostly related to gene expression or modulation of protein function, we indirectly utilise them to determine their effect on upregulated genes as detailed in the next section.

### 3.3.5 Pathway Analysis

#### 3.3.5.1 Transcription Factors

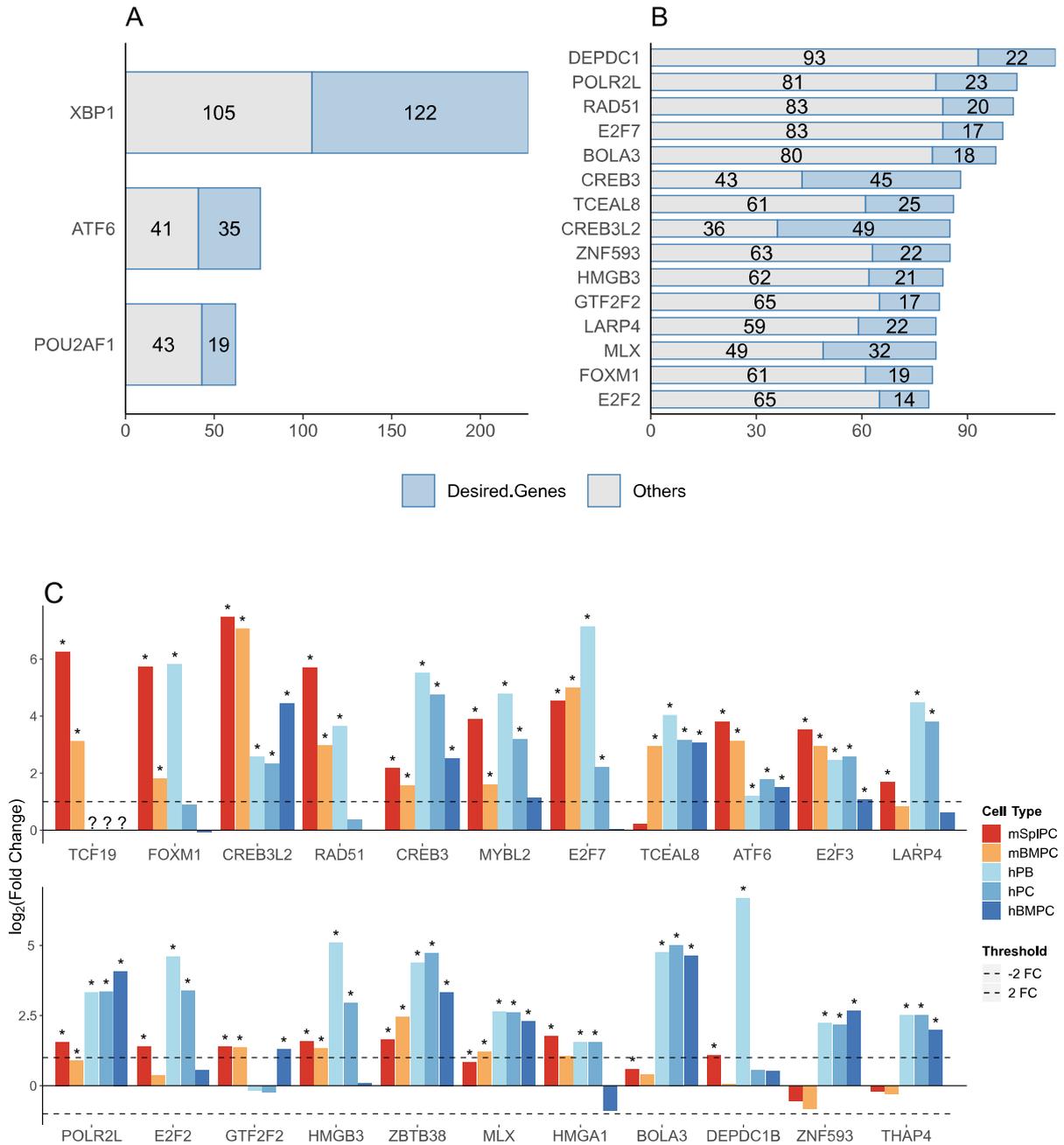
In order to investigate the role of transcription factors in the regulation of DEGs in ASCs as opposed to NBCs we looked at changes in these genes in relation to their downstream targets. Not surprisingly, the key regulators of ER stress, XBP1 and ATF6, as well as the lymphocyte host defense related gene, POU2AF1, were highly enriched in ASCs (Fig 3-21A). XBP1 and ATF6 are well characterised transcription factors implicated in regulating unfolded protein response by upregulating chaperones, protein folding enzymes and the ER degradation pathway [5]. However these TFs only explain the upregulation of 318 unique genes, of which only 149 belong to our desired gene set (Fig 3-21A, in blue). The specific genes upregulated in ASCs by these transcription factors are given in Appendix Table 6-8.

The ARCHS4 database allows us to determine which genes are commonly seen to be co-expressed with transcription factors across hundreds of RNA-Seq profiles. In order to identify potentially novel transcription factors playing a role in the ASC phenotype, we

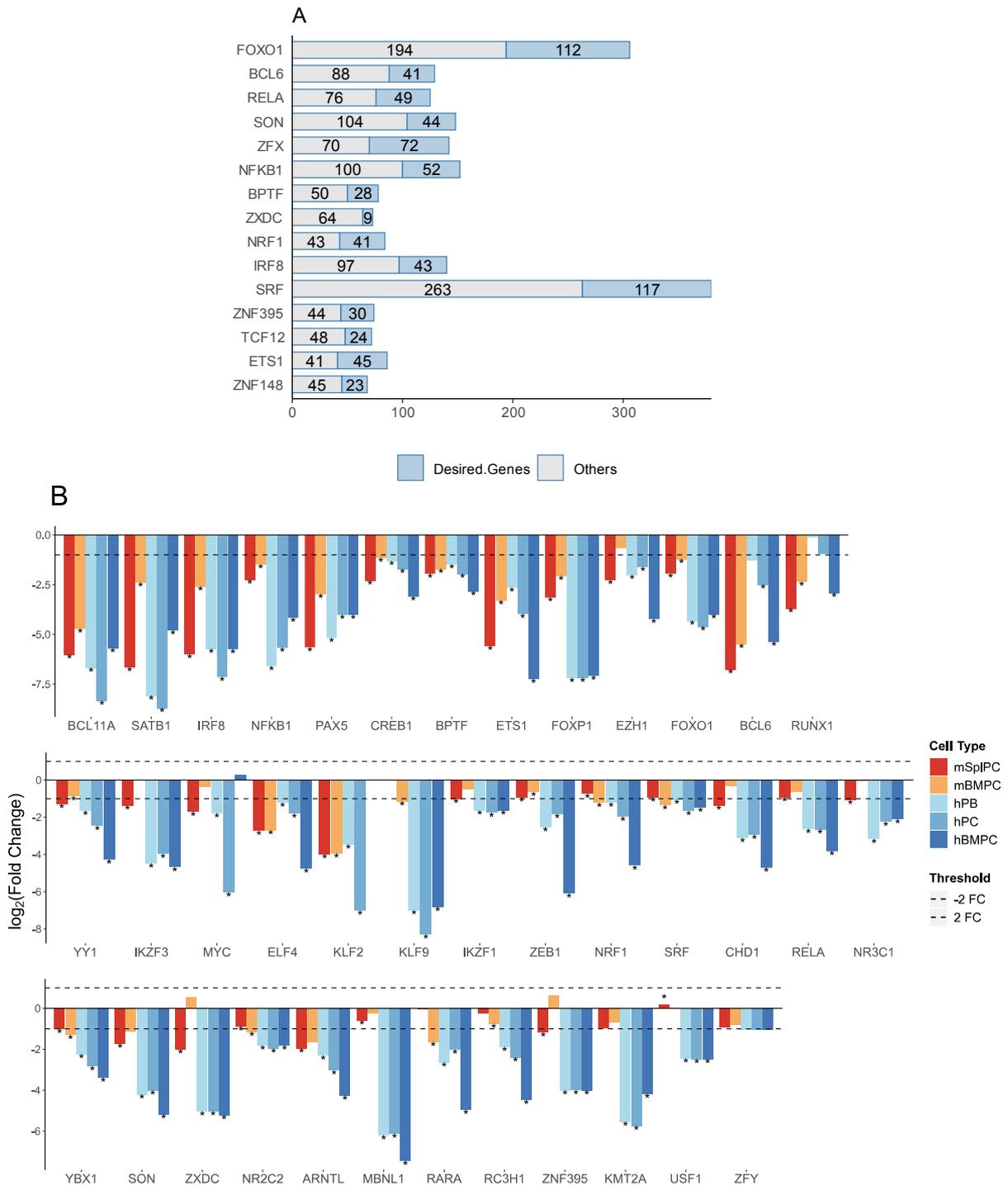
analysed which transcription factors are commonly co-expressed with genes upregulated in ASCs. We found 23 transcription factors that met this criterion. Top 15 coregulated TFs are shown in Fig 3-21B. In order to select the most important candidate among these TFs, we visualised the differential expression of these TFs in Fig 3-21C, to rule out false positives. We note that CREB3L2 was one of the top hits with highest numbers of genes co-expressed with genes belonging to our gene set of interest. This gene has also been highlighted by Shi *et al* as a signature transcription factor for *in vivo* plasma cells[56].

#### *Transcriptional Repressors*

We found 38 transcription factors downregulated in ASCs that were previously found to inhibit the expression of our differentially upregulated gene set (adjusted  $p$ -value < 0.05). Downstream targets of these transcription factors accounted for the expression of 1169 unique DEG of which 381 were among our desired gene set (Fig 3-22A). BCL6 transcription factor is a well known marker of germinal centre B cells, whose downregulation is required for plasma cell differentiation [146]. As expected BCL6 and its upstream enhancer, FOXO1, were the top two TFs consistently downregulated and likewise their repressed targets were upregulated in the ASCs of both mice and humans [147]. The full list of upregulated genes affected by the inhibition of transcription factors are shown in Appendix Table 6-8. SRF, is a TF factor linked to cell proliferation and differentiation. While the perturbation of this regulator appears to affect the largest number of genes of interest, Fig 3-22B illustrates that this gene had relatively low fold change and higher noise than most of the other downregulated TFs. Thus it is unlikely to be a candidate for downstream analysis. In fact the TFs exhibiting greatest downregulation, BCL11A, SATB1, affected genes involved in cell cycle, mitochondrial electron transport respectively (Fig 3-22B and Appendix Table 6-8). Overall, downregulated TF of interest to membrane trafficking were not immediately clear from our results.



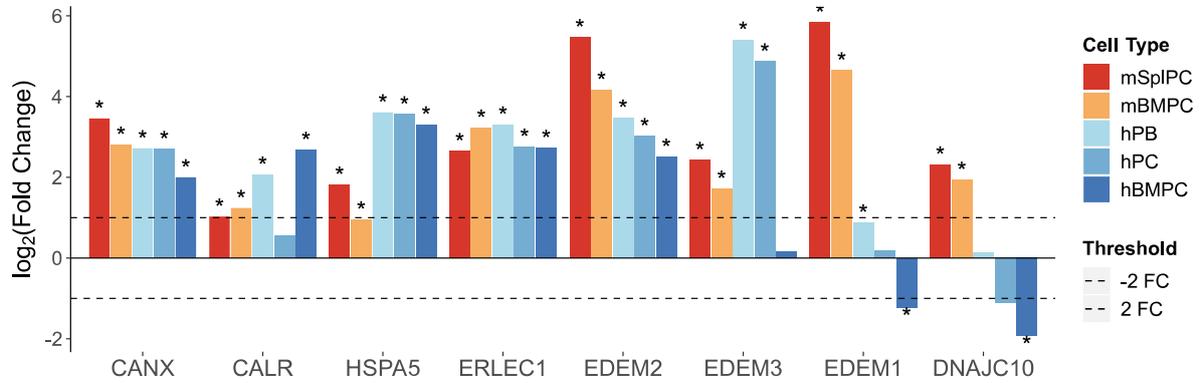
**Fig 3-21** | Bar charts showing potential positive regulators of genes upregulated in ASCs. **A.** Gene upregulation as a result of TF activation, upregulation or gain of function mutation. **B.** TFs found to be commonly co-expressed with relevant gene sets in hundreds of publicly available transcriptomes **A-B.** Transcription factors are ordered by adjusted *p*-value from ranked based ranking (most significant enrichment first). **Desired genes** (in dark blue) are upregulated genes that were enriched for GO summary terms of interest, i.e. ER stress and membrane trafficking related genes as described in Section 3.3.4.2.1. **C.** Differential Expression profile of ASCs compared to NBCs of TFs that are predicted to be co-expressed with considerable number of upregulated genes. Genes are arranged in order of adjusted *p*-value calculated for multi-group comparison. This means the TCF19 would have had highest probability of being truly differentially regulated in ASCs across species, if it had no missing value. (\*) indicate FDR adjusted *p*-value < 0.05. (?) indicates missing data, i.e. no assumption can be made.



**Fig 3-22 | A.** Bar charts showing potential negative regulators of genes upregulated in ASCs as evidenced by TF downregulation, inhibition or loss of function mutation experiments. Transcription factors are ordered by adjusted *p*-value from ranked based ranking (most significant enrichment first). **Desired genes** (in dark blue) are upregulated genes that were enriched for GO summary terms of interest, i.e. ER stress and membrane trafficking related genes as described in Section 3.3.4.2.1. **B.** Differential expression profile of TFs, whose inhibition/downregulation is known to overexpress considerable number of genes upregulated in ASCs. **A-B.** Genes are arranged in order of adjusted *p*-value calculated for multi-group comparison. This means the BCL11A had highest probability of being truly differentially regulated in ASCs across species. (\*) indicate FDR adjusted *p*-value < 0.05.

## 3.4 DISCUSSION

### 3.4.1 EDEM1-ERdj5 complex shows differential regulation in human



**Fig 3-23** | Differential Expression profile of ER chaperones in the ASCs of mice and human compared to NBCs. Asterisks (\*) indicate statistically significant expression among ASCs per species. Most chaperones show DE above the 2-fold change (FC) threshold except EDEM1 and its interactor DNAJC10 (ERdj5).

Due to the considerable upregulation of ER components, we looked into the protein folding machinery of ASCs of mouse as opposed to humans. Our cross-species analysis shows that most chaperones participating in protein folding and degradation of terminally misfolded proteins were consistently upregulated in human and mice ASCs to meet the increased demand for protein biogenesis.

In Chapter 1, Section 1.1.1.1.3, we discussed the role of the EDEM chaperones in mediating the entry of misfolded proteins into the ERAD pathway. Interestingly, according to our cross-platform microarray analysis, EDEM1 isoform was differentially regulated between mice and humans. This gene, implicated in promoting glycoprotein mannosylation, was highly upregulated in mice ASCs but exhibited relatively minor change in human ASCs (Fig 3-23) [14–16].

We note EDEM2 and EDEM3, paralogs of EDEM1, are highly upregulated in both species. EDEM2 has been shown to catalyse the first step of mannosylation and commit misfolded proteins to the ERAD pathway, while EDEM3 and EDEM1 perform redundant

function in mediating the subsequent mannosylation steps [148]. Unlike its paralogs, EDEM1 functions as both a mannosidase and a lectin, whereby it delivers substrates to the ER membrane for retrotranslocation and subsequent degradation [148]. Therefore, it is surprising that unlike mouse EDEM1, the human ortholog does not follow a similar pattern of expression as other lectins such as OS9 and XTP-3B (ERLEC).

The differential regulation of EDEM1 in ASCs across species is supported by the similar pattern of expression of the EDEM1 interactor, DNAJC10 (ERdj5) (Fig 3-23). This co-chaperone reportedly binds to EDEM1 and enhances ERAD activity by catalysing the unfolding of EDEM1 substrates. (Fig 3-23) [21].

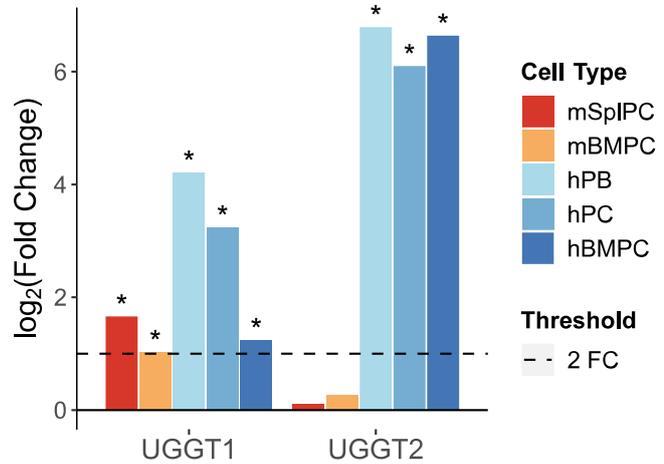
Transcription perturbation analysis showed that EDEM2 and EDEM3 are upregulated in response to XBP1 overexpression but their paralog EDEM1 and its interactor DNAJC10 is unaffected. Although, this explains the muted upregulation in humans it does not account for the considerable upregulation in mice. Together these results suggest that misfolded protein unfolding and delivery to the retrotranslocon mediated by EDEM1 and ERdj5 may be regulated at a post transcriptional level or alternative machinery may perform the function of these proteins in human ASCs. Therefore, further analysis via western blotting or proteomic validation is required to validate these findings.

### 3.4.2 UGGT2 folding checkpoint enzyme is upregulated in human ASCs only

In Chapter 1, Section 1.1.1.2, we discuss genes that mediate the return of misfolded proteins to the folding cycle. We noted that one of these genes, UGGT2 was differentially regulated among human and mice (Fig 3-24). The two known reglucosylation enzymes, UGGT1 and UGGT2, share about 55% similarity in their amino acid sequence. A study in 2010 transiently transfected human UGGT genes into monkey cells and found that UGGT1 expression increases in response to ER stress but not UGGT2 and reasoned that this is due to the UGGT2 protein possibly lacking glycosyltransferase activity [149]. Recently, however, UGGT2 has been shown to

expressly bind and glycosylate proteins in non-native conformations in a similar manner to UGGT1 [150]. However, the specific role of UGGT2 as opposed to the UGGT1 enzyme remains unclear.

As expected, UGGT1 was consistently upregulated in ASCs across species in response to proteotoxic ER stress. Surprisingly, UGGT2 was upregulated as well but was unique to humans only, as the gene expression in the mouse model showed little or no change (Fig 3-24). While this result may contradict Arnold *et al*'s findings, it is important to note that the study had



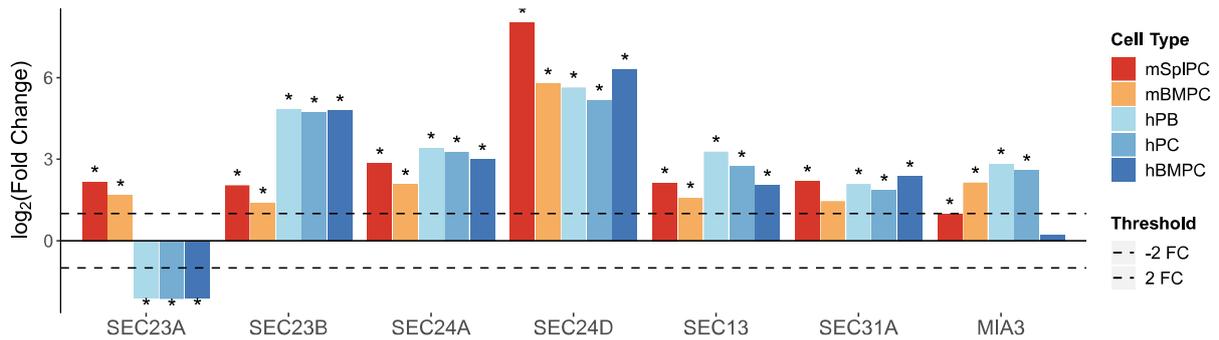
**Fig 3-24** | Expression profile of UGGT genes (microarray).

These genes are known to reglucosylate misfolded proteins in the ER. UGGT1 is consistently upregulated across species.

(\*) indicate FDR adjusted  $p$ -value < 0.05

utilised immortalised fibroblast like cell line from monkey to study human UGGT2 expression. Therefore, the study may not have accurately reflected the physiology of ER stress in human ASCs. Moreover, Takeda *et al* showed that UGGT1 and UGGT2 has similar substrate binding capacity using synthetic substrates in human embryonic kidney cells [150]. We noticed that the binding capacity of UGGT2 has not been studied across species. If UGGT2 is able to detect misfolded substrates specifically incompatible with human, the enzyme may potentially be responsible for mediating species-specific folding in humans. This hypothesis may explain the accumulation of misfolded recombinant proteins in CHO cells, which is a major issue in the biologics industry. Therefore, we propose that UGGT2 is a key candidate for downstream validation by western blotting using mouse and human ASCs.

### 3.4.3 CREB3L2 may be a cargo selector for COPII vesicles



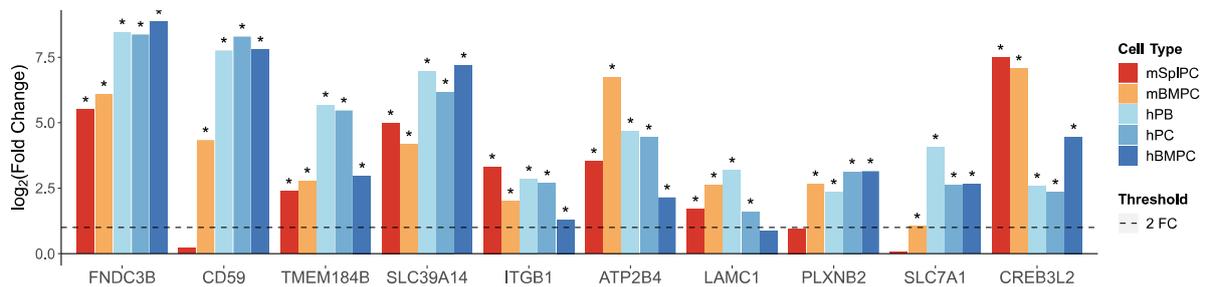
**Fig 3-25** | Differential Expression of COPII vesicles components (microarray). SEC24D clearly stands out compared to other coat proteins. Genes are arranged in order of adjusted  $p$ -value calculated for multi-group comparison.

After proteins are folded, they are packaged into COPII vesicles and transported to the Golgi apparatus for further processing [1]. Fig 3-25 shows the differential expression of COPII coat proteins in ASCs. It is evident that SEC24 isoform, SEC24D, stood out compared to other COPII components. Like ASCs, during hepatic fibrosis hepatic stellate cells enlarge their ER and Golgi apparatus to become a factory for releasing large amounts of  $\alpha$ -smooth muscle actin and collagen I [151, 152]. A recent study reported that SEC24D transcription is activated by the CREB3L2 during hepatic stellate cell differentiation [153]. Interestingly, CREB3L2 was one of the very top ranking genes upregulated in ASCs compared to NBCs across species (Fig 3-21C). In addition, the gene shows co-expression with the highest number of genes belonging to our gene set of interest compared to other TF candidates (Fig 3-21B). Although 28 of its enriched targets, primarily chaperones, are known to be upregulated as a result of XBP1 or ATF6 overexpression, the regulation of 57 unique genes coexpressed with CREB3L2 are less clear.

The bZIP transcription factor, CREB3L2/ BBF2H7 is a transmembrane protein implicated in ER stress responses in chondrogenesis and improved protein secretion through regulation of SEC23A and MIA3 (TANGO1) [154, 155]. Studies in zebrafish mutant

**Table 3-5** | Cargo Proteins co-expressed with CREB3L2

	Size (aa)	Description	Glycosylated?
FNDC3B	1204	Fibronectin	Yes
CD59	128	Antigen	Yes
TMEM184B	407	-	-
SLC39A14	492	Solute Carrier	Yes
ITGB1	798	Integrin	Yes
ATP2B4	1241	Solute Carrier	No
LAMC1	1601	Laminin	Yes
PLXNB2	1838	Plexin	Yes
SLC7A1	692	Solute Carrier	-

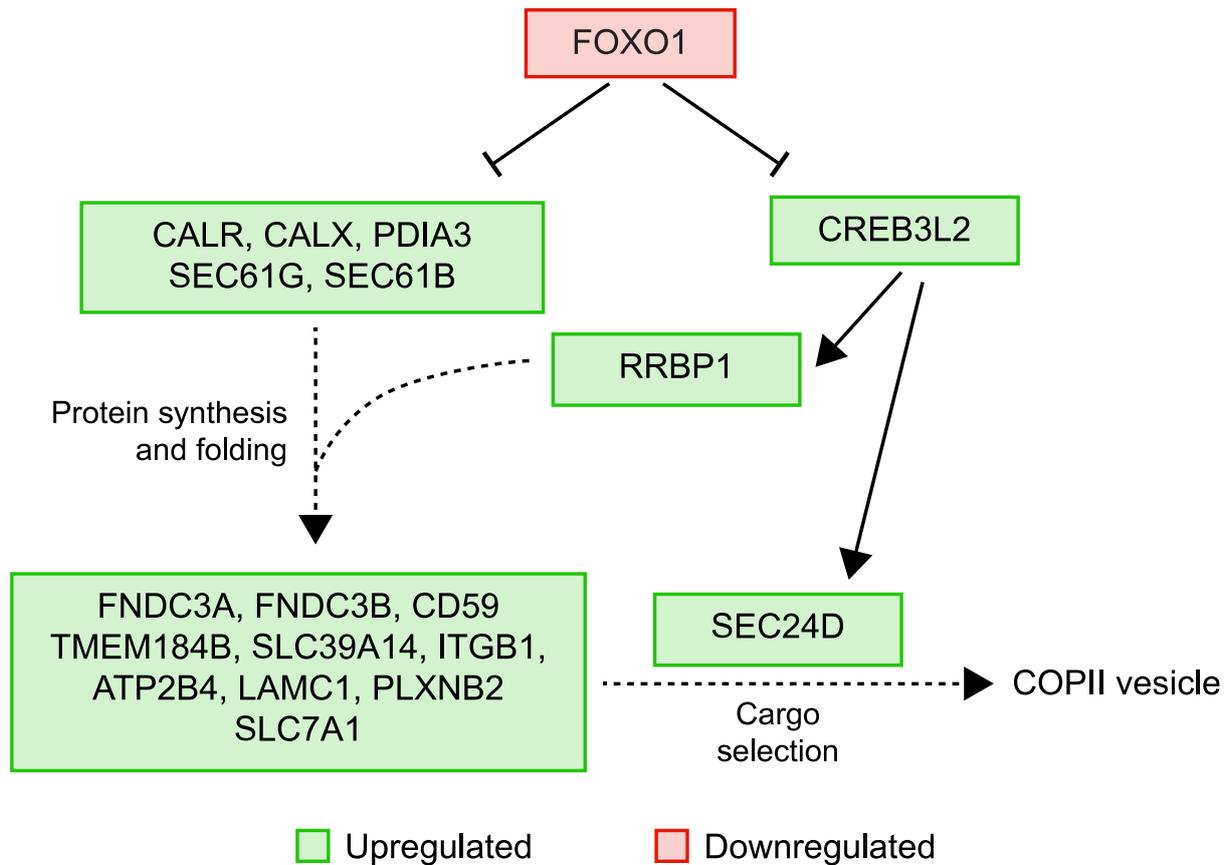


**Fig 3-26** | Differential regulation of cargo proteins in ASCs vs NBCs (microarray). These genes, destined for cell surface, are predicted to be co-regulated with CREB3L2. Genes are arranged in order of adjusted  $p$ -value calculated for multi-group comparison. This means the leftmost gene had highest probability of being truly differentially upregulated in ASCs across species compared to equivalent NBCs. (\*) indicate FDR adjusted  $p$ -value < 0.05

(*feelgood*) and the CREB3L2 knockout in medaka fish has previously shown a marked reduction in ER-to-Golgi transport [156, 157].

SEC24D has been implicated in the selective transport of the GPI anchored transmembrane protein, CD59 [158]. And our analysis shows that in addition to CD59, CREB3L2 is co-expressed with 8 other cargo proteins destined for the cell surface. As shown in Table 3-5, the majority of these cargoes are large glycosylated proteins, whose role is not restricted to chondrogenesis. Due to the diversity of these protein cargo, we hypothesise that CREB3L2 may play a crucial role in the transport of large proteins, by regulating cargo selection of the COPII vesicle. Whether this cargo selection includes antibodies remains unclear. Further studies at the protein level is needed to validate our results.

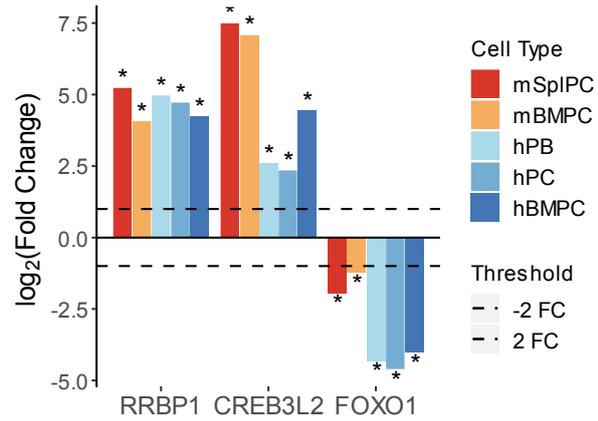
## 3.4.3.1 CREB3L2 may regulate ribosome recruitment to the ER



**Fig 3-27** | Predicted Gene regulatory network for large protein transport in the ER. FOXO1 downregulation potentially removes inhibitory effects on CREB3L2, translocon and protein folding components and explain their upregulation. CREB3L2 overexpression leads to upregulation of SEC24D and potentially the RRBP1 gene. Ribosome recruitment by RRBP1 and increased protein folding and translocation activity lead to synthesis and accumulation of cargo proteins. The recruitment of these cargo into COPII vesicles and transport to the Golgi is likely facilitated by the now abundant SEC24D.

According to existing perturbation experiments, master regulators of ER stress, XBP1 or ATF6, do not affect CREB3L2 expression. Our pathway analysis results show that CREB3L2 might be upregulated as a result of FOXO1 inhibition (Appendix Table 6-8). Upon encountering an antigen, B cells commit to plasma cell differentiation and concurrently downregulate FOXO1 [147]. As FOXO1 inhibition enhances gene expression of translocon components, folding enzymes and chaperones, it is logical to assume that FOXO1 inhibition improves rate of protein biosynthesis. If CREB3L2 is a regulator of large cargo selection, its upregulation may be an indirect response to increased supply of folded proteins in the ER.

Interestingly, however, CREB3L2 gene regulation has been predicted to coincide with the poorly characterised ribosome anchor, RRB1 (p180), whose upregulation was consistent across species in ASCs (Fig 3-28). CREB3L2 may indirectly contribute to enhanced protein synthesis if it acts upstream of RRB1. ASCs exhibit extensive rER compared to their non-secretory precursors to account



**Fig 3-28** | Differential Expression of RRB1 and FOXO1 relative to CREB3L2 in ASCs vs NBCs. While RRB1 and CREB3L2 is upregulated in ASCs, FOXO1 is strongly downregulated. (\*) indicate FDR adjusted  $p$ -value < 0.05

for increased demand for antibody secretion [1]. This may be explained by the upregulation of RRB1, whose presence has been reported to enhance rough membrane proliferation, protein translocation and mammalian secretory phenotype [159–161]. To date, the regulatory mechanisms governing RRB1 expression remains unclear. Our pathway analysis show that RRB1 gene expression most significantly correlates with CREB3L2 (Appendix Table 6-9). Therefore, we hypothesize that CREB3L2 may be acting upstream of RRB1 to enhance protein synthesis. Our proposed regulatory network for CREB3L2 is illustrated in Fig 3-27.

#### 3.4.4 FNDC3B and TMEM184B may be potential biomarkers for ASCs

Among the cargoes coregulated with CREB3L2, we note that the poorly characterised transmembrane proteins, FNDC3B and TMEM184B, and the divalent metal transporter, SLC39A14, were the top upregulated proteins in ASCs across species (Fig 3-26). Therefore, these genes maybe potential biomarkers for ASCs and are good candidates for downstream validation via proteomics or western blotting.

### 3.5 CONCLUSION

In this chapter we carried cross species meta-analysis and merging of microarray data to ASCs. We found that secretory phenotype in ASCs was maintained by strong upregulation of UPR and membrane trafficking. Meanwhile, components showing downregulation were those related to transcription factors and kinases. Specifically, we found robust upregulation of the poorly characterised transcription factor, CREB3L2, and proposed a potential gene regulatory network for this component. Furthermore, we identified potentially novel biomarkers for ASCs.

Our premise for carrying out the cross species meta-analysis was to narrow down robust candidates for downstream analysis assuming inconsistent regulation are poor candidates. However, differences in glycosylation of recombinant proteins in CHO cells are likely to be due to these very inconsistencies between species. As such, we identified differences in the expression of UGGT enzymes and ERDM1-ERdj5 complex between mouse and human ASCs. As we did not have an independent cross-species analysis to refer to, it is unclear whether this cross-species differential regulation is a robust result. In the next chapter we analyse the recently published RNA-Seq data for mice and human ASCs to verify this expression pattern.

We note that candidates identified in this chapter require protein level validation. Our colleagues E Rajan and AWA Aswani have recently generated the proteomic profile of mouse B cells and plasmablasts. In the next chapter, we carry out proteomic analysis to soft validate our candidate genes.

# Chapter 4

## Multi-Omics analysis of ASCs

## 4 MULTI-OMICS ANALYSIS OF ASCS

### 4.1 BACKGROUND

In Chapter 3, we used microarray profiling to study differential expression of genes between antibody secreting cells (ASC) and their non-secreting naïve B cells (NBC) counterparts. While cross species conservation in gene expression patterns indicated the robustness of our results, the reliability of potentially relevant species-specific expression was less clear. In this section, we ascertain the reproducibility of different gene expression patterns by studying equivalent RNA-Sequencing data and whole cell proteome profiling. We have shown that hundreds of genes were robustly upregulated in ASCs as opposed to NBCs. It is not feasible to validate all these genes in our cell biology laboratory. Therefore, we have opted to soft validate most of our results through proteomics.

#### 4.1.1 RNA-Sequencing

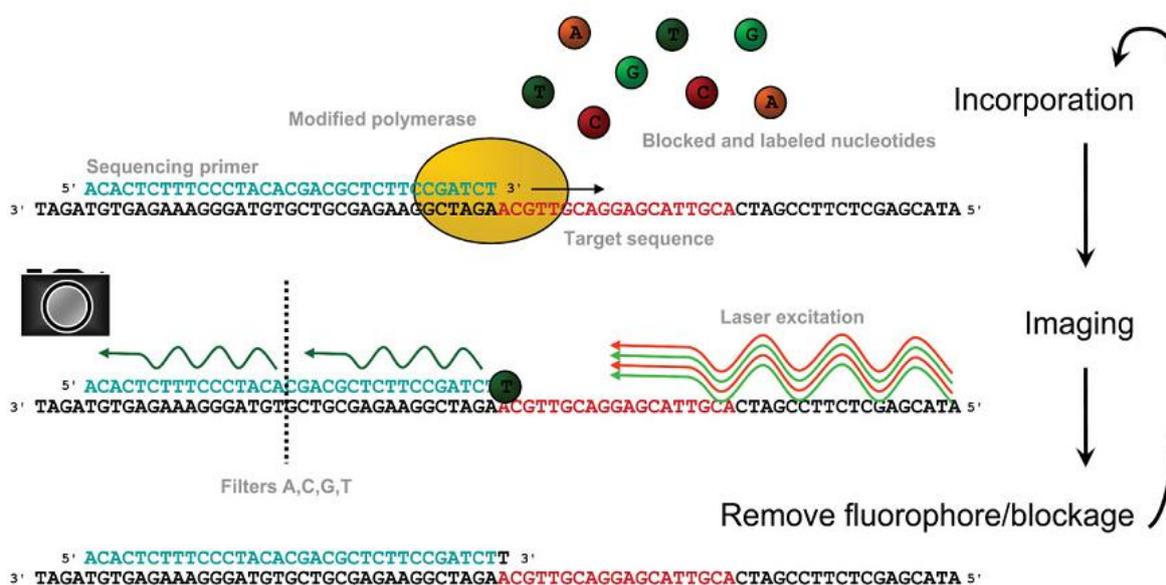
In this chapter we delve into RNA-Seq analysis to obtain robust gene expression patterns in ASCs. As such we review RNA-Sequencing and discuss its advantages over microarrays to underscore how, taken together, results from both platforms can serve as strong evidence for gene expression.

##### 4.1.1.1 History of RNA-Sequencing

Nucleic acid sequencing involves resolving the arrangement of nucleotides in a nucleic acid strand. The Human Genome project was completed using first generation sequencing based on Frederick Sanger's chain termination method [162]. In the late 1990s shot-gun sequencing emerged. This procedure relies on digesting template DNA or RNA into short fragments following cloning and immobilisation of the reverse strand. The complementary strands are then enzymatically extended and each extension is detected via incorporation of fluorescently labelled deoxyribose nucleotide triphosphates (dNTPs). As this process could be done in

parallel for each shotgun fragment it was much faster than Sanger sequencing [163]. This sequencing method would later become the basis of next generation sequencing (NGS) technology available today. NGS began with the sequencing of only bacterial genomes [164], since then the falling costs and improved robustness of analytical methods have allowed the sequencing of 326,287 bacterial species, 3558 Achaea, and 31,093 eukaryotes according to Genomes Online Database as of 2019.

#### 4.1.1.1.1 Illumina Sequencing



**Fig 4-1** | Overview of Illumina Sequencing (Kircher & Kelso 2010)

The RNA-Seq data used in this chapter was generated by Illumina HiSeq 2500. The typical protocol for RNA sample preparation for this instrument requires sample RNA to be reverse transcribed into cDNA and primed with random hexamers [165]. The Illumina instrument then sequences these fragments via a method akin to Sanger's capillary sequencing method where fluorescently labelled dNTP incorporation to target sequence is accompanied with chain termination. The difference is that the system utilizes reversible dye-terminators which are subsequently removed so chain extension and base identification can continue [166]. The HiSeq 2500 in particular can sequence around 44.44 million reads per hour, where each read

is a maximum of 100 nucleotide length [167]. Furthermore, sequencing with this instrument can cost as little as only 0.05 USD per mb today [168, 169]. This makes Illumina one of the choice methods for sequencing complex genomes, counting mRNAs, small RNA and single nucleotide polymorphisms [166].

#### 4.1.1.2 Advantages of RNA-Seq

Unlike microarrays that rely on optical signals, RNA-Seq measures the number of transcripts present in a sample by directly sequencing cRNA or cDNA. The sequenced fragments or reads are then aligned to a reference genome using analysis software to generate the whole cell transcriptome or alternatively *de novo* assembled into a new transcriptome if reference genome is unavailable [170]. As such the process is independent of genome annotations as it is not reliant on predesigned probes. There is mounting evidence that RNA-Seq data is highly reproducible and exhibit very little technical variations compared to hybridisation based platforms [61, 171–174]. Transcripts expressed at high levels tend to oversaturate microarray probes, while low abundance transcripts tend to fall below background noise levels. This is why RNA-Seq analysis outperforms microarrays in detecting very low or very high abundance transcripts [61]. Moreover, a study estimated the accuracy of RNA-Seq and microarray datasets by comparing it to equivalent proteomics data and found that RNA-Seq transcript counts were more representative of absolute transcript levels [175].

#### 4.1.1.3 Limitations of RNA-Seq

Nevertheless, RNA-Seq remains a predictive science due to certain limitations. Complementary DNA or RNA must be fragmented before they can be detected by the RNA-Sequencer. This typically results in short reads some of which do not map to a unique gene. Paralogous genes with high sequence similarity further complicate mapping prediction for such non-unique hits. Therefore, gene assignment for multi-genic mapping remain an unavoidable issue for RNA-Sequencing.

Typically, *de novo* sequencing methods suffer from issues arising from inconsistencies in fragment sampling and sequencing due to preferential fragmentation sites, primer bias and composition of tagged dNTPs in the reaction mixture [176, 177]. While some of these biases can be addressed using computational algorithms, they can lead to some fragments being undetected altogether.

One limitation of sequencing by synthesis is the preference for longer genes. Fragmentation of genes means that longer genes produce a larger number of fragments than smaller ones. Using transcript per million (TPM) measures and TMM normalisation (Section 4.1.2.1.6) account for systematic variation potentially arising from differing transcript composition to an extent but they do not fully address the gene length bias [94]. Furthermore, these normalisations are performed separately per study. This is of particular importance when conducting cross-species analysis. Although, human and mice exons show high conservation, about 68% of human introns have been reported to be 1.5 times larger than mice [178]. This means RNA-Sequencers will be better able to detect these longer human mRNA than their mouse equivalent. This could explain some of the genes detected in humans but not in mice.

We note that in this experiment the human data was generated using single cell transcriptomics while mouse data was generated using bulk transcriptomics. While single cell transcriptomics should theoretically be more accurate in predicting gene expression between cell types, the input sample RNA content and thus the library size tend to be much smaller than in bulk RNA-Seq analysis [179]. This can unavoidably lead to genes detected in mice but not human.

#### 4.1.1.4 Microarrays can complement RNA-Seq analysis

In such cases where genes are undetected in one species but not in the other, microarray evidence can complement RNA-Seq data, as probes for such genes should theoretically ensure correct detection. As microarray probes are a set length, the higher number of fragments in a

longer gene is irrelevant. This is viable as a previous study has reported that differential expression in RNA-Seq and microarray have considerable overlap [171]. It is this overlap that we exploit to refine our search for candidate genes for downstream analysis.

## 4.1.2 RNA-Seq Data Analysis

### 4.1.2.1.1 Retrieving Publicly Available Data

Overall steps involved in RNA-Seq data analysis are summarised in Fig 4-2. In house generated RNA-Sequencing data are typically outputted in FastQ format and can be piped into downstream processes immediately. If using publicly available data, raw RNA-Seq output can typically be downloaded as a compressed Sequence Read Archive (SRA) file from the SRA repository at NCBI or EBI. This compressed format is not usable for downstream analysis and must be converted to FastQ format for downstream processing.

### 4.1.2.1.2 Trim Adapters / Contaminants

RNA library preparation can introduce biases into RNA-Seq output due to possible microbial contamination, polymerase errors and the necessary introduction of random hexamer adapters for reverse transcription [180]. This can further be compounded by errors during sequencing i.e. mistakes in optical detection and phasing errors [180]. Phasing issues in sequencing by synthesis occur when more than one nucleotide get incorporated in the chain during the same

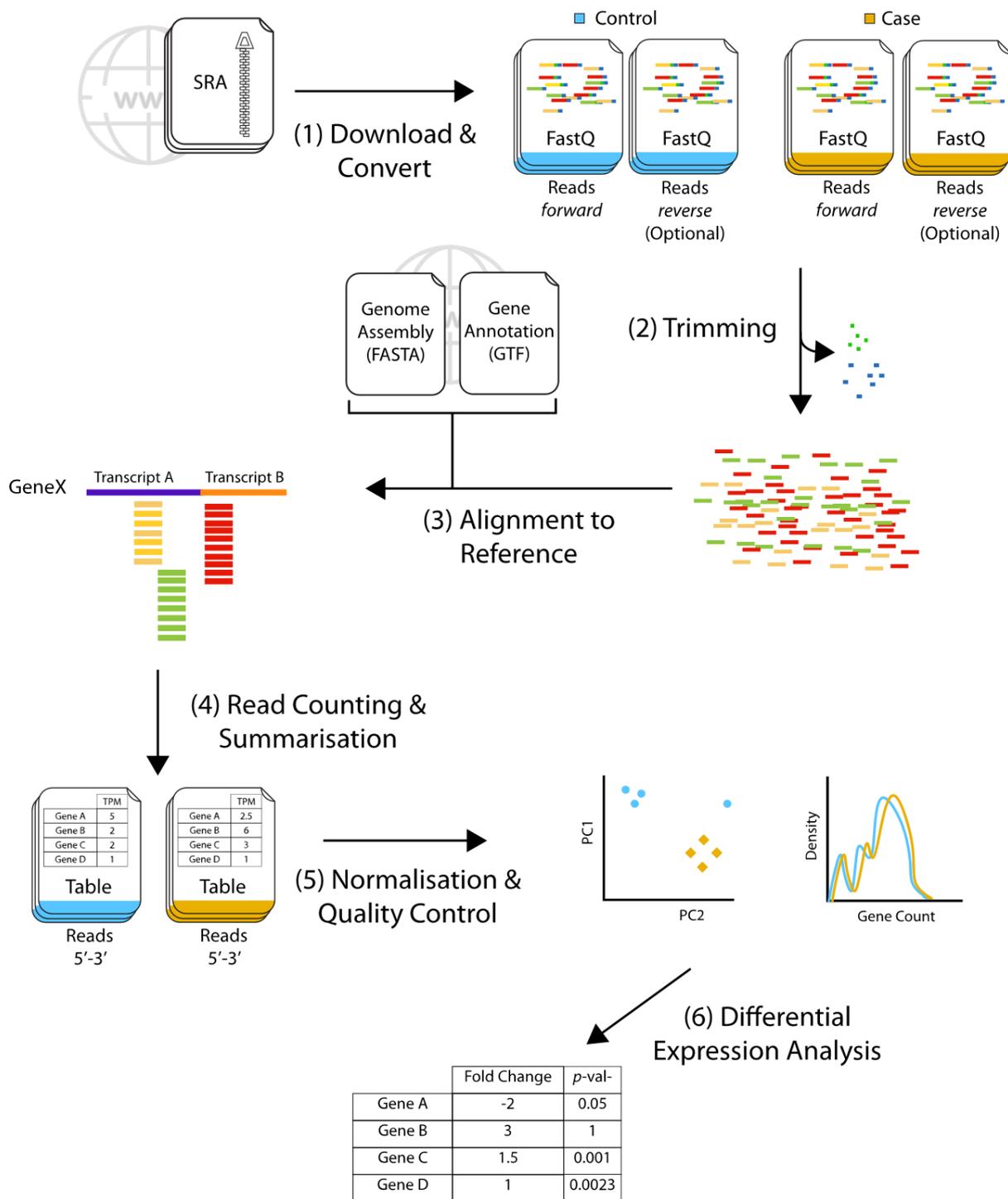


Fig 4-2 | Schematic of RNA-Seq Analysis Workflow. (1) Publicly available RNA-Seq reads are downloaded and converted to a usable format such as FastQ. (2) Contaminants such as adapters and low-quality reads are trimmed away. (3) Reference genome assemblies and gene annotations serve as reference to which trimmed reads are aligned. (4) Reads aligned to transcripts are counted and summarised as gene counts. (5) Counts are normalised and visualised for quality control purposes. (6) Normalised counts are used to calculate differential expression.

phase or the chain terminator does not get removed properly. This leads to poor quality reads that tend to accumulate with sequencing time i.e. at the 3' end of sequence reads [181].

Therefore, after retrieving the raw reads of an RNA-Seq run, they have to be trimmed to remove contaminants and low-quality reads [94].

#### 4.1.2.1.3 **Mapping to Reference Genome**

Once these processing steps are complete, reads must be aligned to a reference genome with accompanying gene annotations in order to identify which gene or transcript the reads belong to. If reference genome does not exist, *de novo* assembly can be performed to generate a new one. However, this step is unnecessary for mice and human genomes as highly curated versions already exist [94].

#### 4.1.2.1.4 **Read Counting and summarising gene expression**

Next reads aligned to specific transcripts must be counted and then summarised for the estimation of gene counts. These gene counts have to then be normalised before performing differential expression analysis [94].

#### 4.1.2.1.5 **Low Count Filtering**

RNA-Sequencing utilises random sampling to quantify a transcriptome. This can directly affect accuracy and typically manifests in genes with low expression [182]. Therefore, to reduce false positives, it is common practise in RNA-Seq workflows to remove these low confidence genes. This not only improves confidence but has also been shown to improve differential expression analysis [182].

#### 4.1.2.1.6 **Normalisation**

In this chapter, we utilise Trimmed Mean of M value (TMM) normalisation to handle systemic variations. Systemic variations typically arise from external factors irrelevant to biological differences, such as sample handling, sequencing depth, etc, that often result in varying sample distributions. The TMM method assumes that most genes in a sample are not differentially expressed. It separately calculates log<sub>2</sub> ratio of case vs control and trims away ratios for differentially expressed genes. Then the average log ratio of non-differentially expressed genes

serve as a scaling factor that is used to transform the counts of each sample, such that differences in sequencing depth and other unwanted variables are accounted for [183].

#### 4.1.2.1.7 Differential Expression

As with microarrays we opted to utilise the limma R package for calculating differential expression in this chapter. Standard limma procedure fits data to a normal distribution. This is not suitable for discrete gene counts generated by RNA-Seq as the variation from the mean is exaggerated with higher counts, and understated for lower counts. As such an underlying distribution cannot be assumed for such data. While logarithmic transformation brings these gene counts to a comparable level, it tends to overcompensate and obscure the importance of highly expressed and overestimate the importance of low-expressed genes. The variance modelling at the observational level (voom) algorithm was created to address this issue. The function logarithmically transforms count data while remembering the importance or weight of these genes, which is then reflected on the calculated fold change and  $p$ -values [184].

#### 4.1.2.1.8 An Error Prone Process

Unlike microarrays, RNA-Seq reads from each run typically take multiple gigabytes of disk space as each individual sample contain millions of reads. Processing each sample can take hours to fully complete on a low performance computer. The entire process can be done using proprietary software or open source tools. We chose the latter option as proprietary software can cost hundreds of dollars each year in subscription.

No single open source tool can carry out the entire RNA-Seq analysis workflow. Starting from processing raw data to differential expression analysis, each step requires one or more different tools, meaning familiarity with several different software across languages is required to complete the analysis. Depending on computational resources available, the whole process can take hours or even days to complete depending on the number and size of samples being studied. The need to use different tools for each step increases the scope for user errors,

and these errors can then be carried over and compounded in the next steps leading to the waste of considerable computational time and power.

#### 4.1.2.1.9 **Defensive Programming**

In order to avoid such issues, it is important to utilise a tool that anticipates errors as soon as it is executed, gracefully fails and then instructs the user on how to recover from the failure.

#### 4.1.2.1.10 **RNA-Seq Pipeline**

In informatic pipelines, the output of one program becomes the input of another. This concept can be applied to RNA-Seq analysis for automation. For example, trimmed reads generated by a trimming software automatically become the input for the alignment software. Using an RNA-Seq pipeline can reduce the complexity in manually executing each tool and maintain processing conformity among all samples passed through the pipeline. Therefore, we aim to create our own RNA-Seq analysis pipeline equipped with defensive programming to streamline our analysis.

### 4.1.3 **Proteomics**

#### 4.1.3.1 **Principles of Proteomics**

Unlike transcriptomics, whole cell proteome profiling relies on the molecular weight of proteins. The basis of proteomic measurements is mass spectrometry, where peptides generated from proteolytic digestion are ionized in gas phase and separated by their mass to charge ratio ( $m/z$ ). The peptide masses are then referenced against a database of *in silico* digested proteins for identification. As it is, a mass spectrometer is unable to handle complex mixtures of different proteins. Therefore, proteomic analysis today utilises tandem mass spectrometry (MS-MS).

MS-MS involves two mass spectrometers arranged in a series. Proteolytically digested and chromatographically separated peptides entering the first mass spectrometer are ionised and separated by their  $m/z$  ratio. Then individual peptides are selected and passed into a

chamber for further fragmentation to lower their complexity. These fragments are again separated by  $m/z$  values by the second mass spectrometer, which then outputs a mass spectra that can be used to computationally predict the peptide sequences, abundances as well as post translational modifications [185, 186]. We have illustrated a typical MS-MS run in Fig 4-1.

#### 4.1.3.1.1 Separation

Complex mixture of peptides entering a mass spectrometer confounds results as the machine is unable to differentiate between two separate peptide ions. Therefore, liquid chromatography techniques are used to separate complex mixtures by specific affinities, such as hydrophobicity, such that the peptides form an affinity gradient. Thus at any given time simpler and purer peptides are sequentially fed to the mass spectrometer [185, 187].

#### 4.1.3.1.2 Ionisation

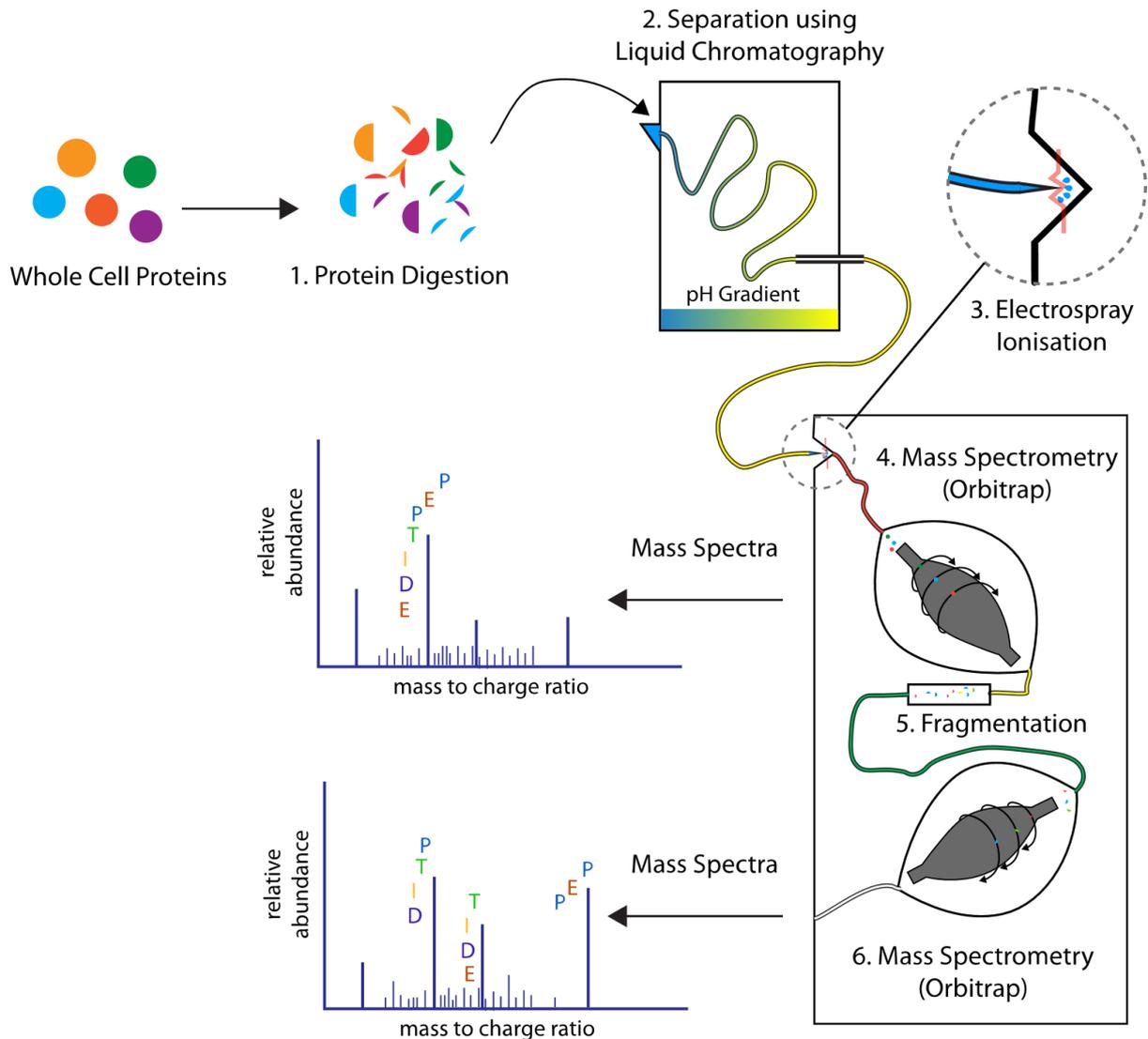
Before peptides can enter the first mass spectrometer they have to be ionised. The instrument utilised in this study was equipped with a liquid chromatographer (LC) coupled to Electrospray ioniser (ESI). Here LC-separated peptide droplets are sprayed via a needle into the first mass spectrometer. As droplets travel, a high voltage is passed through them to ionise the peptides before they reach the mass spectrometer [188].

#### 4.1.3.2 Mass Analyser - Orbitrap

A type of ion trap mass analyser (Orbitrap Elite™ Hybrid Ion Trap-Orbitrap from Thermofisher) was utilised in this study. The core of the orbitrap is charged using a high voltage. As ions enter the orbitrap, the shape and charge of the orbitrap core causes them to oscillate around the core and reach a stable circular orbit. A complex equation then uses the period of oscillation to calculate the mass to charge ratio of the ions [186, 187].

## 4.1.3.3 Fragmentation

After exiting the first spectrometer, peptide ions enter a fragmentation chamber where they are typically bombarded by inert molecules, which causes the peptide to break into smaller



**Fig 4-3** | A Schematic Representation of Proteome Quantification. (1) Whole cell proteins must be isolated, purified and proteolytically digested (usually with trypsin). (2) Resultant peptides are inserted in the liquid chromatographer which separates peptides based on affinities such as hydrophobicity and an affinity gradient is established. (3) Separated peptides are sprayed by a needle at the entrance to the mass spectrometer and ionised on the way via a high electric voltage. (4, 6) Peptide ions entering the orbitrap mass spectrometer travel forward while oscillating around the charged core. The period of oscillation is used to calculate mass to charge ( $m/z$ ) ratio. (5) Peptides with abundances above background noise are sent into the fragmentation chamber where peptides are broken up into smaller fragments using high voltage or collision. (6) Fragmented ions are passed through an Orbitrap mass spectrometer and resultant  $m/z$  ratios are outputted as mass spectra for estimation of peptide abundance, sequence and post translational modifications.

fragments. This specific form of fragmentation is called fast atom bombardment [185, 187]. Fragmented ions are fed to the second mass spectrometer, whose mass spectra output can be used to determine peptide sequence and abundance.

#### 4.1.3.4 Quantification methods

In labelled quantification methods, different samples are mixed and analysed together after they have been labelled with stable isotope labels. The quantification is entirely based on the ratio of isotope labelled peptide pairs, where each member of a pair come from a different sample. Unfortunately, isotope labelling is costly and time consuming as it complicates sample preparation, increases the mass spectrometry run time, necessitates expensive reagents, proprietary software and high sample concentrations [189].

Label free approaches, on the other hand, involve measuring each sample individually. The quantification is like mRNA-Seq in that it compares measurements between samples to determine differential expression. In this case peptide intensities or spectral protein counts between samples are compared [189].

#### 4.1.3.5 Limitation of Proteomics

Despite improvements in reproducibility and robustness of proteomics approaches over the years, its usage continues to lag behind transcriptomics. This can be attributed to several factors.

##### 4.1.3.5.1 Sensitivity

PCR amplification now allows RNA-Sequencing of the level of a single cell. This level of sensitivity cannot be achieved in proteomics as amplification of total cell proteome is near impossible with current technology [190].

#### 4.1.3.5.2 **Proprietary software and less robust algorithms**

Most RNA-Seq analysis software are free and open source and has considerably matured in the recent years. This is not the case for proteomics, where the use of mainly proprietary software has limited the development, streamlining and benchmarking of proteomics algorithms [190].

#### 4.1.3.5.3 **Difficulty in interpretation**

As proteomics is based on measuring ions with the same m/z ratio, different molecules with the same m/z ratio confounds interpretation. This means it is difficult to say for sure if a peak is definitely a protein of interest as it may just as well be a different peptide ion or even a non-peptide ion having the same m/z ratio [190].

#### 4.1.3.5.4 **Limited number of proteins detected**

Unlike mRNA, different proteins have different biochemical properties due to varying amino acid composition, post translational modification, etc. This means conditions for solubilisation can drastically differ across proteins. As such some proteins will be better detected than others. This introduces an added layer of complication in proteomics analysis, as it is near impossible with the current technology to solubilize every protein and attain full coverage of a cell's proteome [191]. While the technology utilised in this project separates proteins into fractions based on their hydrophobicity to overcome some of these issues. The number of fractions resolved for mass spectrometry is still highly limited by sequencing cost and machine availability.

Due to the above mentioned limitation, large-scale proteomics output considerably fewer proteins than equivalent genes detected by transcriptomics[192]. After multiple hypothesis testing, i.e. correction for false discovery, this number drops even further. Meaning the overall number of differentially expressed genes in a comparative analysis can be considerably lower in proteomics experiments simply due to overall limitation in coverage [192].

#### 4.1.3.6 Proteogenomic Analysis may improve results

In studies where thousands of differentially expressed hits are found, such as those in antigen secreting vs non-secreting cells, it is very challenging to validate every hit through low throughput method. Therefore, despite the limitations highlighted in the previous section, proteomics can be a useful tool for soft validation of transcriptomics data. Furthermore, accompanying gene level data can give credence to protein level regulation, despite higher *p*-values, and improve comparative analysis. Nevertheless, this must be done with careful consideration as gene and protein level does not always correlate due to different mRNA or protein turnover rates and translation efficiency.

#### 4.1.4 Investigation of Membrane Trafficking Components

In previous chapter (section 3.4.3), we highlighted the ASC specific upregulation of COPII coat proteins, the RRBP1 ribosome anchor and proposed that CREB3L2 may be involved in the regulation of some of these components. We expand our investigation of known membrane trafficking genes and examine how tethering complexes, SNAREs and other coated vesicles are being regulated according to our proteogenomics analysis

#### 4.1.5 Cell Markers for Antibody Secreting Cells

To generate antibody secreting cells *in vitro*, NBCs must be purified from the spleen, activated by a selected antigen and the resultant ASCs must be isolated from a mixed population of non-secreting B cells and ASCs. In order to improve the yield of laboratory purification of ASCs, and potentially, the targeted destruction of plasma cells in cancer we aim to isolate high confidence surface markers that show consistent upregulation in ASCs across species, across platforms and soft validate them using proteomics.

#### 4.1.6 Data visualisation

In this project we carried out proteogenomic analysis of antibody secreting cells. We see ~1500 genes upregulated in ASCs but highlighted and visualised selected components we deem important for our hypothesis. The remaining data exist as tables of hundreds of genes/proteins that has be to be explored with tools such as Microsoft Excel. Results in this format are difficult for bench biologists to navigate. Web-based tools such as Amazonia! and Genomicscape lets users mine and visualise microarray profiles of the PC lineage [71, 72]. However, these studies incorporate neither RNA-Seq nor proteomics data. Therefore, we aim to create a tool that allows users to visualise microarray, RNA-Seq and proteomics data side by side. By visualising changes in mRNA expression alongside their corresponding protein product, we hope to give biologists a better idea of whether their candidate protein is likely to be a “real” hit and thus potentially improve the cost effectiveness of downstream validation.

#### 4.1.7 AIMS & OBJECTIVES

In the last chapter we created a bioresource through microarray analysis of mouse and human B cell lineage. In this chapter, we expand this analysis and aim to:

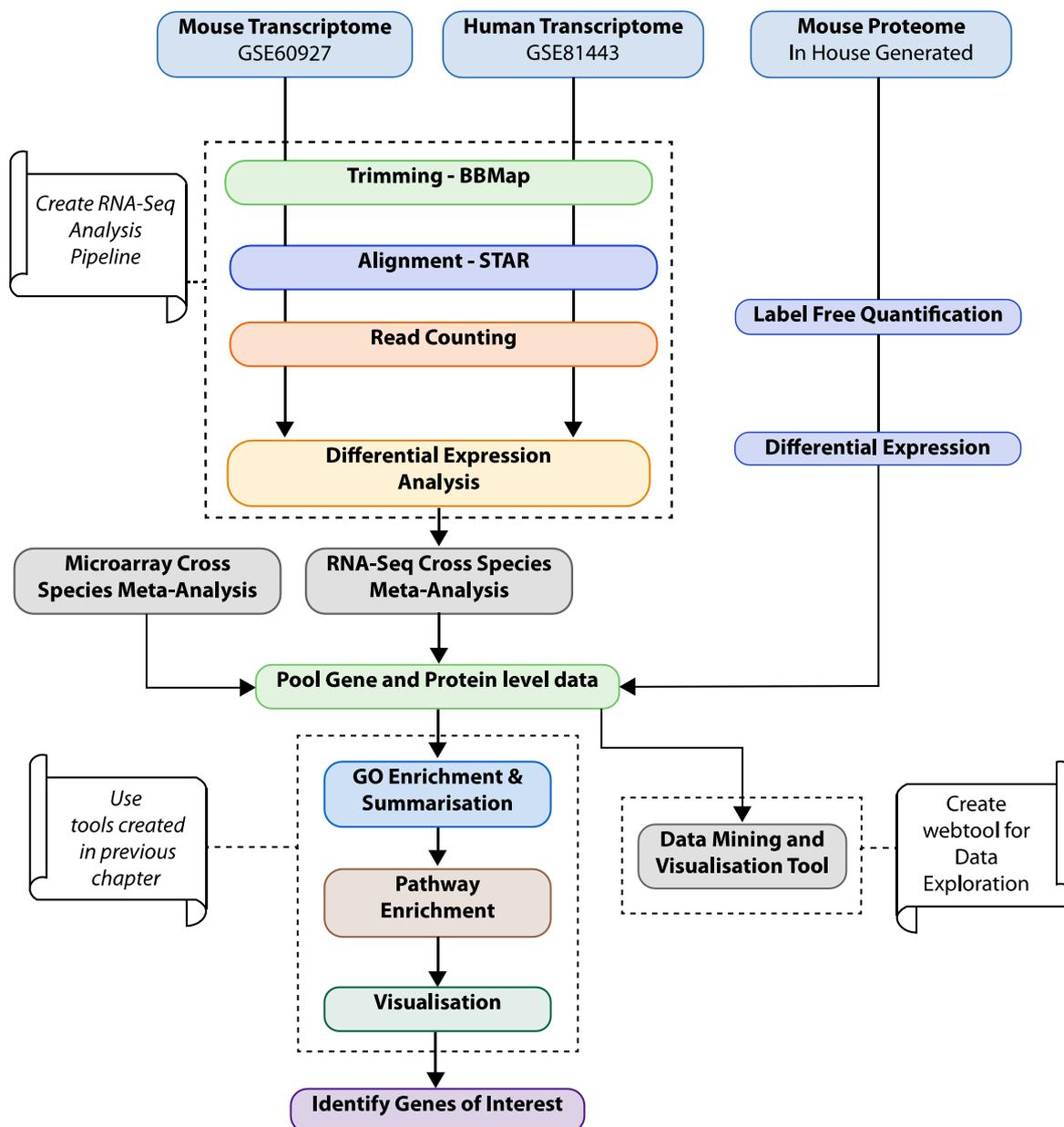
1. Study changes in gene expression in the PC lineage using publicly available RNA-Seq data of mice and human.
2. Investigate the regulation of known membrane trafficking components such as SNAREs, tethering complexes and coated vesicles.
3. Identify consistently expressed markers of ASCs.
4. Carry out proteomic analysis of mouse plasma blast (PB) compared to NBCs
5. Soft validate potential candidates identified in the current and previous chapter using a combination of RNA-Seq, microarray and proteomic results.

In the process of fulfilling these aims we create:

6. An RNA-Seq pipeline or workflow for automated analysis of raw data
7. A data visualisation tool that will allow our colleagues and the wider scientific community to easily explore gene/protein expression between ASCs and NBCs.

## 4.2 METHODS

### 4.2.1 Workflow



**Fig 4-4** | Workflow of multi-omics Analysis. We created an RNA-Seq analysis pipeline to streamline the analysis of publicly available RNA-Seq data of mice and human PC lineage. Resultant differential expression results were merged using orthologs for cross species analysis. The whole cell proteome of mice B cells, and plasmablasts were profiled by our colleagues. We performed proteogenomic analysis using RNA-Seq, microarray and proteomics result to isolate most reproducible hits. Then we used EnrichR mining tool, detailed in the previous chapter (3.2.5.1.1), to perform functional and pathway enrichment. Finally, to allow cell biologists to explore our data easily, we created a web application for viewing fold changes in genes/proteins across species and platforms.

## 4.2.2 RNA-Seq

### 4.2.2.1 Data Source

#### 4.2.2.1.1 Mouse B cell Lineage

**Table 4-1** | Phenotype of mouse RNA-Seq profiles of PC cell Lineage – Illumina HiSeq 2500 (GSE60927)

Sample	Type	Tissue	Stimulus	Day	Markers	Replicates
GSM1493786-7	NBC	Spleen	-	0	B220+ CD21+ CD23+	2
GSM1493800-1	prePB	<i>in vitro</i>	LPS	3	B220+ CD138- Blimp1-GFP+	2
GSM1493802-3	PB	<i>in vitro</i>	LPS	3	B220+ CD138+ Blimp1-GFP+	2
GSM1493794	SplPB	Spleen	-	0	CD138+ Blimp1-GFP-lo	1 (Pooled)
GSM1493795-7	SplPC	Spleen	-	0	CD138+ Blimp-GFP+	3 (Pooled)
GSM1493798	BMPC	Bone Marrow	-	0	CD138+ Blimp-GFP+	1 (Pooled)

Shi *et al* extracted *ex vivo* cell populations (NBCs, SplPBs, SplPCs and BMPCs) from the lymphoid tissues of unstimulated C57BL/6 or Blimp GFP reporter mice using FACSAria (BD Sciences) or MoFlo (Beckman Coulter) flow cytometers. As PCs are rare *in vivo*, cells had to be pooled from the spleens or bone marrows of 3 different Blimp1 reporter mice. These PCs were enriched using anti-CD138 beads (Miltenyi Biotec) before the sorting process. The markers used for sorting these cells are shown in Table 4-1 [56].

For *in vitro* differentiation, resting splenic B cells were extracted from the spleen of Blimp1-GFP reporter mice and enriched using anti-B220 magnetic beads. These cells were immunised with 10µg/ml lipopolysaccharide (LPS) *in vitro*. 3 days post activation preplasmablasts (prePB) were negatively selected for CD138 followed by sorting for B220+ and Blimp1-GFP+ fractions. More mature, *in vitro* generated plasmablasts were extracted using B220+, Blimp1-GFP+ and CD138+ markers [56].

Shi *et al* extracted total RNA using either Qiagen RNeasy Micro or Mini Kits dependent on cell number. Standard Illumina protocol were followed for RNA library preparation [193]. RNA-Sequencing was carried out using Illumina HiSeq 2500 [56].

#### 4.2.2.1.2 Human B cell Lineage

**Table 4-2** | Phenotype of human RNA-Seq profiles of PC cell Lineage – Illumina HiSeq 2500 (GSE81443)

Sample	Type	Tissue	Markers	Replicates (Biological)	Replicates (Technical)
GSM2197438	NBC	Tonsils	CD19+ CD20+ CD27- IgM+ CD2-	1	3
GSM2197435	tPB	Tonsils	CD19+ CD27+ CD38+ CD20- CD2-	1	3
GSM1493798	BMPC	Bone Marrow	CD19- CD138+ CD27+ CD38+	4	3

Lam *et al* isolated NBCs and tPBs from the tonsils of children undergoing elective tonsillectomy and BMPCs from the bone marrow of patients undergoing elective total hip arthroplasty. Prior to sorting, BMPCs were enriched using CD138 microbead (Miltenyi Biotech). All cells were sorted using BD FACS Aria II, LSRII or LSR Fortessa flow cytometer. Markers used for cell sorting are given in Table 4-2. Total RNA was extracted using Qiagen RNeasy Microkit. Sequencing libraries were generated using Clontech Smart-Seq kit and sequenced using Illumina HiSeq 2500 [194].

#### 4.2.2.2 Data Processing

##### 4.2.2.2.1 Download & Trimming

Publicly available RNA-Seq data were downloaded and converted to FASTQ format using the *fastq-dump* command of NCBI's SRA Toolkit. We used *BBDuk* command of the BBTools package for adapter and quality trimming of raw reads. We then sought to remove reads with low quality scores in the 10 base pairs on the 3' end of reads to avoid phasing errors.

##### 4.2.2.2.2 Mapping

After quality trimming, we aligned reads to reference genome using STAR aligner due to its speed, accuracy and robustness against SNPs compared to other aligners [195]. In order to maintain consistency, we used the same genome assemblies and gene annotations as those used for annotating microarray data in the previous chapter (section 3.2.2.4). For mouse RNA-Seq data we used NCBI's mm10 gene assembly, submitted on Dec 2011 and Gencode version 16 gene annotation submitted in Dec 2017. For human RNA-Seq data we used NCBI's hg38

genome assembly submitted on Dec 2013, and Gencode gene annotation version 27 submitted in Jan 2017.

#### 4.2.2.2.3 Read Counting and Summarisation

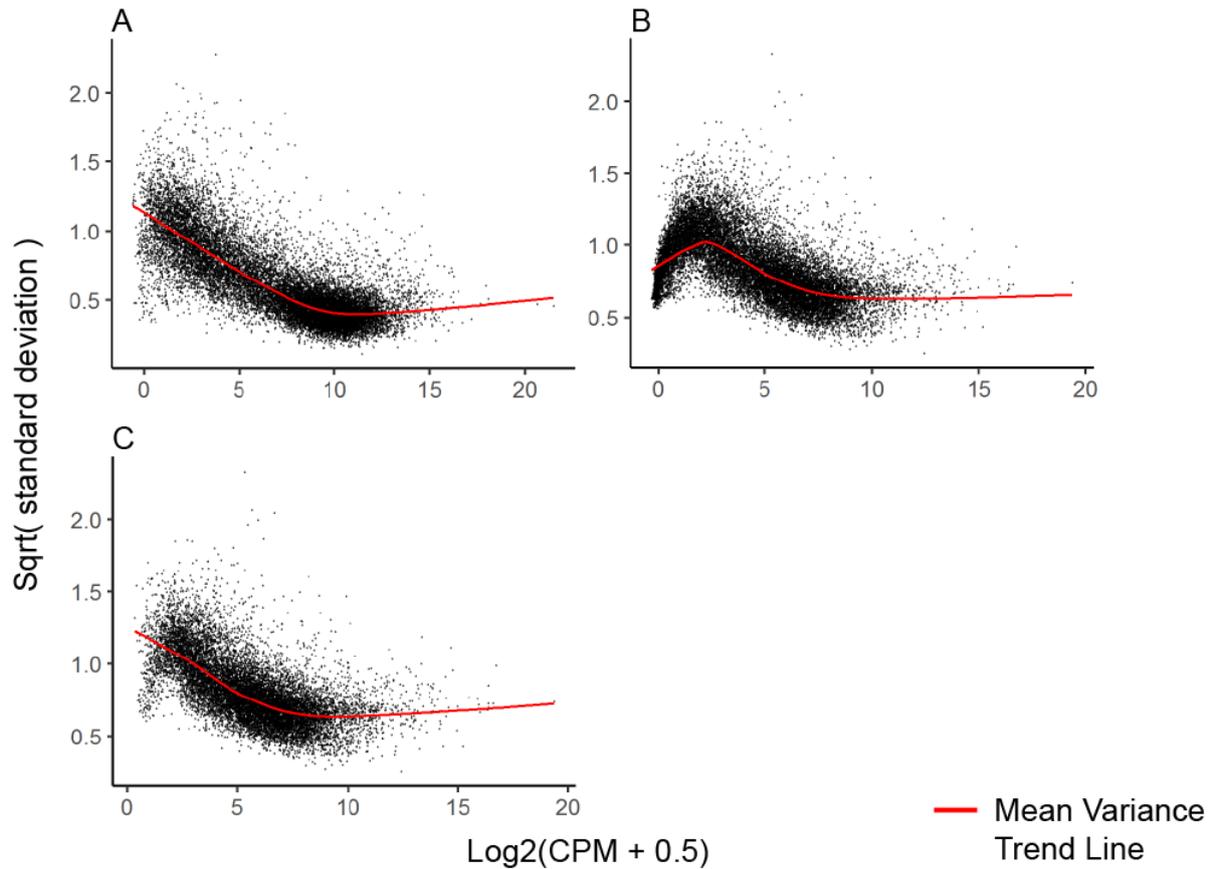
We chose to use the RSEM tool for read counting. The workflow so far has been carried out in the UNIX environment. We now import our data into R and summarise gene counts using the tximport R package.

#### 4.2.2.2.4 Low Count Filtering

The *filterByExpr* function of EdgeR package can identify and remove gene counts whose magnitude was below a statistically significant threshold. This algorithm by default keeps genes with more than 10 counts in a relevant number of samples based on experimental design [196].

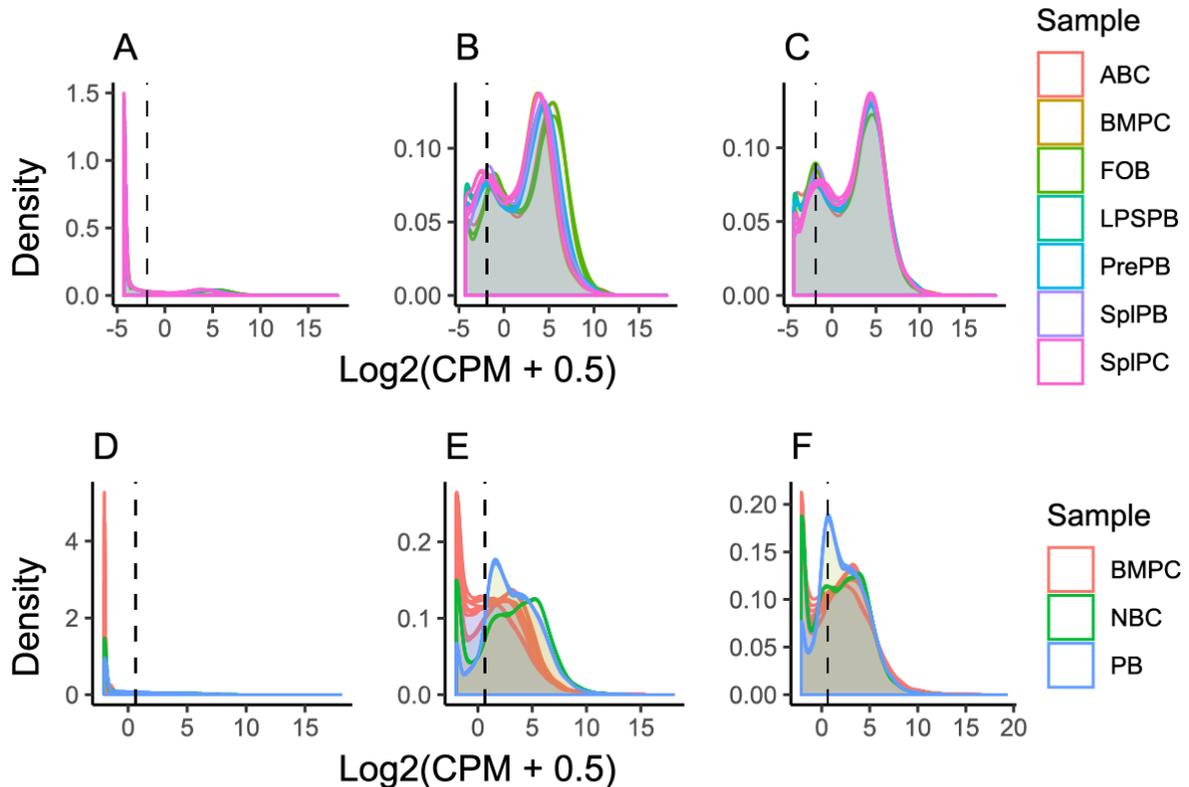
#### *Quality Controlling Low Count Filtering*

To reduce false positives, we remove low confidence genes using the *filterByExpr* function. These genes can be visualised using a mean variance (MV) plot, which charts square root of standard deviation on the y-axis and the log<sub>2</sub> transformed gene counts per million on the x-axis. Typically, well-formed data consisting of replicates from different phenotypes exhibit high biological variation among lower counts and so standard deviation tend to be asymptomatic to the y-axis early on as shown in Fig 4-5A [184]. Low-confidence counts tend to accumulate among lower counts. These tend to break the asymptomatic trend and result in very low standard deviation among lowest counts, which then sharply increase to the expected high standard deviation (Fig 4-5B). These low-confidence genes are concentrated to the left of the chart and generally a certain count threshold is established to filter out these genes.



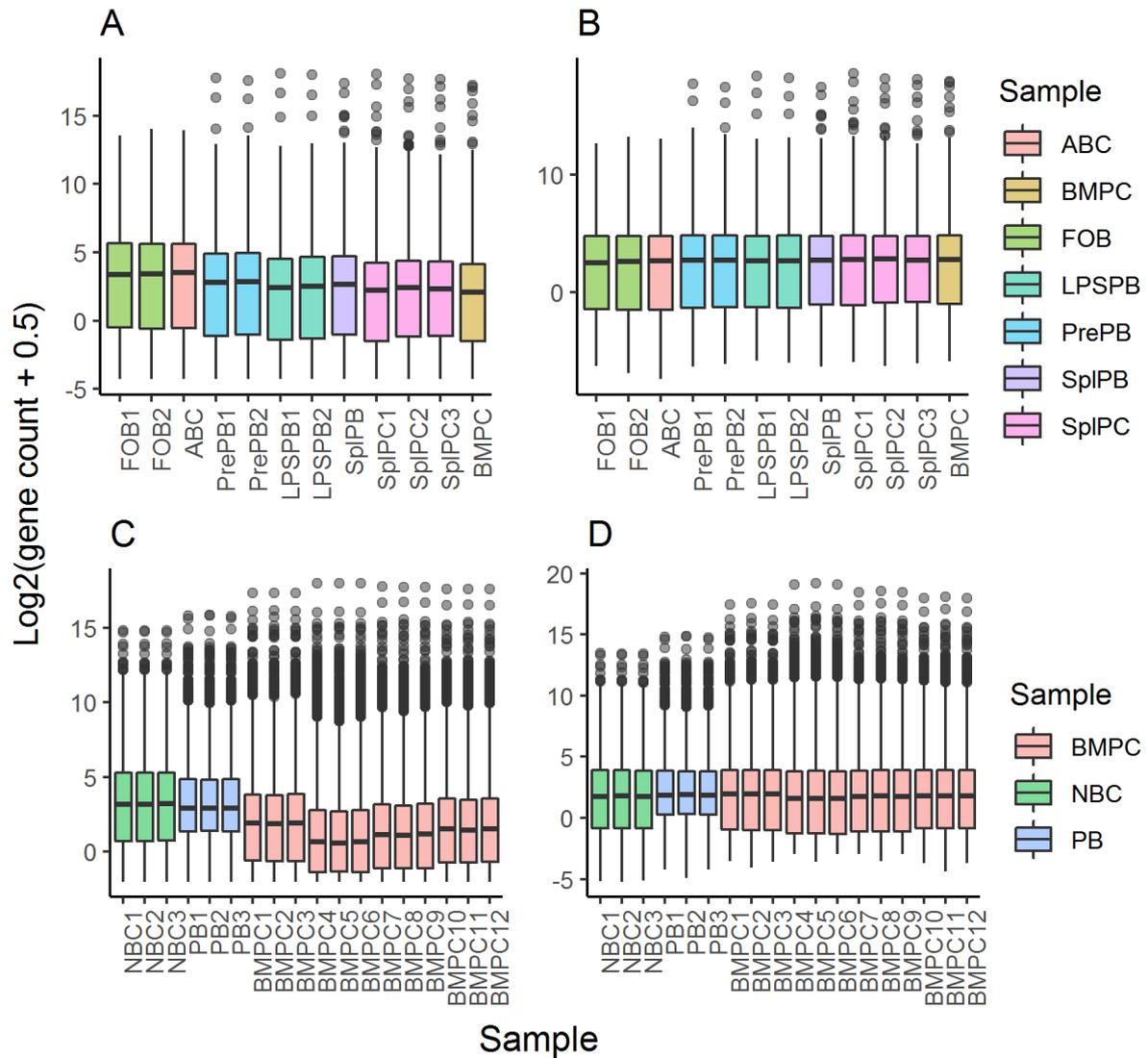
**Fig 4-5** | Mean Variance Trend Plot depicting limma-voom normalisation and low count filtering efficacy. **A.** Mouse RNA-Seq data has been correctly filtered for low counts by *FilterByExpr* algorithm as the standard deviation i.e. biological variation is highest among low count genes (Red line). **B.** MV-plot for Human RNA-Seq shows remnants of low-expression genes as we see very low standard deviation in the leftmost counts, which rises immediately after a count threshold is reached. **C.** Re-adjustment of low count filtering of Human-RNA-Seq samples initially show better results. These samples have been filtered for genes with at least 10 counts per million in a minimum of 4 samples. However, this filtering removes genes that are uniquely expressed in human PBs but not in BMPCs.

Fig 4-5A shows filtered, mouse RNA-Seq data. We observe that filtering by the *FilterByExpr* function has been successful as low counts showed highest standard deviation. However, the filtering of human RNA-Seq samples shown in Fig 4-5B was not as successful as lower counts still showed very little standard deviation, i.e. biological variation. We redid the filtering with different criteria to try and solve this issue. First, we selected for genes with at least 10 counts, (~1.53 counts per million (CPM)) across samples in at least 4 samples. This was 1 sample more than the minimum number of available replicates among all tested groups (See Table 4-2). As shown in Fig 4-5C, the revised filtering criteria seemingly improves the



**Fig 4-6** || Density Distribution Plot of RNA-Seq Counts Before and After Low-Count Filtering. **A-C** Mouse B Cell Lineage. **D-F**. Human B Cell Lineage. **A, D**. Low counts in Unfiltered data overwhelm the count distribution. **B, E**. After filtering, the data distribution is more apparent. **C, F**. Normalisation of the filtered data yields uniform distribution across samples. Dotted line indicates estimated low CPM threshold

filtering output. On closer inspection, we find that this filtering criteria led to the loss of genes that were exclusively expressed in plasmablasts, which had 3 replicates, but not in BMPCs or NBCs. Changing the other relevant criteria, such as minimum gene count, reduces sensitivity toward genes expressed in lower levels in NBCs but high levels in ASCs while the aberrantly low standard deviation among low counts remains the same. Therefore, we opt to retain the results from the standard *FilterByExpr* filtering criteria. The presence of low confidence counts is apparent in the aberrant peak in human PBs in the density distribution plot in Fig 4-6E-F. Nevertheless, we continue with this filtering as we predict that genes showing very low counts in one phenotype but not another are present among this group due to the nature of the biological phenotypes we are looking at.



**Fig 4-7** | Box Plot of RNA-Seq Counts before and after TMM normalisation. **A.** Mice and **C.** human gene counts before normalisation show considerable variation in distribution. After TMM normalisation it is apparent that both **B.** mice and **D.** human gene counts have a uniform distribution.

As shown Fig 4-6B, E, the distributions of gene counts are considerably different across samples for both mice and human datasets. Therefore, we used EdgeR package to perform Trimmed Mean of M value (TMM) normalisation. The efficacy of normalisation procedure is visualised in both Fig 4-6C and Fig 4-7B-D, where a more uniform distribution is apparent.

#### 4.2.2.2.5 Differential Expression Analysis

The *voom* function of limma R package was used to calculate the precision-weight of gene counts. This is based on gene count variation from the global mean, where genes with low

variance received higher weight than those with high variance. These weights were used to correctly fit the normalised data to a linear regression model. The T- and F-statistic were calculated using empirical Bayes moderation provided in the limma R package. Computed  $p$ -values were adjusted using Benjamini & Hochberg method for global false discovery [133]. No fold change threshold was initially enforced on intra-species differential expression results.

#### 4.2.2.3 Pipeline Creation

In order to allow automated sequential processing of multiple RNA-Seq samples we created a pipeline using the packages mentioned in the previous section. Note that this program is not suitable for *de novo* assembly. This tool is available at:

<https://github.com/NabilaRahman/RNA-Seq-Pipeline>

##### 4.2.2.3.1 Dependencies and System Requirements

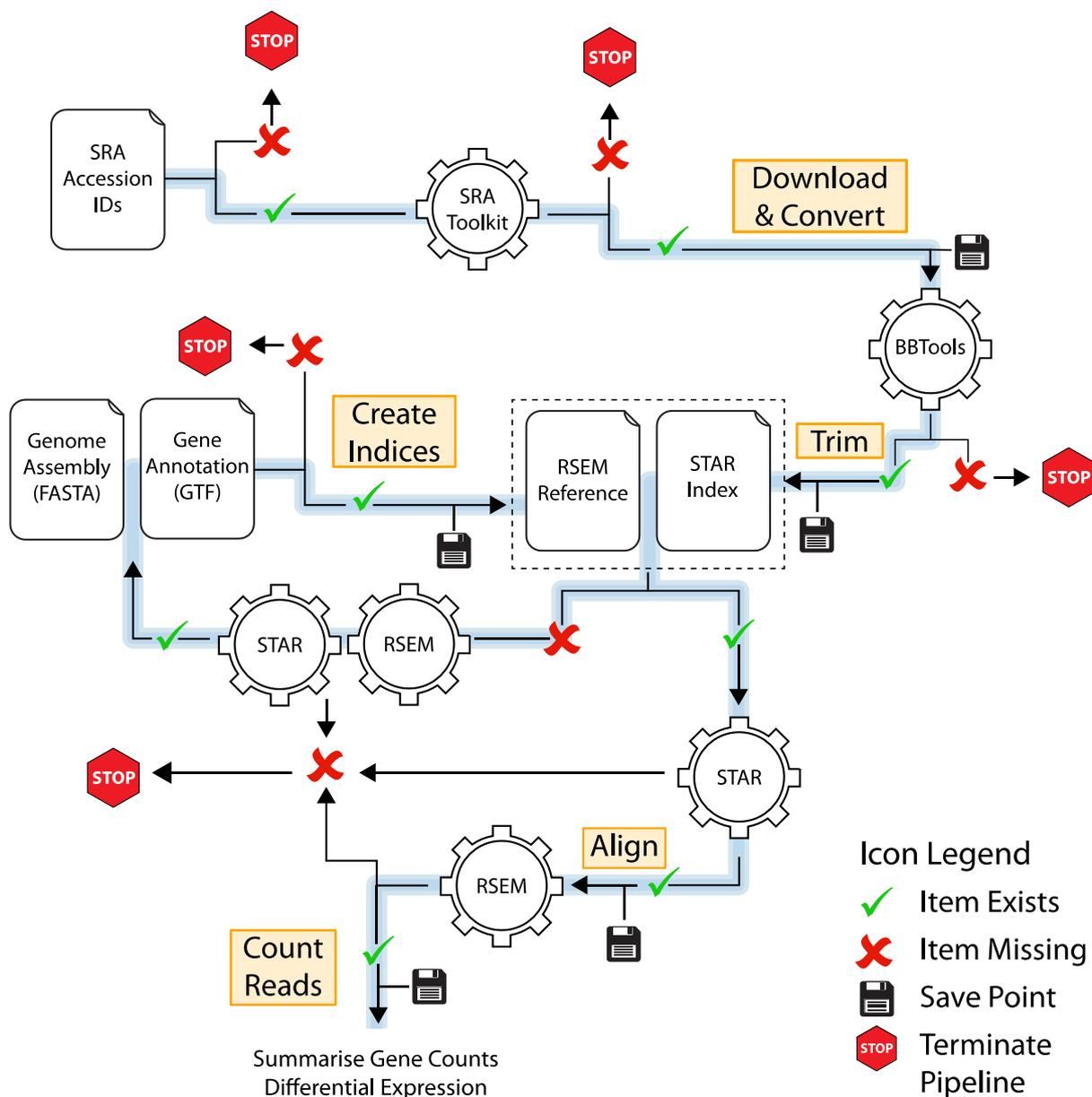
The pipeline was built on Bash version 4.3.48 and R version 3.5.1. It requires a minimum of 16GB Physical Memory but over 32GB is recommended. It is dependent on the following UNIX packages of the given version or higher: SRA Toolkit version 2.9.0, BBTools version 38, STAR version 2.6.9c, RSEM version 1.3.1 and optionally GNU parallel 2017 [195, 197–200]. Furthermore, the pipeline requires the following R packages: tximport, EdgeR, limma, ggplot, ggpubr, reshape2 [196, 201–205].

##### 4.2.2.3.2 Configuration

In order to run the pipeline, the user has to minimally specify (1) the SRA IDs to be processed, (2) location of requisite tools, (3) a Reference Genome in FASTA Format, (4) a gene annotation file in GTF Format, (5) a transcript ID to gene ID conversion table and (5) phenotypes to compare for differential expression analysis. We also provide additional options for advanced users which can be viewed in the *Config.sh* file at the GitHub repository.

STAR and RSEM tools require special indices created from the reference gene assembly and gene annotation file. Users are given the option to specify the location of these indices, otherwise the pipeline can create them on the fly.

## 4.2.2.3.3 Defensive Programming



**Fig 4-8** | Schematic of our RNA-Seq Pipeline equipped with defensive programming and staged processes. For clarity we have placed the check points where they are relevant in the pipeline. However, all checks for missing packages, files or bad parameters are carried out before the pipeline even begins. The pipeline, in blue, will only initiate if there are no errors. After each successful processing step, a “save point” is made. If the program is terminated early, restarting the pipeline will cause the process to run from the last “save point”.

In order to prevent hours wasted on processing bad data, the first thing the pipeline does is anticipate errors in the users input. To do this, *ValidateInput.sh* script is run before performing any of the analysis. This shuts the entire program down if there is: (1) Insufficient RAM available, (2) Missing UNIX tools or reference files, (3) Invalid data e.g. decimals where

integers are expected. Note that the process is not overly stringent with R packages, as processing times after read counting are very quick and can be redone in a matter of minutes.

#### 4.2.2.3.4 Staged Processes

Each stage of RNA-Seq analysis e.g. trimming, alignment or counting can take a long time. Sometimes unexpected errors occur if users terminate the pipeline before it is fully complete. Having to re-run the whole pipeline from scratch can mean the waste of many hours of processing. Therefore, we mark the completion of each processing stage, such that re-running the pipeline allows users to restart the process from the last completed stage. The entire process is visualised in Fig 4-8.

#### 4.2.2.3.5 Output

Once the pipeline is complete users will have, among others: (1) statistics on how many bases have been trimmed for adapters and quality, (2) machine readable (BAM) files containing read mapping to the genome, (3) Gene & Transcript Counts, (4) Density Distribution Plot before and after filtering low counts, (5) Box plot comparing raw and normalised counts, (6) Mean-Variance Trend Plot (7) PCA plot comparing raw and normalised counts (8) Differential Expression Table. Users are advised to quality check outputs 4-7 before accepting the differential expression results.

#### 4.2.2.4 Cross Species Analysis

We repeated the protocol for cross-species data pooling as detailed in Chapter 4. Differential expression (DE) results from mice and human RNA-Seq data were merged using Ensembl ortholog annotations and Global False Discovery adjusted  $p$ -values were averaged. Duplicate ortholog pairs with the lowest average  $p$ -values were retained. Genes with missing orthologs were processed separately.

Voting system was used to determine differentially regulated genes conserved in ASCs across species. We set a threshold of 1.2-Fold Change (FC) to differentiate up/downregulated

genes from those showing no change. This is less stringent than the industry standard 2 FC as (1) RNA-Seq is considerably more robust than microarrays (2) Fold change of relevant ER-Golgi components may occur at a lower threshold than the industry standard (3) Comparison of microarray and RNA-Seq results from different species lend credence to small but consistent fold changes. For cross species analysis, we do not enforce *p*-value cut-offs until after data pooling.

### 4.2.3 Proteomics

#### 4.2.3.1 Data source

The sample preparation steps were wholly carried out by our colleagues E Rajan and AWA Aswani. Spleens were obtained from healthy C57BL/6 mice of 8-10 weeks of age. Spleens were crushed and filtered through a 40 µm cell strainer along with complete RPMI-1640 medium, 50 µM of β-mercaptoethanol and 10% endotoxin-free Foetal Calf Serum, 2mM Penicillin-Streptomycin and L-Glutamine. After preparation of single cell suspension, standard protocol was used to lyse red blood cells.

NBCs were purified using Magnetic-activated cell sorting (MACS) system and mouse B cell Isolation Kit (Miltenyi biotec) following manufacturer's instructions, where cells were negatively selected for CD43, TER119 and CD4. The resultant isolated B cells had a purity of approximately 80%.

To generate antibody secreting plasmablasts, NBCs were incubated with 10µg/ml lipopolysaccharide (LPS) *in vitro*. 3 days post activation, plasmablasts (mPB) were sorted for CD138 using magnetic beads (purity >80%).

CD93 was identified as one of the top highly expressed cluster differentiation markers across ASCs according to our cross-species RNA-Seq analysis. Therefore, our colleagues, isolated another population of plasmablasts (mPB93) which was purified for CD93+ using magnetic beads [56]. Western blot carried out by my colleagues showed that this CD93 fraction

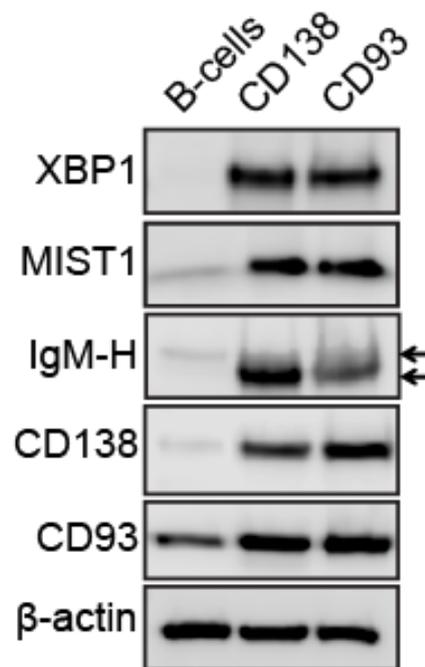
contained CD138, IgM, XBP1, and MIST1, which are known markers of antibody secreting cells (Fig 4-9). Western blotting of CD93 purified fraction alongside CD138 purified fraction shows that the former expressed fewer immunoglobulin M heavy chain (IgM-H) than the latter fraction. Prior to class switching in the spleen, IgM is a hallmark of circulating plasmablasts. Therefore, we predict CD93 purified cells might consist of a less mature fraction of ASCs than those captured by CD138 purification, especially as the CD93 was present at lower levels in Naïve B cells.

NBCs, mPBs and mPB93 samples (4 replicates each) were digested using trypsin by our colleagues and sent to the biOmics: Biological Mass Spectrometry facility, University of Sheffield for shot-gun proteomics. Samples were fractionated by high pH reverse phase chromatography. Each of the 12 fractions generated per sample were analysed by 2-hour MS/MS run using Orbitrap Elite Hybrid Ion Trap-Orbitrap (ThermoFisher) system.

#### 4.2.3.2 Proteomic Data Analysis

Raw data processing and normalisation was externally carried out using MaxQuant and Perseus software by Dr M. Collins.

We calculated differential expression between ASCs and NBCs using limma function in R. To allow comparison with RNA-Seq data, we set the threshold for fold change at 1.2 while Benjamini-Hochberg adjusted *p*-value for multiple group comparison was set to be below



**Fig 4-9** | Western Blot of resting B cells, CD138 purified plasmablasts and CD93 purified cells (E Rajan & AWA Aswani). Although typical markers of antibody secreting cells such as CD138, MIST1 and XBP1 are similar in CD93 and CD138 fractions.

0.05. We considered genes that met these criteria in either mPB or mPB93 as differentially expressed.

#### 4.2.3.3 Proteogenomic Analysis

We combined differential proteomic results with differential transcriptomic results from mice and human B cell lineage, generated by both microarray and RNA-Seq. As we have only profiled the proteome of mouse B cell lineage, we simply utilised mouse gene symbols to add proteome data to the microarray and RNA-Seq results.

#### 4.2.3.4 Functional Analysis

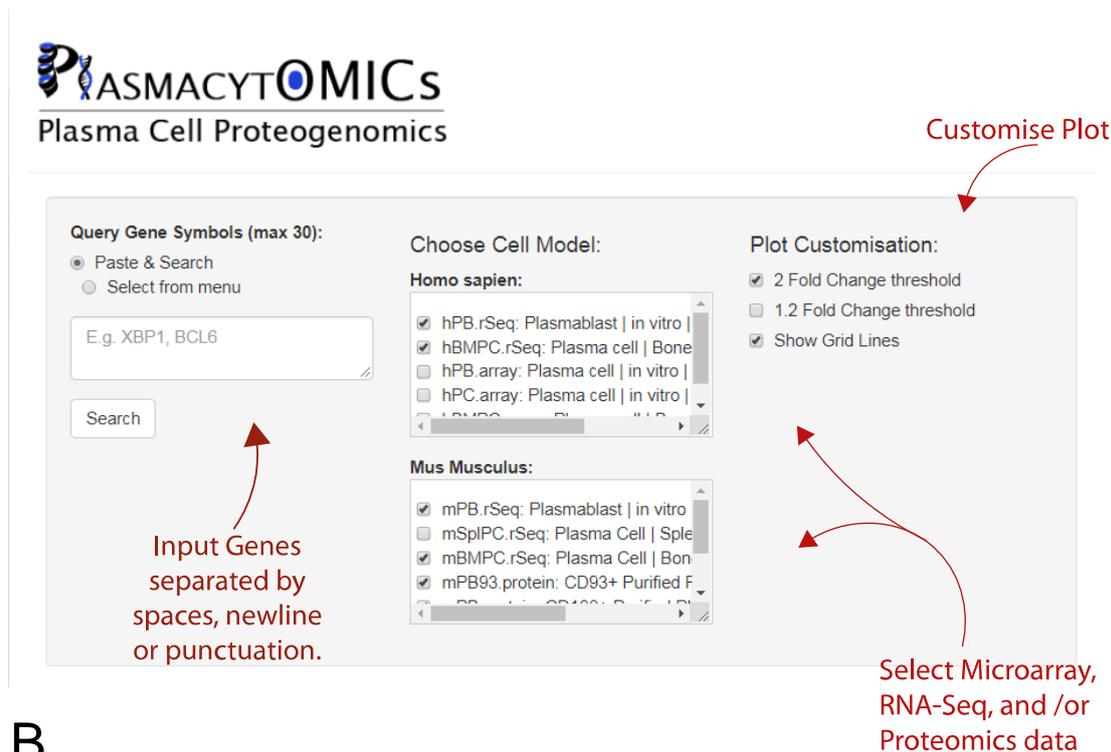
In order to determine the biological functions enriched in differentially expressed genes across species and platforms we utilised our EnrichR-mining-tool detailed in Chapter 3, section 3.2.5.1.

#### 4.2.3.5 Fold Change Visualisation as a Web Application

We wanted to easily visualise how a component was changing across species in the gene and the protein level simultaneously. In order to allow users to explore our results, we retain all unique genes regardless of *p*-value or fold change. We visualised the fold changes of one or more gene using *ggplot2* R package. Using the *shiny* package, we designed a web application to dynamically feed user input, i.e. gene/s of interest, to *ggplot2* and output the resultant plot to the users' browser (Fig 4-10, Fig 4-11). At the time of writing, the app, PlasmacytOMICs is available only on University of Sheffield LAN and is available on request.

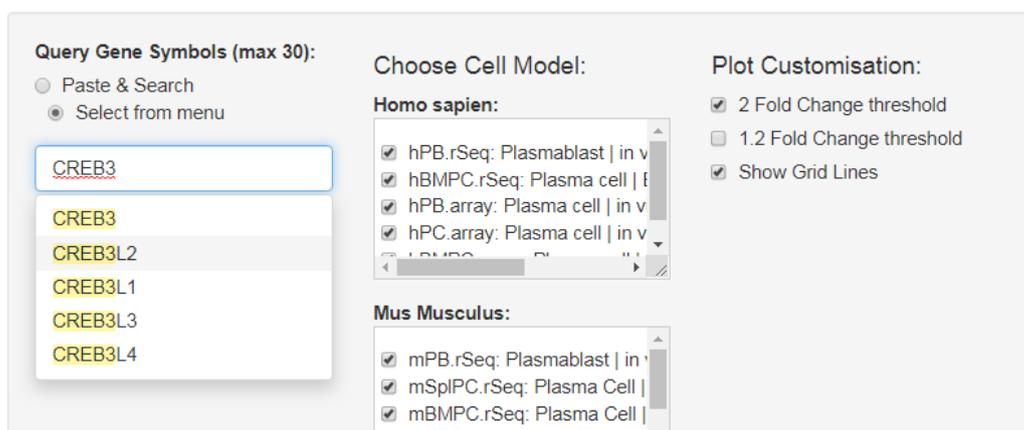
## 4.2.3.5.1 Gene Input

A



B

**PlasmacytOMICS**  
Plasma Cell Proteogenomics



**Fig 4-10** | Screenshots of input fields in the PlasmacytOMICS Web Application. **A.** Upon visiting the application, users are greeted with the input text box, which allows them to type in genes, separated by spaces, tabs, newline or any other punctuation. In addition, users can choose the cell models to display and customise the plot, e.g. showing/ hiding threshold and grid lines. **B.** By clicking on the “Select from menu” radio button, users can access the autocomplete search box. This is useful for finding gene isoforms or for finding genes whose nomenclature the user is unsure of.

This app allows user to input a list of genes separated by spaces, newline or punctuation (Fig 4-10A). The input is then fed into our server and matched to our bioresource. Once inputted genes are found, their fold changes in ASCs vs NBCs are displayed to the client's browser. The app also gives users an autocomplete option (Fig 4-10B). For example, if one is looking for genes of the CREB3 family, they can type in "CREB3", and as shown in Fig 4-10B, other CREB3 isoforms present in the dataset will be displayed. This can be useful for looking up isoforms of a gene or in cases where a user is unsure of the nomenclature of a gene of interest.

#### 4.2.3.5.2 Options

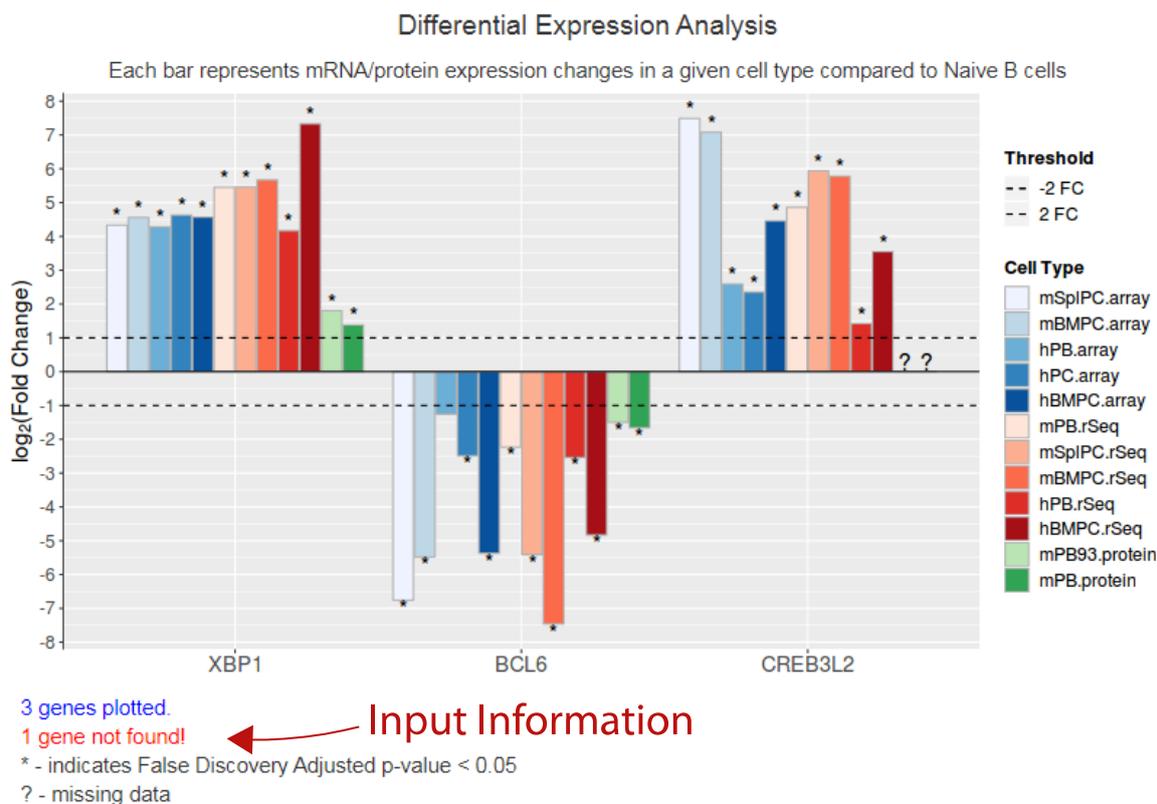
The PlasmacytOMICs app lets users choose which ASC phenotype to plot. ASCs profiled by microarray, RNA-Seq and MS/MS are available (Fig 4-10A). Furthermore, users can customise the outputted plot by showing/hiding gridlines as well as threshold lines indicating 2-fold or 1.2-fold change.

#### 4.2.3.5.3 Output

If inputted genes are not found in the bioresource, users will receive a message saying "No matching data". Otherwise, the inputted genes will be plotted as bar charts displaying the logarithmic fold change in ASCs versus NBCs (Fig 4-11). By default, every available cell type is plotted.

Below the bar chart, the app displays the number of inputted genes matched to our bioresource. As shown in Fig 4-11, we inputted 4 search terms, 3 of which were actual genes and 1 was a dummy input. As a result, the app informs the user that 3 genes inputted was plotted and 1 was not found.

Once users have customised the plot to their liking, they can download the plot in png, pdf, svg or tiff format. The app then downloads a high-quality version of the chart while maintaining the aspect ratio seen on screen. This is useful as *shiny* dynamically alters the width



**Export Plot:**

File Type:

- png ▲
- png
- pdf
- svg
- tiff

Download Plot

Author: Rajan E, Aswani AWA, Peden, AA, 2019, *PlasmacytOMICs*, at \_\_\_\_\_

**Fig 4-11** | Screenshot of an output from PlasmacytOMICs web application. Once users inputted desired genes and pressed search, the fold changes of these genes in selected ASCs compared to NBCs will be plotted as a bar chart. Input information is given below the chart, which tells users how many of their inputted genes are being shown, and how many were not found in our dataset.

of the chart based on the browser window. On that note, this app will dynamically adjust to screen sizes and, therefore, is usable in desktop as well as mobile browsers.

#### 4.2.4 Low Throughput Validation

All validation experiments were carried out by our colleagues E Rajan and AWA Aswani. Resting B cells were extracted, purified and stimulated as detailed in Section 4.2.2.1 to obtain *in vitro* generated PBs. These cell populations were used in Western blotting to validate the protein expression in PBs versus NBCs.



#### 4.3.1.2.1 Upregulated Genes

Fig 4-12A summarises the overlap of differentially upregulated genes in mice and human. We looked at a total of 5 different antibody secreting cell (ASC) types. By considered genes upregulated in at least 3 out of 5 ASCs we narrowed down 2027 genes consistently upregulated across species. A total of 1323 genes were not detected or missing in mice ASCs but were upregulated in humans. ~95% (1268) of these genes were protein coding. Another 482 genes were missing in humans but upregulated in mice and ~92% of these genes were protein coding (Fig 4-12A, in red). Thus, we consider a total of 3832 upregulated genes for downstream analysis.

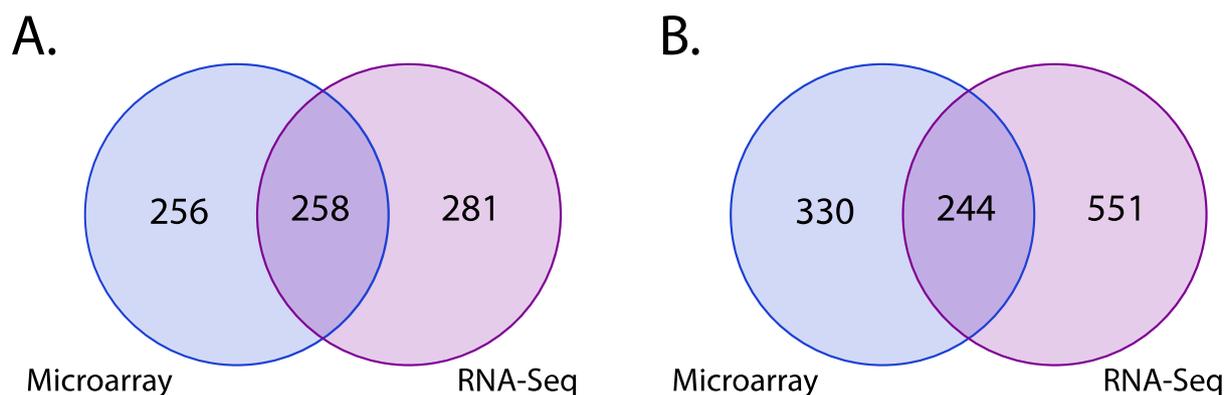
#### 4.3.1.2.2 Downregulated Genes

Fig 4-12B summarises the overlap of differentially downregulated genes across mice and human. By considered genes downregulated in at least 3 out of 5 ASCs we narrowed down 1746 genes consistently downregulated across species. 257 genes were not detected or missing in mice ASCs but were downregulated in humans and ~91% (235) of these genes were protein coding. Another 246 genes were missing in humans but downregulated in mice where ~94% were protein coding genes (Fig 4-12B, in red). A total of 2249 downregulated genes were therefore considered for further analysis. Non coding genes have overall lower homology between species than protein coding ones and could contribute to missing orthology information [206–208]. However, the majority of genes with missing data in either species were protein coding in both up and downregulated sets and unlikely to be explained by poor homology between non-coding genes.

#### 4.3.1.2.3 Contradictory Regulation

In Fig 4-13, we summarise genes that show contradictory regulation in mice as opposed to human ASCs. In the previous chapter, we discussed that while these results may specify diverging regulation in either species these results could contain erroneous results brought on by platform specific limitations or noise. Therefore, we isolated genes that showed differential

regulation between species in both RNA-Seq and microarray. The reproducibility of this expression pattern across platforms thus improve the likelihood of these genes having species specific regulation. Overall, 258 genes were found to be upregulated in mice but were downregulated or had no change in human (Fig 4-13A, Appendix Table 6-7), while 244 were downregulated or showed no change in mice but were upregulated in human (Fig 4-13B, Appendix Table 6-16 ).



**Fig 4-13** | Venn diagram of genes showing cross platform reproducibility in diverging gene regulation across species. **A.** Number of genes upregulated in mice but downregulated or no change in human ASCs. **B.** Number of genes downregulated or no change regulated in mice but upregulated human ASCs.

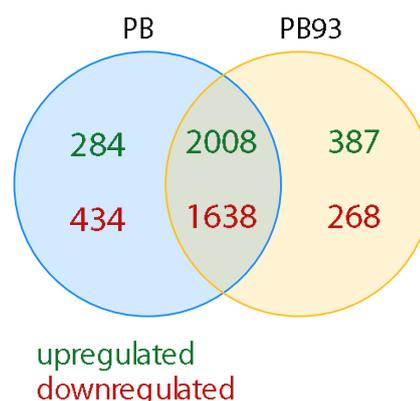
### 4.3.2 Proteomics

My colleagues have recently profiled the proteome of NBCs, CD138 purified (mPB) as well as CD93 purified plasmablasts (mPB93). We wanted to know which of the genes identified in our transcriptomic analysis showed consistent protein level regulation in CD138+ and CD93+ purified cells versus their non-secreting counterpart.

Out of 7958 unique proteins detected by mass spectrometry 2679 genes were uniquely upregulated in either mPB.prot, mPB93.prot or both (FC > 1.2, FDR adjusted  $p$ -value < 0.05). These genes are given in Appendix Table 6-18. Another 2340 genes were downregulated (Appendix Table 6-19).

### 4.3.3 Proteogenomics

With transcriptome and proteome profile of ASCs and NBCs available, we wanted to know how many genes showed conserved regulation across species as evidenced by RNA-Seq and



**Fig 4-14** | Venn diagram of showing overlap of differentially expressed proteins in CD138 and CD93 purified cells. Among 2679 genes showing upregulation in either ASC types, 2008 genes showed consistent upregulation in both mPB and PB93 cells. While 1638 genes were downregulated.



out of 3 platforms, we note that 258 (~90% protein coding) genes, evidenced by RNA-Seq profiling were not detected by proteomics or microarray whatsoever. We predict that limitations of proteomics technology, discussed in Section 4.1.1.3, may play a large part in this as known marker of ASCs e.g. CD138/SDC1 was not detected by proteomic profiling. Due to these limitations we also consider this additional 258 genes evidenced solely by RNA-Seq but missing in both microarray and proteomics datasets. Therefore, we study a total of 1756 upregulated genes in the functional/pathway enrichment analysis. Rank priority was given to genes/proteins expressed in 3 platforms followed by those expressed in 2 platforms and finally genes/proteins missing in microarray/proteomics data.

#### **4.3.3.2 Downregulated genes/proteins across platforms**

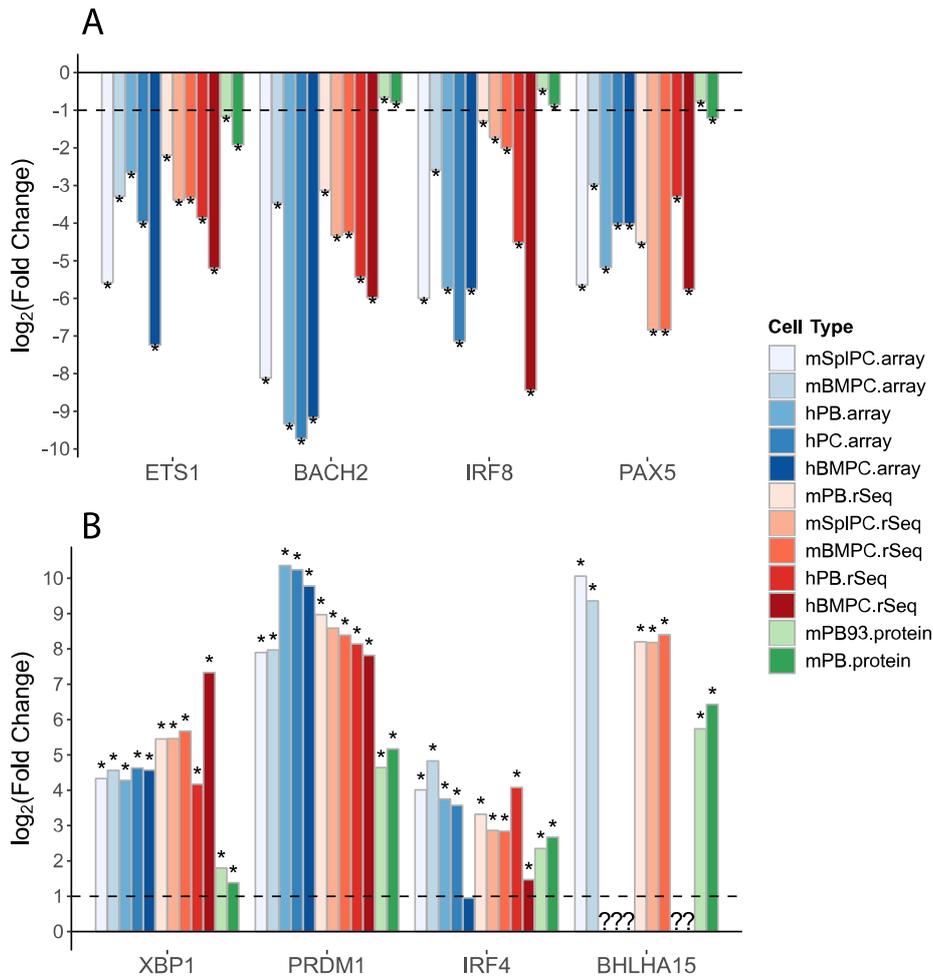
Fig 4-15B, shows that a total of 393 components showed shared downregulation in ASCs vs NBCs in all platforms. While 1093 show consistent downregulation in at least 2 platforms (Table 6-23). After considering 113 downregulated genes (~81% protein coding) evidenced solely by RNA-Seq but missing in the proteomics and microarray dataset, we consider a total of 1599 genes for further analysis.

#### **4.3.3.3 Components potentially regulated post translationally**

It is also interesting to note that ~1300 upregulated and ~1000 downregulated components were unique to proteomics and show poor evidence in the transcriptome of ASCs (Appendix Table 6-24, Table 6-25). We predict that these proteins are potentially regulated via post translational modification.

#### **4.3.4 Verification**

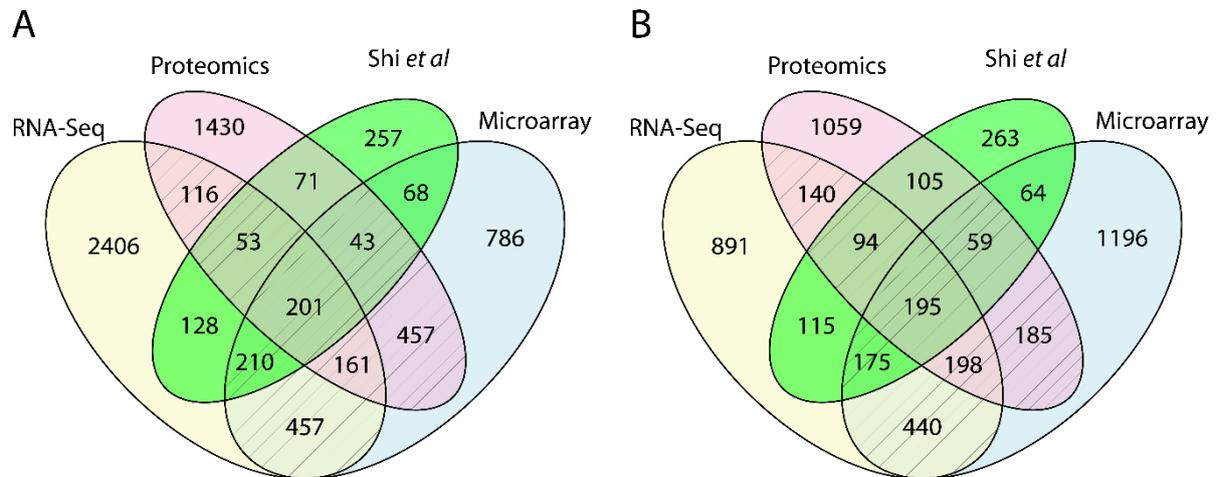
As in the Chapter 3, Section 3.3.3, we determine the integrity of our results by analysing well-established markers of resting B cells and their antibody secreting counterparts. As shown in Fig 4-16, transcription factors that maintain the NBC phenotype, i.e. ETS1, BACH2, IRF8 and PAX5, are downregulated in ASCs across species and in both mice transcriptome and



**Fig 4-16** | Differential Regulation of well characterised ASC and NBC markers according to multi-omics analysis. **A.** Genes that maintain naïve B cell phenotype were downregulated in ASCs. **B.** Genes that maintain ASC phenotype were upregulated in ASCs compared to NBCs. Evidence for BHLHA15/MIST1 was missing in human RNA-Seq and microarray data.

proteome. Likewise, transcription factors that enhance plasma cell differentiation, XBP1, IRF4, PRDM1 are highly upregulated. Measurement for BHLHA15/MIST1 gene was missing in the profile of both human microarray and RNA-Seq dataset, but strongly upregulated in mice. This is not surprising as the probe for this gene is missing in the human microarray GeneChip utilised, and our analysis shows that the RNA-Sequencing depth (number of reads detected) for human ASCs were considerably lower than that of the mouse transcriptome, which we predict is due to limited availability of patient samples as compared to the mouse model. This consolidates the need to incorporate genes missing in either species in the analysis.

## 4.3.5 Known and Novel Signature Genes for ASCs



**Fig 4-17** | Venn diagram showing overlap of reproducible gene/proteins isolated by multi-omics versus those identified by Shi *et al*. Diagonal lines indicate genes we considered for downstream analysis (excluding missing genes). **A.** Upregulated genes. Only 49.9% (517) genes upregulated in Shi *et al*'s data were validated either by our results and considered for further analysis. 1204 genes/ proteins were novel hits uniquely upregulated in ASCs according to our multi-omics analysis. **B.** Downregulated Genes. Only 49% (525) genes downregulated in Shi *et al*'s data were validated either by our multi-platform analysis method. 1044 genes/proteins were novel hits uniquely downregulated in ASCs according to our multi-omics analysis. The labels “RNA-Seq” and “Microarray” in the Venn diagram include cross species meta-analysis, and is prefiltered for genes that are inconsistent across species in the individual platforms.

Shi *et al* isolated 1036 genes that were consistently upregulated in the transcriptome of mouse BMPC, and LPS activated plasmablasts [56]. In this project we reanalysed the raw RNA-Seq data from Shi *et al*'s study and differed our analysis method of the RNA-Seq data mainly in the stringency of the fold change threshold (1.2 FC as opposed to 1.5 FC) and reference genome version and source. Furthermore, we incorporated mouse microarray and MS/MS profiles as well as human microarray and RNA-Seq data.

We wanted to know how many of the genes predicted by Shi *et al* to be consistently differentially regulated genes across ASCs match our results. About 50% of genes regulated in either direction identified by Shi *et al* were validated by our multi-omics analysis as shown in (Fig 4-17). Remaining ~50% genes had contradictory regulation across species or in the mouse proteome. Interestingly, the regulated secretory cargo, PRG2, is one such gene.

We found 1204 novel genes/ proteins uniquely upregulated in ASCs that have not been previously isolated using functional genomics (Appendix Table 6-11) and 1044 novel genes/ proteins were found to be uniquely downregulated (Appendix Table 6-16).

#### 4.3.5.1 Novel Transcription Factors

Shi *et al* identified 103 transcription factors (TF) differentially regulated in ASCs as opposed to NBCs[56]. We wanted to know how many of these were validated by our multi-omics analysis and extract potentially new ones.

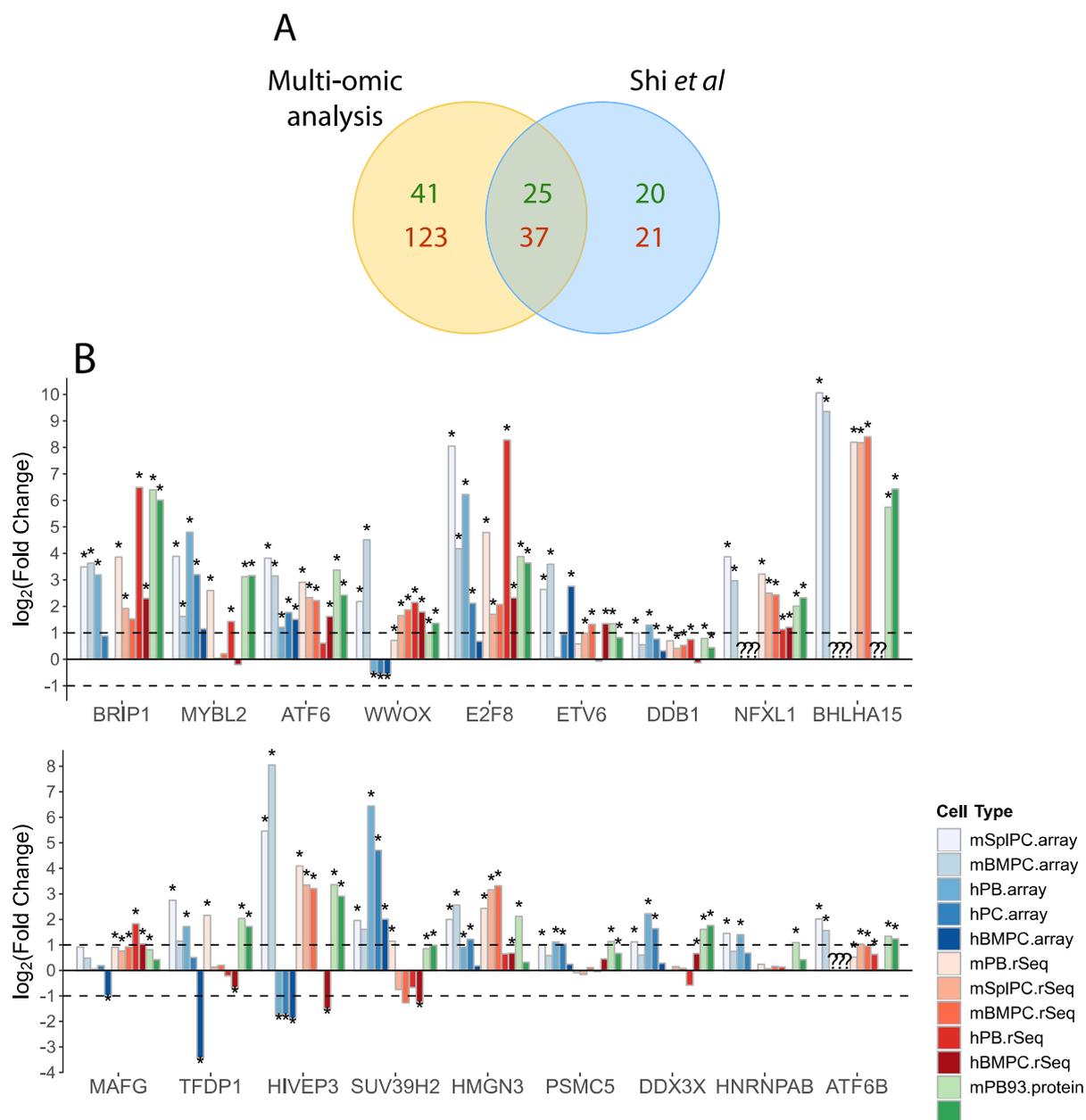
##### 4.3.5.1.1 Upregulated Transcription Factors

24 out of 45 upregulated TFs identified by Shi *et al* were validated by our analysis (Fig 4-18A). This includes well characterised TFs such as XBP1, PRDM1, IRF4 as well as our TF of interest, CREB3L2 highlighted in Chapter 3, Section 3.4.3.

Interestingly, our results showed an additional 41 TFs upregulated in ASCs (Appendix Table 6-13). This gene list included the key regulator of ER stress ATF6, the ASC marker BHLHA15/MIST1 and also BHLHE41, which has been identified as a key repressor of cell proliferation for the generation of terminally differentiated plasma cells in a study published in 2017 [5, 144, 209]. Proteomics results soft validated 18 out of these 41 transcription factors. These genes are shown in Fig 4-18B.

##### 4.3.5.1.2 Downregulated Transcription Factors

37 out of 58 the downregulated TFs identified by Shi *et al* were validated by our analysis (Fig 4-18A). This includes known enhancers of the B cell phenotype including BCL6 and PAX5.

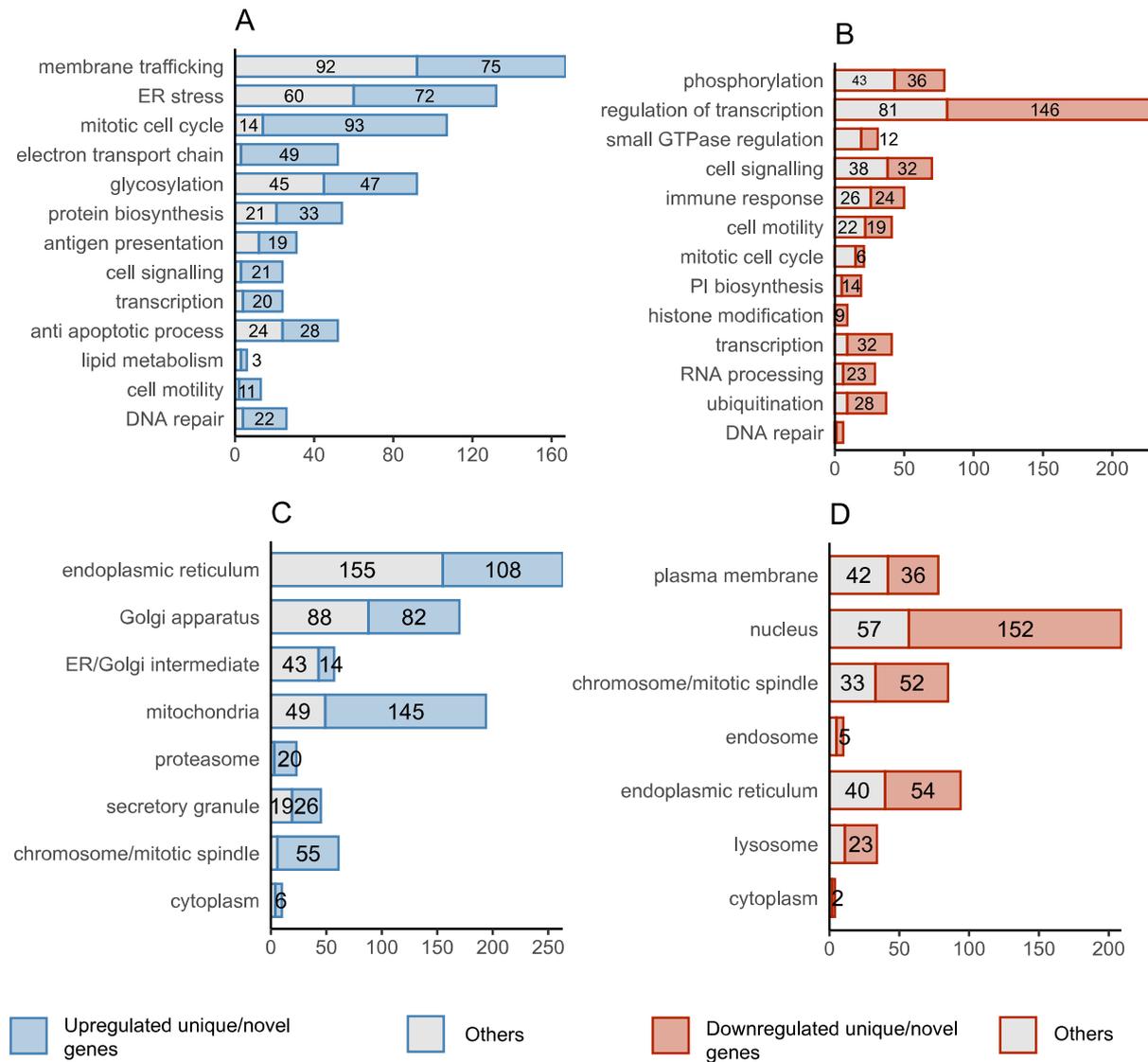


**Fig 4-18** | Transcription factors uniquely isolated by multi-omics analysis. **A.** Venn diagram showing transcription factors differentially regulated in ASCs according to our analysis and those identified by Shi *et al* (red: downregulated, green: upregulated). **B.** Top upregulated TFs uniquely isolated by our multi-omics analysis. The presence of the known marker of ASCs, BHLHA15, and the ER stress regulator, ATF6, demonstrate the integrity of our analysis.

Our results showed an additional 123 TFs downregulated in ASCs, of which 89 were soft validated by proteomics data (Appendix Table 6-14). This gene set includes FOXO1 and RUNX1 genes, which are reported to be responsible for the maturation of B cells [210].

We have highlighted known TFs relating to ASCs and NBC phenotype identified by our multi-omics analysis that were not isolated through RNA-Seq analysis of mice model alone. As this set consists of several genes whose role is less characterised in ASCs, this list could be a useful reference for characterisation of transcription enhancers in ASCs.

### 4.3.6 Functional Analysis

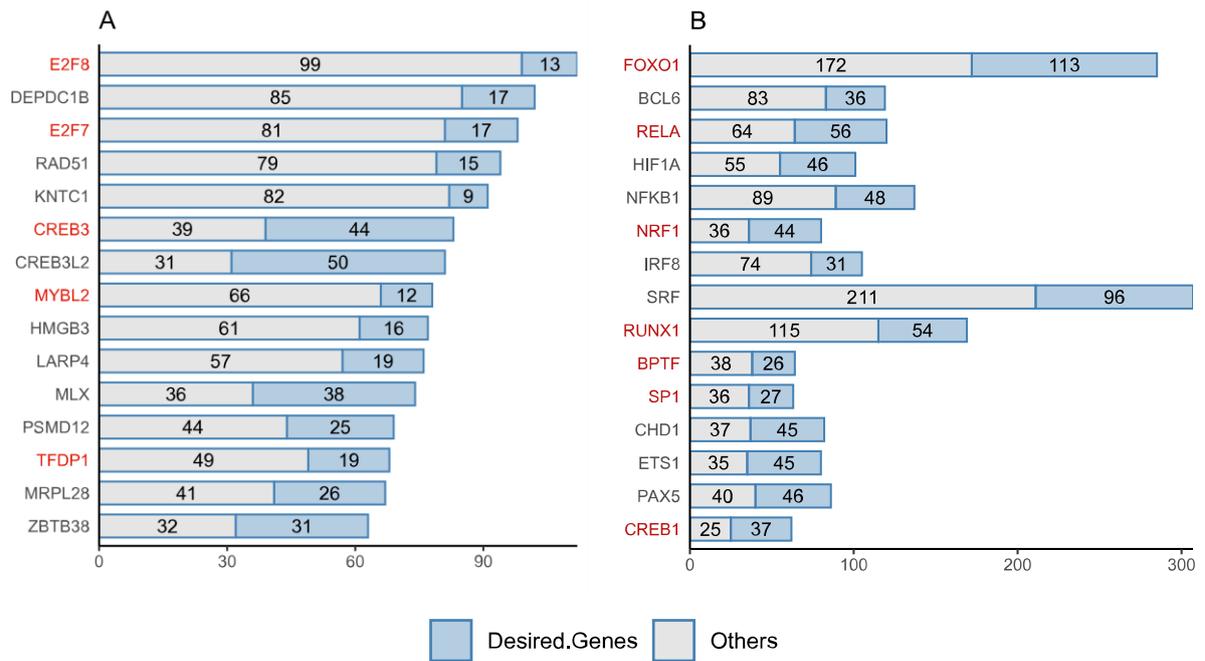


**Fig 4-19 | Summarised GO Enrichment Results for multi-omics analysis. Red/blue highlights the number of genes mapping to GO terms that were uniquely isolated by our proteogenomic analysis A. GO Biological Processes enriched among upregulated genes. B. GO Biological Processes enriched among downregulated genes. C. GO Cellular Components enriched among upregulated Genes. D. GO Cellular Components enriched among downregulated Genes.**

GO enrichment analysis for Biological processes consolidates our findings from the chapter 3, in that upregulated components in ASCs across platforms showed greatest enrichment for components of membrane trafficking and unfolded protein response (UPR) while downregulated genes were primarily related to transcriptional control and protein phosphorylation (Fig 4-19). GO cellular component were likewise enriched for ER and Golgi components for upregulated genes and the nucleus for downregulated genes. Interestingly, genes/proteins we isolated uniquely using multi-omics analysis improves the coverage of genes enriched for Biological function and Cellular Compartments as shown in Fig 4-19. For example, where Shi *et al* will have identified 92 components of membrane trafficking, we find a total of 167 components.

#### 4.3.7 Pathway Analysis

In chapter 3, we highlighted CREB3L2 as the TF with co-expression with the largest number of components related to membrane trafficking/ UPR or machinery localised to ER, Golgi or secretory granules. As shown in Fig 4-20A, ARCHS4 transcription factor co-expression enrichment of our multi-omics analysis consolidates these results. In Fig 4-20A, transcription factors marked in red were those that were uniquely isolated by multi-omics analysis. We do not explore most of these genes in this project as they are not enriched for membrane trafficking components. However, we notice that CREB3 co-regulates with a set of ER-Golgi components that are somewhat distinct from its CREB3L2 isoform. This includes the COPII component SEC23B. Genes enriched for CREB3 and CREB3L2 co-expression are given in the Appendix Table 6-26.



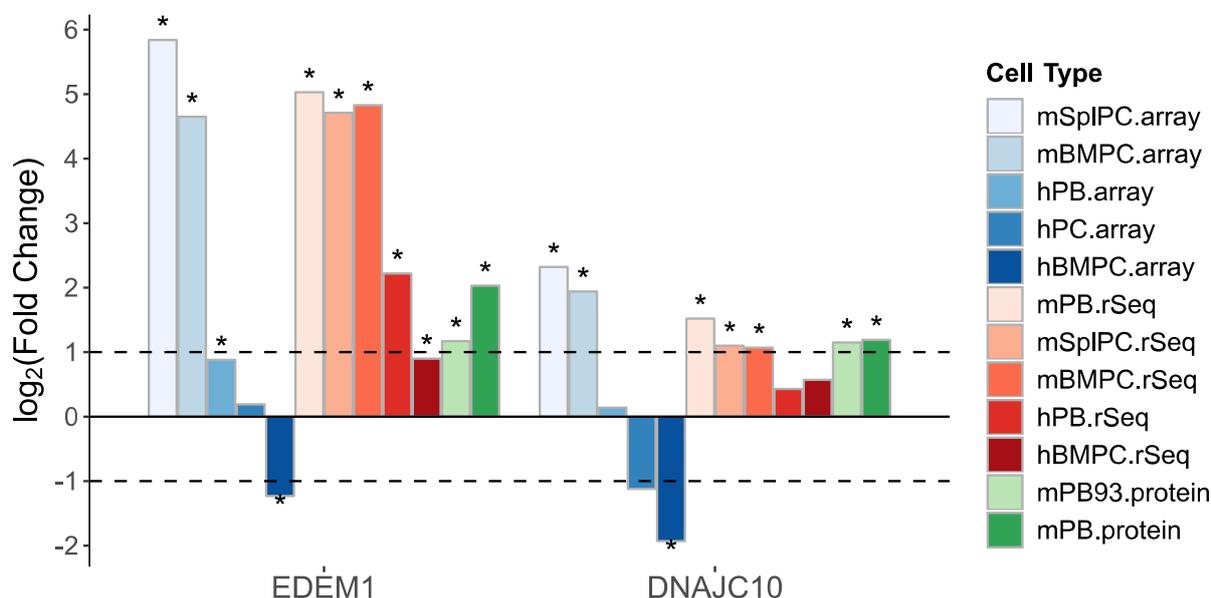
**Fig 4-20** | Transcription factors predicted to co-regulate with genes/proteins upregulated in ASCs according to multi-omics analysis. **A.** Top 15 ARCHS4 Transcription Factor Coexpression enrichment results for TFs and their targets upregulated in ASCs. **B.** Transcription perturbation analysis for TFs and targets downregulated in ASCs. “Desired genes” in blue were components that had GO terms related to ER, Golgi, and membrane trafficking machinery. TFs marked in red were those uniquely identified by our multi-omics analysis. X-axis and bar length indicate number of enriched gene members. TFs are order by FDR adjusted enrichment *p*-value

## 4.4 DISCUSSION

### 4.4.1 EDEM1-ERDJ5 complex does not show diverging regulation

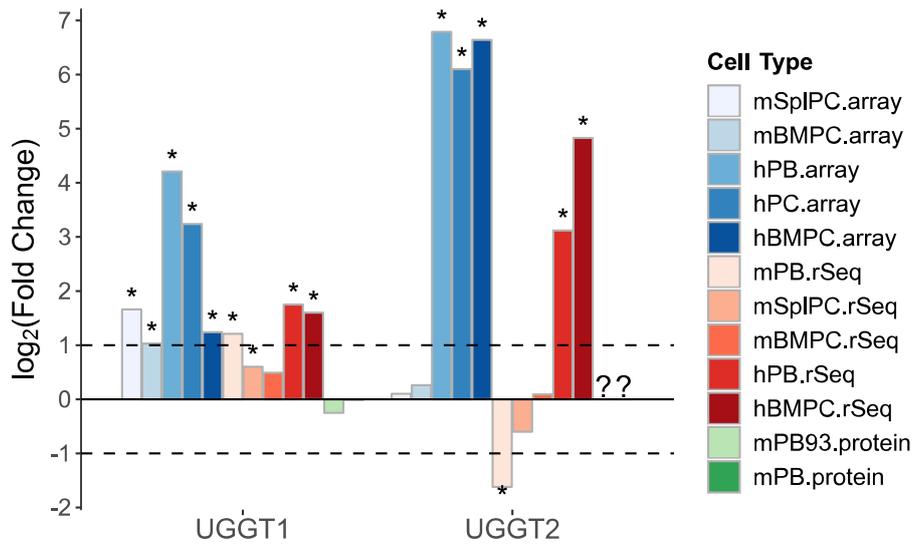
EDEM1-ERDJ5 complex is responsible for the delivery of misfolded protein to the retrotranslocon for degradation. In the Chapter, Section 3.4.1, we hypothesised that this complex was being differentially regulated in mice as opposed to human based on microarray data analysis. However, cross platform proteogenomic study reveals that this is not the case, as EDEM1 shows significant upregulation in human ASCs according to RNA-Seq data.

On the other hand, neither human RNA-Seq nor microarray showed significant in ERDJ5/DNAJC10, while equivalent mice transcriptome was upregulated. However, as the human mRNA detected by RNA-Seq showed a positive tendency, taken together with EDEM1, we predict these genes may not be good candidates for downstream validation. Nevertheless, this analysis demonstrates the utility of multi-omics analysis in identifying noisy data.



**Fig 4-21** | Differential regulation of EDEM1-ERDJ5 complex according to multi-omics analysis. Microarray analysis showed a stark contrast between EDEM1 and ERDJ5/DNAJC10 expression in mice as opposed to human. However, RNA-Seq analysis demonstrate that this is not the case for EDEM1, as it was significantly upregulated in the human transcriptome. The upregulation of ERDJ5/DNAJC10 is debatable as its change is not statistically significant in the ASCs studied in the human RNA-Seq profiles. (?) indicates missing data. (\*) indicates FDR adjusted  $p$ -value < 0.05.

## 4.4.2 UGGT2 may be specifically upregulated in human ASCs



**Fig 4-22** | Differential Regulation of UGGT enzymes according to multi-omics analysis. In human ASCs, both microarray and RNA-Seq results confirm a strong upregulation of UGGT2 gene. In contrast, mouse ASCs show no change in UGGT2 level in either platforms. Unlike UGGT2, its UGGT1 isoform shows mild but significant upregulation in gene level across species, however, protein level evidence showed no significant change. (?) Indicates missing data. (\*) indicates FDR adjusted  $p$ -value < 0.05.

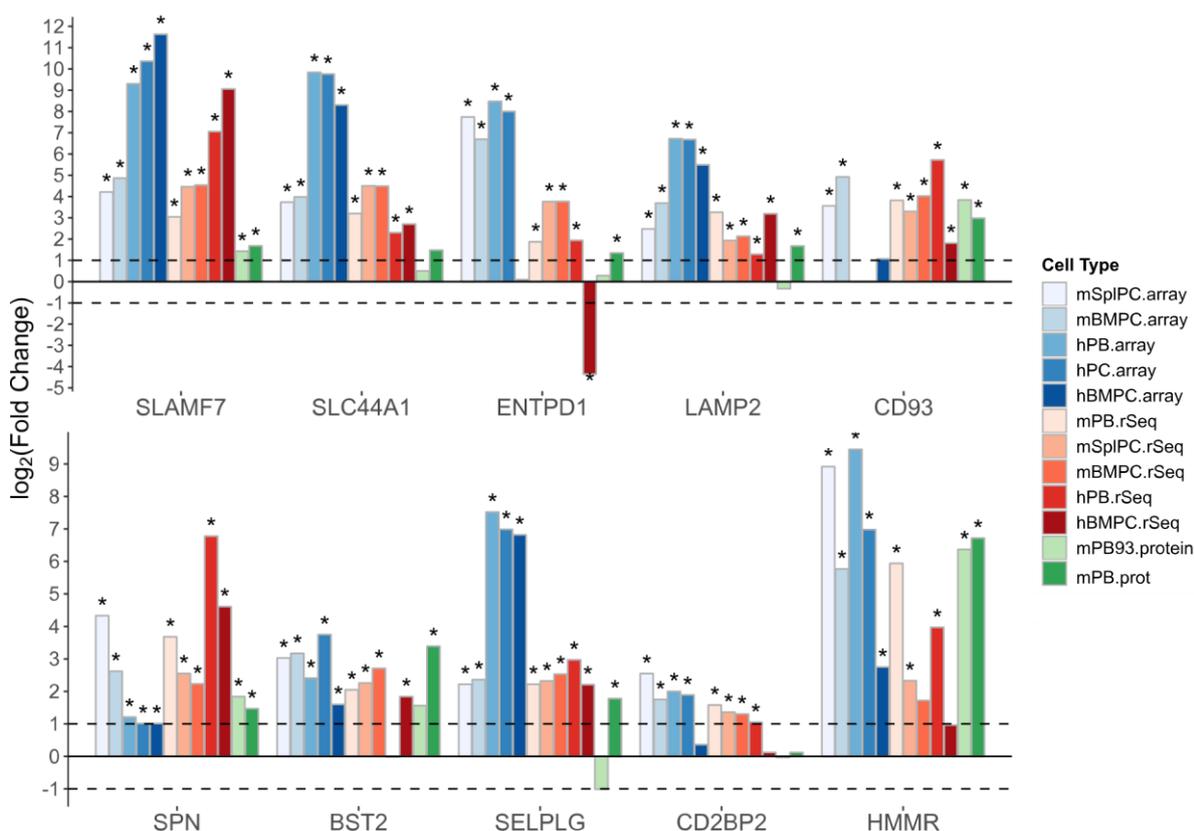
During protein processing in the ER, UGGT enzymes mediate the return of misfolded proteins to the folding cycle. The two known UGGT isoforms, UGGT1 and UGGT2, share about 55% similarity in their amino acid sequence. Takeda *et al* showed that UGGT1 and UGGT2 has similar substrate binding capacity using synthetic substrates in human embryonic kidney cells [150]. We noticed that the binding capacity of UGGT2 has not been studied across species. In Chapter 3, Section 3.4.2, we noted that UGGT2 was specifically upregulated in human ASCs but not mice. New evidence from our cross-species RNA-Seq analysis consolidates this diverging regulation between species (Fig 4-22). Overall, UGGT2 showed 8 to 64-fold upregulation, while UGGT1 showed only 4-fold change.

Cross platform analysis supports our hypothesis that UGGT2 is differentially regulated in humans as opposed to mice. This result has exciting implications in human specific glycosylation. For instance, if UGGT2 can detect misfolded substrates specifically incompatible with human, the enzyme may potentially be responsible for mediating species-

specific folding in humans and thus explain some of the accumulation of misfolded recombinant proteins in CHO cells, which is a major issue in the biologics industry. Therefore, we predict that the protein level validation of UGGT2 would be a promising follow up to this finding.

### 4.4.3 Novel Markers of ASCs

#### 4.4.3.1 CD Markers



**Fig 4-23** | Top 10 CD markers upregulated in ASCs according to multi-omics analysis.

In order to isolate high confidence markers for antibody secreting cells, we utilised an existing list of cluster differentiation (CD) markers for mice and human from Uniprot. Then we looked at which of these genes showed species-conserved upregulation in ASCs versus NBCs according to our proteogenomic analysis. A total of 42 genes matched this criteria and are listed in Appendix Table 6-17 in order of FDR adjusted *p*-value. The top 10 hits are shown in Fig 4-23.

#### 4.4.3.2 CD93

Among them CD93 was particularly interesting as it has not been implicated as a marker that is specific for antibody secreting variants of the B cell lineage. As discussed in Section 4.2.3.1, our colleagues validated the expression of CD93 by Western blot and have generated the proteome of CD93 isolated plasmablasts (Fig 4-9). Although not in the scope of this project, we note that the CD93 isolated cells (mPB93) differentially upregulated 386 genes and differentially downregulated 267 genes compared to CD138 isolated PBs (Appendix Table 6-20, Table 6-21).

We are aware that many non-CD proteins localise to the cell surface, however, study of these proteins was not in the scope of this project and characterisation of these proteins should be performed for further analysis.

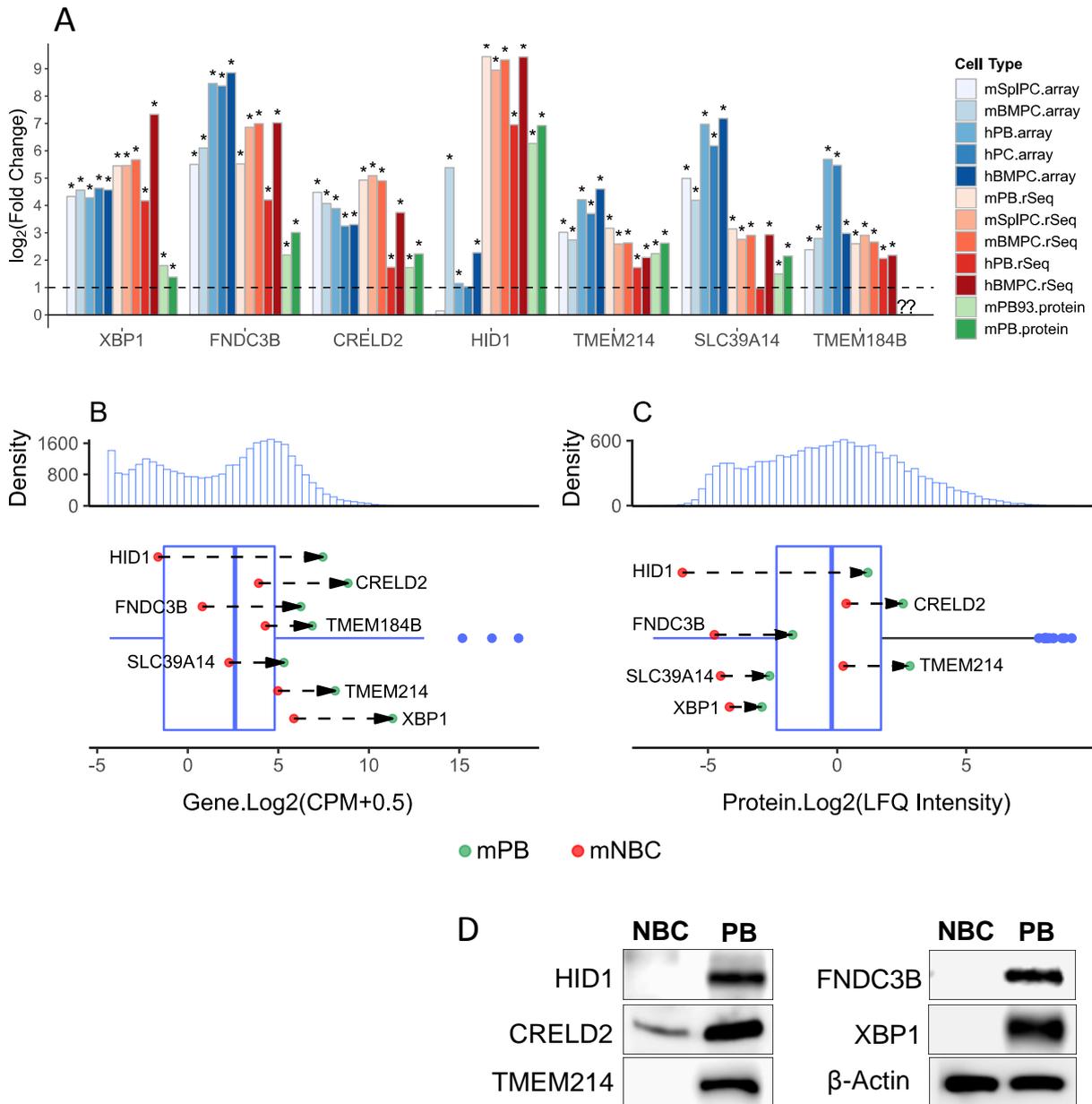
#### 4.4.3.3 Genes strongly upregulated in every cell type studied

An obvious way to isolate ASC markers would be to pick the top hits that show no inconsistencies in any of the cell types we studied. As we mentioned before, these robust hits make up a total of 362 genes given in Appendix Table 6-22. Within this gene set, the less characterised CRELD2 gene was particularly interesting as it is one of top most statistically significant hits ( $p$ -value<0.05). CRELD2 is a ER stress related protein induced by ATF6 and at the time of writing, a 2018 study has been published that shows that CRELD2 deficiency increases the susceptibility of Neuro2a cells to tunicamycin induced ER stress [211]. Due to the reproducible upregulation of this gene in ASCs, as shown in Fig 4-24A-D, this gene may be a promising marker for ASCs.

#### 4.4.3.4 Genes Co-regulated with CREB3L2

We identified 3 potential markers co-regulated with the transcription factor, CREB3L2, as discussed in Chapter 3, Section 3.4.4. These genes: FNDC3B, TMEM184B and SLC39A14, were highly upregulated in microarrays. Multi-omics analysis revealed another 54 genes

upregulated in ASCs, also predicted to be co-regulated with CREB3L2. Among them we note the induction of the poorly characterised TMEM214 transmembrane protein, which is reportedly involved in ER stress mediated apoptosis [212].



**Fig 4-24 | A.** Proteogenomic regulation of potential markers of ASCs.(?) indicates missing data. (\*) indicates FDR adjusted  $p$ -value<0.05. **B-C.** Histograms and equivalent box plot representing normalised distributions of LPS activated mouse plasmablasts and naïve B cells. Expression of specific genes/proteins have been visualised in context of global data distribution. XBP1, a marker for ASCs is used as a positive control. **B.** Normalised transcriptome (RNA-Seq) and **C.** Proteome (MS/MS) of mouse plasmablasts. **D.** Western Blot of HID1, CRELD2, TMEM214, and FNDC3B validate the unique expression of these genes in PBs as opposed to NBCs (AWA Aswani and E Rajan).

As shown in Fig 4-24A, we highlight some of these cargoes and all show consistent upregulation in the transcriptome of mice and human model across platforms. Protein level evidence is available for 3 out of 4 genes. We wanted to know how the expression of these genes and gene products compared to global expression. Therefore, we plotted the transcriptome and proteome of mouse PB and NBCs as histograms and highlighted the relative expression of our components of interest (Fig 4-24BC). Fig 4-24C show that the protein level expression of FNDC3B and SLC39A14 was lower in NBCs and higher in LPS activated PBs compared to the ASC marker, XBP1. This suggests that these genes could serve as novel biomarkers for ASCs. In fact, expression of TMEM214 and FNDC3B in PB as opposed to NBC has been validated by our colleagues using Western blot (Fig 4-24D).

#### 4.4.3.4.1 **FNDC3B**

FNDC3B has been reported to localise to the Golgi network [213]. Aberrant expression of FNDC3B circular RNA has recently been implicated in the reduction of E-cadherin expression, leading to decreased cell-cell adhesion and promotion of cancer metastasis [214–216]. Furthermore, overexpression of this gene induces the cell surface localisation of transforming growth factor  $\beta$  receptor (TGFB1) in cancer cells [213]. Beyond these studies, little is known about FNDC3B, especially in plasma cell physiology. From existing reports, we theorise that FNDC3B may play a role in trafficking of proteins from the Golgi apparatus to the cell surface, especially because its isomer FNDC3A has recently been implicated in the secretion and correct localisation of collagen [217]. We note that FNDC3A is also reproducibly upregulated in ASCs but to a lesser extent than its isomer. Therefore, through our analysis we have identified a potentially novel factor playing a role in post-Golgi trafficking in ASCs. This gene would be an ideal target for knockdown, knockout or overexpression studies to determine its effect on antibody secretion.

#### 4.4.3.4.2 **TMEM184B**

Studies in mice have revealed that TMEM184B localises to recycling endosomes in neuronal cell bodies and plays a role in axon degeneration [218]. Little else is known about this gene. As protein level data was missing, Western blotting of TMEM184B to validate its expression in ASCs versus B cells is recommended for further studies. If results match the transcript regulation, this protein may be a potential novel marker of ASCs and a possible candidate for low throughput functional studies.

#### 4.4.3.4.3 **CRELD2**

CRELD2 is a stress inducible protein that localised to the ER/Golgi apparatus and is also reported to be constitutively exocytosed [219]. This glycoprotein is known to be upregulated in response to misfolded protein accumulation in the rER and reportedly regulated by BiP and ATF6 [219–221]. A previous study has proposed that CRELD2 may potentially function as a protein disulphide isomerase (PDI) based on its specificity for misfolded proteins [221]. Interestingly, exogenous CRELD2 has been shown to enhance the secretion of ECM proteins such as collagen during osteogenic differentiation [222]. According to our multi-omics analysis, this gene shows remarkable upregulation in ASCs and its expression has been validated by our colleagues using Western blot (Fig 4-24A-D). Based on this data we observe another parallel between the trafficking of the bulky cargo, collagen, to antibody secretion as in the case of COPII vesicle loading discussed in Chapter 3, Section 3.4.3. As ASCs do not secrete collagen, we hypothesise that this gene may be a key factor for enhanced secretion of antibodies in PCs.

#### 4.4.3.4.4 **SLC39A14**

ZIP14/ SLC39A14 is a well characterised divalent metal transporter known to localise to the plasma membrane [223]. It has been implicated in the transport and homeostasis of primarily zinc, but also iron, cadmium and manganese [224–226]. As a cell surface membrane protein, SLC39A14 is an ideal marker for cell targeting. Due to its marked upregulation in ASCs vs

NBCs in both protein and gene level, SLC39A14 coupled with other ASC biomarker e.g. CD138, may improve the targeting of ASCs for laboratory purification and potentially targeted destruction in cancer therapy.

#### 4.4.3.4.5 **TMEM214**

A study in 2013 determined TMEM214 to be prevalent in ER membrane fractions and colocalised with the translocon subunit, SEC61 $\beta$  [212]. Overexpression of this gene was found to promote apoptosis while knockdown led to the inhibition of ER stress induced apoptosis by preventing ER recruitment and subsequent cleavage of procaspase 4 [212]. ASCs can survive tremendous proteotoxic ER stress, therefore, TMEM214 induction in ASCs suggest that it may have a hitherto unknown role that is unrelated to apoptosis.

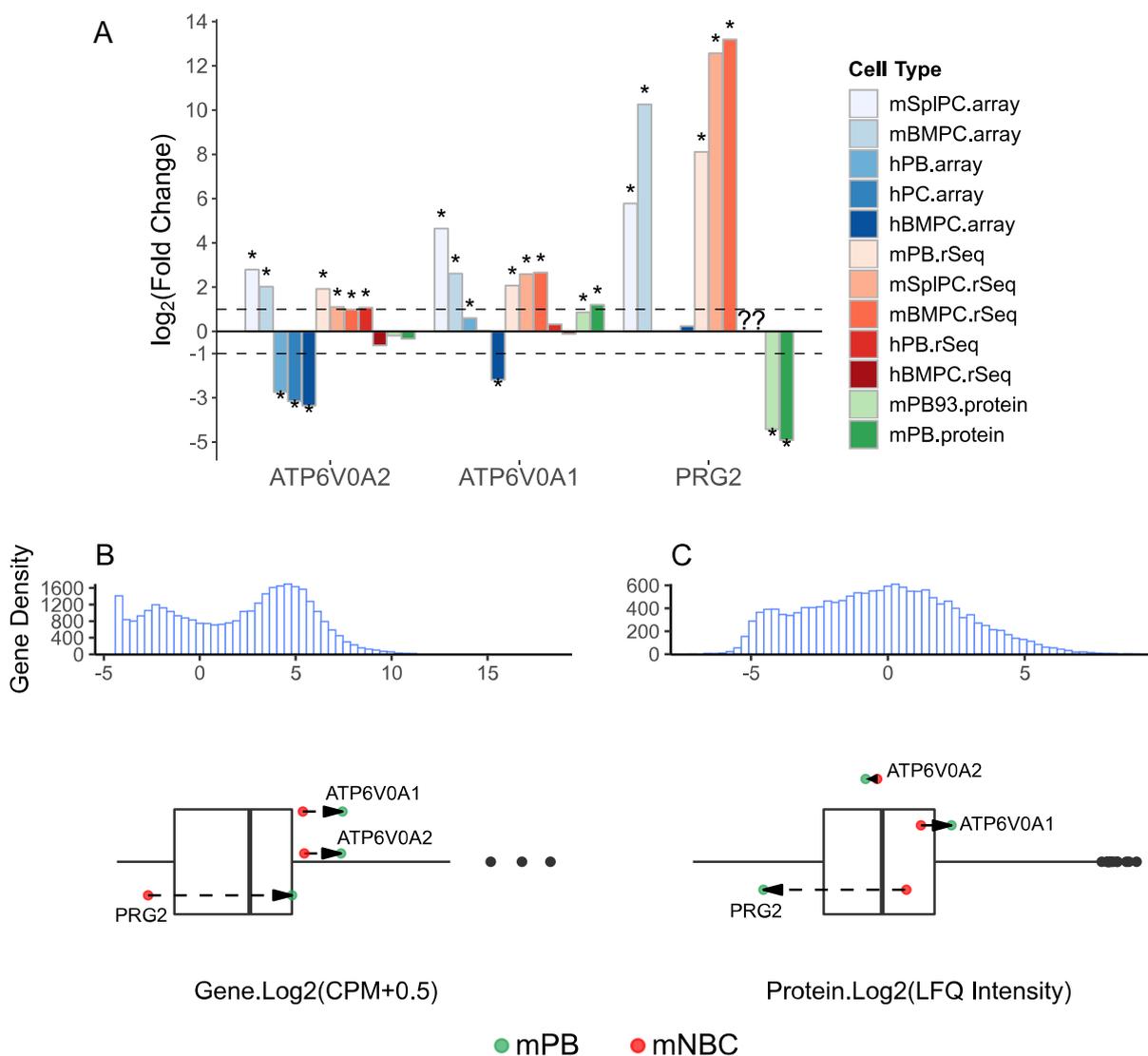
#### 4.4.3.5 **HID1 – a promising marker for ASCs**

Perhaps the clearest marker for ASCs is HID1, which we first identified in the preliminary study in Chapter 2, Section 2.3.1.1. As apparent in Fig 4-24A, this gene showed almost 128 fold consistent upregulation in our proteogenomic analysis. Histogram/boxplot of the mouse proteome shows that change in HID1 protein and gene expression far exceeded those of other markers we have discussed so far. HID1 levels were lower than many other markers in NBCs and encompassed their upregulation in the corresponding PBs (Fig 4-24BC). Western blotting confirms the unique upregulation of this gene in ASC vs NBCs (Fig 4-24D).

HID1 localises to the medial and *trans*-Golgi membrane [227]. HID1 is reported to be a component of dense core vesicles in synaptic transmission and *hid1* null mutant was found to lower the secretion of neurotransmitter cargo [228]. In these dense core vesicles, HID1 is thought to play a role in the prevention of incorrect sorting of peptide cargoes for lysosomal degradation. [229]. Knockdown of HID1 in pancreatic  $\beta$  cells is reported to lead to an increase in the abundance of immature secretory granules and elevate the secretion of immature proteins, i.e. proinsulin [230]. A recent study has further shown that HID1 promotes trans-

Golgi acidification via correct localisation of vacuolar ATP thus promoting dense core vesicle formation [230]. To-date HID1 gene function has been characterised in neurotransmitter and insulin secretion, both of which utilise the regulated secretory pathway. This is why we previously speculated the potential existence of a regulated secretory pathway in ASCs. As we find little evidence of regulated secretory pathway in ASCs, it is unclear what role HID1 is playing in the plasma cell physiology. As we have validated its upregulation in ASCs by both Western blotting and whole cell proteomics, future studies should include knockdown/knockout of this gene to determine its role in ASCs.

## 4.4.4 Review of Genes identified in Preliminary study

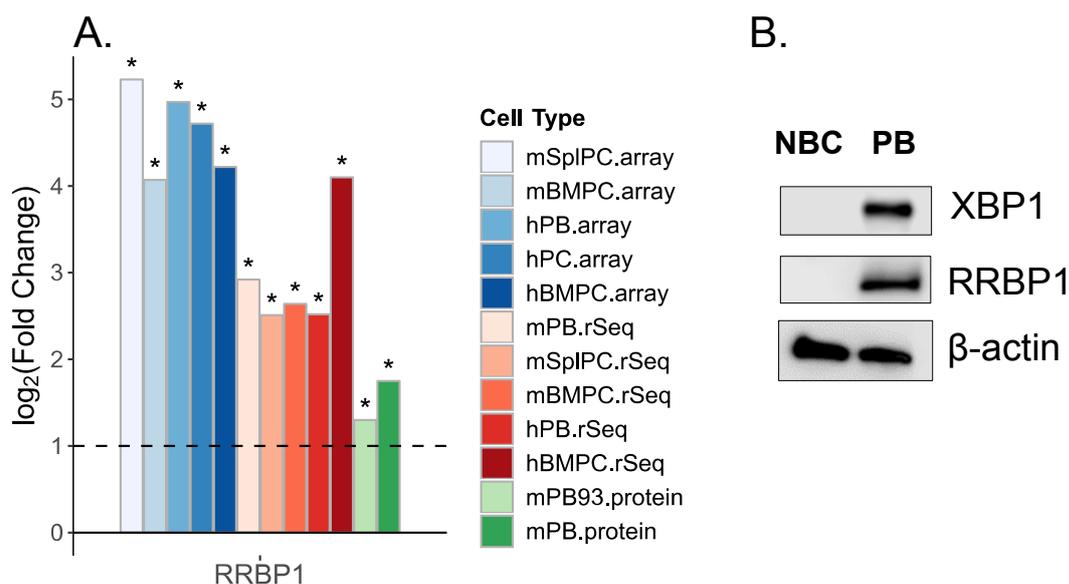


**Fig 4-25 | A.** Proteogenomic regulation of components related to regulated secretory pathway. (?) indicates missing data. (\*) indicates FDR adjusted  $p$ -value < 0.05. **B-C.** Histograms representing normalised distributions of LPS activated mouse plasmablasts and naïve B cells. Expression of specific genes/proteins have been visualised in context of global data distribution. **B.** Normalised transcriptome (RNA-Seq) and **C.** Proteome (MS/MS).

As HID1 is a component of regulated secretion, we hypothesised the presence of this regulated pathway in our preliminary study. We revisit the results of this study to show how cross species meta-analysis have improved our results. We rationalised that regulatory secretory pathway may play a role in ASCs noting the upregulation of HID1 but also the PRG2 cargo, which is transported in granule fractions of eosinophils and neutrophils [104, 109–111]. As further indication we noted the upregulation of vacuolar ATPases, ATP6V0A1 and ATP6V0A2,

which are also implicated in regulated secretion [231]. Proteogenomic meta-analysis shows that PRG2 expression is inconsistent across species and is downregulated in the mouse proteome. Although, ATP6V0A1 showed consistent upregulation in mouse proteome, it exhibited little change in human ASCs. Furthermore, its isoform ATP6V0A2 show very inconsistent and relatively muted regulation in ASCs. Thus, it is likely that our initial hypothesis was incorrect regarding regulated secretion in ASCs. This highlights the importance of multi-omics analysis in filtering out poor hits in the study of core cellular processes, such as secretion. Furthermore, the lack of evidence towards regulated secretion, indicate that the considerable upregulation in HID1 may play a role distinct from the regulated secretory pathway in the plasma cell physiology.

#### 4.4.5 RRBP1 – Ribosome Anchor



**Fig 4-26 | A.** Proteogenomic regulation of RRBP1. This gene is consistently upregulated in every ASC type studied. **B.** Western Blot of RRBP1 show unique expression of RRBP1 in PBs but not NBCs. (?) indicates missing data. (\*) indicates FDR adjusted  $p$ -value < 0.05.

In chapter 3, Section 3.4.3.1, we predicted that the CREB3L2 TF may co-regulate the poorly characterised ribosome anchor, RRBP1/p180. Proteogenomic analysis consolidates this finding as RRBP1 was upregulated in all ASCs studied across platforms, across species, and in both

gene and protein level (Fig 4-26A). The expression of this gene has also been validated by our colleagues using Western blot (E Rajan & AWA Aswani).

#### 4.4.5.1 RRBP1 may stabilise ribosomal RNA

The ribosome is a complex multiprotein machinery made up of over 50 proteins that guide the synthesis of proteins from mRNA molecules. Each ribosome can be split into a small and a large subunit [1]. RRBP1 is thought to interact directly with translationally active ribosomes and promote their assembly at the ER [232]. RRBP1 has also been reported to directly anchor and potentially stabilise mRNA at the ER, independent of any interaction with ribosomes [232, 233]. Notably, our data indicates the vast majority of ribosomal mRNA showed a tendency to be downregulated (Fig 4-27A), whereas, ribosomal proteins exhibited considerable upregulation (Fig 4-27B). As such, we observe that ribosomal proteins may be upregulated in ASCs in a post transcriptional manner. Therefore, we speculate that RRBP1 may recruit ribosomes to the ER by potentially stabilising the mRNA of their subunits, thus enhancing their translation.

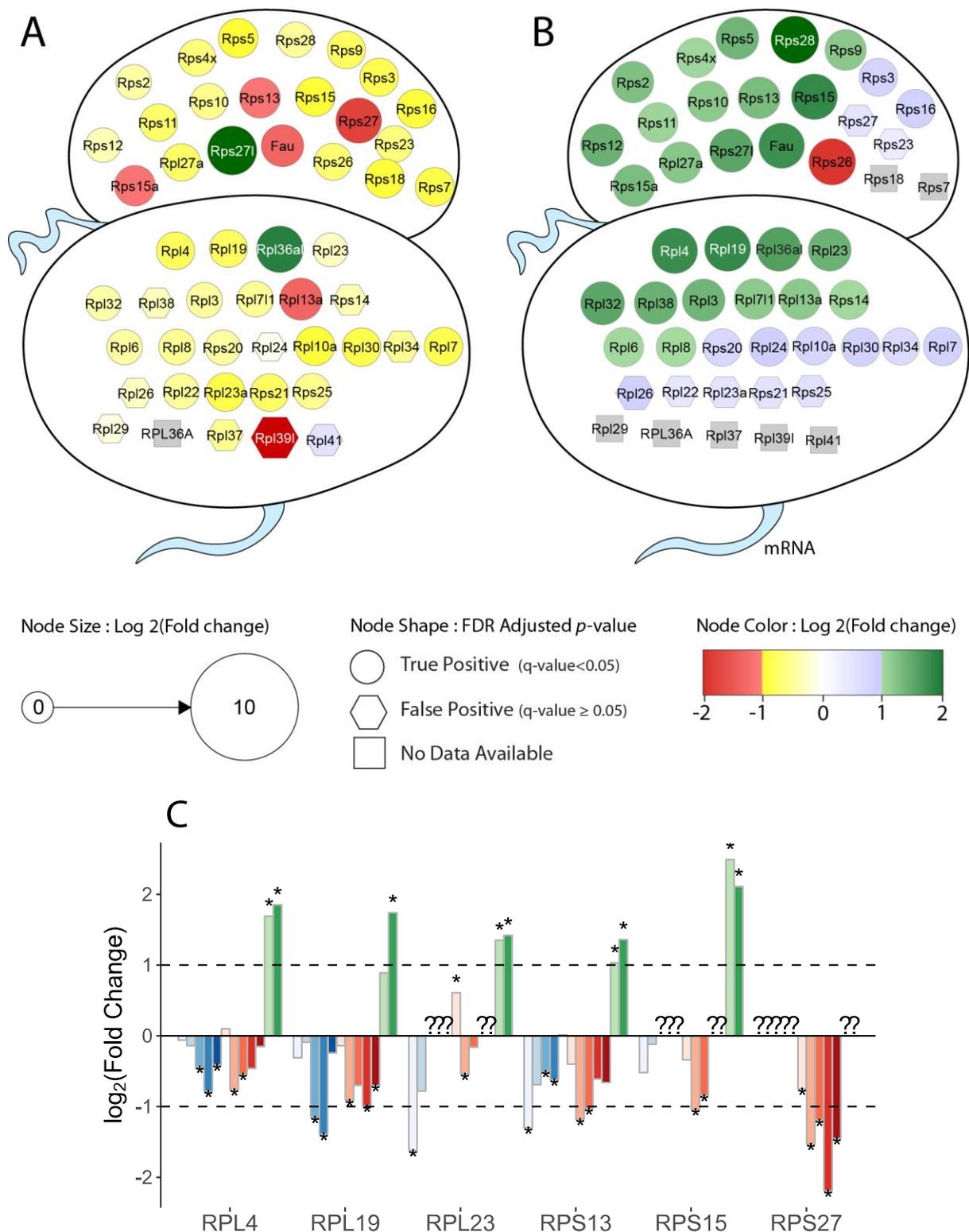
#### 4.4.5.2 RRBP1 may stabilise ribosomal protein

Soluble ribosomal subunits are typically synthesized in excess, where soluble and complexed subunits maintain a state of equilibrium [234]. Turnover of ribosomal subunits is reported to correlate with size and solubility, with soluble or large subunits showing higher rate of degradation than their complexed or lower molecular weight variants, respectively [234]. Excess subunits undergo degradation via ubiquitin-proteasomal degradation [235]. Our data shows marked upregulation in most ribosomal protein subunits in ASCs. While increased mRNA stability via RRBP1 may contribute to this increased expression, an alternative hypothesis is that RRBP1 may stabilise ribosomal proteins by (a) increasing ribosome assembly and thus the proportion of ribosomal subunits in complexed form; or (b) inhibiting

the ubiquitination and thus degradation of ribosomal subunits. Further study is required to ascertain the role of RRBP1 in potentially stabilising ribosomal proteins and/or mRNA.

#### 4.4.5.3 RRBP1 may promote antibody secretion

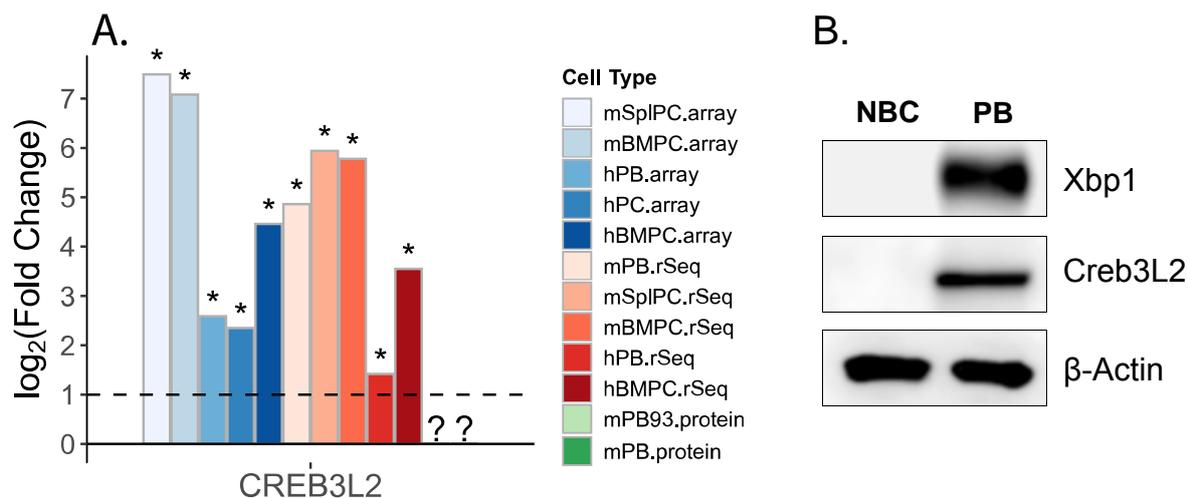
Knockdown of RRBP1 in fibroblasts has previously been reported to impair *trans*-Golgi expansion and ribosome association to the ER [160, 236]. Importantly, this resulted in the specific reduction in the secretion of the bulky protein, collagen, via perturbation of procollagen translation, as mRNA levels were unaffected [160]. As collagen is not secreted by antibody secreting cells, the reproducible upregulation of RRBP1 gene and protein in ASCs leads us to hypothesise that RRBP1 may play a similar role in enhancing antibody translation and, therefore, secretion in plasma cells. This also highlights the potential importance of CREB3L2 as a regulator of antibody secretion, as we predict that this transcription factor may promote the expression of RRBP1.



**Fig 4-27 | A-B.** Proteogenomics regulation of ribosomal subunits in ASCs represented as coloured nodes. **A.** Transcriptomic Regulation of Ribosomal proteins in ASCs. mRNA for most small and large ribosomal subunits show a tendency to be downregulated (yellow/red) in ASCs as opposed to B cells. **B.** Proteomic Regulation of Ribosomal proteins in ASCs. Protein level expression of most small and large ribosomal subunits show a tendency for statistically significant upregulation (purple/green) in ASCs. **C.** Bar chart of fold changes of select ribosomal subunits, show start contrast in protein and gene regulation of ribosomal subunits.

## 4.4.5.4 Cargo selection at COPII vesicles may be mediated by CREB3L2

In the Chapter 3, we also identified CREB3L2 as a regulator of potential cargo selectors for COPII vesicles by studying species conserved regulation in genes as evidenced by microarrays. In this chapter, we determine whether CREB3L2 upregulation is reflected in other platforms and in the protein level.

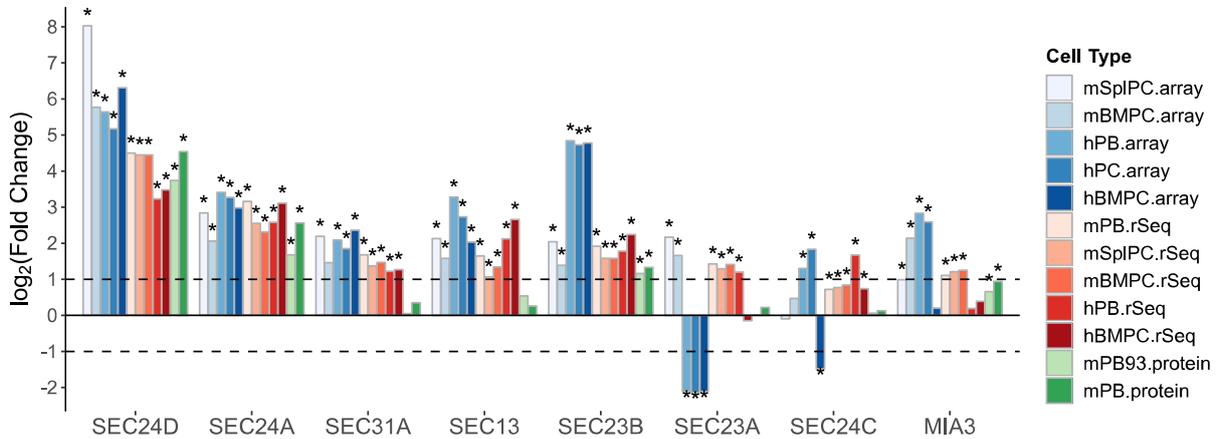


**Fig 4-28 | A.** CREB3L2 gene regulation across platforms and species. CREB3L2 is highly upregulated across mice and human transcriptome of ASCs as opposed to B cells. Protein level evidence was missing in the label free MS/MS output. (?) indicates missing data. (\*) indicates FDR adjusted  $p$ -value  $< 0.05$ . **B.** Western Blot carried out by colleagues E Rajan & AWA Aswani show clear induction of CREB3L2 in mouse CD138+ plasmablasts (PB) but none in equivalent naïve B cells (NBC).

As shown in Fig 4-28, CREB3L2 was consistently upregulated across species in both RNA-Seq and microarray data. Although, protein level data for CREB3L2 was missing MS/MS output, Western blot shows that CREB3L2 is uniquely expressed in 3-day old PBs but not in NBCs. A recently published report inhibited the expression of CREB3 family proteins and sterol regulatory element-binding proteins (SREBPs) by disrupting a site-1 protease using the drug, PF-429242, in ASCs [237]. This resulted in the dramatic reduction in antibody secretion and Al-Maskari *et al* suggests that CREB3L2 may play a key role in this as PF-429242 greatly attenuates its activity. We note that specific perturbation of CREB3L2 has not

yet been performed as the phenotype observed by Al-Maskari *et al* is a combined result of downregulating CREB3 TFs and SREBPs.

#### 4.4.5.5 COPII vesicle components



**Fig 4-29** | Differential Regulation of COPII vesicle components according to multi-omics analysis. (?) indicates missing data. (\*) indicates FDR adjusted  $p$ -value < 0.05.

The inner coat of the COPII complex is composed of SAR1 and the SEC23-SEC24 heterodimer, while the outer coat is formed by SEC13-SEC31 heterotrimer [238]. Previous qRT-PCR studies have proposed the upregulation of SEC31A and SEC24C in ASCs [54]. It is evident from Fig 4-29, that while these genes showed a tendency to be upregulated in both the protein and gene level, the extent of regulation was relatively weak compared to other isoforms of the inner COPII coat, i.e. SEC24A, SEC24D and SEC23B.

##### 4.4.5.5.1 SEC23

CREB3L2 has been implicated in marginal but statistically significant upregulation of SEC23A during hepatic stellate cell differentiation [153]. However, in ASCs this is not the case as this SEC23 isoform showed contradictory regulation in gene level and little or no change in protein level. SEC23B is the likely SEC23 isoform that operates in ASCs as there is evidence of consistent upregulation of SEC23B in both transcriptome and protein level (Fig 4-29). Transcription perturbation analysis shows that the well characterised enhancer of plasma cell

differentiation, XBP1, operates upstream of this gene. In addition, SEC23B is also predicted to be regulated by the upregulated transcription factor, CREB3.

#### 4.4.5.5.2 **SEC24**

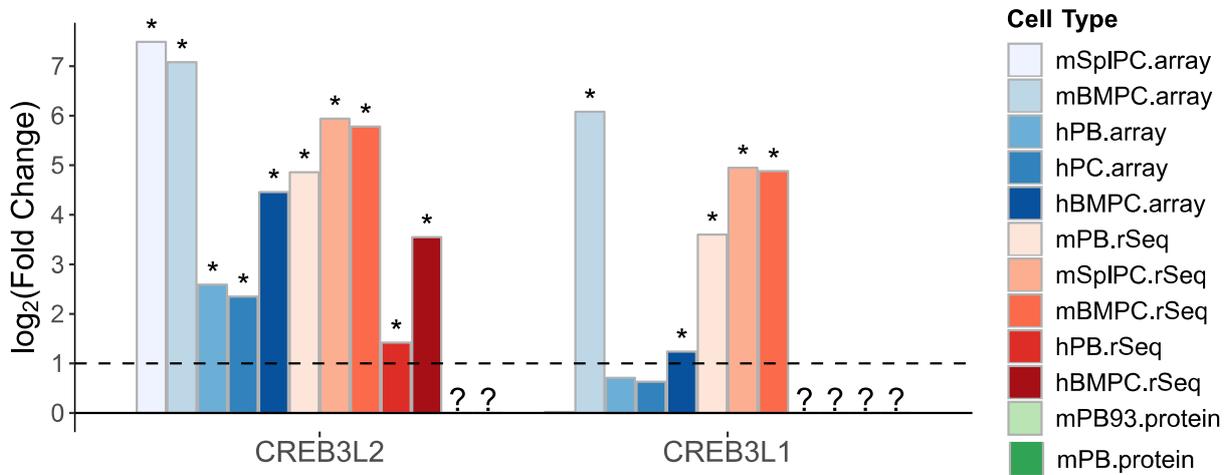
In this chapter, cross platform analysis highlighted a considerable upregulation in SEC24A (Fig 4-29). The upstream regulator for this gene has not been identified, but interestingly, ARCHS4 transcription factor co-expression analysis indicate that SEC24A is a likely co-regulation target for the CREB3L2 TF (Appendix Table 6-9).

We highlighted SEC24D in Chapter 3 as the most upregulated component of coatamer protein complex (COP)II vesicles. Here we show that this regulation is mirrored in the RNA-Seq as well as MS/MS profile of ASCs (Fig 4-29). CREB3L2 knockout studies marginally but significantly reduced the expression of SEC24D during hepatic stellate cell differentiation [153]. The specific overexpression of the CREB3L1 isoform in HeLa cells has been reported to upregulate SEC24D as well [239]. The redundant regulation of SEC24D by CREB3L1 and CREB3L2 may explain why CREB3L2 knockout only had a marginal affect in HeLa cells. Furthermore, it is worth noting that HeLa cells are not naturally optimised for secretion.

Unlike CREB3L2, the CREB3L1 gene is upregulated in mouse ASCs but not in the human model according to microarray profiles. Unfortunately, the equivalent RNA-Seq profile did not detect this gene (Fig 4-30). This contradiction points to CREB3L2, not CREB3L1, being the likely regulator of SEC24D in ASCs. Therefore, we predict that specific upregulation of CREB3L2 may improve protein secretion in ASCs. Nevertheless, as we do not have the full picture of CREB3L1, validation of this gene expression in ASCs vs NBCs is needed to alleviate any concerns of redundant gene regulation.

As SEC23B, SEC24A and SEC24D are the highest upregulated components in antibody secreting cells. We hypothesise that a combination of SEC23B-SEC24A and/or SEC23B-SEC24D may play a role in loading antibodies into COPII vesicles in ASCs. As we have soft

validated their expression by proteomics, for further studies, knockdown/knockout of these genes should be carried out in ASC model to determine their effect on antibody secretion.



**Fig 4-30** | Differential Regulation of CREB3-like isoforms (multi-omics). While CREB3L2 gene expression is consistent across species, CREB3L1 genes show little or no change in the transcriptome of human ASCs. MS/MS could not detect either genes. (?) indicates missing data. (\*) indicates FDR adjusted  $p$ -value < 0.05.

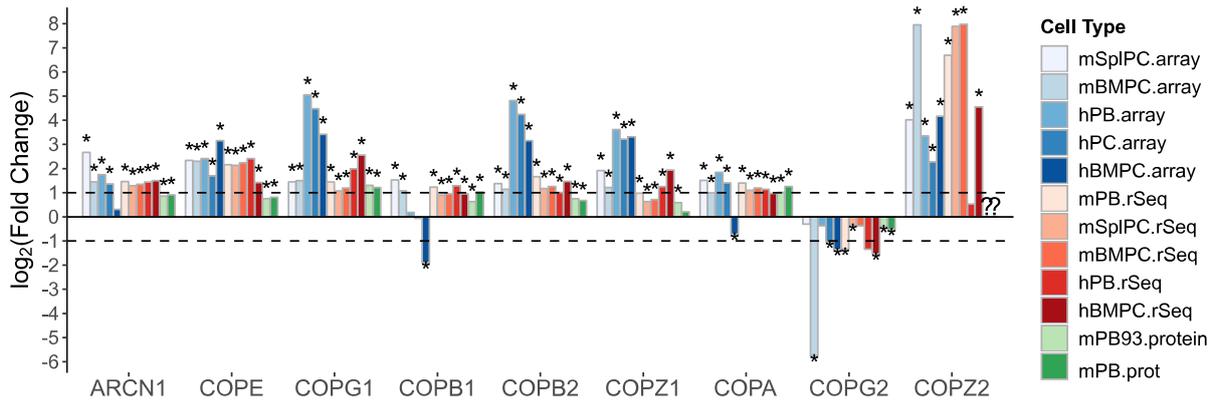
#### 4.4.6 Regulation of Membrane Trafficking Components in ASCs

Previous studies in PCs have shown that SEC31A and SEC24C is upregulated in plasma cells [54]. We have established that most components of the COPII coat including these proteins had a tendency to be upregulated in the ASC phenotype. Our aims were also to investigate how other known membrane trafficking vesicles were regulated in the plasma cell physiology. The summary of our findings for the upcoming sections is illustrated in Fig 4-40, page 186.

##### 4.4.6.1 Coated vesicles

###### 4.4.6.1.1 COPI isomers. ASCs show preference for early Golgi to ER transport route

COPI participates in the retrograde transport of cargo from the Golgi apparatus to the ER. As Golgi-ER retrograde transport is a highly enriched GO term in ASCs, it is not surprising that almost all components of the COPI vesicles show reproducible upregulation in both gene and protein level (Fig 4-31).



**Fig 4-31** | Differential Regulation of COPI coat proteins (multi-omics). COPZ2 gene shows consistently higher upregulation (up to 256-fold) in ASCs of mice and human.

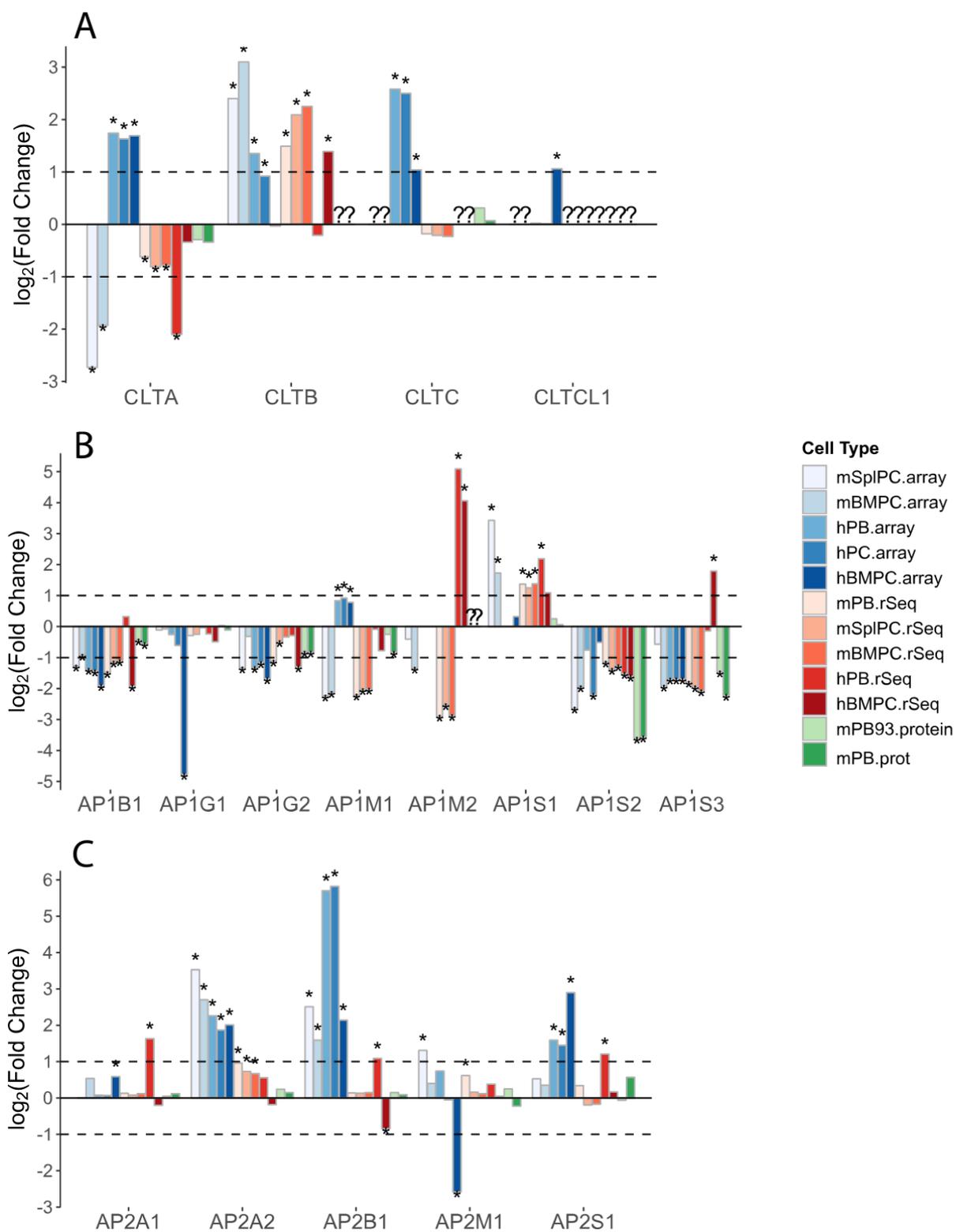
Interestingly, the COPZ2 isoform showed consistently higher gene expression compared to other COPI proteins. Quantitative immunoelectron microscopy has previously shown that COPI isoforms may perform distinct functions in mammalian cells [240]. The COPG1-COPZ2 subunit combination has been reported to localise to the ERGIC and *cis*-Golgi, while the COPG2-COPZ1 may act in the *trans*-Golgi [240]. Our data suggests that ASCs may exhibit a physiological preference for retrograde transport from early Golgi apparatus to the ER, in line with our GO functional analysis. This hypothesis is further substantiated by the small but significant downregulation of COPG2 (Fig 4-31) acting in the *trans*-Golgi network (TGN). Based on interaction with other COPII coat proteins, COPZ2 is believed to have redundant functions in the cell [241]. We, however, note that the cargo specificity for COPZ2 containing vesicles have not been studied. Therefore, as COPZ2 is remarkably upregulated in ASC, we speculate that this subunit may bind to signal sequences/ adaptors on bulky misfolded cargo, such as antibodies, and aggregate them for retrograde transport. Nevertheless, as the protein level data for COPZ2 was missing in the mass spectrometry analysis, we recommend Western blotting to validate this result, before attempting functional analysis.

#### 4.4.6.1.2 Clathrin coated vesicles show no significant regulation in ASCs

Clathrin coated vesicles transport cargo to and from the *trans*-Golgi and are key components of receptor mediated endocytosis and endocytic recycling. Different heterotetrameric adaptor protein (AP) complexes participate in the recruitment of cargo and clathrin during the formation of these vesicles [242]. Typically professional secretory cells couple exocytosis with endocytosis to recycle membranes and prevent secreted cargo from accumulating on the plasma membrane [30]. As such one would expect components of clathrin mediated endocytosis to be upregulated in ASCs due to higher secretory demand. However, our data indicates that clathrin heavy and light chains (Fig 4-32A) and the requisite cargo-sorting AP-2 complex (Fig 4-32C) show no significant change in ASCs as opposed to B cells. We hypothesize that this is likely because the relative expression levels of clathrin vesicles components in NBCs are enough to handle the increased endocytic demand in ASCs.

We note that the AP-1 complex, which mediates sorting of cargo for endocytic recycling, i.e. trafficking from the *trans*-Golgi to recycling endosomes, showed statistically significant downregulation in specific AP-1 $\sigma$  subunits (Fig 4-32B). Although, AP1S2 and AP1S3 is strongly downregulated, AP1S1 shows no significant change in expression in ASCs. As such, we predict that this may an indirect selection for the AP1S1 subunit in the AP-1 mediated vesicles of ASCs. Further validation using Western blotting is needed to validate this result.

We also note that other AP-1 subunits tended to show small but statistically significant downregulation in ASCs, especially in the protein level. This pattern of regulation is surprising as we expect a professional secretory cells to upregulate receptor mediated endocytosis and endocytic recycling [30]. We initially noted that some of these regulations are less than 2-fold change and assumed that the change may not be substantial enough to have a significant impact on the ASC phenotype. However, the consistency in the pattern of downregulation across



**Fig 4-32 | Differential regulation of clathrin coated vesicles (multi-omics).** **A.** Clathrin coat proteins. **B.** Clathrin adaptor protein complex 1 (AP1) that mediate transport between *trans*-Golgi and recycling endosomes. AP1 $\sigma$  isomers (AP1S2, AP1S3) show significant downregulation in ASCs. **C.** Clathrin adaptor protein complex 2 (AP2) mediate receptor mediated endocytosis at the plasma membrane. AP-2 complex shows little or no change in ASCs.

subunits was suspect, especially as this pattern was also seen in associated tethers discussed in the next section.

#### 4.4.6.2 Tethering Factors

Membrane bound vesicles bud from donor membranes, travel to target membrane where Rab GTPases and tethering complexes make initial contact and promote SNARE assembly for the fusion of the incoming vesicle with the target membrane [30]. As tethers play a key role in vesicular trafficking, we take a closer look at the regulation of different tethering factors.

##### 4.4.6.2.1 Post-Golgi tethers

###### *CORVET and HOPS complex show mild downregulation*

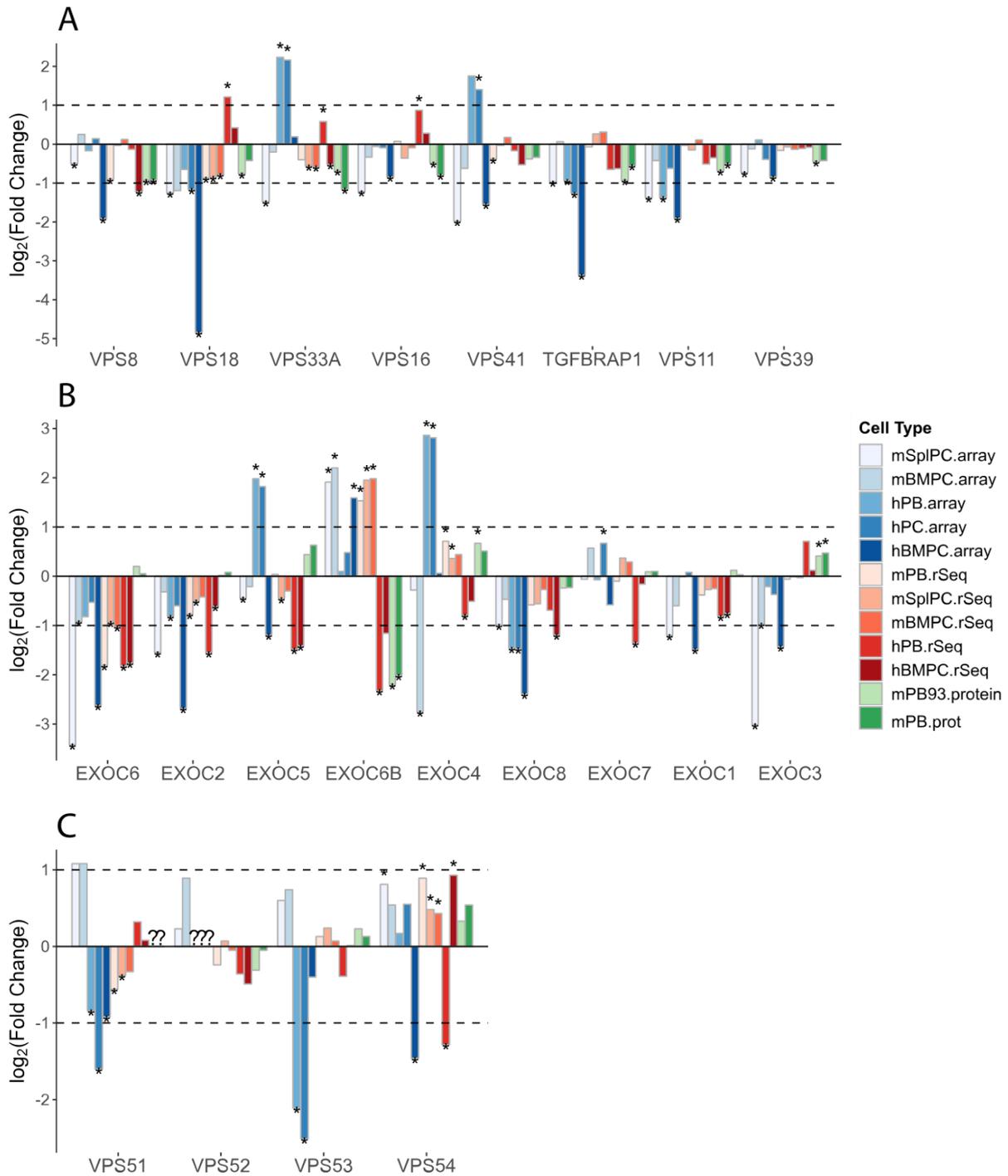
CORVET and HOPS are homologous multi-subunit complexes. CORVET complex resides on the membrane of early endosomes and functions as a tether for incoming vesicles from the *trans*-Golgi and late endosome. HOPS, on the other hand, reside in lysosomes and late endosomes and tether incoming vesicles destined for lysosomal degradation [34]. According to our data, both these tethers show a tendency for mild but statistically significant downregulation in protein level in almost all their subunits (FDR adjusted  $p$ -value < 0.05) although there were some inconsistencies in gene level (Fig 4-33A). As we have previously noted a similar regulatory pattern in components of the clathrin adaptor, AP-1, we suspect that ASC may downregulate forward transport of cargo from the *trans*-Golgi to endosomes and lysosomes to some extent. Further validation is required to confirm these peculiar results.

###### *Exocyst and GARP complex*

Components of the GARP complex act to tether incoming vesicles from recycling endosomes to the *trans*-Golgi. As shown in shows Fig 4-33C, expression of this complex was relatively unaffected by plasma cell differentiation [34]. Likewise, the Exocyst complex, thought to tether incoming vesicles at the plasma membrane from the *trans*-Golgi and recycling endosomes, was also unaffected (Fig 4-33B).

*Post Golgi trafficking of antibodies may utilise a shortcut to the plasma membrane*

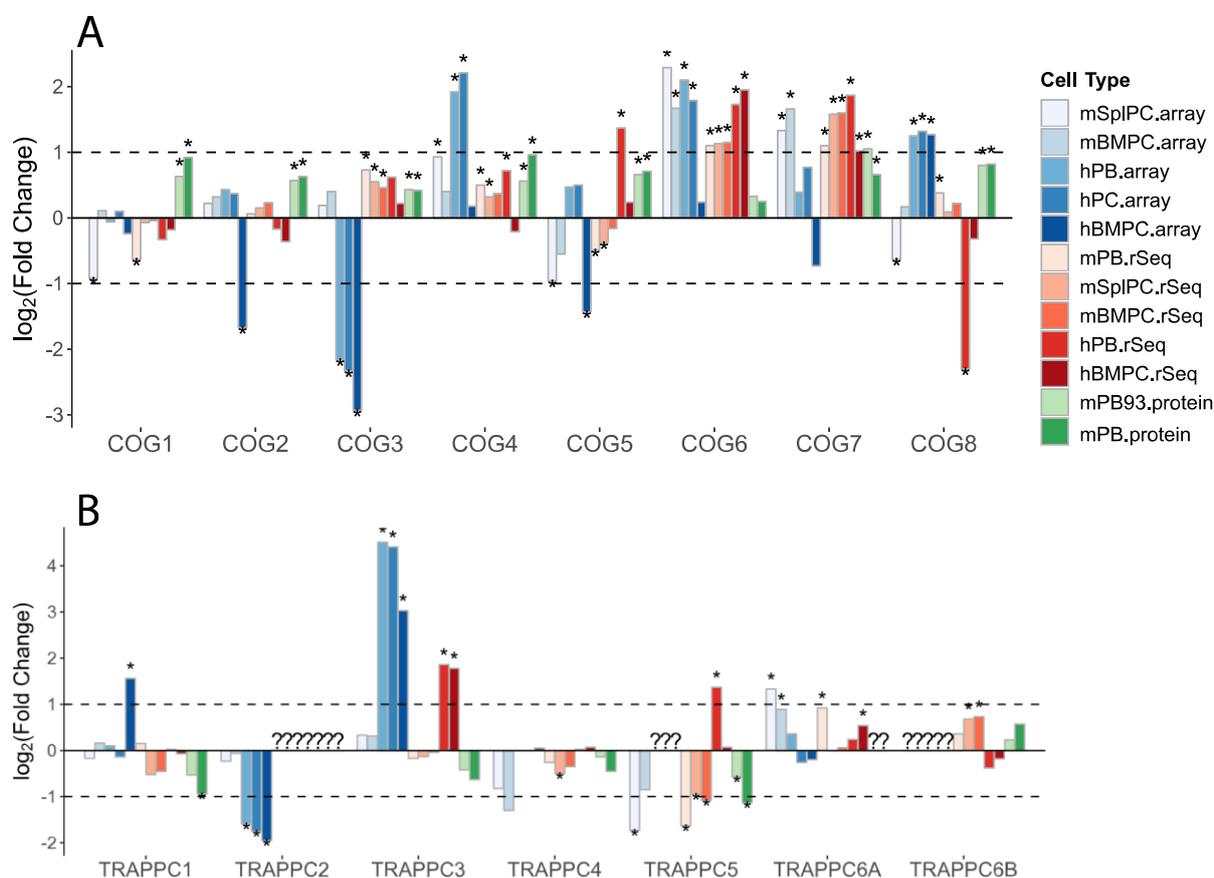
Constitutive cargo in Madin-Darby Canine Kidney (MDCK) cells and macrophages are reportedly trafficked through endosomal compartments and require the Exocyst complex to mediate the fusion of incoming vesicles to the plasma membrane [243, 244]. However, we have shown that the AP-1 adaptor proteins of clathrin coated vesicles and CORVET/HOPS tethering complexes implicated in endocytic recycling show a tendency to be downregulated in ASCs, and the Exocyst complex exhibit no change in expression. A live-imaging study demonstrated that post-Golgi transport in non-polarised cells may not require trafficking through endosomal compartments or tethering via the Exocyst complex [245]. As ASCs are indeed non-polarised, our data supports this hypothesis. Therefore, we predict that antibodies forego trafficking through endosomal compartments and may take an alternate “shortcut” to the plasma membrane. Nevertheless, we recommend performing Western blot of these endosomal recycling components for confirming their expression in ASCs.



**Fig 4-33** | Differential regulation of post-Golgi tethers (multi-omics). **A.** CORVET/HOPS showed mild downregulation in protein level and inconsistent regulation in gene level (less than 2 absolute FC). These two tethering complexes act in the delivery of cargo to lysosomes or the cell surface. **B.** The Exocyst complex showed no significant change in ASCs. This complex tethers vesicles arriving from the TGN to the plasma membrane. **C.** GARP complex showed no significant change in ASCs. This complex acts as a tether for the retrograde transport of cargo from recycling endosomes to the Golgi apparatus.

## 4.4.6.2.2 ER-Golgi Tethers

## COG Complex



**Fig 4-34** | Proteogenomic regulation of components of ER-Golgi tethers. **A.** COG complex acts as a tether for intra-Golgi transport vesicles. COG subunits show consistent upregulation. **B.** Components of TRAPPC1 complex show little or no reproducible change in ASCs.

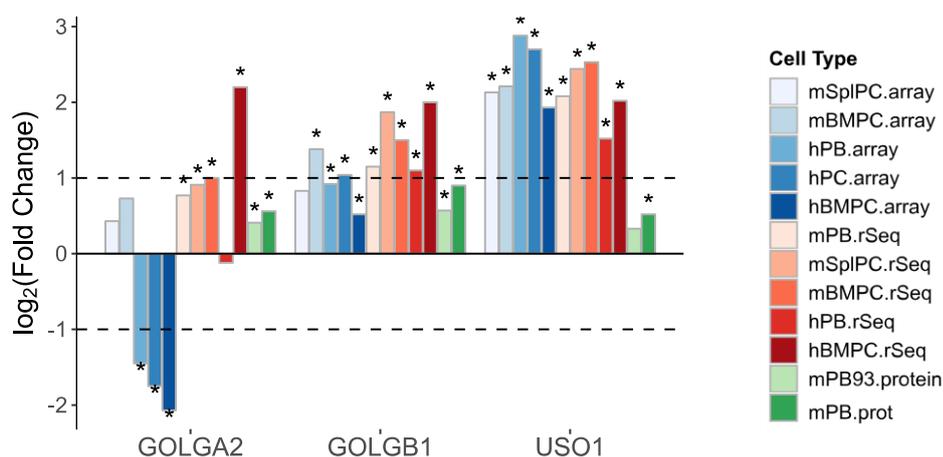
The Conserved Oligomeric Golgi (COG) complex is thought to act as the tether for intra-Golgi retrograde transport vesicles (COPI) and plays a role in the correct localisation of glycosylation enzymes and other Golgi resident components [31]. This tethering complex shows small but consistent upregulation in the protein level, but has somewhat inconsistent gene regulation (Fig 4-34A). Curiously, COG subunits show small but statistically significant regulation reminiscent of CORVET/HOPS complex, but in this case, we observe an overall upregulation in these tethers. This tendency for upregulation possibly reflects the parallel induction of COPI vesicles in ASCs discussed in Section 4.4.6.1.1. Further validation, is required to determine if

the fold changes in the protein level are sufficiently large enough to have any discernible effect on the ASC phenotype.

### TRAPPI

The TRAPPI complex acts at the ERGIC or *cis*-Golgi to tether incoming COPII vesicles from the ER. Based on the considerable upregulation of COPII vesicles in ASCs, we assumed this tether would also show upregulation to reflect the need to tether a greater number of COPII vesicles budding from the ER. However, as shown in Fig 4-34, TRAPPI components show no significant change in ASCs.

### USO1



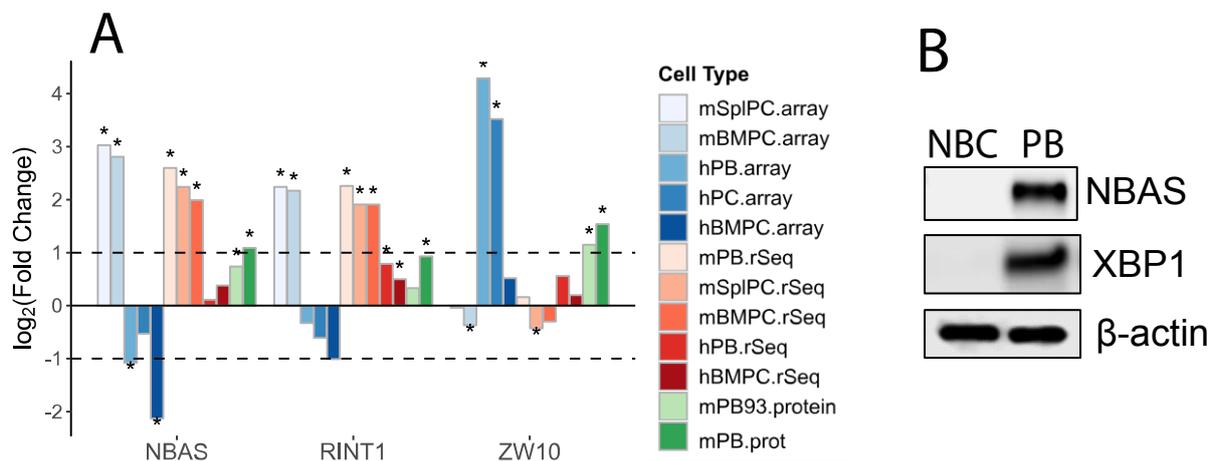
**Fig 4-35** | Differential regulation of USO1 tethering factor and its interactors at the *cis*-Golgi.

Interestingly, USO1/p115, the coiled-coil tethering factor operating at the *cis*-Golgi was considerably upregulated in ASCs (Fig 4-35). In Chapter 1, Section 1.1.5.1.3, we discuss the role of this protein in the tethering of COPII vesicles at the *cis*-Golgi. Based on the lack of change in TRAPPI expression and the distinct upregulation of USO1, we hypothesize that increased number of USO1 proteins may serve to speed up forward trafficking of cargo and potentially lead to the enhancement of antibody secretion.

USO1 is also believed to interact with GOLGA2 and GOLGB1 to mediate intra-Golgi tethering of COPI vesicles. Although, gene expression of GOLGA2 was inconsistent, both

GOLGA2 and GOLGAB1 showed a tendency to be upregulated in ASCs (Fig 4-35). Thus, our data indicates that USO1 mediated tethering may play an active role in intra-Golgi transport within ASCs.

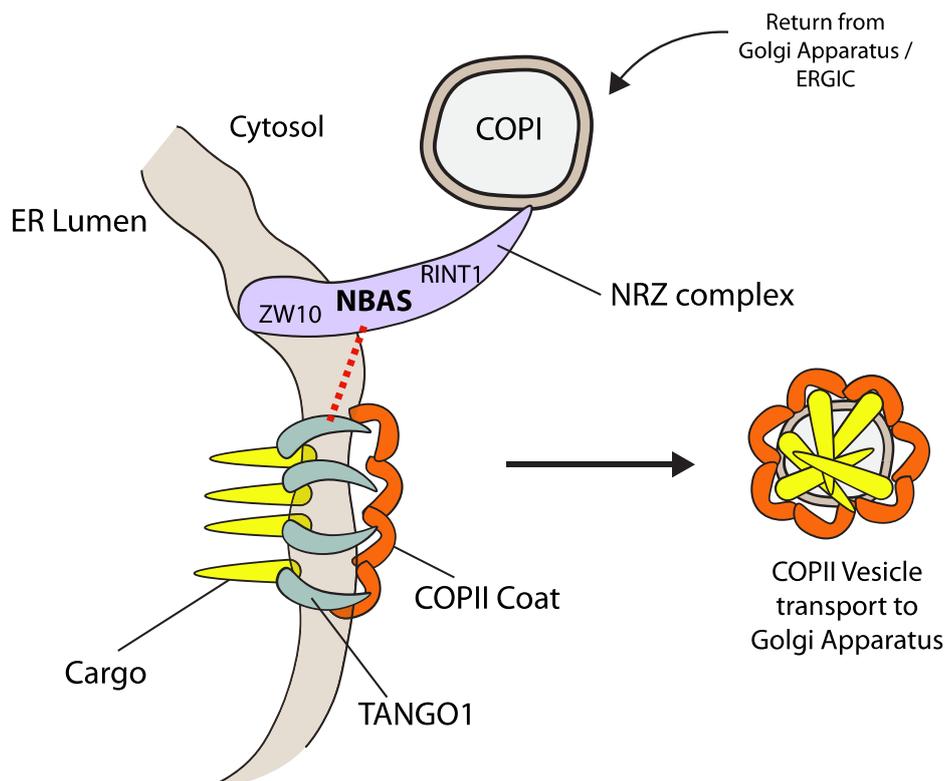
#### NRZ Complex



**Fig 4-36 | A.** Differential Regulation of NRZ subunits in ASCs (multi-omics). Gene level regulation is somewhat inconsistent for NRZ proteins, but all three components show statistically significant upregulation in protein level. **B.** Western blot carried out by our colleagues validates the upregulation of NBAS in ASCs.

The NRZ complex, discussed in Chapter 1, Section 1.1.5.1.4, mediates tethering at the ER by “catching” incoming COPI vesicles returning from the Golgi and thus promoting their fusion with the ER membrane [33].

NRZ was previously thought to only participate in retrograde transport but it has recently been implicated in the loading of collagen into COPII vesicles for antegrade transport from the ER [246]. As shown in Fig 4-37, the tethering of COPI vesicles allows NRZ complex to interact with MIA3/TANGO1, which is a component of the COPII vesicle. Raote *et al* proposed that TANGO1 requires interaction with the large cargo (collagen), COPII inner coat proteins (SEC23- SEC24) as well as the NRZ complex to prevent premature budding and thus promote the loading of collagen into COPII vesicles. Due to the considerable upregulation of COPI and COPII components and the dual role of NRZ in both COPI and COPII vesicle transport we took a closer look at the NRZ complex. Components of NRZ complex, NBAS and



**Fig 4-37** | Schematic diagram of TANGO1 interaction with NRZ complex, COPII coat proteins and the cargo, collagen. This interaction prevents premature budding of COPII vesicles, and allows the loading of collagen into COPII vesicles (based on Raote *et al*'s report).

RINT1, show consistent upregulation in ASCs, while ZW10, showed upregulation mainly in the protein level (Fig 4-36A). This upregulation of NRZ complex in ASCs is interesting as we note that ASCs do not secrete collagen. As in the case of CRELD2 and COPII cargo loading, we find another potential parallel between collagen and antibody trafficking. Thus we predict that the NRZ complex participates not only in the transport of collagen but other bulky cargo such as antibodies.

NBAS is a poorly characterised member of the NRZ tethering complex. This gene shows upregulation in the mice model but somewhat contradictory results in human ASCs. Nevertheless, NBAS showed ~2-fold upregulation in the mouse proteome and our colleagues performed Western blotting to validate this result and found that this gene was markedly upregulated in mouse plasmablasts (Fig 4-36B). Patients exhibiting mutations in NBAS gene have been reported to suffer from multisystem disorders characterised by liver dysfunction,

optic atrophy and, importantly, suffer multiple bone fractures, which suggests a possible dysregulation of collagen secretion [33, 247]. Interestingly, these patients also exhibit hypogammaglobulinemia. As such, we predict that the potential dual of NRZ complex in the COPI and COPII mediated transport of bulky cargoes such as antibodies, and not just collagen, may explain why these patients lack antibody secretion.

#### 4.4.6.3 SNAREs

SNAREs mediate the last step of vesicle docking and fusion to the target membrane. Structural and functional characteristics of SNAREs are discussed in Chapter 1, Section 1.1.6.

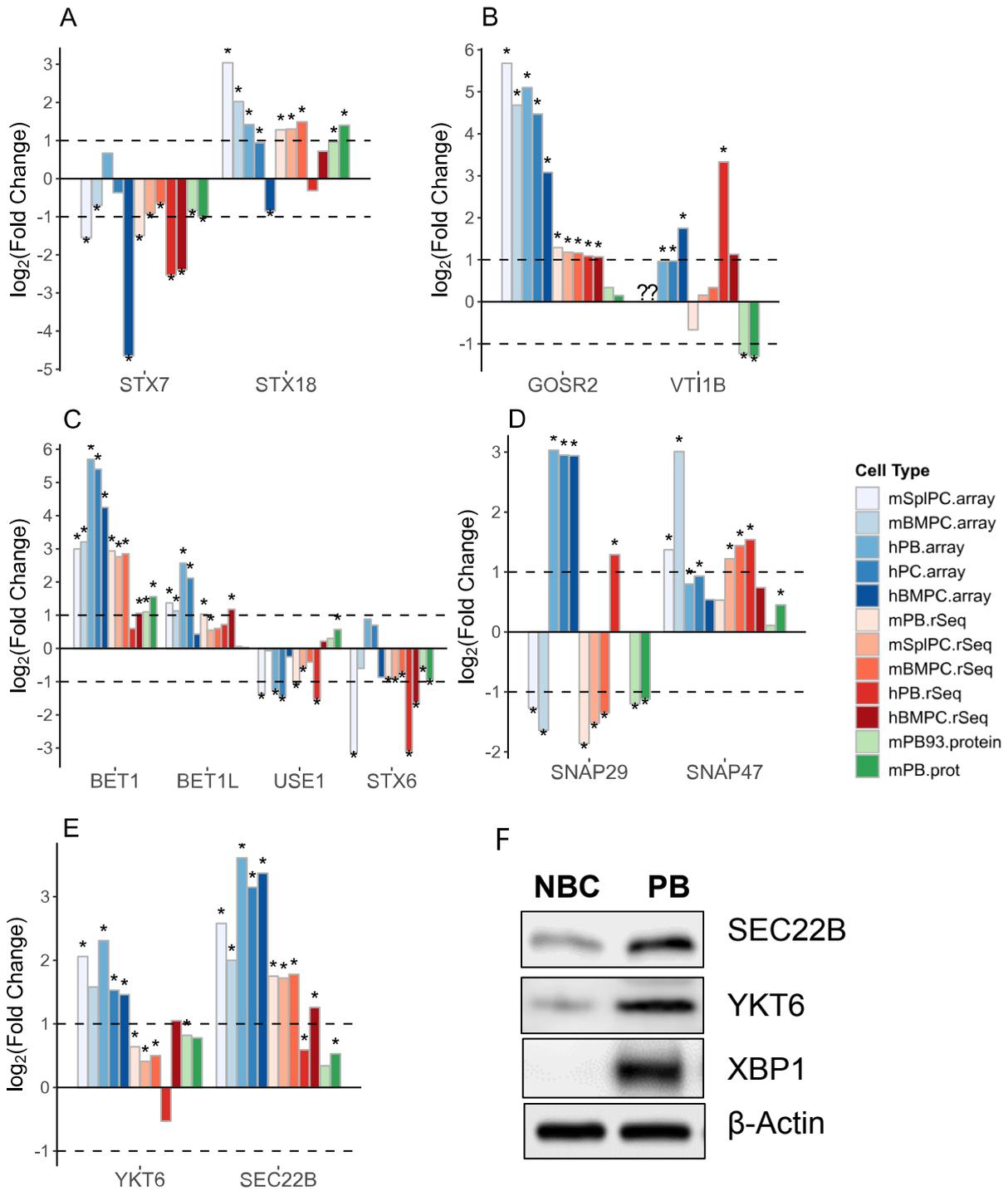
##### 4.4.6.3.1 ER-Golgi trafficking SNAREs show upregulation

###### *Retrograde transport from the Golgi to ER*

The Q SNAREs, syntaxin 18 (STX18), BNIP1 and USE1 each contribute a coiled-coil motif to make up a heterotrimeric t-SNARE. The NRZ complex traps incoming COPI vesicles by binding to the R-SNARE, SEC22B and subsequently promotes the formation of *trans*-SNARE complex at the ER [33]. In accordance with the upregulation in NRZ subunits, we observe consistent upregulation in STX18 and SEC22B in both gene and protein level (Fig 4-38A, E). Our colleagues have validated the expression of SEC22B in ASCs as shown in Fig 4-38F. As USE1 and BNIP show somewhat contradictory results (Fig 4-38C), overall, we predict that ASCs specifically enhance the activity of v-SNAREs participating in the retrograde transport from the Golgi to ER.

###### *Intra-Golgi Retrograde transport*

Retrograde transport of cargo within the Golgi apparatus is mediated by the v-SNARE YKT6 and t-SNAREs consisting of STX5, GOSR2 and BET1L [36]. Fig 4-38B, shows consistent upregulation in GOSR2 and BET1L gene, but neither showed induction in the protein level. YKT6 on the other hand show consistent upregulation in both the gene and protein level in



**Fig 4-38 | Differential Regulation of SNAREs (multi-omics).** **A.** Q<sub>a</sub>-SNAREs. Syntaxin 18 showed consistent upregulation in ASCs while Syntaxis 7 showed consistent downregulation. **B.** Q<sub>b</sub>-SNAREs. GOSR2 displayed consistent upregulation in gene level but little or no significant change in protein level whereas VTI1B showed statistically significant downregulation in protein level. **C.** Q<sub>c</sub>-SNAREs. BET1 was consistently upregulated in ASCs its isoform BET1L showed a similar pattern in gene level but showed no change in protein level. USE1 showed contradictory regulation in gene as opposed to protein level. **D.** Q<sub>bc</sub>-SNAREs. SNAP47 showed small but consistent upregulation in ASCs. SNAP29 showed over 2-fold downregulation in protein level **E.** R-SNAREs. YKT6 and SEC22B show statistically significant upregulation in nearly all ASC types. **F.** Western blot of select SNAREs.

ASCs and its expression has been validated using Western blot (Fig 4-38F). Overall, evidence points to a tendency for ASCs to enhance v-SNARE expression for intra-Golgi transport.

*Anterograde transport from ER to Golgi/ERGIC*

YKT6 or SEC22B typically acts as the v-SNARE for COPII vesicles [36]. In accordance with the remarkable upregulation in COPII coat proteins, these SNAREs were highly induced in ASCs as shown in (Fig 4-38C). At the Golgi apparatus and ERGIC, the Q-SNAREs, STX5, GOSR1/GOSR2 and BET1 constitute the t-SNAREs for incoming vesicles [33]. While GOSR1 showed no significant change, we have previously noted the tendency for SEC22B and YKT6 to be upregulated in ASCs. Based on this data, we hypothesize that v-SNAREs acting in the cis-Golgi are induced to promote ER to Golgi transport in ASCs.

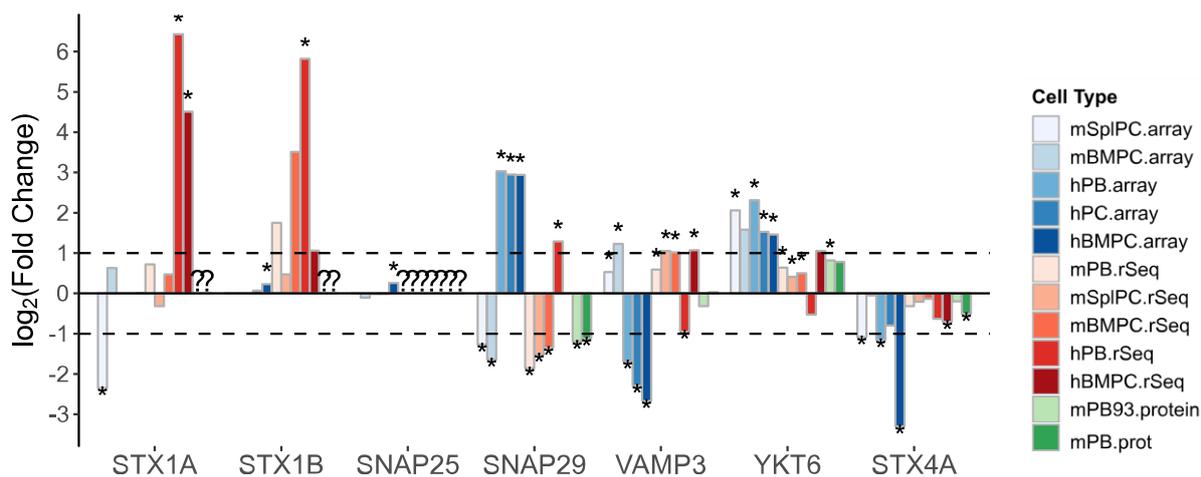
**4.4.6.3.2 Post Golgi SNAREs show no change or downregulation in ASCs**

We note that t-SNAREs subunits (STX7 and VTL1B) mediating fusion of cargo destined for lysosomal degradation showed downregulation in ASCs (Fig 4-38A, E) [40].

The Q-SNARE, Syntaxin 6 (STX6), contributes a coiled-coil motif to the *trans*-SNARE complex at the TGN for fusion of vesicles arriving from the cell surface via endosomes [39]. This gene was downregulated in ASCs as shown in Fig 4-38C. This result consolidates our hypothesis that antibodies may forego trafficking through endosomal compartments and take an alternate “shortcut” to the plasma membrane.

*SNAREs mediating fusion at the plasma membrane*

In this regard, we investigated SNAREs implicated in exocytosis from the *trans*-Golgi to the cell surface. The *trans*-SNARE complex made up of the heterodimeric t-SNAREs, STX1 and SNAP25, and the v-SNARE, VAMP2, is implicated in the spontaneous fusion of synaptic vesicles arriving at the plasma membrane from the TGN [43]. Other SNARE complexes consisting of a combination of the Q-SNAREs STX1, SNAP25/ SNAP29 and the R-SNARE VAMP3 or YKT6 have recently been reported in *Drosophila* [248]. However, the regulation



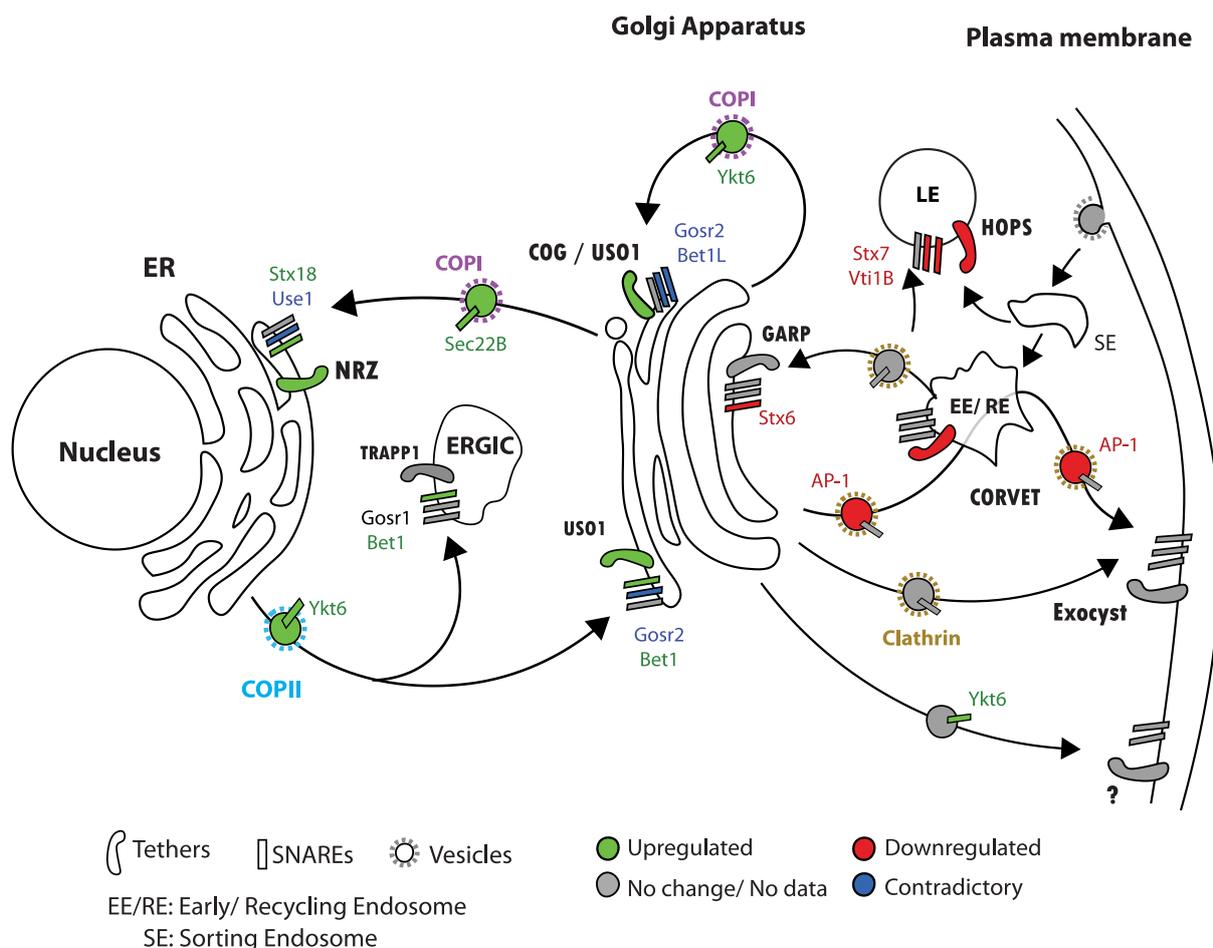
**Fig 4-39** | Differential Regulation of SNARE complex components known to play a role in exocytosis in neurons. Overall, these genes/proteins show inconsistent results, and in most cases, little or no change.

of STX1 isoforms and SNAP25 were inconclusive according to our data and most of these components other than YKT6, either showed a tendency for downregulation (SNAP29, STX4A) or no change (VAMP3) in the protein level (Fig 4-39). As such, we recommend alternative validation techniques such as Western blot of STX1A/B and SNAP25 to determine whether these genes may be upregulated in ASCs alongside YKT6.

#### 4.4.6.4 Summary of differentially regulated membrane trafficking genes

In Fig 4-40, we summarise the regulation of coated vesicles, tethers and SNARE proteins in ASCs. The machinery involved in the forward transport of cargo from the endoplasmic reticulum to the Golgi apparatus show overall upregulation in ASCs. This includes almost all COPII coat proteins, associated v-SNARE as well as the tether, USO1, known to mediate docking and fusion at the *cis*-Golgi.

In retrograde transport from Golgi to ER membrane, COPI coat proteins, their corresponding tether (NRZ) and v-SNAREs show upregulation in ASCs. Likewise, we also see the upregulation of intra-Golgi retrograde transport. Curiously, ASCs seem to preferably upregulate v-SNAREs rather than t-SNAREs in the early secretory pathway. The majority of known t-SNARE complexes do not show consistent upregulation. In fact, only the Q-SNAREs,



**Fig 4-40** | Summary of regulation and coordination of coated vesicles, tethers and SNAREs in ASCs according to multi-omics analysis. Overall, we see consistent upregulation in membrane trafficking components mediating anterograde and retrograde transport from the ER to Golgi as well as intra-Golgi retrograde transport. In contrast, components of post-Golgi transport showed little or no change, whereas tethering complexes and associated SNAREs showed a tendency for downregulation.

BET1 and STX18, acting in anterograde and retrograde transport between ER and Golgi apparatus, respectively, showed reproducible upregulation in ASCs. We, therefore, speculate that upregulation in YKT6 and SEC22B may simply reflect the increased number of COPI and COPII coated vesicles operating in ASCs versus NBCs. Therefore, we hypothesize that ASCs may enhance the number of transport carriers and specific docking machinery operating at the early secretory pathway to coordinate the forward transport of antibodies from the ER to Golgi and recycling of ER membrane components to match this forward momentum.

In contrast, we see little evidence of post-Golgi trafficking via endosomal compartments and clathrin coated vesicles in ASCs. Instead clathrin adaptors, SNAREs and tethers involved

in transport between *trans*-Golgi and plasma membrane via endosomes were downregulated. This is in accordance to previous studies that demonstrated that non-polarized cells likely utilise an alternate pathway for the delivery cargo from the TGN to the cell surface [245]. Based on the upregulation of Ykt6, we hypothesize that this R-SNARE may act as a v-SNARE for antibody containing vesicles acting in post-Golgi transport. However, what t-SNAREs may be mediating the fusion of these vesicles remains unclear. Interestingly, the Exocyst complex implicated in the tethering of vesicles in plasma membrane did not show any significant change in ASCs, suggesting that either the expression of Exocyst complex subunits in NBCs is sufficient to meet the increased secretory demand or perhaps different machinery acts in the tethering of antibody containing vesicles. However, what this machinery might be remains unclear.

## 4.5 CONCLUSION

In this project, we carry out cross species, cross platform comparative analysis of antibody secreting and non-secreting cells. We have demonstrated that combining microarray, RNA-Seq and tandem mass spectrometry data from two different species may help remove noisy hits such as PRG2 and EDEM1, and improve the coverage of differentially expressed genes, including those implicated in membrane trafficking. Cross-platform reproducibility allowed us to isolate genes that may be differentially regulated in mice as opposed to humans, such as UGGT2, and also components that are likely to be regulated on a post-transcriptional level (e.g. ribosomal subunits). Furthermore, by exploiting genes that reproducibly show remarkably high upregulation across platforms, species and in both gene and protein level we were able to identify potentially novel ASC markers (CD93, CRELD2, HID1). These markers may improve the purification and targeting of ASCs in research and medicine. As the less understood CRELD2 gene has previously shown specificity for misfolded protein substrate this gene may

also be of industrial relevance as its targeted overexpression in cell lines used in industry may potentially improve the secretion of biologics.

The primary aim of our project was to utilise our multi-omics bioresource to study membrane trafficking in the plasma cell physiology. According to our results, ASCs specifically increase the expression of genes/proteins related to membrane trafficking and the management of proteotoxic stress at the ER. We observe that ASCs induce the expression of vesicle coats and specific tethers acting in the early but not the late secretory pathway. Furthermore, a noticeably higher induction of specific vesicular coat proteins, COPZ2 and SEC24A/D, led us to speculate that these components may play a role in the sorting of large cargoes (antibody) within COPI and COPII vesicles, respectively. Overall, our data highlights the importance of ER-Golgi trafficking for antibody secretion and raises the important question as to what transport vesicles and tethering factor mediate post-Golgi trafficking of antibodies.

Interestingly, we note the marked induction of a number of genes/proteins that has been proposed to play a role in collagen secretion but has yet to be studied in antibody secretion. The potential parallel role of less characterised machinery, NRZ complex, CREB3L2, CRELD2, RRBP1 and SEC24D, in collagen as well as antibody secretion indicate that these proteins may be key components required for constitutive secretion of bulky cargo in general.

Through characterisation of membrane trafficking in the plasma cell physiology we have demonstrated the utility of our multi-omics analysis. In order to make this resource available to the wider scientific community, we have created a user-friendly web application that lets users visualise the differential regulation of a gene/protein in ASCs versus NBCs, across species, across platforms and in both gene and protein level. By visualising the reproducibility, or conversely, the noisiness of a gene, we predict that this bioresource can help bench biologists make an informed decision in deciding whether to carry out low throughput

validation of gene/protein and thus potentially save cost and time otherwise wasted on noisy hits such as PRG2.

## 4.6 EVALUATION

In this study, we primarily focus on what components are consistently regulated among ASCs in different stages of maturation. While this allows us to find ASC markers, we also discard data that are unique to specific stages of PC differentiation. For example, we noticed that components of cholesterol biosynthesis are enriched in the MS/MS data, where ASCs studied were all 3-day old plasmablasts. Revisiting the microarray and RNA-Seq profile confirm that they are upregulated in the PB transcriptome but not in splenic and bone marrow plasma cells. As such our voting system will have filtered these genes out. Thus, we realised that many genes important for the early stages of plasma cell differentiation will not be captured in our analysis. In order to study genes upregulated in specific stages of PC differentiation, we recommend that multi-omics analysis of PBs, SplPCs and BMPCs be done separately.

In Chapter 3, Fig 3-9, we give an example of kinase enrichment based on previous perturbation experiments. We did not utilise this data because it is difficult to gauge the importance of these kinases without phospho-proteomics data. We note that in addition to the proteomic expression profile of ASCs and NBCs, our colleagues have generated phospho-proteomics data for these cells. However, this was not in the scope of this project. While protein expression profile tells us how many molecules of a protein is present, phospho-proteomics can tell us how active a protein is likely to be in a sample. This can then be linked to the kinase perturbation database we highlighted in Fig 3-9 to pull out potentially important kinases regulating protein activity in ASCs. For future studies, we recommend incorporating this phospho-proteomics data into our bioresource in order to better understand the activity of proteins in the ASC phenotype.

We studied microarray, RNA-Seq and MS/MS profile of ASCs in this project and used a voting system to extract genes whose fold changes are reproducible across these platforms. Our experimental design was firstly based on consistency in regulation between sample types within each platform, and secondly the consistency in regulation between platforms. However, this caused some genes such as RINT1, a member of the NRZ tethering complex, to not appear in the final set of filtered genes despite 9 out of 12 ASC types studied showing statistically significant upregulation. We note that this was simply because in the human microarray, RINT1 showed downregulation. Therefore, in future studies we recommend gene filtering to only be done after pooling the data from all platforms together. Furthermore, we believe the voting system is not “fair”, as the accuracy and importance of data generated from RNA-Seq, microarray and proteomics are not equal. Therefore, alternative methods that consider the weight of distinct platforms could be used to improve our analysis. We note that our online bioresource has every gene found in all platforms and is unaffected by the voting criteria.

#### **4.6.1 Differential Expression calls**

In this study we combined p-values from different experiments by calculating a median, which is not the optimal way to determine the statistically significant of meta-analysis procedures. Here we review alternative approaches that can be used to improve our data and potentially remove the need for the voting system.

##### **4.6.1.1 *t*-statistic based Methods**

*t*-statistic or effect size-based methods utilise *t*-statistics to determine differentially expressed genes. The GeneMeta method, for example, introduced in 2003, assigns a *t*-like statistic to genes within each study to give them weight based on their variances, which are then fit to a hierarchical model to determine intra and cross study variation in these *t*-like statistics. Methods such as this do not take into account the magnitude of differential expression when estimating these effect sizes [249].

#### 4.6.1.2 Fold change-based methods

In 2006, a method called RankProd was introduced that utilises the magnitude of fold change instead of the  $t$ -statistics to determine differential expression. This approach uses a non-parametric approach to make differential expression calls. As such it makes fewer assumptions about the underlying distribution of the data and has been shown to be relatively robust for heterogenous datasets or those lacking sufficient replicates [250].  $T$ -statistic based approaches, discussed earlier, have been found to overestimate  $p$ -values and exhibit more false positives than RankProd. The latter has been found to outperform GeneMeta, especially when studies have fewer replicates or high cross-study variation [251]. Furthermore, the ranking of genes based on RankProd were more reproducible than GeneMeta [251].

#### 4.6.1.3 $p$ -value based methods

Fisher's inverse chi-square method is an easy to use method for combining  $p$ -values from individual datasets [252]. It sums the logarithm of all  $p$ -values and compares it against chi squared distribution. However, this method does not give an estimate of overall fold changes as it only utilises  $p$ -value statistics.  $p$ -value combination using this method has to be performed separately for up- and downregulated gene sets [252]. Furthermore, compared to newer methods this approach is highly dependent on the quality of the original  $p$ -values [251].

A newer  $p$ -value combination method, available in the metaMA package, solves the need to combine  $p$ -values separately for up- and downregulated genes by using a weighted method described by Marot and Mayer [253, 254]. This method was found to outperform both  $t$ -statistic based and fold change-based methods [254].

#### 4.6.1.4 Combining $t$ -statistic and fold change

A novel method utilising both  $t$ -statistic based and fold change ratios has recently become available. This tool, called iMeta, quantile normalises fold changes against  $t$ -statistic to minimise biases arising from differences in distribution. Benchmarking studies show that this

method outperforms other existing methods for determining differential expression in cross study meta-analysis [255].

#### 4.6.2 Alternative meta-analysis method

We describe merging and meta-analysis in Section 2.1.3, which have also been referred to as “early” and “intermediate” merging, respectively [256]. Recently another approach has been highlighted where differential expression and downstream gene set enrichment analysis (GSEA) is performed first and then results from the GSEA is merged by combining  $p$ -values from the GSEA enrichment. This is known as “late” merging [256]. While early merging provided the best results, interestingly, late merging of GSEA enrichment results yield better results than intermediate merging, which we employed for cross-species and cross platform analysis in this study [256]. Note that  $p$ -value combination was carried out using the weighted inverse method available in the MetaMA package[256]. Using this meta-analysis method with EnrichR for future analysis is thus worth consideration.

### 4.7 FUTURE DIRECTIONS

The protein level expression of a number of genes identified in this project has been validated by our colleagues using Western blot and immunofluorescence. Ideally, the next step would have been to knockdown or overexpress these genes in naïve B cells and assess their antibody secreting capacity after activation. Unfortunately, this has proved to be problematic and the results have been inconsistent.

The monoclonal B cell lymphoma, i.29, can be activated by LPS to differentiate into antibody secreting plasma cells and plasma blasts [257]. As transfection of these i.29 cell lines tend to show better results [258], our colleagues intend to utilise i.29 for characterising the candidate genes identified in this study. However, this cancer cell line is not a physiological model of ASCs, therefore, it would be ideal to map the transcriptome differences in ASCs and

i.29 and test their viability as a model for ASCs before attempting the perturbation of candidate genes/proteins identified in this project.

**Word Count: 33,990**

## 5 REFERENCES

1. Alberts Bruce, Johnson Alexander, Lewis Julian, Raff Martin, Roberts Keith WP. *Molecular Biology of the Cell*, 4th Edition. 2002.
2. Palade G. Intracellular aspects of the process of protein synthesis. *Science*. 1975;189:347–58.
3. Schubert U, Antón LC, Gibbs J, Norbury CC, Yewdell JW, Bennink JR. Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature*. 2000;404:770–4. doi:10.1038/35008096.
4. Erbaykent Tepedelen B, Ballar Kirmizibayrak P. Endoplasmic Reticulum-Associated Degradation (ERAD). In: *Endoplasmic Reticulum*. IntechOpen; 2019. doi:10.5772/intechopen.82043.
5. Liu CY, Kaufman RJ. The unfolded protein response. *J Cell Sci*. 2003;116 Pt 10:1861–2. doi:10.1242/jcs.00408.
6. Sano R, Reed JC. ER stress-induced cell death mechanisms. *Biochim Biophys Acta - Mol Cell Res*. 2013;1833:3460–70. doi:10.1016/J.BBAMCR.2013.06.028.
7. Kim I, Xu W, Reed JC. Cell death and endoplasmic reticulum stress: disease relevance and therapeutic opportunities. *Nat Rev Drug Discov*. 2008;7:1013–30. doi:10.1038/nrd2755.
8. Christianson JC, Ye Y. Cleaning up in the endoplasmic reticulum: ubiquitin in charge. *Nat Struct Mol Biol*. 2014;21:325–35. doi:10.1038/nsmb.2793.
9. Araki K, Nagata K. Protein folding and quality control in the ER. *Cold Spring Harb Perspect Biol*. 2012;4:a015438. doi:10.1101/cshperspect.a015438.
10. Guerriero CJ, Brodsky JL. The Delicate Balance Between Secreted Protein Folding and Endoplasmic Reticulum-Associated Degradation in Human Physiology. *Physiol Rev*. 2012;92:537–76. doi:10.1152/physrev.00027.2011.
11. Vembar SS, Brodsky JL. One step at a time: endoplasmic reticulum-associated degradation. *Nat Rev Mol Cell Biol*. 2008;9:944–57. doi:10.1038/nrm2546.
12. Brodsky JL, Werner ED, Dubas ME, Goekeler JL, Kruse KB, McCracken AA. The requirement for molecular chaperones during endoplasmic reticulum-associated protein degradation demonstrates that protein export and import are mechanistically distinct. *J Biol Chem*. 1999;274:3453–60. doi:10.1074/jbc.274.6.3453.
13. Vashist S, Ng DTW. Misfolded proteins are sorted by a sequential checkpoint mechanism of ER quality control. *J Cell Biol*. 2004;165:41–52. doi:10.1083/JCB.200309132.
14. Hosokawa N, Wada I, Hasegawa K, Yorihuzi T, Tremblay LO, Herscovics A, et al. A novel ER - mannosidase-like protein accelerates ER-associated degradation. *EMBO Rep*. 2001;2:415–22. doi:10.1093/embo-reports/kve084.
15. Oda Y, Hosokawa N, Wada I, Nagata K. EDEM as an acceptor of terminally misfolded glycoproteins released from calnexin. *Science*. 2003;299:1394–7. doi:10.1126/science.1079181.
16. Molinari M, Calanca V, Galli C, Lucca P, Paganetti P. Role of EDEM in the release of misfolded glycoproteins from the calnexin cycle. *Science*. 2003;299:1397–400. doi:10.1126/science.1079474.
17. Termine DJ, Moremen KW, Sifers RN. The mammalian UPR boosts glycoprotein ERAD by suppressing the proteolytic downregulation of ER mannosidase I. *J Cell Sci*. 2009;122:976–84. doi:10.1242/JCS.037291.
18. Stolz A, Wolf DH. Endoplasmic reticulum associated protein degradation: A chaperone assisted journey to hell. *Biochim Biophys Acta - Mol Cell Res*. 2010;1803:694–705. doi:10.1016/J.BBAMCR.2010.02.005.
19. Kanehara K, Kawaguchi S, Ng DTW. The EDEM and Yos9p families of lectin-like ERAD factors. *Semin Cell Dev Biol*. 2007;18:743–50. doi:10.1016/J.SEMCDB.2007.09.007.

20. Słomińska-Wojewódzka M, Sandvig K. The Role of Lectin-Carbohydrate Interactions in the Regulation of ER-Associated Protein Degradation. *Molecules*. 2015;20:9816–46. doi:10.3390/molecules20069816.
21. Ushioda R, Hoseki J, Araki K, Jansen G, Thomas DY, Nagata K. ERdj5 is required as a disulfide reductase for degradation of misfolded proteins in the ER. *Science*. 2008;321:569–72. doi:10.1126/science.1159293.
22. Helenius A, Aebi M. Roles of N-Linked Glycans in the Endoplasmic Reticulum. *Annu Rev Biochem*. 2004;73:1019–49. doi:10.1146/annurev.biochem.73.011303.073752.
23. Jensen D, Schekman R. COPII-mediated vesicle formation at a glance. *Journal of Cell Science*. 2011;124:1–4.
24. Miller E, Antony B, Hamamoto S, Schekman R. Cargo selection into COPII vesicles is driven by the Sec24p subunit. *EMBO J*. 2002;21:6105–13. doi:10.1093/emboj/cdf605.
25. Arakel EC, Schwappach B. Formation of COPI-coated vesicles at a glance. 2018. doi:10.1242/209890.
26. Robinson MS. Forty Years of Clathrin-coated Vesicles. *Traffic*. 2015;16:1210–38. doi:10.1111/tra.12335.
27. Royle SJ. The cellular functions of clathrin. *Cell Mol Life Sci*. 2006;63:1823–32. doi:10.1007/s00018-005-5587-0.
28. Park SY, Guo X. Adaptor protein complexes and intracellular transport. *Biosci Rep*. 2014;34. doi:10.1042/BSR20140069.
29. Chia PZC, Gleeson PA. Membrane tethering. *F1000Prime Reports*. 2014;6.
30. Yamaoka M, Ando T, Terabayashi T, Okamoto M, Takei M, Nishioka T, et al. PI3K regulates endocytosis after insulin secretion by mediating signaling crosstalk between Arf6 and Rab27a. *J Cell Sci*. 2016;129:637 LP – 649. doi:10.1242/jcs.180141.
31. Smith RD, Lupashin V V. Role of the conserved oligomeric Golgi (COG) complex in protein glycosylation. *Carbohydr Res*. 2008;343:2024–31. doi:10.1016/j.carres.2008.01.034.
32. Sztul E. Tethering factor P115: a new model for tether-SNARE interactions. *Bioarchitecture*. 2012;2:175–80.
33. Tagaya M, Arasaki K, Inoue H, Kimura H. Moonlighting functions of the NRZ (mammalian Dsl1) complex. *Front Cell Dev Biol*. 2014;2 June:1–9. doi:10.3389/fcell.2014.00025.
34. Balderhaar HJ, Ungermann C. CORVET and HOPS tethering complexes – coordinators of endosome and lysosome fusion. *J Cell Sci*. 2013;126:1307 LP – 1316. doi:10.1242/jcs.107805.
35. Gautreau A, Oguievetskaia K, Ungermann C. Function and regulation of the endosomal fusion and fission machineries. *Cold Spring Harb Perspect Biol*. 2014;6.
36. Daste F, Galli T, Taresté D. Structure and function of longin SNAREs. *J Cell Sci*. 2015;128:4263 LP – 4272. doi:10.1242/jcs.178574.
37. Kloepper TH, Kienle CN, Fasshauer D. An elaborate classification of SNARE proteins sheds light on the conservation of the eukaryotic endomembrane system. *Mol Biol Cell*. 2007;18:3463–71. doi:10.1091/mbc.e07-03-0193.
38. Wang T, Li L, Hong W. SNARE proteins in membrane trafficking. *Traffic*. 2017;18:767–75. doi:10.1111/tra.12524.
39. Tran THT, Zeng Q, Hong W. VAMP4 cycles from the cell surface to the trans-Golgi network via sorting and recycling endosomes. *J Cell Sci*. 2007;120:1028 LP – 1041. doi:10.1242/jcs.03387.
40. Luzio JP, Gray SR, Bright NA. Endosome–lysosome fusion. *Biochem Soc Trans*. 2010;38:1413–6. doi:10.1042/BST0381413.
41. Spessott WA, Sanmillan ML, Kulkarni V V, McCormick ME, Giraudo CG. Syntaxin 4 mediates endosome recycling for lytic granule exocytosis in cytotoxic T-lymphocytes. *Traffic*. 2017;18:442–52. doi:10.1111/tra.12490.

42. Bruce Alberts. *Molecular Biology of Cells*. 6th edition. 1998.
43. Ulloa F, Cotrufo T, Ricolo D, Soriano E, Araújo SJ. SNARE complex in axonal guidance and neuroregeneration. *Neural Regen Res*. 2018;13:386–92. doi:10.4103/1673-5374.228710.
44. Liang T, Qin T, Xie L, Dolai S, Zhu D, Prentice KJ, et al. New Roles of Syntaxin-1A in Insulin Granule Exocytosis and Replenishment. *J Biol Chem*. 2017;292:2203–16. doi:10.1074/jbc.M116.769885.
45. Zhu D, Koo E, Kwan E, Kang Y, Park S, Xie H, et al. Syntaxin-3 regulates newcomer insulin granule exocytosis and compound fusion in pancreatic beta cells. *Diabetologia*. 2013;56:359–69.
46. Tartakoff A, Vassalli P, Détraz M. Comparative studies of intracellular transport of secretory proteins. *J Cell Biol*. 1978;79:694–707. doi:10.1083/jcb.79.3.694.
47. Kantele J, Kantele A, Arvilommi H. Circulating immunoglobulin-secreting cells are heterogeneous in their expression of maturation markers and homing receptors. *Clin Exp Immunol*. 1996;104:525–30. [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list\\_uids=97254529](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list_uids=97254529).
48. Allman D, Pillai S. Peripheral B cell subsets. *Current Opinion in Immunology*. 2008;20:149–57.
49. Bannard O, Cyster JG. Germinal centers: programmed for affinity maturation and antibody diversification. *Current Opinion in Immunology*. 2017;45:21–30. doi:10.1016/j.coi.2016.12.004.
50. Bortnick A, Chernova I, Quinn WJ, Mugnier M, Cancro MP, Allman D. Long-Lived Bone Marrow Plasma Cells Are Induced Early in Response to T Cell-Independent or T Cell-Dependent Antigens. *J Immunol*. 2012;188:5389–96. doi:10.4049/jimmunol.1102808.
51. Hibi T, Dosch H-M. Limiting dilution analysis of the B cell compartment in human bone marrow. *Eur J Immunol*. 1986;16:139–45. doi:10.1002/eji.1830160206.
52. Werner-Favre C, Matthes T, Barnet M, Zubler RH. High IgE secretion capacity of human plasma cells. *Eur J Immunol*. 1993;23:2038–40. doi:10.1002/eji.1830230849.
53. Bromage E, Stephens R, Hassoun L. The third dimension of ELISPOTs: Quantifying antibody secretion from individual plasma cells. *J Immunol Methods*. 2009;346:75–9. doi:10.1016/j.jim.2009.05.005.
54. Kirk SJ, Cliff JM, Thomas JA, Ward TH. Biogenesis of secretory organelles during B cell differentiation. *J Leukoc Biol*. 2010;87:245–55. doi:10.1189/jlb.1208774.
55. Jourdan M, Caraux A, De Vos J, Fiol G, Larroque M, Cognot C, et al. An in vitro model of differentiation of memory B cells into plasmablasts and plasma cells including detailed phenotypic and molecular characterization. *Blood*. 2009;114:5173–81. doi:10.1182/blood-2009-07-235960.
56. Shi W, Liao Y, Willis SN, Taubenheim N, Inouye M, Tarlinton DM, et al. Transcriptional profiling of mouse B cell terminal differentiation defines a signature for antibody-secreting plasma cells. *Nat Immunol*. 2015;16:663–73. doi:10.1038/ni.3154.
57. Kaji T, Ishige A, Hikida M, Taka J, Hijikata A, Kubo M, et al. Distinct cellular pathways select germline-encoded and somatically mutated antibodies into immunological memory. *J Exp Med*. 2012;209:2079–97. doi:10.1084/jem.20120127.
58. Barwick BG, Scharer CD, Bally APR, Jeremy M. hypomethylation and gene regulation. 2017;17:1216–25. doi:10.1038/ni.3519.Plasma.
59. Shi L, Reid LH, Jones WD, Shippy R, Warrington JA, Baker SC, et al. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol*. 2006;24:1151–61.
60. Su Z, Labaj PP, Li S, Thierry-Mieg J, Thierry-Mieg D, Shi W, et al. A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. *Nat Biotechnol*. 2014;32:903–14.
61. Zhao S, Fung-Leung WP, Bittner A, Ngo K, Liu X. Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. *PLoS One*. 2014;9. doi:10.1371/journal.pone.0078644.

62. Etienne W, Meyer MH, Peppers J, Meyer RA. Comparison of mRNA gene expression by RT-PCR and DNA microarray. *Biotechniques*. 2004;36:618–26.
63. Wu C, Huang BE, Chen G, Lovenberg TW, Pocalyko DJ, Yao X. Integrative Analysis of DiseaseLand Omics Database for Disease Signatures and Treatments: A Bipolar Case Study. *Front Genet*. 2019;10:396. doi:10.3389/fgene.2019.00396.
64. Keel BN, Zarek CM, Keele JW, Kuehn LA, Snelling WM, Oliver WT, et al. RNA-Seq Meta-analysis identifies genes in skeletal muscle associated with gain and intake across a multi-season study of crossbred beef steers. *BMC Genomics*. 2018;19:430. doi:10.1186/s12864-018-4769-8.
65. Feder ME, Mitchell-Olds T. Evolutionary and ecological functional genomics. *Nat Rev Genet*. 2003;4:649–55. doi:10.1038/nrg1128.
66. Edfors F, Danielsson F, Hallström BM, Käll L, Lundberg E, Pontén F, et al. Gene-specific correlation of RNA and protein levels in human cells and tissues. *Mol Syst Biol*. 2016;12:883. doi:10.15252/msb.20167144.
67. Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J, et al. Deep proteome and transcriptome mapping of a human cancer cell line. *Mol Syst Biol*. 2011;7:548. doi:10.1038/MSB.2011.81.
68. Lundberg E, Fagerberg L, Klevebring D, Matic I, Geiger T, Cox J, et al. Defining the transcriptome and proteome in three functionally different human cell lines. *Mol Syst Biol*. 2010;6:450. doi:10.1038/msb.2010.106.
69. Romijn EP, Christis C, Wieffer M, Gouw JW, Fullaondo A, van der Sluijs P, et al. Expression clustering reveals detailed co-expression patterns of functionally related proteins during B cell differentiation: a proteomic study using a combination of one-dimensional gel electrophoresis, LC-MS/MS, and stable isotope labeling by amino acids in cell culture (SILAC). *Mol Cell Proteomics*. 2005;4:1297–310. doi:10.1074/mcp.M500123-MCP200.
70. van Anken E, Pena F, Hafkemeijer N, Christis C, Romijn EP, Grauschopf U, et al. Efficient IgM assembly and secretion require the plasma cell induced endoplasmic reticulum protein pERp1. *Proc Natl Acad Sci U S A*. 2009;106:17019–24. doi:10.1073/pnas.0903036106.
71. Kassambara A, Rème T, Jourdan M, Fest T, Hose D, Tarte K, et al. GenomicScape: An Easy-to-Use Web Tool for Gene Expression Data Analysis. Application to Investigate the Molecular Events in the Differentiation of B Cells into Plasma Cells. *PLoS Comput Biol*. 2015;11:1–10.
72. Le Carrouer T, Assou S, Tondeur S, Lhermitte L, Lamb N, Reme T, et al. Amazonia!: An Online Resource to Google and Visualize Public Human whole Genome Expression Data. *The Open Bioinformatics Journal*. 2010;;5–10.
73. Bard F, Casano L, Mallabiabarrena A, Wallace E, Saito K, Kitayama H, et al. Functional genomics reveals genes involved in protein secretion and Golgi organization. *Nature*. 2006;439:604–7. doi:10.1038/nature04377.
74. Simpson JC, Joggerst B, Laketa V, Verissimo F, Cetin C, Erfle H, et al. Genome-wide RNAi screening identifies human proteins with a regulatory function in the early secretory pathway. *Nat Cell Biol*. 2012;14:764–74. doi:10.1038/ncb2510.
75. Xiao S, Chen YC, Buehler E, Mandal S, Mandal A, Betenbaugh M, et al. Genome-scale RNA interference screen identifies antizyme 1 (OAZ1) as a target for improvement of recombinant protein production in mammalian cells. *Biotechnol Bioeng*. 2016;113:2403–15. doi:10.1002/bit.26017.
76. Jego G, Robillard N, Puthier D, Amiot M, Accard F, Pineau D, et al. Reactive plasmacytoses are expansions of plasmablasts retaining the capacity to differentiate into plasma cells. *Blood*. 1999;94:701–12.
77. Underhill GH, Kolli KP, Kansas GS. Complexity within the plasma cell compartment of mice deficient in both E- and P-selectin: implications for plasma cell differentiation. *Blood*. 2003;102:4076–83. <http://www.bloodjournal.org/content/102/12/4076.abstract>.
78. RK B, Balabanov R, Han X, Al E. Association of nonmyeloablative hematopoietic stem cell transplantation with neurological disability in patients with relapsing-remitting multiple sclerosis.

- JAMA. 2015;313:275–84. <http://dx.doi.org/10.1001/jama.2014.17986>.
79. Grillo-Lopez AJ. Rituximab: an insider's historical perspective. *Semin Oncol*. 2000;27 6 Suppl 12:9–16.
80. Pettersson M, Sundstrom C, Nilsson K, Larsson LG. The hematopoietic transcription factor PU.1 is downregulated in human multiple myeloma cell lines. *Blood*. 1995;86:2747–53. <http://www.bloodjournal.org/content/86/7/2747.abstract>.
81. Biologics Global Market Opportunities And Strategies To 2021. <https://www.reportlinker.com/p05482361/Biologics-Global-Market-Opportunities-And-Strategies-To.html>. Accessed 27 May 2019.
82. Kim JY, Kim Y-GG, Lee GM. CHO cells in biotechnology for production of recombinant proteins: current state and further potential. *Appl Microbiol Biotechnol*. 2012;93:917–30. doi:10.1007/s00253-011-3758-5.
83. Walsh G. Biopharmaceutical benchmarks 2018. *Nat Biotechnol*. 2018;36:1136–45. doi:10.1038/nbt.4305.
84. Krämer O, Klausung S, Noll T. Methods in mammalian cell line engineering: from random mutagenesis to sequence-specific approaches. *Appl Microbiol Biotechnol*. 2010;88:425–36. doi:10.1007/s00253-010-2798-6.
85. Todd DJ, McHeyzer-Williams LJ, Kowal C, Lee A-H, Volpe BT, Diamond B, et al. XBP1 governs late events in plasma cell differentiation and is not required for antigen-specific memory B cell development. *J Exp Med*. 2009;206:2151–9. doi:10.1084/jem.20090738.
86. Fischer S, Handrick R, Otte K. The art of CHO cell engineering: A comprehensive retrospect and future perspectives. *Biotechnol Adv*. 2015;33:1878–96. doi:10.1016/j.biotechadv.2015.10.015.
87. Manz RA, Thiel A, Radbruch A. Lifetime of plasma cells in the bone marrow. *Nature*. 1997;388:133–4.
88. Reinhart D, Damjanovic L, Kaisermayer C, Kunert R. Benchmarking of commercially available CHO cell culture media for antibody production. *Appl Microbiol Biotechnol*. 2015;99:4645–57. doi:10.1007/s00253-015-6514-4.
89. Nutt SL, Hodgkin PD, Tarlinton DM, Corcoran LM. The generation of antibody-secreting plasma cells. *Nat Rev Immunol*. 2015;15:160–71. doi:10.1038/nri3795.
90. Wols HAM, College B, Forest L. Plasma Cells. *Encycl Life Sci*. 2001;:1–8. doi:10.1038/npg.els.0004030.
91. Beißbarth T. Microarray: Quality Control, Normalization and Design. 2005. [http://compdiag.molgen.mpg.de/ngfn/docs/2005/sep/beissbarth\\_cDNA\\_QCPP.pdf](http://compdiag.molgen.mpg.de/ngfn/docs/2005/sep/beissbarth_cDNA_QCPP.pdf). Accessed 1 May 2019.
92. Lee MLT, Kuo FC, Whitmore GA, Sklar J. Importance of replication in microarray gene expression studies: Statistical methods and evidence from repetitive cDNA hybridizations. *Proc Natl Acad Sci U S A*. 2000;97:9834–9.
93. Bryant PA, Smyth GK, Robins-Browne R, Curtis N. Technical Variability Is Greater than Biological Variability in a Microarray Experiment but Both Are Outweighed by Changes Induced by Stimulation. *PLoS One*. 2011;6:e19556. doi:10.1371/journal.pone.0019556.
94. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, et al. A survey of best practices for RNA-seq data analysis. *Genome Biol*. 2016;17:13. doi:10.1186/s13059-016-0881-8.
95. Taminau J, Lazar C, Meganck S, Nowé A. Comparison of Merging and Meta-Analysis as Alternative Approaches for Integrative Gene Expression Analysis. *ISRN Bioinforma*. 2014;2014:1–7. doi:10.1155/2014/345106.
96. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostat*. 2007;8:118–27. doi:10.1093/biostatistics/kxj037.
97. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression

- analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43:e47.
98. Lifemap Sciences, Weizermann Institute of Science. GeneAlaCart. 2016.
99. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4:44–57.
100. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 2009;37:1–13.
101. Binder JX, Pletscher-Frankild S, Tsafou K, Stolte C, O'Donoghue SI, Schneider R, et al. COMPARTMENTS: unification and visualization of protein subcellular localization evidence. *Database* . 2014;2014. doi:10.1093/database/bau012.
102. Hu Y, Flockhart I, Vinayagam A, Bergwitz C, Berger B, Perrimon N, et al. An integrative approach to ortholog prediction for disease-focused and other functional studies. *BMC Bioinformatics.* 2011;12:357.
103. Hummer BH, de Leeuw NF, Burns C, Chen L, Joens MS, Hosford B, et al. HID-1 controls formation of large dense core vesicles by influencing cargo sorting and trans-Golgi network acidification. *Mol Biol Cell.* 2017;28:3870–80. doi:10.1091/mbc.E17-08-0491.
104. Popken-Harris P, Checkel J, Loegering D, Madden B, Springett M, Kephart G, et al. Regulation and Processing of a Precursor Form of Eosinophil Granule Major Basic Protein (ProMBP) in Differentiating Eosinophils. *Blood.* 1998;92:623–31. <http://www.bloodjournal.org/content/92/2/623.abstract>.
105. Kelly RB. Pathways of protein secretion in eukaryotes. *Science* (80- ). 1985;230:25–32.
106. Moore H-PH. Factors Controlling Packaging of Peptide Hormones into Secretory Granules. *Ann N Y Acad Sci.* 1987;493:50–61. doi:10.1111/j.1749-6632.1987.tb27180.x.
107. Capoccia BJ, Lennerz JKM, Bredemeyer AJ, Klco JM, Frater JL, Mills JC. Transcription factor MIST1 in terminal differentiation of mouse and human plasma cells. *Physiol Genomics.* 2011;43:174–86. doi:10.1152/physiolgenomics.00084.2010.
108. Pin CL, Bonvissuto AC, Konieczny SF. Mist1 expression is a common link among serous exocrine cells exhibiting regulated exocytosis. *Anat Rec.* 2000;259:157–67.
109. Abu-Ghazaleh RI, Gleich GJ, Prendergast FG. Interaction of eosinophil granule major basic protein with synthetic lipid bilayers: a mechanism for toxicity. *J Membr Biol.* 1992;128:153–64.
110. Giembycz MA, Lindsay MA. Pharmacology of the eosinophil. *Pharmacol Rev.* 1999;51:213–340.
111. Nakatsu F, Ohno H. Adaptor protein complexes as the key regulators of protein sorting in the post-Golgi network. *Cell Struct Funct.* 2003;28:419–29.
112. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics.* 2013;14:128.
113. Kuleshov M V, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* 2016;44:W90-7.
114. Rosenthal N, Brown S. The mouse ascending: perspectives for human-disease models. *Nat Cell Biol.* 2007;9:993–9. doi:10.1038/ncb437.
115. Makalowski W, Zhang J, Boguski MS. Comparative analysis of 1196 orthologous mouse and human full-length mRNA and protein sequences. *Genome Res.* 1996;6:846–57. <http://www.ncbi.nlm.nih.gov/pubmed/8889551>. Accessed 3 Jun 2019.
116. Mural RJ, Adams MD, Myers EW, Smith HO, Miklos GLG, Wides R, et al. A Comparison of Whole-Genome Shotgun-Derived Mouse Chromosome 16 and the Human Genome. *Science* (80- ). 2002;296:1661–71. doi:10.1126/science.1069193.
117. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res.* 2003;31:e15–e15. doi:10.1093/nar/gng015.
118. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Speed TP. Exploration,

- normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*. 2003;4:249–64. doi:10.1093/biostatistics/4.2.249.
119. Wu Z, Irizarry RA, Gentleman R, Martinez-Murillo F, Spencer F. A Model-Based Background Adjustment for Oligonucleotide Expression Arrays. *J Am Stat Assoc*. 2004;99:909–17. doi:10.1198/016214504000000683.
120. Mosteller F. *Data analysis and regression : a second course in statistics*. Reading, Mass.: Reading, Mass.; 1977.
121. Robson S. Probe-Level Data Normalisation: RMA and GC-RMA. 2005. [https://warwick.ac.uk/fac/sci/moac/people/students/2003/sam\\_robson/usergroups/rmavsmas5/](https://warwick.ac.uk/fac/sci/moac/people/students/2003/sam_robson/usergroups/rmavsmas5/). Accessed 15 Jul 2019.
122. Lazar C, Meganck S, Taminau J, Steenhoff D, Coletta A, Molter C, et al. Batch effect removal methods for microarray gene expression data integration: a survey. *Brief Bioinform*. 2013;14:469–90. doi:10.1093/bib/bbs037.
123. Nygaard V, Rødland EA, Hovig E. Methods that remove batch effects while retaining group differences may lead to exaggerated confidence in downstream analyses. *Biostatistics*. 2016;17:29–39.
124. Liu H, Zeeberg BR, Qu G, Koru AG, Ferrucci A, Kahn A, et al. AffyProbeMiner: A web resource for computing or retrieving accurately redefined Affymetrix probe sets. *Bioinformatics*. 2007;23:2385–90. doi:10.1093/bioinformatics/btm360.
125. Arloth J, Bader DM, Röh S, Altmann A. Re-Annotator: Annotation pipeline for microarray probe sequences. *PLoS One*. 2015;10:1–13. doi:10.1371/journal.pone.0139516.
126. Barbosa-Morais NL, Dunning MJ, Samarajiwa SA, Darot JFJ, Ritchie ME, Lynch AG, et al. A re-annotation pipeline for Illumina BeadArrays: Improving the interpretation of gene expression data. *Nucleic Acids Res*. 2009;38.
127. Saka E, Harrison BJ, West K, Petruska JC, Rouchka EC. Framework for reanalysis of publicly available Affymetrix® GeneChip® data sets based on functional regions of interest. *BMC Genomics*. 2017;18:875. doi:10.1186/s12864-017-4266-5.
128. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10. doi:10.1186/gb-2009-10-3-r25.
129. Luckey CJ, Bhattacharya D, Goldrath AW, Weissman IL, Benoist C, Mathis D. Memory T and memory B cells share a transcriptional program of self-renewal with long-term hematopoietic stem cells. *Proc Natl Acad Sci*. 2006;103:3304–9. doi:10.1073/pnas.0511137103.
130. Benson MJ, Aijo T, Chang X, Gagnon J, Pape UJ, Anantharaman V, et al. Heterogeneous nuclear ribonucleoprotein L-like (hnRNPLL) and elongation factor, RNA polymerase II, 2 (ELL2) are regulators of mRNA processing in plasma cells. *Proc Natl Acad Sci*. 2012;109:16252–7. doi:10.1073/pnas.1214414109.
131. Jensen GS, Andrews EJ, Mant MJ, Vergidis R, Ledbetter JA, Pilarski LM. Transitions in CD45 isoform expression indicate continuous differentiation of a monoclonal CD5+ CD11b+ B lineage in Waldenstrom’s macroglobulinemia. *Am J Hematol*. 1991;37:20–30. <http://www.ncbi.nlm.nih.gov/pubmed/1708944>. Accessed 3 Jun 2019.
132. Chung JB, Sater RA, Fields ML, Erikson J, Monroe JG. CD23 defines two distinct subsets of immature B cells which differ in their responses to T cell help signals. *Int Immunol*. 2002;14:157–66. doi:10.1093/intimm/14.2.157.
133. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1995;57:289–300. doi:10.2307/2346101.
134. Mudge JM, Harrow J. Creating reference gene annotation for the mouse C57BL6/J genome assembly. *Mamm Genome*. 2015;26:366–78. doi:10.1007/s00335-015-9583-x.
135. Jourdan M, Caraux A, Caron G, Robert N, Fiol G, Rème T, et al. Characterization of a transitional preplasmablast population in the process of human B cell to plasma cell differentiation. *J Immunol*.

- 2011;187:3931–41. doi:10.4049/jimmunol.1101230.
136. Mahtouk K, Moreaux J, Hose D, Rème T, Meissner T, Jourdan M, et al. Growth factors in multiple myeloma: a comprehensive analysis of their expression in tumor cells and bone marrow environment using Affymetrix microarrays. *BMC Cancer*. 2010;10:198. doi:10.1186/1471-2407-10-198.
137. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PLoS One*. 2011;6:e21800. doi:10.1371/journal.pone.0021800.
138. Resnik P. Semantic Similarity in a Taxonomy: An Information-Based Measure and its Application to Problems of Ambiguity in Natural Language. *J Artif Intell Res*. 1999. doi:10.1613/jair.514.
139. Lin D. An Information-Theoretic Definition of Similarity. In: *ICML*. 1998.
140. Rearick D, Prakash A, McSweeney A, Shepard SS, Fedorova L, Fedorov A. Critical association of ncRNA with introns. *Nucleic Acids Res*. 2011;39:2357–66. doi:10.1093/nar/gkq1080.
141. Pang Y, Mao C, Liu S. Encoding activities of non-coding RNAs. *Theranostics*. 2018;8:2496–507. doi:10.7150/thno.24677.
142. Nutt SL, Taubenheim N, Hasbold J, Corcoran LM, Hodgkin PD. The genetic network controlling plasma cell differentiation. *Semin Immunol*. 2011;23:341–9. doi:https://doi.org/10.1016/j.smim.2011.08.010.
143. Dengler HS, Baracho G V, Omori SA, Bruckner S, Arden KC, Castrillon DH, et al. Distinct functions for the transcription factor Foxo1 at various stages of B cell differentiation. *Nat Immunol*. 2008;9:1388–98. doi:10.1038/ni.1667.
144. Yeung CCS, Mills JC, Hassan A, Kreisel FH, Nguyen TT, Frater JL. MIST1—a novel marker of plasmacytic differentiation. *Appl Immunohistochem Mol Morphol AIMM*. 2012;20:561–5. doi:10.1097/PAI.0b013e31824e93f2.
145. Chalmers F, van Lith M, Sweeney B, Cain K, Bulleid NJ. Inhibition of IRE1 $\alpha$ -mediated XBP1 mRNA cleavage by XBP1 reveals a novel regulatory process during the unfolded protein response. *Wellcome open Res*. 2017;2:36. doi:10.12688/wellcomeopenres.11764.2.
146. Basso K, Dalla-Favera R. BCL6. master regulator of the germinal center reaction and key oncogene in B Cell lymphomagenesis. In: *Advances in Immunology*. Academic Press Inc.; 2010. p. 193–210.
147. Amin RH, Schlissel MS. Foxo1 directly regulates the transcription of recombination-activating genes during B cell development. *Nat Immunol*. 2008;9:613–22.
148. Ninagawa S, Okada T, Sumitomo Y, Kamiya Y, Kato K, Horimoto S, et al. EDEM2 initiates mammalian glycoprotein ERAD by catalyzing the first mannose trimming step. *J Cell Biol*. 2014;206:347–56. doi:10.1083/jcb.201404075.
149. Arnold SM, Fessler LI, Fessler JH, Kaufman RJ. Two homologues encoding human UDP-glucose:glycoprotein glucosyltransferase differ in mRNA expression and enzymatic activity. *Biochemistry*. 2000;39:2149–63.
150. Takeda Y, Seko A, Hachisu M, Daikoku S, Izumi M, Koizumi A, et al. Both isoforms of human UDP-glucose:glycoprotein glucosyltransferase are enzymatically active. *Glycobiology*. 2014;24:344–50.
151. Friedman SL, Sheppard D, Duffield JS, Violette S. Therapy for fibrotic diseases: Nearing the starting line. *Sci Transl Med*. 2013;5.
152. Friedman SL. Hepatic Stellate Cells: Protean, Multifunctional, and Enigmatic Cells of the Liver. *New York*. 2010;88:125–72. doi:10.1152/physrev.00013.2007.Hepatic.
153. Tomoishi S, Fukushima S, Shinohara K, Katada T, Saito K. CREB3L2-mediated expression of Sec23A/Sec24D is involved in hepatic stellate cell activation through ER-Golgi transport. *Sci Rep*. 2017;7:7992. doi:10.1038/s41598-017-08703-6.
154. Saito A, Hino S, Murakami T, Kanemoto S, Kondo S, Saitoh M, et al. Regulation of endoplasmic reticulum stress response by a BBF2H7-mediated Sec23a pathway is essential for chondrogenesis. *Nat Cell Biol*. 2009;11:1197. http://dx.doi.org/10.1038/ncb1962.

155. Ishikawa T, Toyama T, Nakamura Y, Tamada K, Shimizu H, Ninagawa S, et al. UPR transducer BFB2H7 allows export of type II collagen in a cargo- and developmental stage- specific manner. *J Cell Biol.* 2017;216:1761–74. doi:10.1083/jcb.201609100.
156. Melville DB, Montero-Balaguer M, Levic DS, Bradley K, Smith JR, Hatzopoulos AK, et al. The feelgood mutation in zebrafish dysregulates COPII-dependent secretion of select extracellular matrix proteins in skeletal morphogenesis. *Dis Model Mech.* 2011;4:763–76. doi:10.1242/dmm.007625.
157. Kondo S, Saito A, Hino S -i., Murakami T, Ogata M, Kanemoto S, et al. BFB2H7, a Novel Transmembrane bZIP Transcription Factor, Is a New Type of Endoplasmic Reticulum Stress Transducer. *Mol Cell Biol.* 2007;27:1716–29. doi:10.1128/MCB.01552-06.
158. Bonnon C, Wendeler MW, Paccaud JP, Hauri HP. Selective export of human GPI-anchored proteins from the endoplasmic reticulum. *J Cell Sci.* 2010;123:1705–15.
159. Becker F, Block-Alper L, Nakamura G, Harada J, Wittrup KD, Meyer DI. Expression of the 180-kD ribosome receptor induces membrane proliferation and increased secretory activity in yeast. *J Cell Biol.* 1999;146:273–84. doi:10.1083/jcb.146.2.273.
160. Ueno T, Tanaka K, Kaneko K, Taga Y, Sata T, Irie S, et al. Enhancement of procollagen biosynthesis by p180 through augmented ribosome association on the endoplasmic reticulum in response to stimulated secretion. *J Biol Chem.* 2010;285:29941–50.
161. Benyamini P, Webster P, Meyer DI. Knockdown of p180 eliminates the terminal differentiation of a secretory cell line. *Mol Biol Cell.* 2009;20:732–44. doi:10.1091/mbc.e08-07-0682.
162. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A.* 1977;74:5463–7. doi:10.1073/pnas.74.12.5463.
163. Mitra RD, Church GM. In situ localized amplification and contact replication of many individual DNA molecules. *Nucleic Acids Res.* 1999;27:e34–e34. doi:10.1093/nar/27.24.e34.
164. Huang C, Barker SJ, Langridge P, Smith FW, Graham RD. Zinc Deficiency Up-Regulates Expression of High-Affinity Phosphate Transporter Genes in Both Phosphate-Sufficient and -Deficient Barley Roots. *Plant Physiol.* 2000;124:415–22. doi:10.1104/pp.124.1.415.
165. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods.* 2008;5:621–8.
166. Kircher M, Kelso J. High-throughput DNA sequencing--concepts and limitations. *Bioessays.* 2010;32:524–36. doi:10.1002/bies.200900181.
167. Buermans HPJ, den Dunnen JT. Next generation sequencing technology: Advances and applications. *Biochim Biophys Acta - Mol Basis Dis.* 2014;1842:1932–41. doi:https://doi.org/10.1016/j.bbadis.2014.06.015.
168. Pettersson E, Lundeberg J, Ahmadian A. Generations of sequencing technologies. *Genomics.* 2009;93:105–11.
169. Core Next-Generation Sequencing. Comparison of Illumina Sequencers. 2019. <https://ngsc.med.upenn.edu/faqs-mini/HeadToHead.html>.
170. Nagalakshmi U, Waern K, Snyder M. RNA-seq: A method for comprehensive transcriptome analysis. *Current Protocols in Molecular Biology.* 2010; SUPPL. 89.
171. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res.* 2008;18:1509–17. doi:10.1101/gr.079558.108.
172. Bottomly D, Walter NAR, Hunter JE, Darakjian P, Kawane S, Buck KJ, et al. Evaluating gene expression in C57BL/6J and DBA/2J mouse striatum using RNA-Seq and microarrays. *PLoS One.* 2011;6.
173. Sîrbu A, Kerr G, Crane M, Ruskin HJ. RNA-Seq vs dual- and single-channel microarray data: sensitivity analysis for differential expression and clustering. *PLoS One.* 2012;7:e50986. doi:10.1371/journal.pone.0050986.
174. Poulin J, Tasic B, Hjerling-Leffler J, Trimarchi JM, Awatramani R, Harbom LJ, et al. Gene

- expression profiling in human neurodegenerative disease. *Nat Neurosci.* 2016;38:1131–41. doi:10.1371/journal.pone.0017820.
175. Fu X, Fu N, Guo S, Yan Z, Xu Y, Hu H, et al. Estimating accuracy of RNA-Seq and microarrays with proteomics. *BMC Genomics.* 2009;10:161. doi:10.1186/1471-2164-10-161.
176. Hansen KD, Brenner SE, Dudoit S. Biases in Illumina transcriptome sequencing caused by random hexamer priming. *Nucleic Acids Res.* 2010;38:e131–e131. doi:10.1093/nar/gkq224.
177. McIntyre LM, Lopiano KK, Morse AM, Amin V, Oberg AL, Young LJ, et al. RNA-seq: technical variability and sampling. *BMC Genomics.* 2011;12:293. doi:10.1186/1471-2164-12-293.
178. Batzoglou S, Pachter L, Mesirov JP, Berger B, Lander ES. Human and mouse gene structure: Comparative analysis and application to exon prediction. *Genome Res.* 2000;10:950–8.
179. Chen G, Ning B, Shi T. Single-cell RNA-seq technologies and related computational data analysis. *Frontiers in Genetics.* 2019;10 APR.
180. Fuller CW, Middendorf LR, Benner SA, Church GM, Harris T, Huang X, et al. The challenges of sequencing by synthesis. *Nat Biotechnol.* 2009;27:1013–23. doi:10.1038/nbt.1585.
181. Dohm JC, Lottaz C, Borodina T, Himmelbauer H. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Res.* 2008;36:e105.
182. Sha Y, Phan JH, Wang MD. Effect of low-expression gene filtering on detection of differentially expressed genes in RNA-seq data. *Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS.* 2015;2015-Novem:6461–4. doi:10.1109/EMBC.2015.7319872.
183. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 2010;11:R25. doi:10.1186/gb-2010-11-3-r25.
184. Law CW, Chen Y, Shi W, Smyth GK. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* 2014;15:R29. doi:10.1186/gb-2014-15-2-r29.
185. Coon JJ, Syka JEP, Shabanowitz J, Hunt DF. Tandem Mass Spectrometry for Peptide and Protein Sequence Analysis. *Biotechniques.* 2005;38:519–23. doi:10.2144/05384TE01.
186. Zubarev RA, Makarov A. Orbitrap Mass Spectrometry. *Anal Chem.* 2013;85:5288–96. doi:10.1021/ac4001223.
187. Clark DP, Pazdernik NJ. Chapter 9 - Proteomics. In: Clark DP, Pazdernik NJBT-B (Second E, editors. Boston: Academic Cell; 2016. p. 295–333. doi:https://doi.org/10.1016/B978-0-12-385015-7.00009-0.
188. Wilm M. Principles of electrospray ionization. *Mol Cell Proteomics.* 2011;10:M111.009407-M111.009407. doi:10.1074/mcp.M111.009407.
189. Zhu W, Smith JW, Huang C-M. Mass spectrometry-based label-free quantitative proteomics. *J Biomed Biotechnol.* 2010;2010:840518. doi:10.1155/2010/840518.
190. Sidoli S, Kulej K, Garcia BA. Why proteomics is not the new genomics and the future of mass spectrometry in cell biology. *J Cell Biol.* 2017;216:21 LP – 24. doi:10.1083/jcb.201612010.
191. Helbig AO, Heck AJR, Slijper M. Exploring the membrane proteome—Challenges and analytical strategies. *J Proteomics.* 2010;73:868–78. doi:https://doi.org/10.1016/j.jprot.2010.01.005.
192. Pascovici D, Handler DCL, Wu JX, Haynes PA. Multiple testing corrections in quantitative proteomics: A useful but blunt tool. *Proteomics.* 2016;16:2448–53.
193. Illumina. TruSeq Stranded mRNA Reference Guide. 2017.
194. Lam WY, Becker AM, Kennerly KM, Wong R, Curtis JD, Llufrío EM, et al. Mitochondrial Pyruvate Import Promotes Long-Term Survival of Antibody-Secreting Plasma Cells. *Immunity.* 2016;45:60–73. doi:10.1016/j.immuni.2016.06.011.
195. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29:15–21. doi:10.1093/bioinformatics/bts635.
196. Robinson MD, McCarthy DJ, Smyth GK. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2009;26:139–40.

doi:10.1093/bioinformatics/btp616.

197. Team STD. SRA Toolkit. <http://ncbi.github.io/sra-tools/>. Accessed 13 Feb 2019.
198. Bushnell B. BBtools. <http://sourceforge.net/projects/bbmap/>. Accessed 13 Feb 2019.
199. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011;12:323. doi:10.1186/1471-2105-12-323.
200. Tange O. GNU Parallel - The Command-Line Power Tool. ;login: The USENIX Magazine. 2011;36:42–7. <http://www.gnu.org/s/parallel>.
201. Sonesson C, Love MI, Robinson MD. Differential analyses for RNA-seq transcript-level estimates improve gene-level inferences. *F1000Research*. 2015;4.
202. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47.
203. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York; 2016. <https://ggplot2.tidyverse.org>.
204. Kassambara A. ggpubr: “ggplot2” Based Publication Ready Plots. 2019. <https://cran.r-project.org/package=ggpubr>.
205. Wickham H. Reshaping Data with the {reshape} Package. *J Stat Softw*. 2007;21:1–20. <http://www.jstatsoft.org/v21/i12/>.
206. Ma L, Bajic VB, Zhang Z. On the classification of long non-coding RNAs. *RNA Biology*. 2013;10:924–33. doi:10.4161/rna.24604.
207. Diederichs S. The four dimensions of noncoding RNA conservation. *Trends in Genetics*. 2014;30:121–3.
208. Ulitsky I, Shkumatava A, Jan CH, Sive H, Bartel DP. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell*. 2011;147:1537–50. doi:10.1016/j.cell.2011.11.055.
209. Kreslavsky T, Vilagos B, Tagoh H, Poliakova DK, Schwickert TA, Wöhner M, et al. Essential role for the transcription factor Bhlhe41 in regulating the development, self-renewal and BCR repertoire of B-1a cells. *Nat Immunol*. 2017;18:442–55. doi:10.1038/ni.3694.
210. Boller S, Grosschedl R. The regulatory network of B-cell differentiation: a focused view of early B-cell factor 1 function. *Immunol Rev*. 2014;261:102–15. doi:10.1111/imr.12206.
211. Oh-hashii K, Fujimura K, Norisada J, Hirata Y. Expression analysis and functional characterization of the mouse cysteine-rich with EGF-like domains 2. *Sci Rep*. 2018;8:12236. doi:10.1038/s41598-018-30362-4.
212. Li C, Wei J, Li Y, He X, Zhou Q, Yan J, et al. Transmembrane Protein 214 (TMEM214) mediates endoplasmic reticulum stress-induced caspase 4 enzyme activation and apoptosis. *J Biol Chem*. 2013;288:17908–17.
213. Cai C, Rajaram M, Zhou X, Liu Q, Marchica J, Li J, et al. Activation of multiple cancer pathways and tumor maintenance function of the 3q amplified oncogene FNDC3B. *Cell Cycle*. 2012;11:1773–81. doi:10.4161/cc.20121.
214. Lin C-H, Lin Y-W, Chen Y-C, Liao C-C, Jou Y-S, Hsu M-T, et al. FNDC3B promotes cell migration and tumor metastasis in hepatocellular carcinoma. *Oncotarget*. 2016;7:49498–508. doi:10.18632/oncotarget.10374.
215. Zhong Z, Zhang H, Hong M, Sun C, Xu Y, Chen X, et al. FNDC3B promotes epithelial-mesenchymal transition in tongue squamous cell carcinoma cells in a hypoxic microenvironment. *Oncol Rep*. 2018;39:1853–9.
216. Hong Y, Qin H, Li Y, Zhang Y, Zhuang X, Liu L, et al. FNDC3B circular RNA promotes the migration and invasion of gastric cancer cells via the regulation of E-cadherin and CD44 expression. *J Cell Physiol*. 2019;234:19895–910.
217. Liedtke D, Orth M, Meissler M, Geuer S, Knaup S, Köblitz I, et al. ECM alterations in Fndc3a

- (Fibronectin Domain Containing Protein 3A) deficient zebrafish cause temporal fin development and regeneration defects. *Sci Rep.* 2019;9.
218. Bhattacharya MRC, Geisler S, Pittman SK, Doan RA, Weihl CC, Milbrandt J, et al. TMEM184b Promotes Axon Degeneration and Neuromuscular Junction Maintenance. *J Neurosci.* 2016;36:4681–9. doi:10.1523/JNEUROSCI.2893-15.2016.
219. Oh-hashii K, Kunieda R, Hirata Y, Kiuchi K. Biosynthesis and secretion of mouse cysteine-rich with EGF-like domains 2. *FEBS Lett.* 2011;585:2481–7. doi:10.1016/j.febslet.2011.06.029.
220. Oh-hashii K, Koga H, Ikeda S, Shimada K, Hirata Y, Kiuchi K. Role of an ER stress response element in regulating the bidirectional promoter of the mouse CRELD2 - ALG12 gene pair. *BMC Genomics.* 2010;11:664. doi:10.1186/1471-2164-11-664.
221. Hartley CL, Edwards S, Mullan L, Bell PA, Fresquet M, Boot-Handford RP, et al. Armet/Manf and Creld2 are components of a specialized ER stress response provoked by inappropriate formation of disulphide bonds: implications for genetic skeletal diseases. *Hum Mol Genet.* 2013;22:5262–75. doi:10.1093/hmg/ddt383.
222. Zhang J, Weng Y, Liu X, Wang J, Zhang W, Kim SH, et al. Endoplasmic Reticulum (ER) Stress Inducible Factor Cysteine-Rich with EGF-Like Domains 2 (Creld2) Is an Important Mediator of BMP9-Regulated Osteogenic Differentiation of Mesenchymal Stem Cells. *PLoS One.* 2013;8:e73086. doi:10.1371/journal.pone.0073086.
223. Iyengar V, Pullakhandam R, Nair KM. Coordinate expression and localization of iron and zinc transporters explain iron-zinc interactions during uptake in Caco-2 cells: implications for iron uptake at the enterocyte. *J Nutr Biochem.* 2012;23:1146–54.
224. Tominaga K, Kagata T, Johmura Y, Hishida T, Nishizuka M, Imagawa M. SLC39A14, a LZT protein, is induced in adipogenesis and transports zinc. *FEBS J.* 2005;272:1590–9.
225. Liuzzi JP, Aydemir F, Nam H, Knutson MD, Cousins RJ. Zip14 (Slc39a14) mediates non-transferrin-bound iron uptake into cells. *Proc Natl Acad Sci U S A.* 2006;103:13612–7.
226. Fujishiro H, Yano Y, Takada Y, Tanihara M, Himeno S. Roles of ZIP8, ZIP14, and DMT1 in transport of cadmium and manganese in mouse kidney proximal tubule cells. *Metallomics.* 2012;4:700–8.
227. Wang L, Zhan Y, Song E, Yu Y, Jiu Y, Du W, et al. HID-1 is a peripheral membrane protein primarily associated with the medial- and trans- Golgi apparatus. *Protein Cell.* 2011;2:74–85.
228. Mesa R, Luo S, Hoover CM, Miller K, Minniti A, Inestrosa N, et al. HID-1, a new component of the peptidergic signaling pathway. *Genetics.* 2011;187:467–83.
229. Yu Y, Wang L, Jiu Y, Zhan Y, Liu L, Xia Z, et al. HID-1 is a novel player in the regulation of neuropeptide sorting. *Biochem J.* 2011;434:383–90.
230. Du W, Zhou M, Zhao W, Cheng D, Wang L, Lu J, et al. HID-1 is required for homotypic fusion of immature secretory granules during maturation. *Elife.* 2016;5.
231. Tran HT, Hurley BA, Plaxton WC. Feeding hungry plants: The role of purple acid phosphatases in phosphate nutrition. *Plant Sci.* 2010;179:14–27. doi:http://dx.doi.org/10.1016/j.plantsci.2010.04.005.
232. Ueno T, Kaneko K, Sata T, Hattori S, Ogawa-Goto K. Regulation of polysome assembly on the endoplasmic reticulum by a coiled-coil protein, p180. *Nucleic Acids Res.* 2012;40:3006–17. doi:10.1093/nar/gkr1197.
233. Cui XA, Zhang H, Palazzo AF. P180 promotes the ribosome-independent localization of a subset of mRNA to the endoplasmic reticulum. *PLoS Biol.* 2012;10.
234. Dice JF, Schimke RT. Turnover and exchange of ribosomal proteins from rat liver. *J Biol Chem.* 1972;247:98–111. <https://www.jbc.org/content/247/1/98.short>. Accessed 13 Feb 2020.
235. Sung MK, Reitsma JM, Sweredoski MJ, Hess S, Deshaies RJ. Ribosomal proteins produced in excess are degraded by the ubiquitin-proteasome system. *Mol Biol Cell.* 2016;27:2642–52.
236. Ueno T, Kaneko K, Katano H, Sato Y, Mazitschek R, Tanaka K, et al. Expansion of the trans-

- Golgi network following activated collagen secretion is supported by a coiled-coil microtubule-bundling protein, p180, on the ER. *Exp Cell Res.* 2010;316:329–40. doi:10.1016/j.yexcr.2009.11.009.
237. Al-Maskari M, Care MA, Robinson E, Cocco M, Tooze RM, Doody GM. Site-1 protease function is essential for the generation of antibody secreting cells and reprogramming for secretory activity. *Sci Rep.* 2018;8:14338. doi:10.1038/s41598-018-32705-7.
238. Matsuoka K, Orci L, Amherdt M, Bednarek SY, Hamamoto S, Schekman R, et al. COPII-coated vesicle formation reconstituted with purified coat proteins and chemically defined liposomes. *Cell.* 1998;93:263–75.
239. Fox RM, Hanlon CD, Andrew DJ. The CrebA/Creb3-like transcription factors are major and direct regulators of secretory capacity. *J Cell Biol.* 2010;191:479 LP – 492. doi:10.1083/jcb.201004062.
240. Nickel W, Sohn K, Bünning C, Wieland FT, Reckmann I, Meissner I, et al. p23, a major COPI-vesicle membrane protein, constitutively cycles through the early secretory pathway. *Proc Natl Acad Sci U S A.* 1997;94:11393–8. doi:10.1073/pnas.94.21.11393.
241. Futatsumori M, Kasai K, Takatsu H, Shin H-W, Nakayama K. Identification and Characterization of Novel Isoforms of COP I Subunits. *J Biochem.* 2000;128:793–801. doi:10.1093/oxfordjournals.jbchem.a022817.
242. Martzoukou O, Amillis S, Zervakou A, Christoforidis S, Diallinas G. The AP-2 complex has a specialized clathrin-independent role in apical endocytosis and polar growth in fungi. *Elife.* 2017;6. doi:10.7554/eLife.20083.
243. Ang AL, Taguchi T, Francis S, Fölsch H, Murrells LJ, Pypaert M, et al. Recycling endosomes can serve as intermediates during transport from the Golgi to the plasma membrane of MDCK cells. *J Cell Biol.* 2004;167:531 LP – 543. doi:10.1083/jcb.200408165.
244. Murray RZ, Stow JL. Cytokine Secretion in Macrophages: SNAREs, Rabs, and Membrane Trafficking. *Front Immunol.* 2014;5:538. doi:10.3389/fimmu.2014.00538.
245. Rivera-Molina F, Toomre D. Live-cell imaging of exocyst links its spatiotemporal dynamics to various stages of vesicle fusion. *J Cell Biol.* 2013;201:673–80. doi:10.1083/jcb.201212103.
246. Raote I, Ortega-Bellido M, Santos AJ, Foresti O, Zhang C, Garcia-Parajo MF, et al. TANGO1 builds a machine for collagen export by recruiting and spatially organizing COPII, tethers and membranes. *Elife.* 2018;7. doi:10.7554/eLife.32723.
247. Stauffer C, Haack TB, Köpke MG, Straub BK, Kölker S, Thiel C, et al. Recurrent acute liver failure due to NBAS deficiency: phenotypic spectrum, disease mechanisms, and therapeutic concepts. *J Inherit Metab Dis.* 2016;39:3–16. doi:10.1007/s10545-015-9896-7.
248. Gordon DE, Chia J, Jayawardena K, Antrobus R, Bard F, Peden AA. VAMP3/Syb and YKT6 are required for the fusion of constitutive secretory carriers with the plasma membrane. *PLOS Genet.* 2017;13:e1006698. doi:10.1371/journal.pgen.1006698.
249. Choi JK, Yu U, Kim S, Yoo OJ. Combining multiple microarray studies and modeling interstudy variation. *Bioinformatics.* 2003;19 Suppl 1:i84–90. doi:10.1093/bioinformatics/btg1010.
250. Hong F, Breitling R, McEntee CW, Wittner BS, Nemhauser JL, Chory J. RankProd: A bioconductor package for detecting differentially expressed genes in meta-analysis. *Bioinformatics.* 2006;22:2825–7. doi:10.1093/bioinformatics/btl476.
251. Hong F, Breitling R. A comparison of meta-analysis methods for detecting differentially expressed genes in microarray experiments. *Bioinformatics.* 2008;24:374–82. doi:10.1093/bioinformatics/btm620.
252. Fisher RA. *Statistical Methods for Research Workers.* Springer, New York, NY; 1992. p. 66–70.
253. Marot G, Mayer CD. Sequential analysis for microarray data based on sensitivity and meta-analysis. *Stat Appl Genet Mol Biol.* 2009;8:Article 3. doi:10.2202/1544-6115.1368.
254. Marot G, Foulley J-L, Mayer C-D, Jaffrezic F. Moderated effect size and P-value combinations for microarray meta-analyses. *Bioinformatics.* 2009;25:2692–9. doi:10.1093/bioinformatics/btp444.
255. Cho H, Kim H, Na D, Kim SY, Jo D, Lee D. Meta-analysis method for discovering reliable

- biomarkers by integrating statistical and biological approaches: An application to liver toxicity. *Biochem Biophys Res Commun.* 2016;471:274–81.
256. Kosch R, Jung K. Conducting gene set tests in meta-analyses of transcriptome expression data. *Res Synth Methods.* 2019;10:99–112. doi:10.1002/jrsm.1337.
257. Sitia R, Rubartelli A, Hammerling U. Expression of 2 immunoglobulin isotypes, IgM and IgA, with identical idiotype in the B cell lymphoma I.29. *J Immunol.* 1981;127:1388 LP – 1394. <http://www.jimmunol.org/content/127/4/1388.abstract>.
258. Wong L-Y, Brulois K, Toth Z, Inn K-S, Lee S-H, O'Brien K, et al. The product of Kaposi's sarcoma-associated herpesvirus immediate early gene K4.2 regulates immunoglobulin secretion and calcium homeostasis by interacting with and inhibiting pERP1. *J Virol.* 2013;87:12069–79. doi:10.1128/JVI.01900-13.

## 6 APPENDIX

Table 6-1 | Manually Curated Summary of Overlapping GO Biological Process Terms.

	Summary Term	Description	Enriched Gene Count
Upregulated	Membrane Trafficking	retrograde vesicle-mediated transport, Golgi to ER	36
		ER to Golgi vesicle-mediated transport	42
		transmembrane transport	35
		neutrophil degranulation	78
		establishment of protein localization	10
	ER stress	IRE1-mediated unfolded protein response	32
		ER unfolded protein response	16
		response to ER stress	22
		ubiquitin-dependent ERAD pathway	21
		retrograde protein transport, ER to cytosol	9
		proteasome-mediated ubiquitin-dependent protein catabolic process	39
		ERAD pathway	8
		NIK/NF-kappaB signalling	18
		tumour necrosis factor-mediated signalling pathway	26
		cellular response to oxidative stress	20
	glycosylation	ATF6-mediated unfolded protein response	6
		protein N-linked glycosylation via asparagine	20
		protein N-linked glycosylation	14
	electron transport chain	dolichyl diphosphate biosynthetic process	6
		dolichol-linked oligosaccharide biosynthetic process	8
		mitochondrial translational elongation	32
		mitochondrial electron transport, NADH to ubiquinone	22
		mitochondrial translational termination	31
	antigen presentation	heme biosynthetic process	8
		mitochondrial respiratory chain complex I assembly	27
		antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent	22
	mitotic cell cycle	antigen processing and presentation of exogenous peptide antigen via MHC class II	23
		antigen processing and presentation of peptide antigen via MHC class I	10
		anaphase-promoting complex-dependent catabolic process	29
		positive regulation of ubiquitin-protein ligase activity involved in regulation of mitotic cell cycle transition	27
		negative regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle	26
		regulation of cellular amino acid metabolic process	19
		chromosome segregation	17
		negative regulation of G2/M transition of mitotic cell cycle	18
		sister chromatid cohesion	25
		nucleobase-containing small molecule catabolic process	8
		regulation of transcription involved in G1/S transition of mitotic cell cycle	10
		spindle organization	9
		branched-chain amino acid catabolic process	9
		SCF-dependent proteasomal ubiquitin-dependent protein catabolic process	19
		mitotic cytokinesis	11
	microtubule-based movement	17	
protein biosynthesis	regulation of transcription from RNA polymerase II promoter in response to hypoxia	19	
	tRNA aminoacylation for protein translation	18	
anti apoptotic process	negative regulation of apoptotic process	57	
Downregulated	phosphorylation	protein phosphorylation	81
		peptidyl-serine phosphorylation	35
		negative regulation of myosin-light-chain-phosphatase activity	5
	response to DNA damage	cellular response to DNA damage stimulus	53
	actin cytoskeleton organization	cortical actin cytoskeleton organization	10
	RNA processing	mRNA splicing, via spliceosome	55
		RNA processing	20
		RNA secondary structure unwinding	16
	regulation of transcription	positive regulation of transcription from RNA polymerase II promoter	129
		negative regulation of transcription, DNA-templated	82
		positive regulation of transcription, DNA-templated	87
		regulation of DNA-templated transcription, elongation	7
		negative regulation of transcription from RNA polymerase II promoter	86
		regulation of transcription from RNA polymerase II promoter	60
	PI biosynthesis	phosphatidylinositol biosynthetic process	24
	I-KB kinase/NF-KB signalling	I-kappaB kinase/NF-kappaB signalling	18
	ubiquitination	protein polyubiquitination	47
		protein ubiquitination involved in ubiquitin-dependent protein catabolic process	31
	transcription	transcription from RNA polymerase II promoter	54
	focal adhesion assembly	regulation of focal adhesion assembly	8
small GTPase regulation	regulation of small GTPase mediated signal transduction	32	

**Table 6-2** | Manually Curated Summary of Overlapping GO Cellular Component Terms

	Summary Term	Description	Enriched Gene Count
Upregulated	endoplasmic reticulum	endoplasmic reticulum membrane	152
		endoplasmic reticulum	165
		integral component of endoplasmic reticulum membrane	36
		oligosaccharyltransferase complex	9
	mitochondria	mitochondrion	193
		mitochondrial inner membrane	85
		mitochondrial matrix	72
		mitochondrial respiratory chain complex I	23
		mitochondrial ribosome	13
		mitochondrial large ribosomal subunit	18
		mitochondrial small ribosomal subunit	10
		transport vesicle	30
	ER/Golgi intermediate	endoplasmic reticulum-Golgi intermediate compartment	23
		Golgi apparatus	116
	Golgi apparatus	Golgi membrane	73
		COPI vesicle coat	9
		integral component of Golgi membrane	14
		proteasome complex	18
	proteasome	proteasome core complex	10
		chromosome/mitotic spindle	13
chromosome/mitotic spindle	spindle microtubule	13	
	condensed chromosome outer kinetochore	6	
	kinetochore	20	
	mitotic spindle	14	
secretory granule	azurophil granule lumen	22	
	ficolin-1-rich granule lumen	27	
Downregulated	nucleus	nuclear speck	65
		nucleolus	123
		PML body	20
		nuclear body	53
	cytoplasm	I-kappaB/NF-kappaB complex	5
		beta-catenin destruction complex	7

**Table 6-3 | GO Biological Process enrichment of Upregulated Genes in ASCs across species (Microarray).**

Term	Adj <i>p</i> -value	Gene Symbol
ER stress (139)	6.0E-14	FKBP14, SHC1, GOSR2, MYDGF, SEC61A1, EXTL2, SERP1, SEC61G, ASNA1, LMNA, SSR1, DNAJB9, SEC61B, SEC62, SEC63, SEC31A, XBP1, HSPA5, WFS1, GFPT1, SYVN1, WIPI1, PDIA6, PDIA5, ARFGAP1, YIF1A, ERN1, DNAJC3, CTDSP2, DNAJB11, PREB, KDELR3, EDEM3, VCP, DERL3, EDEM2, DERL1, BHLHA15, CREB3, TBL2, SELENOS, CTH, ATF6, RNF121, ERO1A, PDIA3, CEBPB, UBA5, FAM129A, TXNDC11, ATP2A2, HERPUD1, PDIA4, HSP90B1, ERP44, TMEM33, TMX1, DDRGK1, CREB3L2, FLOT1, TRIB3, P4HB, SEL1L, ERLIN1, UBE2J1, ERLEC1, PSMC4, JKAMP, FBXO6, STT3B, UFD1, PSMD10, ANAPC16, PSMD14, BUB1B, TNFAIP1, PSMD8, PSMB6, PSMD9, PSMB7, RNF115, UCHL1, PSMB4, CDC23, PSMB5, PSMD4, PSMB2, HECTD3, CDC26, SPOP, PSMD1, GTSE1, UBE2C, BFAR, PSMB8, PSMA5, PSMA6, CDC34, PSMA4, GID4, PSMA1, TBL1XR1, CDK1, RNF181, STUB1, CLOCK, RNF187, MAD2L1, RHBDD1, DNAJB12, TRIM25, UBE2M, TNFRSF13B, TRADD, FOXO3, TXNDC17, TNFSF13B, PYCARD, TNFRSF17, CDIP1, LIMS1, VKORC1L1, GPX3, GSTP1, TXNL1, NUDT2, GPX7, HTRA2, PYCR1, PYCR2, PRDX6, PRDX3, PRDX2, PRDX1, SESN2, CYCS, CPEB2, ATF6B, CALR
membrane trafficking (173)	3.2E-12	ARF4, COPB2, COPA, TMED10, KDELR1, STX18, KIF11, SCFD1, RACGAP1, LMAN2, TMED3, TMED2, TMED7, KIFAP3, TMED9, SURF4, GOLPH3L, COPZ2, KIF23, KIF22, ARFGAP3, COPZ1, ARFGAP1, TAPBP, ARCN1, CENPE, KIF18A, GOLPH3, RER1, KIF4A, KDELR2, KDELR3, COPG1, SEC22B, ERGIC2, COPE, DYNC112, LMAN2L, ATL3, SAR1B, GOSR2, USO1, DCTN3, TEX261, MIA3, IER3IP1, LMAN1, CD59, YKT6, SEC31A, RAB2A, SEC24A, COG6, DYNLL1, BET1L, YIF1A, YIF1B, SEC23IP, PREB, GAS6, SEC24D, PSMD10, VCP, ABCD3, PSMD14, ABCB6, SLC44A1, SLC41A2, SEL1L, DERL1, MAGT1, PSMD8, PSMB6, PSMD9, ERLEC1, PSMB7, PSMB4, PSMB5, PSMD4, PSMB2, PDZD11, PSMD1, LRRC8D, ABCC3, RALBP1, ABCA5, SLC33A1, ERLIN1, EIF2S2, PSMB8, PSMA5, PSMA6, PSMA4, PSMA1, PEX3, PSMC4, CYFIP1, RAB3D, AP2A2, MPO, PYCARD, ALAD, GLIPR1, PNP, NAPRT, LAMP2, TIMP2, GUSB, GYG1, CTSD, PGM1, CTSB, HGSNAT, TICAM2, FCER1G, ANXA2, CD300A, ATP6AP2, NDUFC2, COMMD3, MIF, CKAP4, DNAJC3, C16ORF62, PKM, NPC2, DNAJC5, CSNK2B, S100A9, LTF, CSTB, CD63, GRN, GSTP1, MVP, FGL2, PRCP, PGRMC1, PRDX4, RAPIA, CREG1, ORMDL3, MLEC, STOM, METTL7A, AGA, ATG7, ATP6V1D, ARSB, VAT1, HSPA8, TMEM30A, FUCA1, FUCA2, RAB27A, GGH, PA2G4, PRDX6, DERA, RAB10, BST2, ERP44, IMPDH1, GLB1, GLA, SELENOS, PLK1, RDX, CRIPT, MDM2, CEP55, HINFP, LIMS1
electron transport chain (67)	1.8E-08	NDUFB9, NDUFB7, NDUFA11, NDUFB6, NDUFA12, NDUFB11, NDUFB5, NDUFB3, NDUFB1, FOXRED1, NDUFV3, NDUFA9, NDUFA8, TIMMDC1, NDUFA3, NDUFA2, NDUFA1, NDUFC2, NDUFAF8, NDUFS8, NDUFS7, NDUFS5, NDUFAF4, NDUFS4, NDUFAF1, NDUFAF2, NDUFAF1, MRPS15, MRPS14, MRPS11, MRPL18, MRPS12, MRPS34, MRPL17, MRPL36, MRPL58, MRPL37, MRPL15, MRPL34, MRPL35, MRPL13, MRPL57, MRPL4, MRPL42, MRPL20, MRPL40, CHCHD1, TSFM, MRPS24, MRPL27, MRPS2, MRPS18A, MRPL48, MRPS5, MRPS18C, MRPL53, MRPL50, MRPL51, AURKAIP1, ALAD, ALAS1, FECH, ATP1F1, HMBS, CPOX, FXN, NFE2L1
mitotic cell cycle (113)	2.8E-07	PSMD10, PSMD14, ANAPC16, BUB1B, AURKB, AURKA, PSMD8, PSMB6, PSMD9, PSMB7, CCNB1, PSMB4, CDC23, PSMB5, PSMD4, PSMB2, CDC26, PSMD1, SKP2, UBE2C, PLK1, PSMB8, PSMA5, PSMA6, PSMA4, PSMA1, PSMC4, CDK1, MAD2L1, FBXO5, OAZ1, AZIN1, CDT1, CENPW, SPAG5, KIF11, BRCA1, SKA3, KNSTRN, SKA1, SKA2, CENPE, CENPF, NTMT1, NUF2, NEK2, OIP5, TOP1, SPC25, CDCA5, CDCA8, CENPA, NUP43, BUB1, CENPU, SEC13, FBXW7, KNL1, KIF22, SGO2, KIF18A, CENPI, BIRC5, SPC24, NUDT9, ENTPD1, ENTPD4, ENTPD5, ENTPD7, NUDT1, NUDT18, NUDT5, PCNA, RRM2, CDC45, CCNE1, GFII1, CDC6, TYMS, HINFP, ASPM, TTK, AUNIP, MCCC2, BCKDK, HIBADH, AUH, BCAT1, HSD17B10, HIBCH, ACAT1, BCAT2, FBXO6, ANLN, SNX18, RACGAP1, KIF4A, NUSAP1, KIF23, STAMBP, KIF20A, CEP55, DYNC112, CLTA, AP2B1, AP2A2, KIF18B, KIFC1, KIF13B, KIF1B, KIFAP3, EPAS1, VEGFA
Glycosylation (38)	4.0E-07	OST4, VCP, ST6GAL1, RPN2, DERL3, RPN1, ALG5, SYVN1, MAGT1, UBE2J1, LMAN1, FUT8, DAD1, OSTC, STT3A, STT3B, NUDT14, ST6GALNAC4, ST3GAL2, ST6GALNAC6, DPAGT1, ALG8, ENTPD5, MOGS, TMEM165, TMEM258, DOLPPI, MAN1C1, PGM3, NUS1, SRD5A3, DOLK, DHDDS, ALG9, PQLC3, ALG14, ALG2, ALG3
antigen presentation (46)	4.5E-04	PDIA3, PSMD10, ITGB5, PSMD14, PSMB8, TAPBP, PSMD8, PSMA5, PSMB6, PSMD9, PSMA6, PSMB7, PSMB4, PSMA4, PSMB5, PSMD4, PSMA1, PSMB2, PSMC4, PSMD1, CALR, SEC22B, DYNC112, SEC13, FCER1G, SEC24A, SAR1B, DCTN3, CLTA, KIF23, KIF11, CTSV, KIF22, AP2B1, DYNLL1, AP2A2, CENPE, KIF18A, RACGAP1, KIF4A, CANX, KIFAP3, SEC24D, CTSD, SEC31A, ERAP1
Protein biosynthesis (13)	3.1E-03	CARS, YARS, DARS, TARS, SARS, EPRS, AIMP2, PPA2, PPA1, MARS, GARS, IARS, AARS
Anti apoptotic process (57)	4.2E-03	AVEN, ARF4, CITED2, APIP, MPO, SH3RF1, MYDGF, HMG5, CCND2, CASP3, PIM2, ERC1, PRKCI, ANXA5, MIF, DNAJC3, HAX1, BIRC5, NAA38, ATF5, GAS6, FXN, LTF, PSMD10, SHC1, GSTP1, NOL3, MTDH, FAM129B, HSP90B1, PRDX3, SOCS2, PRDX2, DDRGK1, TAF9B, MCL1, PRELID1, XBP1, NPM1, STIL, HSPA5, RHBDD1, PLK1, ASNS, DHCR24, BFAR, PA2G4, CFLAR, VEGFA, GOLPH3, IL2RB, PPT1, CDK1, FAS, BCL2L2, MAD2L1, BCL2L1

(#) Number in brackets indicate total number of enriched genes



**Table 6-5 | GO Cellular Component enrichment of Upregulated Genes in ASCs across species (Microarray)**

Term	Adj <i>p</i> -value	Gene Symbol
endoplasmic reticulum (260)	4E-25	ERO1A, DPAGT1, VKORC1L1, KDELRL, PLOD3, PLOD1, TM7SF2, FADS2, NSDHL, TMEM147, SLC39A7, SEC62, FADS1, SEC63, ALG8, ALG9, ALG5, ALG2, ALG14, ALG3, MOGS, ACSL4, ALG11, ATG13, PDIA6, PDIA5, DHDSD, ERN1, SELENON, SELENOS, TMX1, PREB, KDELRL2, KDELRL3, ATF6, RNF121, COPB2, RTN3, COPA, VCP, TMED10, RETSAT, RPN2, ABHD4, LMAN2L, SAR1B, INSIG2, PRN1, SEL1L, MGST2, TMTC2, HTRA2, MTDH, HSP90B1, APOO, PORCN, EBP, DDRGK1, HMOX1, SSR1, MLEC, KDSR, SEC31A, NUS1, SDF2L1, HSPA5, WFS1, RHBDD1, SRD5A3, SURF4, ERLIN1, DHCR24, DOLK, TMEM110, EXT1, NCEH1, GGCX, CDK1, STT3A, DHCR7, COPE, PIGU, PIGT, PIGN, GPAA1, TECR, ATP2A2, MIA3, IER3IP1, HERPUD1, SEC61A1, SERP1, SEC61G, TMED3, TMED2, SEC61B, SDF2, TMED7, TMED9, SEC13, COPZ2, ELOVL6, LSS, COPZ1, TAPBP, TMEM33, DAD1, CANX, RFT1, PIGK, PIGM, SEC22B, PIGF, MGLL, ATF6B, HM13, GOSR2, STX18, DERL1, FITM2, SELENOK, AGPAT1, LMF1, LMAN1, EXTL2, ASNA1, DNAJB9, CD59, SPTSSA, CDIPT, LNP, APOL2, SEC11C, SEC24A, GALNT2, SYVN1, DHRS3, YIF1A, ARCN1, DHRS4, YIF1B, RAB10, ERP44, SPCS2, CD4, SPCS1, DHRS9, SCD, DOLPPI, COPG1, SEC24D, LPIN1, GOLT1B, TXNDC11, APMAP, CREB3L2, CAPN2, SIL1, PDIA3, MAP2K2, ENTPD5, DNAJB12, P3H1, CNPY2, TEX2, MR1, PDIA4, CPED1, TPST2, MINPPI, TMEM258, DNAJB11, PPIB, CES2, LRPAP1, LRRC59, FKBP2, ABCB6, HADC3, GIGYF2, EDEM3, NDFIP2, XBP1, CRTAP, VCAM1, EDEM2, BFAR, TOR1B, TMEM230, MANF, CCDC88A, EHD4, TBL2, CALU, EBPL, PSMG1, ACO1, CALR, STT3B, CPEB4, CD320, TRAM1, CCDC47, UFPS2, MAGT1, MAN1A2, UFD1, TMED1, IKBIP, PSENE1, JAGN1, SSR2, SIGMAR1, GPX7, CKAP4, DNAJC3, CREB3, DNAJC1, RCN1, HAX1, TMEM214, PEX3, GRN, ATL3, USO1, RRBP1, CLN6, CLN5, PGRMC1, FAM213B, SRP72, ORMDL2, ORMDL3, EMC2, MAN1C1, STOM, AGA, YKT6, SEC23B, UGGT1, TMEM30A, ERAP1, YIPF5, SRP68, DHRS1, TMC01, VMP1, YIPF6, STUB1, P4HB, PTPN2, SLC35B3, SLC35B1, DERL3, TEX261, ZMPSTE24, BSCL2, OST4, KRTCAP2, OSTC
Mitochondria (240)	5E-20	DHFR2, ECI2, NUDT1, CISD1, GLS, GCSH, NFU1, BSG, VPS35, EARS2, MCCC2, CPT1A, STARD5, CMC2, SLC11A2, ATG13, CLPX, NIF3L1, MTHFD2, SCO2, CLPP, BOLA3, FDX1, UQCRC1, AURKAIP1, ALDH7A1, PIF1, GCDH, ABCB6, MRPL18, ATP5J, MRPL15, ATP5H, HADC3, PRDX3, MRPL20, APOO, NTHL1, PCK2, PRELID1, NDFIP2, APOOL, XRCC3, RMDN3, PYCR1, PYCR2, NDUFAF8, NDUFAF4, NDUFAF2, ALDH18A1, AARS, MRPS15, CLIC4, FEN1, FASTKD1, SLC44A1, MRPS11, MRPL36, MRPL37, IDE, MRPL42, MAPILC3B, HINT2, AIFM1, AUH, AIFM2, ACADM, DLGAP5, ATP1F1, SIRT4, MRPS2, USMG5, MRPS5, UCHL5, IMMP1L, ALDH5A1, NME6, PTRH2, SDHAF1, PTRH1, ECHS1, GSTP1, MRPS34, IFI6, MRPL58, MALSU1, NOL3, TYMS, TMEM70, SAMM50, PDF, BRD8, YKT6, WWOX, TIMMDC1, MDH1, MDH2, MTX2, DHRS4, NUDT9, GATM, NDUFAB1, PNPLA4, FAS, MTCH2, ACAA2, NDUFA11, NDUFA12, CPOX, COX6A1, ALDH1L2, PYCARD, CHCHD1, MAP2K2, SDHC, LIG3, SDHD, TRNT1, ATP5J2, HTRA2, AK3, ACAT1, BCL2L13, NACC2, NDUFV3, MCL1, DNAJC19, PDHA1, GOT1, UQC22, GOT2, IDH2, COQ9, CRLS1, COQ3, SCCPDH, UQCRCQ, CDK1, CYCS, ACO1, ACO2, MXD1, BCL2L1, ALAS1, ETFA, ATP5G1, DHTKD1, NADK2, OPA3, ME2, COA7, TSFM, NDUFC2, GLRX2, REXO2, MUT, SMDT1, HAX1, NDUFS8, PKM, NDUFS5, DNAJC5, PCCB, NDUFS4, MTFR1, BCAT1, SLC25A11, FXN, BCAT2, NDUFB9, NDUFB7, NDUFB6, NAXE, POLDIP2, NDUFB5, NDUFB1, PDHB, MFF, HSD17B10, FOXRED1, C6ORF203, STOM, SLC25A20, SLC25A23, NDUFA9, NDUFA8, NDUFA3, NDUFA1, GSTZ1, BCKDK, PNKD, GOLPH3, RAD51, MRPS14, MRPS12, MRPL34, MRPL35, MRPL4, MRPL40, TIMM17B, TIMM29, MRPS24, MRPS18A, MRPL48, MRPS18C, MRPL53, MRPL50, MRPL51, MTFP1, NDUFB11, NDUFB3, MRPL17, UQCRC10, MRPL13, COX5A, MRPL57, TMEM65, SQOR, NDUFA2, MRPL27, NDUFAF1, SLC25A35, HIBADH, NUDT2, ETFB, TFB1M, MCEE, HIBCH, GPT2, BTD, PPA2, NDUFS7, LACTB2, GARS, RNASEL, ACSS2, FECH, ETFDH, CCNB1, CBR4
ER/Golgi intermediate (45)	3E-10	COPB2, COPA, TMED10, KDELRL, USO1, GOLIM4, RABEPK, TMEM187, GOLGA5, TMED3, TMED2, CD59, AP3S1, TMED7, YKT6, TMED9, YIPF2, YIPF3, CAV2, SURF4, M6PR, COPZ2, COPZ1, ARCN1, TRIP11, KDELRL2, KDELRL3, COPG1, SEC22B, COPE, LRPAP1, HSPA5, PDIA6, YIF1A, YIF1B, LMAN1, ERP44, SEC23IP, LMAN2, NUCB1, TMED1, P4HB, UGGT1, TMED5, PTPN2
Golgi apparatus (164)	4E-08	GOLIM4, PLOD3, MANEA, HID1, GOLGA3, GOLGA4, GOLGA5, CAPN2, UNC13B, RAB2A, SLC30A7, SLC30A6, MAP2K2, AKRIE2, F2R, FNDC3A, ESCO2, LAX1, SCAMP2, TPST2, PREPL, KDELRL2, ATF6, LRPAP1, PARM1, TMED10, ABCB6, FAMI14A1, LMAN2L, ARL3, GLT8D1, ARL1, SLC38A10, GIGYF2, GLG1, ARHGAP21, FUT8, MYO6, NDFIP2, YES1, VCAM1, ATP8B2, SLC35E1, CAV2, PTCH1, RAB39B, CDC6, EXT1, ACER2, BST2, CCDC88A, B4GAT1, PPT1, IFT27, PSMG1, ACO1, ITM2B, COPE, TP73, ITM2C, DPY30, IER3IP1, LITAF, BICD1, MAN1A2, TRIM68, TMED3, TMED2, SDF4, TMED1, TMED7, PSENE1, TMED9, SLC39A11, PPHLN1, ARFGAP3, BET1L, PLSCR1, B3GNT9, CTTN, TMEM214, RER1, GORASP2, IMPAD1, TMEM5, SLC29A3, LEPROT, ATF6B, GOSR2, USO1, NOD2, SELENOK, CLCN3, TM9SF4, TMEM165, CLN5, LMAN2, CDIPT, YKT6, SEC23B, WWOX, YIPF2, YIPF3, GALNT2, QPCTL, YIPF5, PDE4DIP, PLEKHA8, RAB10, ATAT1, SEC23IP, GOLPH3, GLB1, TRIP11, COPG1, GLA, SLC35B3, CHPF, KDELRL1, HS2ST1, NDST1, C1GALT1C1, BSG, PMPA1, ST6GAL1, COG6, ENTPD4, CHPF2, GOLPH3L, COPZ2, NOSIP, WIPI1, COPZ1, TAPBP, CREB3, KDELRL3, SEC22B, ST6GALNAC4, RNF121, ST6GALNAC6, COPB2, COPA, APOO, CHST11, CHST12, MAN2A1, DSEL, ST8SIA4, MGAT3, MAN1C1, CD59, MGAT1, ST3GAL2, B3GAT2, SURF4, SLC33A1, B3GALT6, ARCN1, C6ORF89, SLC35B1, TEX261, YIF1A, YIF1B, UNC50
Proteasome (18)	3E-05	PSMD10, VCP, PSMD14, PSMB8, PSMD8, PSMB6, PSMA5, PSMB7, PSMA6, PSMB4, PSMB5, PSMA4, PSMD4, PSMB2, PSMA1, PSMC4, PSMD1, PAAF1
chromosome/mitotic spindle (37)	1E-04	SPAG5, ARL3, PLK1, KIF11, SKA3, SKA1, AURKB, AURKA, SKA2, TBL1XR1, KIF4A, CDK1, KIFAP3, CENPE, CENPF, BUB1B, CDT1, SEC13, CENPW, ANAPC16, KIF22, DYNLL1, KNSTRN, CENPI, TRAPPC12, NEK2, NUP43, BUB1, MAD2L1, CKAP2L, KIF23, CDC14B, ATAT1, TPX2, RACGAP1, ESPL1, ECT2
secretory granule (44)	3E-03	VAT1, VCP, GRN, ANXA2, FUCA1, FUCA2, GGH, PA2G4, MPO, PRDX6, PYCARD, DNAJC3, NPC2, GLB1, IMPDH1, NAPRT, CREG1, PSMD1, AGA, GUSB, GLA, ARSB, CSTB, PSMD14, MVP, GSTP1, FGL2,

ALAD, PSMB7, PNP, PRDX4, TIMP2, GYG1, ATG7, CTSD, PGM1, CTSB, HSPA8, COMMD3, MIF, DERA, PSMA5, PKM, CSNK2B

**Table 6-6 | GO Cellular Component enrichment of Downregulated Genes in ASCs across species (Microarray)**

Term	p-value	Gene Symbol
nucleus (245)	3E-05	DAZAP2, RBM25, DDX46, PHF7, PNN, RBM4, SART1, DUSP10, RSRC1, POLI, SRRM2, PNISR, BNIP3L, TCF12, THOC1, DYRK1A, PRPF40A, PPP4R3B, REXO4, UHMK1, ATXN2L, DDX39A, SYF2, KAT6A, CRY2, BMP2K, SRSF3, LUC7L3, FAM76B, NCBP3, MAML3, PPIG, SRSF6, RAF1, CBL1, SLU7, SRSF8, FYTDD1, AP5Z1, RBM8A, YTHDC1, WBP4, OSGEP, TAF5L, BCLAF1, WRN, LPXN, SNRPB2, SF3B1, RBM39, SF3A1, CPSF6, SMURF2, WTAP, CSNK1A1, PUM1, MAPK14, SON, STK17A, APEX1, SPRTN, ACIN1, TRIP12, CDK12, CDK13, GORAB, HNRNPR, IKZF1, TTF1, PWP1, RPS19, MYC, OSBP, KAT7, MAF1, RPS10, MED1, NOM1, AGTPBP1, SCAF11, DIMT1, NGDN, CIRBP, WDR75, DNTTIP2, VRK1, SDHA, FBXO11, MIF4GD, THAP2, ILF3, NEIL2, SLBP, ZCCHC7, L3MBTL3, DHX9, SNU13, CDKN2AIP, FXR1, ATXN3, EMG1, CAMTA1, BRX1, MDN1, SPTBN1, ZC3H15, BCL9L, PLK3, ZBTB14, PNO1, IQSEC1, NUP153, TBCA, ELL3, PUM3, MOB1B, ABTB1, RPS25, DIAPH2, GNL3L, UBLCP1, PLEKHM1, NOL11, BRWD1, SETD4, PHF2, ZCCHC11, CHD3, KRI1, CHD2, RBM3, XPO1, AKAP11, XPO6, RBM34, UTP15, DDX54, AEN, URB1, TERF1, SIRT1, ZFX, SEPT2, RBL2, ZEB2, PPM1B, RCN2, SDHAF2, KDM7A, NUFIP1, DDX27, DDX24, NIP7, RPL11, HEATR1, DDX21, CAPG, MALTI1, RELA, PNMA1, TAF1C, ZNF506, SAP30L, ATXN1L, MKNK2, TRA2A, RPL18, ZNF346, E2F5, SUZ12, PCGF5, CRIPT, DDX31, YY1AP1, NIN, CDKN2AIPNL, BMT2, TTC3, PTPN6, CIAPIN1, ERCC6, NOP53, MPHOSPH8, RB1, CIITA, MORC3, SPI00, MRE11, UBE2I, SATB1, SMC6, SENP2, PIAS2, PIAS1, KLHL20, SKI, ZMYM2, ELF4, SUMO2, EIF3E, ATR, RIF1, ICE2, TESK2, TCF20, AFF3, CDC14A, CAPN7, PAPOLG, UIMC1, SH3BGR13, NELFA, RAB8B, KLF11, USP7, TERF2, NCOR2, ELF2, SLTM, KCTD13, GEMIN4, ZNF350, RNF34, DTX1, STK4, ZBTB4, FAAP20, ADD1, PCNP, TP53BP1, BANP, TOE1, ATF7IP, CREBBP, NFATC1, INO80, USP28, POU6F1, CLK2, SETX, GON4L, AGO4, FBXL3, RETREG1, SUGP2
cytoplasm (12)	4E-03	NFKBIA, REL, RELA, NFKB1, RELB, GSK3B, APC, CSNK1A1, SIAH1, AXIN1, CTNNB1, DACT1

**Table 6-7 | Differentially upregulated genes that are known/ predicted to localise to ER, Golgi, proteasome and secretory granules. Includes genes enriched for ER stress response, anti apoptotic processes, membrane trafficking, glycosylation, and antigen presentation as indicated by GO enrichment analysis.**

Gene Symbol (Total 534)
ERO1A, DPAGT1, VKORC1L1, KDELR1, PLOD3, PLOD1, TM7SF2, FADS2, NSDHL, TMEM147, SLC39A7, SEC62, FADS1, SEC63, ALG8, ALG9, ALG5, ALG2, ALG14, ALG3, MOGS, ACSL4, ALG11, ATG13, PDIA6, PDIA5, DHDDS, ERN1, SELENON, SELENOS, TMX1, PREB, KDELR2, KDELR3, ATF6, RNF121, COPB2, RTN3, COPA, VCP, TMED10, RETSAT, RPN2, ABHD4, LMAN2L, SAR1B, INSIG2, RPN1, SEL1L, MGST2, TMTC2, HTRA2, MTDH, HSP90B1, APOO, PORCN, EBP, DDRGK1, HMOX1, SSR1, MLEC, KDSR, SEC31A, NUS1, SDF2L1, HSPA5, WFS1, RHBDD1, SRD5A3, SURF4, ERLIN1, DHCR24, DOLK, TMEM110, EXT1, NCEH1, GGCX, CDK1, STT3A, DHCR7, COPE, PIGU, PIGT, PIGN, GPAA1, TECR, ATP2A2, MIA3, IER3IP1, HERPUD1, SEC61A1, SERP1, SEC61G, TMED3, TMED2, SEC61B, SDF2, TMED7, TMED9, SEC13, COPZ2, ELOVL6, LSS, COPZ1, TAPBP, TMEM33, DAD1, CANX, RFT1, PIGK, PIGM, SEC22B, PIGF, MGLL, ATF6B, HM13, GOSR2, STX18, DERL1, FITM2, SELENOK, AGPAT1, LMF1, LMAN1, EXTL2, ASNA1, DNAJB9, CD59, SPTSSA, CDIPT, LNP, APOL2, SEC11C, SEC24A, GALNT2, SYVN1, DHRS3, YIF1A, ARCN1, DHRS4, YIF1B, RAB10, ERP44, SPCS2, CD4, SPCS1, DHRS9, SCD, DOLPPI, COG1, SEC24D, LPIN1, GOLT1B, TXNDC11, APMAP, CREB3L2, CAPN2, SIL1, PDIA3, MAP2K2, ENTPD5, DNAJB12, P3H1, CNPY2, TEX2, MR1, PDIA4, CPED1, TPST2, MINPP1, TMEM258, DNAJB11, PPIB, CES2, LRPAP1, LRRC59, FKBP2, ABCB6, HACD3, GIGYF2, EDEM3, NDFIP2, XBPI, CRTAP, VCAM1, EDEM2, BFAR, TOR1B, TMEM230, MANF, CCDC88A, EHD4, TBL2, CALU, EBFL, PSMG1, ACO1, CALR, STT3B, CPEB4, CD320, TRAM1, CCDC47, UFSP2, MAGT1, MAN1A2, UFD1, TMED1, IKBIP, PSENEN, JAGN1, SSR2, SIGMAR1, GPX7, CKAP4, DNAJC3, CREB3, DNAJC1, RCN1, HAX1, TMEM214, PEX3, GRN, ATL3, USO1, RRBPI, CLN6, CLN5, PGRMC1, FAM213B, SRP72, ORMDL2, ORMDL3, EMC2, MAN1C1, STOM, AGA, YKT6, SEC23B, UGGT1, TMEM30A, ERAP1, YIPF5, SRP68, DHRS1, TMC01, VMPI, YIPF6, STUB1, P4HB, PTPN2, SLC35B3, SLC35B1, DERL3, TEX261, ZMPSTE24, BSCL2, OST4, KRTCAP2, OSTC, GOLIM4, RABEPK, TMEM187, GOLGA5, AP3S1, YIPF2, YIPF3, CAV2, M6PR, TRIP11, SEC23IP, LMAN2, NUCB1, TMED5, MANEA, HID1, GOLGA3, GOLGA4, UNC13B, RAB2A, SLC30A7, SLC30A6, AKR1E2, F2R, FNDC3A, ESCO2, LAX1, SCAMP2, PREPL, PARM1, FAM114A1, ARL3, GLT8D1, ARL1, SLC38A10, GLG1, ARHGAP21, FUT8, MYO6, YES1, ATP8B2, SLC35E1, PTCH1, RAB39B, CDC6, ACER2, BST2, B4GAT1, PPT1, IFT27, ITM2B, TP73, ITM2C, DPY30, LITAF, BICD1, TRIM68, SDF4, SLC39A11, PPHLN1, ARFGAP3, BET1L, PLSCR1, B3GNT9, CTTN, RER1, GORASP2, IMPAD1, TMEM5, SLC29A3, LEPROT, NOD2, CLCN3, TM9SF4, TMEM165, WWOX, QPCTL, PDE4DIP, PLEKHA8, ATAT1, GOLPH3, GLB1, GLA, CHPF, HS2ST1, NDST1, CIGALTIC1, BSG, PMEP1, ST6GAL1, COG6, ENTPD4, CHPF2, GOLPH3L, NOSIP, WPI1, ST6GALNAC4, ST6GALNAC6, CHST11, CHST12, MAN2A1, DSEL, ST8SIA4, MGAT3, MGAT1, ST3GAL2, B3GAT2, SLC33A1, B3GALT6, C6ORF89, UNC50, PSMD10, PSMD14, PSMB8, PSMD8, PSMB6, PSMA5, PSMB7, PSMA6, PSMB4, PSMB5, PSMA4, PSMD4, PSMB2, PSMA1, PSMC4, PSMD1, PAAF1, VAT1, ANXA2, FUCA1, FUCA2, GGH, PA2G4, MPO, PRDX6, PYCARD, NPC2, IMPDH1, NAPRT, CREG1, GUSB, ARSB, CSTB, MVP, GSTP1, FGL2, ALAD, PNP, PRDX4, TIMP2, GYG1, ATG7, CTSD, PGM1, CTSB, HSPA8, COMMD3, MIF, DERA, PKM, CSNK2B, FKBP14, SHC1, MYDGF, LMNA, GFPT1, ARFGAP1, CTDSP2, BHLHA15, CTH, CEPB, UBA5, FAM129A, FLOT1, TRIB3, UBE2J1, ERLEC1, KAMPE, FBXO6, ANAPC16, BUB1B, TNFAIP1, PSMD9, RNF115, UCHL1, CDC23, HECTD3, CDC26, SPOP, GTSE1, UBE2C, CDC34, GID4, TBL1XR1, RNF181, CLOCK, RNF187, MAD2L1, TRIM25, UBE2M, TNFRSF13B, TRADD, FOXO3, TXNDC17, TNFRSF13B, TNFRSF17, CDIP1, LIMS1, GPX3, TXNLI, NUDT2, PYCR1, PYCR2, PRDX3, PRDX2, PRDX1, SESN2, CYCS, CPEB2, NUDT14, PGM3, PQLC3, ITGB5, DYNC1I2, FCER1G, DNT3, CLTA, KIF23, KIF11, CTSV, KIF22, AP2B1, DYNLL1, AP2A2, CENPE, KIF18A, RACGAP1, KIF4A, KIFAP3, AVEN, ARF4, CITED2, APIP, SH3RF1, HMGNS, CCND2, CASP3, PIM2, ERC1, PRKCI, ANXA5, BIRC5, NAA38, ATF5, GAS6, FXN, LTF, NOL3, FAM129B, SOCS2, TAF9B, MCL1, PRELID1, NPM1, STIL, PLK1, ASNS, CFLAR, VEGFA, IL2RB, FAS, BCL2L2, BCL2L1

**Table 6-8** | Tables shows differentially **upregulated** genes in ASCs that are known to be upregulated as a result of TF overexpression/induction/gain of function mutation.

Transcription Factor	Adj <i>p</i> -value	Desired Genes (Upregulated)	Other genes (Upregulated)
XBP1 Overexpression	7E-49	ERO1A; SLC35B1; TXNDC17; PNP; SLC39A7; PDIA3; MOGS; PDIA6; PDIA5; PDIA4; SELENOS; DNAJB11; PREB; KDELR3; PPIB; LRRC59; COPB2; FKBP2; COPA; VCP; RPN2; PSMD14; UBA5; RPN1; SLC38A10; ZMPSTE24; HSP90B1; EBP; MAN2A1; HMOX1; SSR1; EDEM3; SDF2L1; HSPA5; WFS1; SURF4; EDEM2; PRDX6; MANF; OSTC; CALU; STT3A; RNF181; CALR; COPE; ITM2C; ARF4; MIA3; MAGT1; HERPUD1; SEC61A1; SERP1; SEC61G; TMED3; TMED2; SEC61B; SDF2; PSENNEN; TMED9; SSR2; DYNLL1; ARFGAP3; PSMA5; DNAJC3; HAX1; GORASP2; DAD1; GSTP1; USO1; DERL1; RRBP1; SELENOK; LMAN1; LMAN2; DNAJB9; CDIPT; SEC23B; SEC11C; YIPF5; SYVN1; SRP68; YIF1A; ERP44; SPCS2; SPCS1; NUCB1; P4HB; SEC24D; FAM129B; KDELR1; CITED2; CCND2; AP3S1; CTSD; CTSB; FCER1G; ANXA5; PKM; NPC2; LTF; GRN; CEBPB; PRDX1; CREG1; PRELID1; XBP1; VCAM1; IL2RB; ITM2B; TM7SF2; BSG; KIFAP3; PGM1; KIF22; TBL1XR1; RTN3; TM9SF4; PGRMC1; SPOP; TMEM230; PPT1; COPG1	LGALS1; HDLBP; MT2A; IER3; ARMCX3; SELENOM; NUPR1; NFXL1; ISG20; GPR180; CLPTM1L; NOP10; CCDC127; MRPS12; SSR4; SSR3; NME1; CRELD2; BCAT1; ALDH9A1; SELENOF; GLRX; SLC3A2; TM9SF3; SRM; TCEAL8; TMEM208; RABAC1; EMC7; AAMP; TCEAL9; SLC16A1; RRM2; TMEM176B; NDUFA1; ATOX1; CDK2AP2; FAM98A; LGALS3BP; IGHM; IFITM1; COX6A1; JCHAIN; SPN; IGHG1; LGALS1; MRC1; ATP6AP2; AKR1A1; TALDO1; TOX2; S100A9; HIST1H2BC; C1QB; SELPLG; POMP; ABRACL; COX7A2; OAZ1; COX5A; NDUFV3; CYB5A; NDFIP1; MDH2; NDUFA2; DNAJC15; DAP; RNF149; CD28; RGS10; MSRB1; GRINA; NRP1; PLEKHF1; QPRT; SRP54; GPHN; DTWD1; DEPTOR; IMPA2; BASP1; NMRK1; ARFIP2; TNS3; MITF; ACYP2; HDHD3; GNL3; SLC7A5; HIST1H2BE; LAP3; HIST1H2BG; NUP62CL; SLC26A2; CLDND1; ETFDH; PRCP; GPNMB; UAP1L1; TAF12; FAH; BSPRY; TM9SF2; PSAT1; GALK2
ATF6 Induction	1E-11	ERO1A; HERPUD1; RNF115; TMED2; GYG1; RAB2A; PDIA3; ANXA5; PDIA6; PDIA4; PSMA5; DNAJC3; PSMA4; CANX; DNAJB11; GRN; VCP; PSMD14; SAR1B; ARL1; HSP90B1; PRDX3; APOO; PSMB4; PRDX1; EMC2; DNAJB9; SPTSSA; HSPA8; NPM1; HSPA5; MANF; EHD4; CALU; CYCS	ETFA; GCSH; MT2A; LAMP2; VPS35; ACADM; IER3; HIST1H1C; GSTO1; SSR3; NEAT1; IMMPL1; PPA1; SUB1; HIST1H2BC; PCNA; NDUFB6; POMP; NDUFB3; RPS27L; SELENOF; PDHB; MFF; OAZ1; TMEM70; ACAT1; RAPIA; PDLIM5; MCTS1; CYB5A; NDFIP1; PDHA1; MDH1; GOT1; MTX2; DSTN; EIF2S2; AZIN1; CCT6A; QDPR; RCAN1
POU2AF1 Overexpression	8E-09	PNP; SERP1; HMGNS5; PSMA5; PSMA6; PSMD14; PSMB6; SOCS2; PGRMC1; PSMB4; XBP1; NPM1; SYVN1; DHRS3; PSMC4; PQLC3; CDK1; MAD2L1; RNF187	IFITM1; NCAPG2; HMGB3; MKI67; GCSH; PBK; BLNK; PHGDH; FBXO5; CHCHD1; CPT1A; GAS2L3; RFC4; GNL3; CCNA2; CCNE2; MZB1; CKS2; KIF20A; LXN; PCNA; NDUFB6; UHRF1; CDCA5; MRPL18; CDCA8; COX7A2; TYMS; FAM117A; SRM; CCNB2; CCNB1; MAT2A; E2F2; CCT7; QSER1; SLC16A1; RRM2; MDH2; SHCBP1; CCT6A; POU2AF1; ADA

**Table 6-9** | TF co-expression results for genes differentially upregulated genes in ASCs

Transcription Factor	EnrichR Adj <i>p</i> -value	Genes of Interest	Others
TCF19	8E-12	KIF11; KIF23; KIF22; PKM; LRRC59; RACGAP1; YKT6; PRELID1; CDC6; PA2G4; KIF4A; CALR	FEN1; CSE1L; NCAPG2; MCM10; MRPL37; MKI67; FOXM1; IPO4; LMNB1; EXO1; MYBL2; TK1; CEP55; CCT3; PRMT1; IFRD2; CAD; TUBG1; SND1; CDC25A; CCNA2; MELK; SAPCD2; ESPL1; TMEM106C; KIF20A; TTLL12; MCM6; GART; DTL; PCNA; RNASEH2A; CDCA5; NCAPG; TYMS; SLC7A1; NCAPH; UNG; SRM; RECQL4; GANAB; CDC45; POC1A; RAD54L; CCT7; CTNNAL1; PLK4; CDT1; CENPU; CDKN2C; RRM2; SPAG5; HMGA1; RRP1B; SHCBP1; ZWINT; TPX2; ANLN; RAD51; CARM1; NCAPD2; CLPTMIL; EIF3C
CREB3L2	1E-16	TRAM1; CHPF; KDELR1; PLOD3; PLOD1; GOLGA3; NDST1; SEC61A1; TIMP2; CAPN2; SLC39A7; PDIA3; CKAP4; PDIA4; TMEM214; GORASP2; IMPAD1; CTDSP2; CANX; KDELR2; KDELR3; COPA; FAM114A1; SHC1; SEL1L; TEX261; RRPB1; TM9SF4; FAM129B; HSP90B1; LMAN1; MLEC; SSR1; CD59; UGGT1; SEC31A; CRTAP; SEC24A; HSPA5; WFS1; GALNT2; SURF4; SYVN1; ARCN1; EXT1; CALU; CALR; COPG1; P4HB	ITGB1; CLIC4; RAPH1; HDLBP; LAMC1; TRAM2; ALDH1L2; AFF1; TNS3; C11ORF24; ARHGEF12; ZBTB38; SSR3; FNDC3B; SLC39A14; SND1; PCYOX1; H6PD; EPAS1; ABHD2; SLC7A1; TM9SF3; ITPRIL2; LIMA1; GANAB; PRRC1; PGRMC2; TMEM184B; ATP2B4; ELL2; DAP; CPD; ASXL2; PLXNB2; ALDH18A1; NFE2L1
RAD51	3E-27	BUB1B; ESCO2; PSMD14; PRDX1; UBE2C; PLK1; CDC6; PA2G4; KIF4A; PSMG1; KIF11; KIF23; KIF22; PSMA6; PSMA4; BIRC5; PSMB6; PSMB2; NPM1; MAD2L1	CSE1L; GMNN; NUDT1; TMEM97; UBE2L3; SMC2; CHEK1; RUVBL1; NUSAP1; BANF1; OIP5; FBXO5; CDC25A; MELK; CCNE2; CCNE1; CDCA2; RNASEH2A; CDCA3; CDCA5; NCAPG; CDCA8; SKA3; MRPL13; NCAPH; CCNB1; EIF4EBP1; LYAR; PLK4; YARS; CDT1; FAM136A; ZWINT; TPX2; DIAPH3; UBE2T; PAK11P1; NCAPG2; MCM10; MRPL37; BRCA1; ATP5G1; CKS1B; EXO1; NUF2; PBK; MYBL2; TK1; TCF19; CCT3; TSM; RFC3; RFC4; ACTL6A; NME1; HAUS1; CCNA2; MCM6; GART; DTL; DTMYK; PCNA; PRIM1; TYMS; AURKB; UNG; PSMC3IP; HIRIP3; CDC45; POC1A; RAD54L; CCT7; BUB1; GINS1; CENPU; GINS2; PPIL1; RRM2; SPAG5; SHCBP1; AUNIP; SPC24; CDKN3
CREB3	3E-18	ARF4; CLTA; PSMD8; MYDGF; SEC61A1; NSDHL; TMEM147; TMED3; TMED2; PSENE1; RAB2A; SEC13; ANXA2; ANXA5; COPZ1; BSCL2; GORASP2; DAD1; CSNK2B; KDELR2; KDELR3; COPB2; FKBP2; COPA; FAM114A1; HM13; GOSR2; DCTN3; ARL1; PSMB6; PSMB7; ASNA1; SEC23B; SEC31A; PRELID1; YIPF3; YIPF5; YIF1A; TMC01; ATAT1; PSMC4; CALU; CYCS; COPG1; COPE	SRP54; DUSP14; RABGEF1; TRMT112; LGALS1; MRPL40; GUK1; PPM1; ARFIP2; NANS; PELO; C11ORF24; GLRX3; SARS; FTS1; NDUFS8; TMX2; AKIP1; AURKAIP1; ERGIC3; PHPT1; PAFAH1B3; CUTA; MAGEH1; CD63; MAGED1; TULP3; SLC3A2; MFF; PLD3; TMEM205; EMC4; RABAC1; EMC7; ATP6V1D; AAMP; POLR2L; MDH1; SLC31A1; RMDN3; DSTN; MARS; MAGED2
MYBL2	1E-07	BUB1B; KIF11; KIF22; CLN6; RACGAP1; PRELID1; PLK1; PYCR1; CDC6; PA2G4; KIF4A; UBE2M	FEN1; CSE1L; NCAPG2; MRPL37; MKI67; FOXM1; IPO4; LMNB1; MRPL4; EXO1; PHGDH; TCF19; CCT3; PRMT1; CAD; MRPS2; SND1; CDC25A; CCNA2; SAPCD2; ESPL1; KIFC1; KIF20A; TTLL12; MCM6; RABL6; PCNA; CDCA5; NCAPG; TYMS; NCAPH; AURKB; UNG; SRM; RECQL4; GANAB; CDC45; NT5DC2; RAD54L; CCT7; BUB1; YARS; CDT1; RRM2; SPAG5; RRP1B; ZWINT; TPX2; CARM1; NCAPD2; CLPTMIL; EIF3C; AARS
RIOK2	5E-04	GOLT1B; MAGT1; IER3IP1; SLC30A7; ALG8; PSMA5; PSMA4; PSMA1; TMX1; PSMD14; PRDX3; PSMB6; PSMB5; NPM1; TXNL1; SPCS1; OSTC; MAD2L1	SUV39H2; PAK11P1; CSE1L; GMNN; THYN1; CCNC; EPRS; MRPL42; NUDCD1; PPAT; USP46; ACTL6A; SDHD; GTF2F2; MRPS18C; THAP1; NME1; HAUS1; MRPL50; ANAPC13; TXNDC9; NDUFB6; GTF3C6; RPE; PDHB; MRPL13; CRYZ; TCEAL8; ORC3; ZNF420; DNAJC19; MDH1; CCT6A; NDUFAF4; NOP10; MPHOSPH6
E2F7	3E-25	BUB1B; GTSE1; RACGAP1; STIL; PLK1; CDC6; PA2G4; KIF4A; CALU; CDK1; KIF11; KIF23; KIF22; PKM; BIRC5; NPM1; MAD2L1	CSE1L; GMNN; MKI67; SMC2; CHEK1; NUSAP1; NEK2; FBXO5; CDC25A; MELK; KIF20A; BLM; CDCA2; RNASEH2A; CDCA5; NCAPG; CDCA8; TARS; SKA3; NCAPH; KIAA1524; ECT2; PLK4; HMGA1; ZWINT; CCT6A; TPX2; ANLN; DIAPH3; UBE2T; IARS; CLIC4; FEN1; NCAPG2; MCM10; BRCA1; BRCA2; FOXM1; CKS1B; LMNB1; EXO1; NUF2; PBK; MYBL2; CEP55; TCF19; RFC4; CKAP2L; ACTL6A; TUBG1; CCNA2; ASPM; ESPL1; KIFC1; DEPDC1; GARS; MCM6; NUP62CL; GART; DTL; FAM72A; PCNA; PRIM1; TYMS; SLC7A1; UNG; RECQL4; CDC45; KDELC2; RAD54L; CCT7; CTNNAL1; BUB1; GINS1; CENPU; GINS2; RRM2; SPAG5; RRP1B; SHCBP1; CENPI; TSRI; NCAPD2
TCEAL8	4E-17	DPY30; IER3IP1; TXNDC17; CDC26; PSMA6; PSMA4; PSMA1; TMEM258; TBL1XR1; TMX1; DAD1; PIGF; PSMD10; PSMD14; PRDX3; PSMB6; APOO; EXTL2; PSMB5;	HDCC2; NDUFA12; PAK11P1; GMNN; THYN1; BRK1; CCNC; MRPL36; HIKESHI; COX6A1; CKS1B; GCSH; TRMT112; MRPL42; YAE1D1; PHACTR2; HMGN3; TMEM14C; GLRX3; USMG5; SDHD; MRPS18C; ATP5J2;

**Table 6-9** | TF co-expression results for genes differentially upregulated genes in ASCs (cont.)

Transcription Factor	EnrichR Adj <i>p</i> -value	Genes of Interest	Others
		PSMB2; NPM1; SPCS2; SPCS1; OSTC; MAD2L1	HAUS1; MRPL50; TMEM256; MRPL51; SDHAF3; SUB1; NDUFS5; NDUFS4; CKS2; BCAT1; ANAPC13; NDUFB6; GTF3C6; POMP; NDUFB3; ATP5J; ABRACL; COX7A2; MRPL15; UQCR10; PDHB; MRPL13; POLR2I; DNAJC19; TIMM8B; MDH1; UQCC2; NDUFA3; NDUFA2; MRPL27; CRIPT; NDUFA1; FAM136A; NDUFAF4; NDUFAB1; NDUFAF2; NOP10; MPHOSPH6
ATF6	1E-03	GOLGA4; CREB3L2; TMED5; SLC30A7; ANXA2; ANXA5; FNDC3A; DNAJC3; CANX; PIGK; SEC22B; COPA; ATL3; FAM114A1; SAR1B; SEL1L; FAM129A; ARL1; LMAN1; SSR1; CD59; SEC31A; EDEM3; SEC24A; ERAP1; CALU; TRIP11; FAS; SEC24D	NRP1; RAPH1; IDE; EPRS; LAMP2; PHACTR2; TMED4; ZBTB38; SSR3; FNDC3B; LRRC41; HEATR5A; ATG4A; GBA; TOR1AIP2; ABHD5; ITPR1L2; SLC17A5; TMEM176B; MGA; ELL2; OSTM1; REEP3; CPD
E2F3	1E-03	BUB1B; UGGT1; NUS1; NPM1; CDC6; PA2G4; KIF4A; MAD2L1	FEN1; NCAPG2; MCM10; MKI67; IPO4; LMNB1; LAPTM4B; EXO1; MYBL2; FBXO5; TCF19; RFC3; RFC4; CAD; CDC25A; CCNA2; ESPL1; KIFC1; DEPDC1; MCM6; TOP1; GART; DTL; PCNA; CDCA5; NCAPG; TARS; TYMS; SLC7A1; UNG; RECQL4; CDC45; RAD54L; BUB1; PLK4; GINS1; CDT1; CENPU; SLC16A1; SPAG5; RRP1B; ZWINT; TPX2; EIF3J; NCAPD2
LARP4	1E-14	CCDC47; BUB1B; KIF11; GOLGA4; KIF23; CANX; USO1; ZMPSTE24; HSP90B1; PRDX3; LMAN1; PRDX1; HSPA8; NPM1; YES1; SEC24A; CDC6; PA2G4; SEC23IP; KIF4A; CALU; MAD2L1	ABCD3; PAK1IP1; CSE1L; NCAPG2; CCNC; MCM10; IDE; EPRS; CDC73; EXO1; PPAT; GNPAT1; SKP2; DLGAP5; CCT3; CSNK2A1; DESI2; SSR3; CDC25A; GNL3; MELK; MTHFD2; DEPDC1; GARS; MTFR1; MCM6; GART; DTL; CDCA2; GMP5; NCAPG; HSD17B4; TARS; HNRNPLL; SKA3; RPAP3; TM9SF3; KIAA1524; CCNB1; CCT7; ECT2; CTNNAL1; BUB1; QSER1; PLK4; PP1L1; SLC16A1; SPAG5; RRP1B; FAM136A; BLOC1S5-TXNDC5; CCT6A; TPX2; ANLN; DIAPH3; PSAT1; TSR1; IARS; FAM98A
FOXM1	4E-14	BUB1B; KIF11; BSG; SLC39A7; GTSE1; KIF23; KIF22; PKM; BIRC5; LRRC59; RACGAP1; YKT6; PLK1; PYCR1; CDC6; PA2G4; KIF4A; CALR; UBE2M	FEN1; CSE1L; NCAPG2; MKI67; IPO4; LMNB1; EXO1; PBK; PHGDH; MYBL2; CEP55; TCF19; CCT3; PRMT1; CAD; SND1; CDC25A; CCNA2; MELK; SAPCD2; ESPL1; KIFC1; DEPDC1; KIF20A; TLL12; MCM6; GART; RABL6; PCNA; CDCA3; CDCA5; NCAPG; FOXK2; TYMS; SLC7A1; NCAPH; AURKB; SRM; RECQL4; CCNB1; GANAB; NT5DC2; RAD54L; CCT7; ECT2; BUB1; CDT1; RRM2; SPAG5; HMG1A; RRP1B; SHCBP1; ZWINT; CCT6A; TPX2; ANLN; KIF18B; CARM1; CLPTM1L; EIF3C; AARS
POLR2L	8E-28	TXNDC17; PSMD8; TMEM147; MIF; TMEM258; PPIB; FKBP2; PRDX4; OST4; PRELID1; RNF181; COPE; DPY30; TMEM187; MYDGF; SEC61B; PSENEN; TMED9; DAD1; NAA38; PSMB6; PSMB7; PSMB2	CCDC167; NDUFA11; NDUFA12; CISD1; NUDT1; HIKESHI; COX6A1; MT2A; LGALS1; BANF1; CHCHD1; TMEM14C; ATP6V0E1; TRAPPC2L; MRPS18C; ATP5J2; TMEM256; BOLA3; BOLA2; AURKAIP1; PHPT1; CUTA; RNASEH2C; NDUFB11; MRPL17; ATP5J; COX7A2; ATP5H; MRPL13; MRPL20; EIF4EBP1; DNAJC19; COA4; UQCC2; MRPL27; LSM2; COPS6; UQCRQ; NDUFAF2; NOP10; MRPS15; MRPS12; BRK1; MRPL36; ATP5G1; CKS1B; TRMT112; UBL5; TMEM223; TIMM17B; MRPS24; ATP1F1; USMG5; NME1; RCN3; MRPL53; MRPL51; NDUFS8; NDUFS5; CD63; NDUFB7; NDUFB6; POMP; NDUFB3; RPS27L; NDUFB1; UQCR10; MALSU1; HSD17B10; POLR2E; RABAC1; POLR2I; POLR2J; TIMM8B; NDUFA8; NDUFA3; NDUFA2; NDUFA1; ATOX1; FKBP1A; NDUFAB1
E2F2	1E-13	BUB1B; KIF11; PNP; GTSE1; ESCO2; KIF22; BIRC5; ST6GALNAC4; FOXO3; UBE2J1; PRDX2; STIL; CDC6; BCL2L1	FEN1; NCAPG2; CPOX; MCM10; BRCA1; MKI67; BRCA2; FOXM1; SMC2; LMNB1; EXO1; NUSAP1; MYBL2; FBXO5; KMT5A; TCF19; RFC4; CKAP2L; BPGM; CDC25A; CCNA2; MMS22L; ASPM; ALDH5A1; CCNE2; ESPL1; KIFC1; HMBS; MCM6; DTL; BLM; PCNA; RNASEH2A; CDCA3; FECH; CDCA5; ARHGAP19; NCAPG; CDCA8; TYMS; NCAPH; SKA1; AURKB; FAM117A; RECQL4; BRIP1; CDC45; POC1A; RAD54L; DCAF12; PLK4; GINS1; CDT1; CENPU; CDKN2C; RRM2; EIF2AK1; XRCC3; SHCBP1; ZWINT; TPX2; KIF18B; UBE2T; NCAPD2; SPC24
GTF2F2	4E-15	DPY30; TXNDC17; CDC26; PSMA5; PSMA4; TMX1; PSMD14; PSMB6; PSMB7; EXTL2; PSMB5; PSMB2; PRDX1; NPM1; OSTC; CYCS; MAD2L1	SUV39H2; PAK1IP1; CSE1L; GMNN; HIKESHI; TMEM97; COX6A1; CKS1B; GCSH; NUDCD1; PPAT; RIOK2; OIP5; CHCHD1; TMEM14C; RFC3; RFC4; GLRX3; USMG5; MRPS18C; NME1; ATP5J2; HAUS1; MRPL50; MRPL51; SDHAF3; CCNE1; BOLA3; PTRH2; NDUFS5; CKS2; PCNA; NDUFB6; GTF3C6; POMP; NDUFB3; MRPL15; UQCR10; MALSU1; SKA3; MRPL13; SKA2; TCEAL8; MRPL20; CCNB1; LYAR; ATG5;

**Table 6-9** | TF co-expression results for genes differentially upregulated genes in ASCs (cont.)

Transcription Factor	EnrichR Adj <i>p</i> -value	Genes of Interest	Others
			DNAJC19; COA4; TIMM8B; CENPU; GINS2; PPIL1; CENPW; SLC16A1; FAM136A; EIF2S2; CCT6A; DIAPH3; NDUFAF4; NDUFAB1; NDUFAF2; NOP10; CDKN3; MPHOSPH6
HMGB3	1E-15	BUB1B; KIF11; TXNDC17; HS2ST1; GTSE1; ST6GAL1; KIF23; PSMA4; BIRC5; PSMD14; PSMB5; PSMB2; PRDX1; NPM1; UBE2C; PLK1; CDC6; CENPE; KIF4A; CDK1; MAD2L1	PAK1IPI; CSE1L; GMNN; MCM10; TMEM97; CKS1B; PPAT; CHEK1; NUF2; NUSAP1; PWWP2A; PHGDH; BANF1; OIP5; DLGAP5; RFC3; RFC4; CSNK2A1; CKAP2L; CDC25C; KNSTRN; CDC25A; NME1; HAUS1; CCNA2; SGO2; ASPM; MELK; NDUFS5; DEPDC1; CKS2; KIF20A; ALDH7A1; CDCA2; RNASEH2A; CDCA3; NCAPG; TARS; TTK; SKA3; NCAPH; AURKB; UNG; AURKA; CCNB2; KIAA1524; CCNB1; ECT2; BUB1; GINS1; CENPU; GINS2; PPIL1; SLC16A1; RRP1B; FAM136A; ZWINT; TPX2; CENPF; PSAT1; UBE2T; CDKN3
ZBTB38	8E-09	TRAM1; TNFRSF13B; TXNDC11; SEC61A1; CREB3L2; TIMP2; CAPN2; ERN1; TMEM214; IMPAD1; KDELR3; GAS6; ATL3; FAM114A1; SEL1L; RRBP1; UBE2J1; HSP90B1; LMAN1; MAN2A1; TNFRSF17; CD59; UGGT1; EDEM3; SEC24A; TMEM30A; ATP8B2; CAV2; ERAP1; EXT1; CALU; FAS; SEC24D	ITGB1; CLIC4; SLC35F5; HDLBP; LAMC1; TRAM2; ALDH1L2; GLS; MFSD6; LGALS1; CSRPI; ANXA6; IGLC1; TNS3; DUSP5; ARHGEF12; SSR3; FNDC3B; TWSG1; IRF4; SLC41A2; TM9SF3; CAMSAP2; ITPRIPL2; LIMA1; ATXN1; OSBPL3; ATP2B4; ELL2; IGLL5; DPY19L1; CPD; CLPTM1L; SERINC3; NFE2L1
MLX	1E-14	SLC35B1; PSMD8; MYDGF; PNP; SSR2; ANXA5; ALG3; SCAMP2; PSMA5; RER1; CSNK2B; PPIB; RNF121; GRN; HM13; MVP; GOSR2; RPN1; MGST2; PRDX3; PSMB6; PSMB7; EBP; PSMB5; PSMB2; MGAT1; PRELID1; PA2G4; YIF1A; MANF; SPCS1; OSTC	MRPS15; ALAS1; PAK1IPI; MRPL36; LSM10; GMPPB; AIFM2; BANF1; SLC39A4; C11ORF24; TCF19; ATP6V0E1; GSTO1; IFRD2; MRPS18A; TALDO1; SDHD; TUBG1; NME1; SCO2; BOLA2; UQCRC1; SLC25A11; CD63; ECHS1; ACSS2; NAXE; GBA; POMP; MRPL17; GLRX; OAZ1; EIF4EBP1; POLR2E; CCT7; PCK2; COA4; MDH2; TMEM176B; GLMP; C20ORF24; KIAA2013; NUDT22; FKBP1A; BCKDK; COPS6; GALE; GALM; MSRB1
DEPDC1	3E-35	BUB1B; GTSE1; PSMD14; HSP90B1; RACGAP1; HSPA8; STIL; YES1; UBE2C; PLK1; CDC6; PA2G4; KIF18A; KIF4A; CDK1; KIF11; KIF22; PSMA4; BIRC5; NPM1; CENPE; MAD2L1	CSE1L; GMNN; EPRS; MKI67; SMC2; PPAT; NUSAP1; NEK2; OIP5; FBXO5; CDC25C; KNSTRN; CDC25A; SGO2; MELK; MTHFD2; KIF20A; BLM; CDCA2; CDCA3; CDCA5; GMPS; NCAPG; CDCA8; TARS; SKA3; NCAPH; CCNB2; KIAA1524; CCNB1; ECT2; QSER1; PLK4; ZWINT; CCT6A; TPX2; ANLN; DIAPH3; UBE2T; IARS; SUV39H2; FEN1; NCAPG2; MCM10; BRCA1; BRCA2; FOXM1; CKS1B; LMNB1; EXO1; NUF2; PBK; GNPAT1; MYBL2; CEP55; DLGAP5; CCT3; RFC3; RFC4; DESI2; CKAP2L; CCNA2; ASPM; SAPCD2; ESPL1; KIFC1; CKS2; MCM6; GART; DTL; FAM72A; PCNA; TTK; CENPA; AURKB; AURKA; POC1A; CCT7; CTNNAL1; BUB1; E2F7; GINS1; CENPU; SLC16A1; RRM2; SPAG5; RRP1B; SHCBP1; CENPF; CENPI; TSR1; NCAPD2; CDKN3
HMGA1	2E-03	KDELR1; CLTA; BSG; SLC39A7; TMED9; PKM; VCP; HM13; ASNA1; LMNA; PRELID1; CRTAP; PLK1; PYCR1; PA2G4; CALR; BCL2L1	FEN1; NCAPG2; FOXM1; IPO4; LMNB1; MRPL4; MYBL2; TCF19; PRMT1; IFRD2; CAD; MRPS2; SND1; CCNA2; SLC04A1; KIFC1; FSCN1; TLL12; CDCA5; UNG; SRM; RECQL4; BCL2L12; GANAB; NT5DC2; POLR2E; CCT7; CDT1; SLC52A2; FKBP1A; TPX2; DAP; CARM1; NCAPD2; EIF3C
BOLA3	5E-24	TXNDC17; CDC26; MIF; TMEM258; OSTC; RNF181; DPY30; SEC61B; DYNLL1; PSMA5; PSMA4; DAD1; NAA38; PSMB6; PSMB7; PSMB5; PSMB2; SPCS1	HDCC2; NDUFA11; NDUFA12; GMNN; CISD1; HIKESHI; COX6A1; UBE2L3; GCSH; BANF1; HMGN3; CHCHD1; TRAPPC2L; ACOT13; MRPS18C; ATP5J2; BOLA2; PHPT1; CUTA; NDUFB11; MRPL18; MRPL17; ATP5J; COX7A2; ATP5H; MRPL13; MRPL20; LYAR; DNAJC19; COA4; UQCC2; MRPL27; UQCRCQ; NDUFAF4; UBE2T; NDUFAF2; NOP10; MRPS15; MRPL36; ATP5G1; CKS1B; TRMT112; UBL5; RFC4; MRPS24; GLRX3; ATP1F1; USMG5; NME1; HAUS1; MRPL51; NDUFS8; NDUFS5; CKS2; NDUFB9; NDUFB7; NDUFB6; GTF3C6; POMP; NDUFB3; RPS27L; NDUFB1; ABRACL; UQCRC10; MALSU1; MRPL57; EMC4; POLR2I; POLR2J; POLR2L; TIMM8B; GINS2; NDUFA8; CENPW; NDUFA3; NDUFA2; NDUFA1; ATOX1; NDUFAB1; MMACHC
ZNF593	1E-16	GPAA1; PSMD8; MYDGF; TMEM147; SEC61B; NUDT14; ALG3; MIF; MINPP1; DNAJC1; RER1; CSNK2B; NAA38; FXN; PRELID1; SDF2L1; B3GALT6; YIF1A; YIF1B; PSMG1; COPE; CD320	MRPS15; NDUFA11; PYCRL; MRPS12; NUDT1; MRPL36; MRPL37; MRPL34; COMT; ATRAID; ATP5G1; UBE2L3; MRPL4; TRMT112; NTMT1; C9ORF16; OIP5; MRPS24; IFRD2; ATP1F1; MRPS2; GLRX2; NME1; NDUFS8; CLPP; BOLA2; PTRH1; HAGH; AURKAIP1; RNASEH2C; DTYMK; ECHS1; NDUFB7; NAXE; NDUFB11; MRPS34; MALSU1; COX5A; MRPL57; LAGE3; UNG; SRM; MRPL20; EXOSC4; MCRIP2; NTHL1; EIF4EBP1;

**Table 6-9** | TF co-expression results for genes differentially upregulated genes in ASCs (cont.)

Transcription Factor	EnrichR Adj <i>p</i> -value	Genes of Interest	Others
			POLR2E; POC1A; KCNN4; POLR2I; POLR2J; POLR2L; CDT1; MDH2; IDH2; HPS6; NMRAL1; LSM2; BCKDK; UQCRC1; NDUFB1; GALK1
THAP4	7E-05	PSMD8; SEC61A1; BSG; SDF4; SLC39A7; PGM1; MAP2K2; SIGMAR1; MOGS; CDC34; PKM; TMEM214; YKT6; PRELID1; SURF4; PYCR1; YIF1A; NUCB1; STUB1; P4HB; COPE; UBE2M	CLTB; HDLBP; MRPL37; IPO4; MRPL4; MYBL2; CCT3; PRMT1; IFRD2; MRPS2; SND1; SAPCD2; CPTP; UQCRC1; TTLL12; AURKAI1; RABL6; POLDIP2; MRPS34; SRM; RECQL4; GANAB; NT5DC2; CLPTM1; EIF4EBP1; POLR2E; CCT7; MDH2; SLC52A2; GOT2; KIAA2013; CARM1; ACO2; EIF3C; NFE2L1

**Table 6-10** | Tables shows differentially **upregulated** genes in ASCs that are known to be upregulated as a result of TF inhibition/downregulation/loss of function mutation

Transcription Factor	Adj <i>p</i> -value	Genes of Interest	Others
FOXO1 Knockout Knockdown siRNA	2.1E-36	F2R; ESCO2; FKBP2; STIL; SDF2L1; UBE2C; PLK1; CDC6; BST2; PQLC3; KIF4A; CDK1; KIF11; LITAF; KIF23; KIF22; BIRC5; CEBPB; FAM129B; CENPE; SEC24D; MAD2L1; TRAM1; CITED2; DPY30; UFSP2; MPO; MANEA; KRTCAP2; SERP1; SEC61G; TRIM25; AP3S1; TMED7; GYG1; PDIA3; TMED9; CKAP4; SCAMP2; DNAJC3; HAX1; PKM; CANX; KDELR2; PPIB; RPN1; FOXO3; TMEM165; HSP90B1; EXTL2; ST8SIA4; MLEC; SSR1; SPTSSA; MANF; SPCS2; SPCS1; OSTC; CYCS; DHCR7; P4HB; TMED2; SEC61B; ANXA5; NOSIP; CNPY2; PSMD14; DCTN3; HSPA5; B3GALT6; YIF1B; CHPF; ATP2A2; PLOD1; BSG; CREB3L2; CTSB; FNDC3A; PDIA6; IMPAD1; LRPAP1; TMED10; PARM1; RPN2; DERL1; LMAN1; MAN2A1; STOM; CD59; CRTAP; GALNT2; SURF4; GFPT1; VEGFA; ARCN1; CALU; STT3A; ACO1; ITM2B; LIMS1; PSMA5; PSMA4; TMX1; DNAJB11; SAR1B; PSMB7; LMFNA; MCL1; SEC11C; NPM1; VMP1; STT3B	IFITM1; ZFAND4; LRR1; GMNN; PRDM1; MKI67; SMC2; IGHG1; LGALS1; ALCAM; SLC16A6; NEK2; FBXO5; ZBP1; GAS2L3; MREG; KNL1; CDC25C; RUNX2; SGO2; MELK; CCNE2; CCNE1; GPR160; KIF20A; ASF1B; CDCA2; EPAS1; CDCA5; NCAPG; CDCA8; UAP1; TAPBPL; NCAPH; ITPRIPL2; CCNB2; BRIP1; CCNB1; SLAMF7; ECT2; PLK4; MLKL; ISG15; ELL2; ISG20; TPX2; KIF18B; UBE2T; ADA; GPSM2; KCNK6; EIF4E3; SOWAHC; NCAPG2; JCHAIN; CKS1B; LMNB1; ACOXL; NUF2; PBK; TK1; IGLC2; CEP55; DLGAP5; TCF19; RFC3; RFC4; CKAP2L; PRSS16; CCNA2; ESPL1; DEPDC1; CKS2; CRELD2; UHRF1; GLRX; TTK; AURKB; PSMC3IP; CXCR3; RAD54L; IL12RB1; FNIP2; BUB1; E2F7; GINS2; CENPW; RRM2; SPAG5; GZMB; SHCBP1; DPY19L1; RAD51; CENPI; SPC24; CDKN3; SPC25; GF11; ETFB; ATP6V0E1; ANKRD46; ELMOD2; ATP6AP2; MED7; ARM CX3; PTRH2; S100A9; DTYMK; PCNA; RPS27L; ABRACL; TYMS; TM9SF3; TSPAN31; MAT2A; METTL6; E2F3; DCAF12; ANKRD28; MDH1; DSTN; DNAJC15; CCT6A; KLHL9; FKBP1A; OSTM1; PSAT1; ACAA2; FAM45A; ETFA; NENF; PDCCD2L; HINT2; BASP1; OPA3; FGFR1OP2; PHGDH; TXNL4A; MRPS24; METRN; ACOT13; H2AFV; UCHL5; SAP30; MRPL53; SDHAF1; UQCRC1; HIST1H2BH; C1QB; ANAPC13; RNASEH2C; NDUFB5; SELENOF; MCRIP2; TMEM205; UAP1L1; RABAC1; LCMT1; TMEM238; RGS10; EEF1AKMT1; GPR19; HSPB6; HDLBP; SLC7A11; LAMC1; CSRPI; LMO4; ACTA2; RCN3; MYADM; BCAT1; PCYOX1; GANAB; PGRMC2; ATXN1; METTL7A; MPZL1; ASS1; SERINC3; CLIC4; RABGEF1; RGS1; DUSP5; SSR3; GNL3; ETV6; SUB1; PRIM1; CYB5A; PDHA1; LARP4; RPIA; TCEA1
BCL6 Knockout	1.6E-29	BUB1B; NSDHL; HMG5; PNP; CAPN2; F2R; PDIA4; TMX1; ATF6; HACD3; SOCS2; EBP; PRDX4; RACGAP1; PRDX1; ST8SIA4; STIL; PLK1; FUCA2; ASNS; EHD4; PQLC3; CDK1; ITM2C; KIF11; CASP3; CTSB; KIF23; DNAJC1; TMEM33; CEBPB; FGL2; FAM129A; CLCN3; PSMB2; TAF9B; YKT6; SEC11C; CENPE; JKAMP; MAD2L1	CYFIP1; NAB1; CPOX; PRDM1; MKI67; SMC2; RCBTB2; ALCAM; LGALS1; RGS1; LAMP2; SLC16A6; TNFSF10; SLC39A4; IER3; CPT1A; ENTPD1; GAS2L3; CISH; SDHD; CCNE2; ASF1B; LXN; CDCA2; MAGED1; MRPL18; NCAPG; LY96; SAMSN1; NCAPH; CCNB2; ATXN1; ECT2; N4BP1; GOT1; MLKL; OSBPL3; RAB27A; DSTN; FAM136A; TPX2; CPD; BMPR1A; LGALS3BP; NRP1; NCAPG2; LAMC1; LMNB1; IMPA2; NUF2; DLGAP5; DUSP5; RFC4; SSR4; ETV6; CCNA2; LMBRD2; CKS2; DEPDC1; DTL; CD63; RNASEL; SLC43A3; PRIM1; GLRX; HSPA13; STX11; TSPAN31; NFIL3; GMDS; EPCAM; CXCR3; BUB1; TCEAL9; PPIL1; CDKN2C; RRM2; MDH1; SLC31A1; NDUFA1; ATP2B4; GZMB; SHCBP1; HIPK2; DHRS7; RAD51; CENPI; PSAT1

**Table 6-10** | Tables shows differentially **upregulated** genes in ASCs that are known to be upregulated as a result of TF inhibition/downregulation/loss of function mutation (cont.)

Transcription Factor	Adj <i>p</i> -value	Genes of Interest	Others
RELA Knockdown	1.9E-25	SLC35B1; PSMD8; TMEM147; BSG; MAP2K2; MIF; PPIB; COPB2; VCP; RPN2; PSMD14; ZMPSTE24; PRDX3; PRDX2; PRDX4; PRDX1; HSPA8; PLK1; ASNS; DHCR24; PRDX6; MANF; CYCS; PSMG1; ARF4; CLTA; ATP2A2; SEC61A1; TMED2; UFD1; TMED9; SEC13; PSMA4; HAX1; PKM; DAD1; BIRC5; DCTN3; PSMB7; PSMB5; SRP72; PSMB2; ASNA1; ARCN1; GOLPH3; PSMC4; SCD; P4HB; GLA	MTCH2; UBE2L3; BASP1; RUVBL1; TXNL4A; MCCC2; CSNK2A1; CDC25A; MTHFD2; SUB1; UQCRC1; MAGED1; NDUFB11; TARS; OAZ1; YARS; EIF2B3; PDHA1; GOT2; DSTN; EIF2S2; CCT6A; AIMP2; TPX2; UQCRCQ; CDK4; IARS; AARS; LGALS3BP; ARPC1A; MRPS12; AKR1B1; ETFA; LAMC1; CKS1B; AIFM1; TIMM17B; PHGDH; CCT3; SSR4; GSTO1; HSBP1; PRMT1; ATP6AP2; NDUFC2; TUBG1; GNL3; NME1; NDUFS5; CKS2; GARS; ECHS1; PCNA; NDUFB6; PDHB; UQCR10; HSD17B10; COX5A; SRM; AURKA; PDF; CLPTM1; POLR2E; RABAC1; POLR2I; CCT7; NDUFA9; TIMM8B; GINS2; MDH1; MDH2; FKBP1A; PSAT1; NDUFAB1; ATXN10; MPHOSPH6
SON Knockdown	4.2E-24	TMEM147; PSMD4; FADS1; ALG5; TMX1; CSNK2B; RTN3; FKBP2; PRDX1; ASNS; BST2; PPT1; CDK1; CCDC47; TECR; GUSB; SEC13; ANXA2; SIGMAR1; DYNLL1; PSMA6; HAX1; PSMA1; CSTB; PSMB6; PSMB7; PSMB4; YIF1A; SPCS1; SCD; STUB1; P4HB; CEBPB; FAM114A1; SAR1B; HERPUD1; NSDHL; CAPN2; STOM; CD59; SEC61B; HSPA5; ARFGAP3; RCN1	ECI2; HMG3; TALDO1; ACOT13; NEAT1; MELK; BOLA2; UQCRC1; AURKAIP1; HIST1H2BD; ASF1B; CUTA; RNASEH2A; GMPS; MRPL17; MRPL15; OAZ1; ACAT1; GSTM3; ISG15; ZWINT; UQCRCQ; CDK4; FEN1; ARPC1A; AKR1B1; ETFA; ATRAID; MRPL40; IMPA2; PHGDH; HIST1H1C; CCDC51; RFC4; PRMT1; AKR1A1; TUBG1; MZT2A; NDUFS7; PCCB; TMEM106C; MCM6; CRELD2; DTL; SLC25A11; ECHS1; PCNA; NDUFB5; PRIM1; NDUFB1; UQCR10; PDHB; TYMS; MRPL57; UNG; ANKMY2; SAMM50; TCEAL9; NDUFA9; TIMM8B; GINS2; NQO2; RRM2; MDH1; MDH2; NDUFA2; MARS; NDUFAB1; TCEA1; PFKM; CLIC4; NXN; RPS27L; TARS; MOSPD1; COX7A2; UAP1; GLRX; TMEM97; COX6A1; AURKA; LIMA1; MT2A; MAP1LC3B; CCNB1; CSRP1; NFIL3; ZMIZ1; BASP1; BANF1; CDR2; MT1HL1; IER3; FDPS; DNAJC15; REXO2; SLC39A14; CCT6A; SLC7A5; DAP; PPA2; FAM98A; GCLM; ALDH9A1
ZFX Knockout	2.0E-17	ERO1A; PIGT; SLC35B1; ATP2A2; MPO; SEC61A1; SERP1; PSMD1; TMED2; TMED7; CTSD; PSENEN; PDIA3; FCER1G; ENTPD4; ALG3; ELOVL6; MIF; PDIA4; DAD1; CANX; RTN3; GRN; HSP90B1; LMAN1; PSMB4; PSMB2; PRDX1; CREG1; EMC2; ST8SIA4; SSR1; SEC31A; SDF2L1; ASNS; PA2G4; SRP68; SEC23IP; SPCS2; OSTC; CTH; PPT1; STT3A; PSMG1; CALR; GLA; ARF4; CITED2; PLOD3; MAGT1; LITAF; PSMD8; CCND2; CAPN2; AP3S1; ANXA2; ANXA5; F2R; CKAP4; RCN1; TMX1; NPC2; COPB2; ABHD4; RPN1; SELENOK; VAT1; VMP1; CALU; P4HB; ITM2B; ITM2C	IFITM1; SLC7A11; MRPL35; MT2A; TMEM223; RIOK2; IFNAR2; RFC4; SSR3; AKR1A1; SLC7A5; SUB1; NDUFS4; GARS; LAP3; CRELD2; BCAT1; GART; NDUFB6; FAM206A; SELENOF; WDR61; CSF2RB; SLC3A2; SAMS1; SLC7A1; TSPAN31; GANAB; NT5DC2; DARS; SEPT11; GOT1; MDH2; GOT2; GLMP; UCK2; DAP; CDK4; INTS7; ALDH18A1; ATXN10; GCLM; SERINC3; MMACHC; ITGB1; NRP1; HMGB3; LAPTM4B; LGALS1; CSRP1; BASP1; LAMP2; PHLDA1; HMG3; IER3; GSTO1; FNDC3B; ACTA2; TUBB2B; CLDN12; MYADM; FSCN1; ATP6V1A; ABRACL; TCEAL8; TCEAL9; MCTS1; DSTN; FKBP1A; CLPTM1L
NFKB1 Inactivation	9.4E-16	ARF4; CLTA; KIF11; SEC61G; GTSE1; FADS1; DYNLL1; PSMA6; PSMA4; PSMA1; NPC2; PSMD14; SELENOK; RACGAP1; PSMB2; PRDX1; HSPA8; UBE2C; DHCR24; GOLPH3; CDK1; CYCS; ITM2B; MAD2L1; TRAM1; HERPUD1; SESN2; TMED2; PDIA3; TMX1; CANX; DNAJB11; TRIB3; CEBPB; RPN2; HACD3; TMEM165; HSP90B1; PGRMC1; PRDX4; NPM1; ARCN1; SCD; CTH; PSMG1; CASP3; PRKCI; NOSIP; BIRC5; PSMD10; RTN3; PSMB6	FEN1; NDUFA12; CSE1L; GMNN; COX6A1; CKS1B; MT2A; MAP1LC3B; NUF2; PBK; NUSAP1; NEK2; FBXO5; DLGAP5; MT1HL1; CMC2; ARMXC3; CCNA2; SGO2; TMX2; NDUFS5; MT1F; DEPDC1; CKS2; KIF20A; CRK; PCNA; POMP; NDUFB3; ARHGAP19; NCAPG; CDCA8; COX7A2; TTK; MRPL13; OAZ1; CENPA; AURKA; CCNB1; FDPS; CENPW; RRM2; NDUFA1; SHCBP1; TPX2; ANLN; KIF18B; EVI2A; UBE2T; NOP10; CDKN3; SPC25; TSEN15; TXNDC12; TMEM97; SLC7A11; EPRS; LMNB1; MRPL42; PHGDH; CCT3; RFC3; RFC4; ARMC1; SLC7A5; PPA1; MTHFD2; TIPRL; GARS; TOP1; DTL; TARS; SELENOF; SLC3A2; TYMS; SLC7A1; HSPA13; ISOC1; CCT7; PCK2; YARS; DSTN; EIF2S2; ZWINT; CCT6A; RPIA; MARS; TCEA1; IARS; AARS; TMEM41B; OIP5; MELK; VGLL4; NDUFB11; BIK; GPN3; REEP5; PSAT1; NDUFAB1

**Table 6-10** | Tables shows differentially **upregulated** genes in ASCs that are known to be upregulated as a result of TF inhibition/downregulation/loss of function mutation (cont.)

Transcription Factor	Adj <i>p</i> -value	Genes of Interest	Others
BPTF Knockout	4.7E-15	CITED2; LITAF; TMEM147; PSMD4; SEC61G; SDF2; PSENEN; F2R; DYNLL1; PSMA4; NPC2; DAD1; ATF5; GAS6; CSTB; GRN; ANAPC16; DCTN3; HACD3; FAM129B; PRDX2; PSMB7; PSMB4; PSMB2; PRDX1; OST4; NPM1; SPCS1	NRP1; CLIC4; MRPS12; BRK1; ETFB; COX6A1; ATRAID; UBL5; GLIPR1; CSRP1; LAMP2; PHGDH; CHAC1; CHCHD1; HIST1H1C; ATP6V0E1; SSR4; GSTO1; AKR1A1; ACOT13; NME1; ACTA2; TMEM256; SMDT1; ZC2HC1A; NDUFS7; HIST1H2BE; HIST1H2BH; NUPR1; HIST1H2BC; CUTA; NDUFB6; POMP; PRCP; UQCR10; TCEAL8; EMC4; ATP6V1D; TCEAL9; POLR2L; MCTS1; TIMM8B; UQCC2; NDUFA2; NDUFA1; ZWINT; ATOX1; SNX18; RGS10; HIST1H4H
ZXDC Depletion	1.4E-14	PIGU; BUB1B; KIF11; TNFSF13B; COMMD3; BIRC5; CDC6; CENPE; KIF4A	FEN1; GF11; GMNN; NUDT1; MCM10; BRCA1; BRCA2; FOXM1; PIK3CG; SMC2; IMPA2; EXO1; PBK; BLNK; CHAC2; OIP5; MYBL2; FBXO5; DLEU2; TK1; TCF19; CDC25A; CCNA2; CCNE2; SDHAF3; GPR160; KIF20A; MCM6; DTL; ASF1B; ZNF670; BLM; RNASEH2A; CDCA3; UHRF1; CDCA5; PRIM1; NCAPG; CDCA8; TTK; TYMS; CENPA; NCAPH; SKA1; SKA2; LIMA1; BRIP1; CCNB1; CDC45; RAD54L; CTNNA1; GINS1; CDT1; CENPU; RRM2; ZWINT; TPX2; COQ3; CMSS1; RAD51; POLE2; SPC24; CDKN3; SPC25
NRF1 Knockout	5.3E-14	ITGB5; SLC35B1; SEC61A1; SERP1; SEC61G; TMED3; FBXO6; SEC61B; SLC39A7; CTSD; CTSB; PDIA3; TMED9; ENTPD5; ANXA5; COMMD3; PDIA6; RER1; DNAJB11; ATF5; GAS6; GRN; TMED10; RPN2; ABHD4; GSTP1; RPN1; RRBP1; HSP90B1; PRDX1; LMNA; HSPA5; SURF4; SYVN1; DHCR24; MANF; BST2; SPCS1; CALR; P4HB; ITM2B	LGALS3BP; ACAA2; ALAS1; GMNN; THY1; NENF; PEPD; COMT; COX6A1; LGALS1; TNFAIP8L1; BLNK; SLC39A4; AKR1A1; TALDO1; NEAT1; HDHD3; HAGH; MCM6; CRELD2; AVPI1; ECHS1; PCNA; RPS27L; SELENOF; COX7A2; COX5A; ACAT1; TSPAN31; MRPL20; POLR2L; DNAJC19; GSTM3; CYB5A; MDH2; TMEM176B; TMEM176A; GOT2; HMGA1; ATOX1; GALE; SERINC3; MSRB1
IRF8 Knockout	6.2E-13	DPAGT1; PIGU; TECR; GOLIM4; BUB1B; MPO; MANEA; TNFSF13B; APMAP; TMED3; GYG1; SLC39A11; WIPI1; ESCO2; KIF22; PLSCR1; BIRC5; LTF; MGST2; ZMPSTE24; PLK1; KIF4A; MAD2L1; APIP; TXNDC17; HMGN5; SSR2; ALG5; ALG14; ALG3; MGLL; GLT8D1; TEX261; EBP; FUT8; ASNA1; SLC33A1; PA2G4; ARCN1; EHD4; CDC26; SURF4; BET1L	NRP1; FEN1; EIF4E3; NDUFA11; GF11; RAB3D; ETFA; FOXM1; PPP3CB; MCEE; FNTB; PBK; PHLDA1; TRAPPC2L; CISH; SGO2; TMEM256; SAPCD2; GPR160; PGP; CD63; RNASEH2C; DTYMK; CDCA3; CDCA8; ABHD5; NCAPH; AURKB; AURKA; CCNB1; MARVELD1; TMEM205; NDUFV3; ANKRD28; DARS; NQO2; DSTN; BCKDK; KIF18B; FMNL2; CDKN3; SPC25; MPHOSPH6; AHCYL2; DENND5B; TSEN15; BHLHE41; TPGS2; GCSH; ABHD12; CEP55; TMED4; PCMTD1; ENTPD1; GLRX3; CMC2; AARSD1; HCFC2; SAP30; TMX2; AKIP1; TMEM106B; GEMIN6; PHPT1; ASF1A; HIST1H2BC; NDUFB6; LY96; TARS; MRPL57; SKA2; TSPAN5; PIP5K1B; RPRD1A; LYAR; MRPL17; MOSPD1; LAGE3; JCHAIN; EPM2AIP1; TNFSF10; HMGN3; HERC6; HIST1H1C; CENPU; SLC10A7; PRMT1; MTX2; RRP1B; TUBG1; REXO2; ATOX1; UCK2; GNAQ; CKS2; LCORL; TMEM19
SRF Knockdown Mutation Knockout	7.6E-13	GOLT1B; FKBP14; BUB1B; KIF11; TNFAIP1; TIMP2; FADS1; COPZ1; BIRC5; SAR1B; SOCS2; HMOX1; UBE2C; PLK1; GGH; CDC6; KIF4A; MAD2L1; KDELR1; SLC35B1; PSMD8; PNP; PSMD4; TRIM25; GTSE1; CTSB; PDIA3; SEC13; SSR2; F2R; ELOVL6; KIF22; DYNLL1; CKAP4; HAX1; UNC50; KDELR3; PPIB; LRRC59; TMED10; PSMD14; GSTP1; PSMB6; PSMB5; ASNA1; MCL1; ASNS; CALU; CDK1; MPO; KRTCAP2; SEC61G; PSENEN; FCER1G; DHRS3; PSMC4; FAM213B; HECTD3; EBPL; CSTB; GRN; FKBP2; PSMB2; HSPA5; RNF181; HERPUD1; UCHL1; HSPA8; NPM1; YES1; PGRMC1; PRDX1; VCP; CEBPB; GLG1; SURF4; ERO1A; TRAM1; LITAF; BSG; CTSD; ANXA2; ANXA5; TAPBP; COPB2; CITED2; COG6; PSMA6; TECR; GPX3; CDC34; NAA38; ABHD4; PRDX6; TMC01; CALR; ITM2B; ITM2C;	LGALS3BP; ENO2; NUDT5; METRNL; FOXM1; CKS1B; LMNB1; DHTKD1; MT2A; NUF2; PBK; NUSAP1; OIP5; MYBL2; DLEU2; TK1; CEP55; DLGAP5; MT1HL1; DUSP5; CKAP2L; LINC00294; ATP1B1; SLC7A5; SGO2; MELK; SAPCD2; ESPL1; CLDN12; MT1F; DEPDC1; CKS2; FSCN1; KIF20A; SLC26A2; CDCA3; UHRF1; CDCA5; POMP; ARHGAP19; CDCA8; TTK; CENPA; PAPSS1; CCNB1; LRRC8D; QSER1; RRM2; SPAG5; SLC31A1; DNAJC12; TPX2; ANLN; CENPF; NDUFAF4; POLR3D; NCAPD2; SPC25; BTG2; CLIC4; ACAA2; MRPS14; ETFA; ETFB; COX6A1; UBE2L3; SMC2; MRPL4; TMEM223; TIMM17B; ACADM; TXNL4A; SLC39A4; PSPH; SS18; GLRX3; AKR1A1; H2AFV; GNL3; SMDT1; MRPL51; NDUFS5; UQCRC1; GARS; ERGIC3; BCAT1; NDUFB9; NDUFB11; NDUFB5; PDHB; MRPL13; CCNB2; SLC25A20; ECT2; NDUFA9; NDFIP1; PDHA1; IDH2; NDUFA2; EIF2S2; AZIN1; ATOX1; TM9SF2; SP3; B9D1; BLMH; SERINC3; EIF3C;

**Table 6-10** | Tables shows differentially **upregulated** genes in ASCs that are known to be upregulated as a result of TF inhibition/downregulation/loss of function mutation (cont.)

Transcription Factor	Adj <i>p</i> -value	Genes of Interest	Others
		SEC31A; PSMG1; COPA; ARL3; RPN1; PSMB4; SERP1; SDF2; ST6GAL1; PRELID1; SPCS1; DAD1; OSTC; DNAJB11; CYCS; PIGF; HMG5; RACGAP1; ALG3	OGT; ECI2; BRK1; HDCC3; LAPTM4B; GLIPR1; LGALS1; HINT2; CHCHD1; MRPS24; CCNA2; NDUFS7; PHPT1; HIST1H2BC; CUTA; NDUFB7; DMAC1; PRIM1; NDUFB3; RPS27L; COX7A2; OAZ1; COX5A; NT5DC2; NDUFV3; NDUFA3; NDUFA1; DAP; OSTM1; RGS10; CDKN3; IFITM1; PAK1IP1; CPOX; LAMC1; SLC16A6; IL6R; TCF19; DTL; TYMS; SLC17A5; LYAR; GINS1; BIK; RPIA; CDK4; ALAS1; SLC44A1; CSRPI; BASP1; ATP6VID; NDUFA8; MDH2; GLMP; ISG15; ASS1; RNF149; CDK2AP2; CARS; EPRS; MKI67; MRPL35; CDH1; CPT1A; ACOT13; TMX2; MTHFD2; ATP6V1A; BLM; SLC3A2; MFF; AURKB; UCK2; PSAT1; MARS; UBE2T; AARS; HIBADH; CISD1; BRCA2; VPS35; CCT3; ACTL6A; HMBS; PCNA; TARS; HSD17B10; POLR2E; RDX; TCEA1; IARS; NOP10; TOP1; CHMP6; AURKA; POLR2L; ITGB1; ANXA6; GSTO1; ACTA2; MYADM; TCEAL9; DSTN; HDLBP; IER3; SDHD; DDX19A; MAGED1; PLD3; MDM2; ATXN10; FEN1; SLC35F5; ALCAM; RGS3; PRKACA; HIST1H1C; ATP6V0E1; ECH1; TUBB2B; GPR160; UQCRI0; TIMM8B; F8A1; MDH1; REEP5; ACO2; GRINA; EIF4EBP1; SMARCB1; TLE3; PRMT1; MRPL15; UNG; PRPSAP1; GANAB; SAMM50; AAMP; FDPS; QDPR; COPS6; IPO11; GMNN; JCHAIN; NTMT1; ARFIP2; CYB5A; NDUFC2; ETNK1; POU2AF1; CLPTM1L; SI00A9; FERMT3; TMEM41B; HELB
ADNP Deficiency	1.4E-12	ERO1A; PSMD8; BSG; PDIA3; F2R; CDC34; TMX1; CSNK2B; COPA; PSMD14; SAR1B; ARL3; RPN1; ARL1; SSR1; MLEC; PRELID1; UBE2C; PSMG1; CALR; STT3B; ITM2C; TECR; CLTA; TMED2; AP3S1; SEC61B; SEC13; GPX3; KIF22; DYNLL1; HAX1; PKM; NPC2; GSTP1; DCTN3; PSMB6; PSMB7; PGRMC1; ASNA1; TMC01; SPCS2; PSMC4	CYFIP1; IPO11; CSE1L; GMNN; MSANTD4; MKI67; COX6A1; SMC2; LAPTM4B; BASP1; BANF1; TXNL4A; CHCHD1; MRPS18A; SAP30; MELK; ARM CX2; OAZ1; PRPSAP1; QDPR; CLPTM1L; NOP10; GALK1; LMNB1; NDUFC2; H2AFV; GNL3; PTRH2; CKS2; GARS; BCAT1; NDUFB7; UHRF1; ABRACL; SLC3A2; MESD; COX5A; DNPEP; MAT2A; NT5DC2; CCT7; POLR2L; TCEAL9; CYB5A; SLC16A1; RRM2; NDUFA3; NDUFA1; TM9SF2; KLHL9; RAD51; RNF149; PSAT1; CARM1
ZNF395 Knockdown	1.3E-11	ARF4; KDELRL1; TXNDC17; TMED3; TIMP2; SEC61B; NUDT14; PGM1; CTSB; TMED9; ANXA2; COMMD3; CKAP4; PSMA5; PSMA6; NPC2; DAD1; CSTB; LEPROT; CEBPB; PSMB6; PSMB7; ARHGAP21; PRDX1; OST4; TMEM230; DERA; PPT1; RNF181; ACO1	LGALS3BP; SRP54; AKR1B1; MAP1LC3B; LGALS1; BASP1; LAMP2; PHLDA1; IER3; TMEM14C; HIST1H1C; ATP6V0E1; DUSP3; GSTO1; ATP6AP2; FNDC3B; IL18; TALDO1; REXO2; PLA2G16; NDUFS8; OAF; NDUFS5; GARS; NDUFB5; TACSTD2; MRPL17; RPS27L; LY96; COX7A2; GLRX; OAZ1; LIMA1; TCEAL9; NDUFA8; NDUFA1; ISG15; ATOX1; FKBP1A; ISG20; DHRS7; UQCRI0; FAM98A; NOP10
ETS1 shRNA	2.0E-09	ERO1A; TRAM1; PIGF; CHPF; GPAA1; TECR; PLOD1; NDST1; FADS2; SEC61A1; TIMP2; CAPN2; CTSB; PDIA3; ANXA2; ANXA5; PDIA6; SELENON; KDELRL2; TRIB3; GAS6; PPIB; LRRC59; GRN; LEPROT; COPA; VCP; ATL3; FAM129B; HSP90B1; PRDX3; PRDX2; LMAN1; PRDX1; HMOX1; CD59; MCL1; CAV2; GFPT1; CALU; STT3A; DHCR7; CALR; COPG1	ITGB1; LGALS3BP; NRP1; CLIC4; ARPC1A; HDLBP; EPRS; DCAF7; MT2A; PHGDH; PHLDA1; TNS3; ABCC3; DUSP5; DESI2; FNDC3B; FSCN1; TOP1; ERGIC3; RABL6; MAGED1; TACSTD2; UAP1; OAZ1; PLD3; TM9SF3; GANAB; DNPEP; MDH2; SLC52A2; HMGA1; CCT6A; FKBP1A; MARS; PLXNB2; SDC1; CALM3; ATXN10; SERINC3; NFE2L1; GRINA
RUNX1 Knockout	1.3E-08	ERO1A; DPY30; SERP1; AP3S1; F2R; CNPY2; DYNLL1; PDIA6; PSMA5; PSMA4; PKM; DYNC112; PSMD14; SAR1B; TMEM165; PSMB7; PGRMC1; PRDX1; EMC2; SPTSSA; NPM1; MANF; ARCN1; SPCS2; CDK1; MAD2L1; TNFRSF13B; SLC35B1; SSR2; BHLHA15; LAX1; KDELRL3; LRRC59; DERL3; LMAN2; TNFRSF17; SEC11C; SDF2L1; HSPA5; PYCR1; ASNS; TXNDC17; ARL1; PSMB6; PTCH1; CYCS; SDF2; ANXA2; ANXA5; PPIB; HSP90B1; FAM213B; VAT1; CALU;	CSE1L; HMGB3; MSANTD4; GCSH; LAPTM4B; NUF2; VPS35; SMIM15; SSR3; MRPS18C; HAUS1; ARM CX3; IMMP1L; ARM CX2; PPA1; CKS2; ASF1A; PCNA; POMP; ABRACL; COX7A2; LAGE3; SKA2; TCEAL8; CCNB1; TCEAL9; NDUFA9; CYB5A; MTX2; RDX; NDUFA1; DSTN; SHCBP1; FMNL2; CDK4; ITM2A; SPC25; IGHM; FEN1; GF11; LRR1; NCAPG2; ALDH1L2; PIK3BG; JCHAIN; CKS1B; GLIPR1; EXO1; MRC1; PBK; PHGDH; CHAC1; PHLDA1; ACOT13; SND1; NME1; SLC7A5; MTHFD2; SLC04A1; KIFC1; FAM173A; CRELD2; SMPDL3B; FAM72A; SELPLG; SLC1A4; NFIL3;

**Table 6-10** | Tables shows differentially **upregulated** genes in ASCs that are known to be upregulated as a result of TF inhibition/downregulation/loss of function mutation (cont.)

Transcription Factor	Adj <i>p</i> -value	Genes of Interest	Others
		TNFAIP1; TMEM147; TMED3; GPX3; MOGS; TRIB3; ATF5; RPN1; FAM129B; XBP1; CALR	GMDS; CCR10; TG; EVI2A; PSAT1; POU2AF1; ALDH18A1; IFITM1; SLC39A8; PRMT1; CCNA2; LAP3; RPS27L; HNRNPLL; NDFIP1; SLC16A1; PDHA1; BLMH; MKI67; TMEM223; BASP1; GNL3; SUB1; TIMM8B; NDUFA2; EIF2S2; ZWINT; GPN3; CENPF; BTG2; CLIC4; CSRP1; IER3; HIST1H1C; ABCC3; DUSP5; NEAT1; REXO2; TUBB2B; ZC2HC1A; GARS; TRIB1; TACSTD2; SLC3A2; SLC7A1; SRM; LRRC8D; KCNN4; LCM1; OSBPL3; CCT6A; AIMP2; MARS; IARS
YBX1 siRNA	1.8E-08	ERO1A; ATP2A2; FADS2; BSG; CAPN2; CTSD; FADS1; PGM1; CTSB; PDIA3; TMED9; ALG9; ANXA2; GPX3; ANXA5; PDIA6; LSS; PDIA4; TAPBP; PSMA6; PKM; LRRC59; VCP; TMED10; HM13; FAM129B; FUT8; PSMB2; PRDX1; LMNA; MCL1; HSPA8; NPM1; SEC24A; HSPA5; GALNT2; GFPT1; SURF4; DHCR24; ARCN1; VEGFA; NCEH1; STT3A; DHCR7; P4HB; CALR; STT3B; SEC24D; PLOD3; HS2ST1; NDST1; TIMP2; SELENON; GRN; HACD3; HSP90B1; CAV2; PPT1	ITGB1; CLIC4; CSE1L; ECI2; HDLBP; EPRS; DCAF7; MT2A; LGALS1; ANXA6; PHLDA1; IER3; CCT3; DUSP5; SSR3; FNDC3B; ATP1B1; DNAJC5; GARS; TOP1; S100A9; UAP1; E2F7; MPZL1; YARS; PPIL1; HMGGA1; DSTN; CCT6A; HOOK1; MARS; PLXNB2; CALM3; IARS; EIF3C; NFE2L1; LAMC1; LAPTM4B; LAMP2; KIF1B; TNS3; ATP6V0E1; ARHGEF12; NCBP2; SUB1; TRIB1; CRK; NXN; NLN; PPP2CB; PGRMC2; CDKN2C; REEP3; GRINA
CREB1 Depletion	2.4E-08	ARF4; TRAM1; PLOD3; SEC61G; CAPN2; UFD1; AP3S1; CTSB; ANXA2; SSR2; MIF; DYNLL1; PDIA6; TMEM258; SELENOS; PPIB; CEBPB; HACD3; HSP90B1; SOCS2; PRDX4; SEC31A; MCL1; HSPA8; STIL; SDF2L1; HSPA5; YIPF5; GFPT1; DHCR24; ARCN1; CALU; P4HB; LPIN1	IFITM1; TMEM41B; ARPC1A; TMEM181; CPOX; TMEM97; ENO2; TMEM263; TRMT112; LAPTM4B; MAP1LC3B; LGALS1; PHACTR2; TRIM27; IER3; CMC2; NEAT1; SLC7A5; TMX2; CKS2; SAMSN1; TSPAN31; RAPIA; PDZD11; NDUFV3; RABAC1; RAB4A; RRM2; RAB27A; TM9SF2; UQCQR; NDUFAB1
PAX5 siRNA Knockdown	3.4E-08	CHPF; ITGB5; CCDC47; APMAP; BSG; TMED3; TIMP2; TMED2; PMPA1; CTSD; CTSB; PDIA3; TMED9; ANXA2; SSR2; PDIA6; PDIA4; PKM; CANX; PPIB; LRPAP1; GRN; VCP; TMED10; RPN2; HM13; RRBPI; GLG1; HSP90B1; PSMB4; PRDX1; LMAN2; STOM; CD59; SPTSSA; NPM1; HSPA5; GALNT2; M6PR; CALU; CALR; STT3B; P4HB; RTN3; SHC1; SEL1L; SERP1; RACGAP1; AP2B1; CTDSP2; TNFRSF13B; CITED2; TMED1; GPX7; WIPI1; LAX1; LEPROT; MVP; GLT8D1; UBE2J1; CHST12; TNFRSF17	LGALS3BP; ITGB1; NRP1; LAPTM4B; BASP1; LAMP2; ATP6V0E1; SSR4; HSBP1; ATP1B1; SND1; SLC7A5; TOP1; CD63; IFI6; TACSTD2; SLC3A2; MESD; GANAB; EPCAM; ISG15; FKBP1A; SCCPDH; REEP5; EIF3C; CYB561; NFE2L1; SLC44A1; LIMA1; CCNB1; ALCAM; NEK2; METTL7A; CCT3; DUSP3; ARHGEF12; SLC11A2; TM9SF2; ANLN; TP53I3; CKS2; CALM3; GARS; ICAM2; IFI35; ETFB; PIK3CG; JCHAIN; LAMP3; ANXA6; KIF13B; ACP2; TNS3; HERC6; PCMTD1; ZBP1; ENTPD1; DESI1; ZBTB38; HRASLS2; CARD16; TRIB1; VGLL4; HIST1H2BD; SLC05A1; CD99L2; STX11; NFIL3; CCR10; CRIPT; ISG20; EVI2A; HIST1H4H; GALM; TMEM19
ZEB1 siRNA	9.4E-08	PLOD1; NDST1; FADS2; BSG; SEC62; FADS1; PDIA3; ACSL4; FNDC3A; PDIA6; PDIA4; KDELR2; PPIB; COPB2; RTN3; COPA; VCP; TMED10; RPN2; SHC1; RPN1; GLG1; HSP90B1; PRDX1; SSR1; MCL1; SEC31A; HSPA8; HSPA5; SURF4; DHCR24; BST2; CALU; CALR; ITM2B; TRAM1; ITGB5; ATP2A2; MIA3; SEC61A1; UCHL1; TIMP2; TMED2; CTSD; CTSB; ANXA2; CKAP4; RCN1; PKM; CANX; GRN; RRBPI; FAM129B; LMAN1; STOM; CD59; NPM1; TMEM30A; P4HB	IFITM1; HDLBP; ALCAM; BASP1; LAMP3; ATP1B1; MAGED1; NDUFV3; RDX; HMGGA1; ISG15; SPART; SYNJ1; MDM2; PLXNB2; IARS; NFE2L1; ITGB1; LGALS3BP; CLIC4; SLC7A11; GLIPR1; CSRP1; PHACTR2; ARHGEF12; CD63; IFI6; SLC3A2; TM9SF3; LIMA1; GANAB; MAT2A; APOL1; PDLIM5; SLC16A1; TM9SF2; EIF3C; GRINA
BCL11A Knockout	2.0E-07	ALAD; PNP; BSG; SEC62; TMX1; BIRC5; VCP; ABCB6; FOXO3; CLCN3; PRDX2; CREG1; UBE2C; CDK1; CPEB4	BTG2; CPOX; IFI35; MKI67; PEPD; SMC2; CKS1B; NADK2; MAP1LC3B; NUF2; PBK; NUSAP1; CHAC2; TMEM14C; IFRD2; TALDO1; CCNA2; PLA2G16; CCNE2; HMBS; DEPDC1; CKS2; MTFR1; HIST1H2BC; NDUFB9; FECH; SLC43A3; CDCA8; SLC3A2; MFF; FAM117A; CCNB2; CCNB1; KCNN4; ECT2; CYB5A; SLC16A1; RRM2; EIF2AK1; ISG15; ARID3A; EIF2S2; AZIN1; SP3; GCLM; SERINC3; SSBP3; CDKN3; GRINA
MYC Silencing Knockdown	4.1E-07	KDELR1; PLOD3; MAGT1; UCHL1; SEC61G; BSG; CAPN2; PDIA3; TMED9; ANXA2; ANXA5; ARFGAP3; PDIA6; PSMA4; PKM; NPC2; CSNK2B; DNAJB11; PPIB; PSMD10; DYNC112; PSMD14; SHC1;	ITGB1; CYFIP1; BRK1; LAMC1; MT2A; MAP1LC3B; CSRP1; ANXA6; HMGN3; IER3; CCT3; ATP6V0E1; TALDO1; ARMCX2; FSCN1; CRELD2; ALDH9A1; CD63; PCNA; TM2D2; MAGED1; SELENOF; CRYZ; PPP2CB; TCEAL9; GINS1;

**Table 6-10** | Tables shows differentially **upregulated** genes in ASCs that are known to be upregulated as a result of TF inhibition/downregulation/loss of function mutation (cont.)

Transcription Factor	Adj <i>p</i> -value	Genes of Interest	Others
		HACD3; HSP90B1; PSMB4; MCL1; HSPA5; TMEM230; OSTC; CALR; ITM2B; ITM2C; LIMS1; ARF4; TIMP2; PDIA4; PSMA6; DAD1; CEBPB; CD59; TMCO1; P4HB	HMGAI; DSTN; ISG15; ZWINT; FKBP1A; RCAN1; MAGED2; ATXN10; GCLM; ARPC1A; LGALS1; BASP1; MT1HL1; ATP1B1; GBP1
IKZF3 Knockdown	1.2E-06	CHPF; ITGB5; TNFSF13B; PYCARD; RNF115; PNP; PIM2; KIFAP3; MAP2K2; TAPBP; TPST2; PREPL; MGLL; TMED10; PARM1; MVP; FGL2; HTRA2; RRBPI; FAM129B; ORMDL3; MGAT3; ST3GAL2; CRTAP; PTCH1; CFLAR; DHRS3; ITM2C	LGALS3BP; SLC44A1; TMEM131L; TRAM2; PIK3CG; LGALS1; ZMIZ1; BASP1; TNFSF10; TIMM17B; TNS3; PLS1; DUSP5; CD300A; REXO2; SND1; RRAGD; LAP3; CARD16; TRIB1; S100A9; BLVRA; ATP6V1A; STAU2; PAPSS1; STK3; LIMA1; CXCR3; RABAC1; MPZL1; APOBEC3G; IDH2; BSPRY; GALM; TMEM19; MSRB1
KLF9 Deficiency	2.1E-06	KDELR1; TMEM147; CCND2; CASP3; ANXA2; MIF; PSMA5; NPC2; DAD1; CSNK2B; GRN; VCP; ABHD4; GSTP1; RPN1; PSMB6; PSMB4; PSMB5; HSPA8; PSMB8	NDUFA11; CISD1; ETFB; COX6A1; MT2A; LGALS1; BANF1; ACADM; SLC39A4; CBR1; AKR1A1; TALDO1; H2AFV; SDHC; NME1; ACTA2; SMDT1; NDUFS8; NDUFS7; PPA1; FDX1; UQCRC1; MPND; PFDN6; ANAPC13; NDUFB9; NDUFB7; SLC3A2; OAZ1; COX5A; PAPSS1; PRPSAP1; SAMM50; DNPEP; NDUFV3; CDR2; GSTM3; NDUFA8; MDH2; NDUFA3; NDUFA1; ACO2; ADA
KLF2 Knockdown	2.3E-06	TMEM147; PSMD4; PSMD1; PIM2; ALG8; ANXA2; DYNLL1; PSMA6; PSMA1; HACD3; PSMB6; PSMB4; PSMB5; PRDX1; HSPA8; PSMB8; TMEM230; BST2; PPT1; STUB1	INTS13; FASTKD1; NDUFA12; MRPL36; ETFA; MRPL42; RGS1; NUSAP1; ME2; HMGN3; TMEM14C; CCT3; PRMT1; TALDO1; SDHD; HAUS1; SCO2; NDUFS5; RRAGD; CKS2; TOP1; ATP6V1A; NDUFB9; CLDND1; NDUFB3; IFI6; COX7A2; OAZ1; CRYZ; CCT7; NDUFA8; RRM2; MDH1; UQCC2; NDUFA3; ISG15; RRP1B; DNAJC15; EIF2S2; COPRS; RPIA; RNF149; PFKM; AARS
NR2C2 Knockdown	2.5E-06	DCTN3; TXNDC17; PRDX2; PSMB4; BSG; AP3S1; OST4; PRELID1; ANXA2; ANXA5; MIF; PRDX6; STUB1	CUTA; ZNF593; ECHS1; PCNA; NDUFA11; NAXE; NDUFB11; AKR1B1; RPS27L; NDUFB1; COX7A2; MRPL37; UQCR10; COX6A1; OAZ1; ATRAID; UBL5; LGALS1; RAP1A; MCRIP2; CHAC2; BANF1; SSR4; GSTO1; NDUFA3; NDUFA1; ATOX1; NME1; MZT2A; UQCRC1; UQCRC1; MCM6; PHPT1
ARNTL Knockdown	3.7E-06	ALAD; CHST12; TIMP2; UBE2C; PLK1; ESCO2; KIF22; CTH; CDK1; BIRC5	LGALS3BP; CDCA3; SOWAHC; CDCA5; CDCA8; HDHC3; AURKB; IGKC; NFIL3; IMPA2; NUF2; NUSAP1; PBK; BUB1; CENPW; RRM2; GOT1; MLKL; FAH; HDHD3; SHCBP1; CCNA2; ACTA2; MELK; MTHFD2; TMEM56; HIST1H4H; HIST1H2BH; SPC24; ASF1B; SPC25
EZH1 shRNA	1.3E-05	LITAF; MPO; FADS2; SERP1; APMAP; GUSB; CTSD; PDIA3; SSR2; PDIA6; LSS; PKM; NPC2; CANX; PPIB; LRPAP1; GRN; COPA; RPN2; ATF6B; GSTP1; SLC38A10; GLG1; MTDH; HSP90B1; UBE2J1; CREG1; LMNA; SSR1; MLEC; STOM; UGGT1; MCL1; HSPA5; SURF4; M6PR; DHCR24; CFLAR; STT3A; NUCB1; P4HB; CALR; COPE	ITGB1; SLC44A1; WIPF1; GFI1; HDLBP; TXNDC12; EPRS; MKI67; DCAF7; LMNB1; LGALS1; BASP1; LAMP2; ANXA6; IL6R; CPT1A; ITGA4; MTHFD2; LCP2; CD63; H6PD; ABHD2; HSD17B4; ARHGAP18; OAZ1; PLD3; MAT2A; METTL7A; YARS; GSTM3; SEPT11; AZIN1; TM9SF2; CENPF; NCAPD2; CALM3; PTPN7; FERMT3
ELF4 Knockout	1.4E-05	GRN; ARL3; MAGT1; HSP90B1; PRDX3; PRDX2; PSMB4; SERP1; PIM2; SEC61B; AP3S1; HSPA8; TMED9; HAX1; CTH; CANX; P4HB; RNF187; FLOT1; TRIM25; CTSD; PSENEN; ST6GAL1; F2R; MOGS; PSMA5; CTDSP2; UNC50; PPIB; RTN3; RPN1; ZMPSTE24; PSMB5; SPTSSA; MCL1; PRELID1; NPM1; SRP68; SPCS1; PSMC4; IMPDH1; ITM2B	MFF; ALDH1L2; OAZ1; STX11; AIFM1; BASP1; NDUFV3; PRKACA; IER3; CYB5A; NDFIP1; CISH; ATP6AP2; SSR3; GZMB; ELL2; CCT6A; GATM; ARMCX2; MTHFD2; PSAT1; ITM2A; NOP10; FEN1; WIPF1; IFI35; JCHAIN; SPN; ACADM; IGLC2; AKR1A1; TALDO1; ATP1B1; PLA2G16; NDUFS7; CCNE1; S100A9; SELPLG; PCNA; DMAC1; NDUFB11; HSD17B4; PLD3; CIAO1; FAM117A; PRPSAP1; CXCR3; KCNN4; ATP6V1D; MCTS1; MDH2; DNAJC15; DAP; CDK4; CD28; RGS10; BLMH
YY1 Knockdown	1.8E-05	PLOD3; PSMD8; PYCARD; BSG; PGM3; SDF4; FBXO6; SEC62; CTSB; RNF121; MVP; PSMB5; MGAT1; SEC31A; WFS1; PQLC3; GLA	CCDC167; PECR; NUDT1; NENF; PPP3CB; MAP1LC3B; PRADC1; PBK; NUSAP1; OIP5; HERC6; TRAPPC2L; MRPS24; DESI2; CARHSP1; SELENOM; NDUFS7; MT1F; CUTA; LY96; SLC1A4; SKA1; TMEM70; BCL2L13; E2F2; LCMT1; SLC31A2; RAB27A; COQ9; NMRAL1; ISG20; SPSB2; PNKD; NCAPD2
CHD1 Knockout	3.6E-05	ITGB5; GPAA1; PLOD3; ATP2A2; SEC61A1; BSG; PMEPA1; TMED2; CTSD;	CLIC4; ARPC1A; HDLBP; DCAF7; ALCAM; PHGDH; TMED4; HIST1H1C; CCT3; SND1;

**Table 6-10** | Tables shows differentially **upregulated** genes in ASCs that are known to be upregulated as a result of TF inhibition/downregulation/loss of function mutation (cont.)

Transcription Factor	Adj <i>p</i> -value	Genes of Interest	Others
		PDIA3; SSR2; PDIA4; HAX1; PKM; TMEM33; CTDSP2; CANX; TRIB3; PPIB; ATF6; COPA; VCP; RPN2; MVP; RPN1; SEL1L; FAM129A; HACD3; PRDX2; LMNA; MLEC; HSPA8; HSPA5; PYCR1; P4HB; CALR; RNF187	SLC7A5; UQCRC1; ERGIC3; S100A9; CD63; ABHD2; SLC3A2; OAZ1; SLC7A1; FAM83H; GANAB; MDH2; HIPK2; CDK4; MARS; PLXNB2; NCAPD2; IARS; ACO2; AARS; GRINA
FOXP1 Knockdown	7.3E-05	CSTB; PSMD14; UBE2J1; PSMB6; TMEM147; HSPA5; UBE2C; DYNLL1; PSMA4; SPCS1; CDK1; NAA38; BUB1B; RACGAP1; ANXA5; ASNS; DERA; CENPE; TBL1XR1; ITM2B	NDUFB9; MTCH2; PCNA; NDUFA12; LRR1; COX7A2; COX6A1; OAZ1; JCHAIN; CKS1B; UBL5; BASP1; LRRC8D; SMCO4; DLEU2; CCT7; CHCHD1; CDR2; POLR2L; GINS1; DARS; RRM2; TALDO1; NME1; CCT6A; CCNA2; TMX2; BOLA3; BOLA2; NDUFAB1; TCEA1; GARS; NOP10; CDKN3; SLC44A1; NCAPG2; NCAPG; AK3; BRK1; SAMSN1; TMEM263; PAPSS1; ALCAM; NUSAP1; BLNK; ACADM; ME2; TMEM14C; CENPU; RFC3; RFC4; H2AFV; NEAT1; SHCBP1; ZWINT; HAUS1; ASPM; EML6; MELK; SUB1; POU2AF1; TCF4; IARS; ATXN10; OGT
RC3H1 siRNA	1.2E-04	PIGT; HERPUD1; PSMD8; CTSD; PSENNEN; PDIA3; ANXA5; DYNLL1; PDIA6; PDIA4; RCN1; TMX1; PPIB; LRPAP1; GRN; TMEM165; HSP90B1; PRDX2; PGRMC1; PSMB4; PRDX1; HSPA8; PRELID1; GALNT2; DHCR24; PSMC4; IFT27; CALR; P4HB; ITM2B; RNF187	SMARCB1; GMNN; HDLBP; AKR1B1; CCNC; COX6A1; UBL5; GUK1; EIF5B; HSBP1; H2AFV; CKS2; TTLL12; TOP1; ERGIC3; RABL6; ECHS1; NAXE; POLDIP2; COX7A2; TYMS; OAZ1; DPCD; HSD17B10; TM9SF3; CCNB1; POLR2E; BRD8; MDH2; UQCRCQ; CDK4; PLXNB2; MAGED2
SUZ12 Knockdown	1.5E-04	VCP; RRBP1; PRDX2; BSG; MLEC; CTSD; PGM1; HSPA5; SURF4; SIGMAR1; PYCR1; DHCR24; AP2B1; PDIA6; PKM; IMPDH1; CANX; CALR; P4HB; RNF187	ATP6V1A; MAGED1; HSD17B4; FOXM1; TM9SF3; LMNB1; MYBL2; EIF5B; MDH2; DESI2; H2AFV; CSNK1E; LARP4; TUBG1; SND1; CDC25A; SLC7A5; TXLNA; FSCN1; CALM3; BLMH; TOP1; EIF3C; S100A9
USF1 Knockdown	1.9E-04	CITED2; FKBP14; TECR; PYCARD; MAP2K2; GOLPH3L; FNDC3A; LAX1; PDIA4; PKM; COPA; LEPROT; INSIG2; MYO6; PA2G4; GLA	NRP1; SLC44A1; CEP19; NUDT1; DCAF7; JCHAIN; SMCO4; ENTPD1; IL18; TUBG1; LZTFL1; TWSG1; IRF4; CPTP; MZB1; TRIB1; PAFAH1B3; VGLL4; BCAT2; CUTA; FOCAD; ECHS1; MAGED1; SLC05A1; GSTT1; FOXRED1; SRM; NFIL3; NT5DC2; GBP1; PCK2; AZIN1; LPAR5; UBAP2; GALK1
IKZF1 Mutation	2.0E-04	KDELR1; AP3S1; COG6; KIF22; PKM; GSTP1; STOM; SPTSSA; NPM1; HSPA5; PYCR2; VMP1; ITM2B; ITM2C	LGALS3BP; ITGB1; ACAA2; HDLBP; BRK1; UBE2L3; LAPTM4B; WDR92; PHLDA1; CPT1A; GLRX3; PRMT1; LMO4; AKR1A1; TUBG1; INPP4A; SLC7A5; NDUFB5; CCNB2; IGKC; MAT2A; NT5DC2; ISOC1; DCAF12; TMEM176B; TMEM176A; GPN3; DAP; RSPH1; CDK4; PSAT1; RGS10; UBAP2; ACO2; EIF3C
RARA Knockdown	6.2E-04	ARF4; PMEPA1; TMED9; ANXA2; PLSCR1; PSMA4; NPC2; CSTB; GRN; CEBPB; TMED10; RPN2; EMC2; XBP1; DHCR24; DHRS3; PSMB8; BST2; SCD; GLB1; P4HB; ITM2B	LGALS3BP; IFITM1; LAPTM4B; BASP1; PHACTR2; TMEM189-UBE2V1; ATP6V0E1; PPA1; GPR160; S100A9; HIST1H2BD; HIST1H2BC; CD63; EPAS1; IFI6; SELENOF; ABRACL; OAZ1; EPCAM; ATP6V1D; MDH2; BIK; RDX; DSTN; ISG15; SCCPDH; REEP5; CYB561
MBNL1 shRNA	2.6E-03	VCP; SEL1L; HSP90B1; GOLGA4; PRDX1; CAPN2; FADS1; MCL1; NPM1; ANXA2; HSPA5; CAV2; ACSL4; DHCR24; PDIA6; ARCNI; PDIA4; VMP1; DNAJC1; PSMA4; PKM; ACO1; CALR	ARPC1A; HDLBP; TARS; EPRS; LIMA1; MT2A; ALCAM; ANXA6; ECT2; DLGAP5; ANKRD28; FDPS; DARS; ARHGEF12; ZBTB38; RDX; RCAN1; REEP5; SUB1; GARS; BCAT1
KMT2A Knockout	2.6E-03	ERO1A; DPAGT1; TRAM1; PSMD8; PMEPA1; TMED7; NPM1; ANXA2; ANXA5; F2R; SIGMAR1; KIF23; MIF; DHRS1; PRDX6; PKM; HID1; FADS1; ASNS; VEGFA	CYFIP1; HSPB6; CSE1L; MKI67; SMC2; CCNB2; MT2A; ALCAM; BASP1; ME2; ECT2; TNS3; BUB1; IER3; TMEM176B; ZBTB38; TMEM176A; CCNA2; ANLN; UCK2; TUBB2B; DPY19L1; PSAT1; MYADM; SLC44A1; ALDH1L2; AURKB; DUSP14; CEP55; GAS2L3; GOT1; IDH2; ATP2B4; ACTA2; ALDH18A1; IARS; ITM2A
SATB1 Knockdown	6.3E-03	ABC6; SAR1B; MANEA; ERC1; PGM1; SEC24A; ATP8B2; F2R; ELOVL6; KIF23; DERA; VEGFA; RCN1	RBM47; CDCA3; EPAS1; TULP3; SLC05A1; COPS7A; COX5A; DUSP16; BCL2L13; MCEE; IMPA2; MAP7; GNPNT1; BPNT1; SLC25A20; HMGN3; CTNNAL1; NDUFA9; CPT1A; GSTO1; IDH2; ATP2B4; RAB27A; ATP1B1; ETV6; DHRS7; MRPL51; UQCRCQ; PCCB; NCAPD2; NUP62CL; GPR19

**Table 6-10** | Tables shows differentially **upregulated** genes in ASCs that are known to be upregulated as a result of TF inhibition/downregulation/loss of function mutation (cont.)

Transcription Factor	Adj <i>p</i> -value	Genes of Interest	Others
NR3C1 Knockdown	9.1E-03	INSIG2; MVP; DERL1; RRBP1; PLOD1; NOD2; TNFSF13B; FUT8; LMAN2; ORMDL3; TNFRSF17; LNPB; PGM1; JAGN1; PSMB8; TAPBP; NCEH1	NDUFB9; CDCA2; TXNDC9; CSE1L; QPRT; ARHGAP19; TTK; SKA1; TRIM69; HERC6; PCMTD1; HIST1H1C; CCDC51; CCT3; CBR1; TIMMDC1; APOBEC3G; HRASLS2; DNAJC12; HCFC2; CARHSP1; TMEM106C; FSCN1; ZNF557; GALM; OGT; VGLL4

**Table 6-11** | Novel components uniquely **upregulated** in ASCs identified by proteogenomic analysis.

Gene Symbol (Total 1204)
KIF4, JCHAIN, TMEM97, CLPB, SELENOK, FOCAD, TJP2, NCAPG, PVR, EVI2A, PA2G4, PPFIBP2, PI4K2B, CDC45, ALDH7A1, GLA, FADS1, PIGK, CCDC167, GMDS, CHEK1, CDC20, DTL, MYDGF, NME2, TPX2, FAH, MGAT1, TECR, GOT1, PSAT1, TUBA1C, NT5DC2, CDCA3, KIF20A, HMMR, GSTO1, CKAP2L, IMPA2, FAM129A, AURKA, SLC1A4, MRPL37, TECPR2, SPC25, TK1, CDCA2, IQGAP3, TYMS, BRIPI, ST7, PFKM, NAPA, ARFGF3, GLCC11, WWOX, KIF22, FLT3, ESCO2, ENTPD7, SAPCD2, ALDH1L2, E2F8, KIF11, UBE2C, NEIL3, TMEM263, RAB39B, PIK3R6, FTSJ1, KIF18B, MELK, OAT, NAA20, DIAPH3, TNFRSF13B, MCM10, KIF2C, MLKL, GAS2L3, PLXNB2, COPZ1, PIGT, SLC7A1, HJURP, PBK, ARL1, GLCE, RAD54B, TXLNA, LRR1, SLC43A1, SLC30A6, IBTK, YIF1B, TMED5, CCNE1, DHCR24, CINP, ALG8, DNAJB12, PCLAF, TMEM33, SOWAHC, NXN, HASPIN, GNL3, EIF2A, ETV6, TXNDC15, MSANTD4, AMPD2, CYP20A1, PSMB2, RHEB, PRDX1, MRPS34, SPC24, PLXNA1, NME1, DDB1, EYA3, MRPL13, TRIM32, ATG4A, LY75, MINPP1, MRPL20, ANAPC5, GSPT1, BCAT1, NDUFAF2, UQCC2, CCDC88A, PSMB6, HEXIM2, MCRIP2, MRPS2, RDH11, COG4, AUH, PSMB5, ORMDL2, MRPL17, PABPC4, TMED1, PDE4DIP, NDUFB9, AAMP, CLPTM1, VPS25, EIF3C, SPG20, CLDND1, EIF2B4, PSMB7, TMEM41B, DCUN1D5, SCFD2, ADRM1, LNPB, SLC6A4, CALML4, DNAAF1, SCIN, CFAP46, RGS5, HIST1H2BG, TMEM167, CD59A, MCTS2, PLPP5, UBL5, GMPPA, NACC2, SLC04A1, PISD, SLC5A2, SELENOS, TRP73, MGLL, TCEAL9, SSPN, TBC1D24, GPR155, TXLNB, GMPPB, EEF1AKMT3, FAM129B, GTSE1, EXO1, CEP170B, VEGFA, TMEM205, BSCL2, OSBPL3, TACC2, NCAPH, GCAT, CCNE2, THBD, MVB12B, NUGGC, CST3, GM20388, PLCD3, SKA3, TWSG1, KCNMA1, GLMP, CPTP, ALDH3B1, SOCS2, MARVELD1, PCDH15, CCNB1, LMNA, PKM, BMI1, SLC1A5, PTPN21, PLEKHF1, PEX11G, RAB2A, CLIC5, CSGALNACT1, YES1, CDK1, DLGAP5, SGO1, PRSS16, KIFC1, NRP1, SUB1, MYO6, LAPTM4B, MANSC1, CFLAR, PRR11, PLK1, BOLA2, NOL3, NEK2, CDC6, CORO2A, ASF1B, BRCA1, CCNA2, UNC13B, CKB, ENTPD4, RAD51, CAPN2, SGO2A, PSRC1, MAGED2, GPRC5D, TOX2, CEP55, GPX3, NCKAP1, SPRED3, AURKB, EPAS1, CST6, GNG12, AB12, FAM114A1, ASS1, CCNB2, MAOA, IL15RA, RQCQL4, CDCA8, H2-T23, LRP8, ESPL1, ITGB6, BIRC5, GPR19, ABLIM2, FKBP14, ANLN, NCAPD2, PCBD1, FUOM, KCTD21, GINS2, CACNG6, KLC3, MAD2L1, TM9SF4, EHHADH, MELTF, EIF3J1, LDLRAP1, UBE2T, SETD3, MGS2, TTK, DHFR, CD274, ARHGAP11A, SCAMP2, FAS, SDF2, SLC7A3, ORC1, PDIA5, GPNMB, PKP2, ASRGL1, MYBL2, ETL4, TICRR, AIG1, CYB561, BMP6, BUB1, CNKSR1, BIK, IFT122, SEC61G, ATXN1, TRIP13, DCLRE1A, THYN1, H2-K1, MT2, LTBP3, BMPR1A, SEC14L2, CDKN3, TCTN3, TMEM141, STT3B, NEK6, SHCBP1, HIRIP3, ENPP6, ATP2B4, XK, CERCAM, FAM83D, NFIL3, DCAF12, CDK2AP2, SPOP, DTYMK, LTBR, FAM162B, GINS1, LGMN, CDC25C, PYPARD, EPN2, MBOAT2, EMC4, KDELC1, ULBP1, MAN1C1, CLN3, ECT2, DUSP26, PCGF5, GOLGA5, CCR10, COMMD3, ACP2, 5730409E04RIK, PHLDB1, MGAT3, MPI, ASPM, PCX, ZNHIT1, NCALD, DEPTOR, PNPO, EMP1, TMEM106C, SLC39A8, 1810055G02RIK, HSPB6, SYNDIG1L, ZC2HC1A, RAB4A, UBE2S, SLC7A11, YIF1A, TMEM256, SERINC3, SHB, SMOX, NDUFS5, PCNA, SWI5, LEPROT, DSTN, DARS, COX17, ITPRIPL2, ATRAD, AUNIP, SUMF2, GM17018, MTFR2, MOGS, PTGR1, KAZALD1, HIBCH, MGAT4B, TMEM255A, ARMCX2, NOD2, WDR62, PGAM2, ANG, 1110032A03RIK, MCPH1, ST14, SLC22A15, GIGYF2, CAMSAP1, ARL3, EDF1, ITFG1, PIPOX, PIF1, SLC11A2, KIF14, KNSTRN, RNF181, GALK2, FTL1, LCA5, TMED7, 1110004E09RIK, B9D1, SEPT2, SMPDL3B, PTER, NMRK1, GAS8, STXBPI1, CEBPB, UNC50, RPS6KA2, AKR1E1, MVB12A, CENPI, RRAGD, WARS, ABHD5, CXADR, TOX3, RHPN2, GM20425, CRYL1, DSCC1, UGGT1, KIF23, SLC29A3, SPAG5, TEX35, POLE2, MORN2, STARD3NL, MS4A3, DHCR7, RSPH1, TG, PRAD21, KNTC1, LIMS1, EHD4, ITGA3, MAP2K2, CRELD1, 6430548M08RIK, DOLK, DLG3, CD63, ALG3, QTRT1, CMC1, 1700047117RIK2, MARC2, TARSL2, CSNK2A1, ETFA, ACSS2, RIMKLA, RNASE4, TESC, HELB, RER1, NDUFA4, AIFM2, HYKK, LOXL3, SEPT8, XRCC3, TMEM107, UNG, GOLGA1, PLXND1, CUTA, FECH, MCEE, YTHDF1, SHISA4, IPO4, PIMREG, ALG9, NDEL1, ETNK1, SHQ1, UBE2L3, GOT2, FAF2, GTF2A1, CARD10, DEPDC1B, NODAL, IER3IP1, PRR5, HMGB3, MSRB1, MITF, PVT1, TMED4, FAM221A, NUDT9, IFRD2, NEU3, AVP1, CENPE, FOXRED1, CYB5D2, METTL6, ASNA1, ZFPL1, PSMC2, ECH1, DCAF10, IPO5, TM7SF2, GART, GOLPH3L, GEMIN5, LPIN1, GPR15, BLVRA, TRMT10A, RABGGTA, GDE1, IL2RB, PLOD1, EXTL2, NADK2, HEATR5A, ADAP1, COMT, PIGG, HAGH, PSMA5, ACO2, IER3, PPM1H, AU040320, LPCAT3, E2F7, PRC1, NDUFV3, ATG5, CEP19, EB13, KCNAB2, FAHD2A, ANAPC13, RHBDD3, MAP3K20, VSIR, NEO1, NUDCD2, TFDPI1, SHPK, S100BPB, AKR1B10, LZTFL1, ST8SIA4, ACADVL, NELFE, C1QBP, COL7A1, SSSX2IP, TBC1D7, TBCD, E2F6, SMIM7, HSD17B14, CMSS1, CDCA5, CENPF, PUSL1, ATP5G3, ACOXL, PSME3, FOXO3, MICALL2, UMPS, LAS1L, MTHFSD, HIST1H2BC, MUC1, DAAMI, IFNA1, LMAN2L, UBXN8, BLMH, CNTNAP1, SELENOM, THNSL2, GRM1, PIEZO1, NT5C2, MIS12, PTCH1, RFT1, GNAS, UFD1, SAMM50, TMTC2, CDC34, PRDX2, HIBADH, FAM45A, PICK1, ETFB, COX7A2, DNAJB13, PSMC1, CASC4, FAM173A, IL12RB1, ATP1F1, RAB23, ISOC1, ISCA2, PSMD11, SLC35F2, GEMIN6, FITM2, UQCQR, LYAR, ATP6V1D, DRG2, IL6RA, TMEM219, SLC45A4, TIMP2, EIF4G1, SDHAF4, PLCD1, TXNRD1, GSTM4, HINFP, SLC25A25, ZFP428, CLPX, PDZD11, DNAJC28, LRRC41, TMEM70, POP1, MRPS7, HIVEP3, ANKRD46, ACO1, DNAJA3, SURF1, RFC3, MRPS12, SAP30, CSTB, WBP1, RND2, TFB1M, MND1, ZFP825, ATF6B, UROD, CSF2RB2, KDELC2, FLNB, SNAP47, NGLY1, PSMC3, CTAGE5, PAPS51, TNFRSF10B, XCR1, ANKRD6, HADC3, XKR8, ATP6AP2, SVIL, MTX1, CBWD1, MKRN2, PRMT7, KLHDC8B, DOCK4, SLC35B2, IFT43, EMC9, DNP1, BFAR, OPA3, ISG15, EIF2B1, PRKCI, ELOC, TRPT1, SLFN9, CASKIN2, TUBG1, SVSIG10, BRCA2, SUV39H2, REEP5, DCTN3, NDUFB6, LNP1, PRDX6, PRIM2, TKTL1, NDUFB11, TIMM17A, RAB13, DHRS3, ABCF2, NIFK, EIF2AK3, STIL, ATP5J, UBE2G1, AHS1, NDUFA13, NUPR1, IKBP, MAT2A, CRYZ, SIGMAR1, GINS3, NUDT5, GUK1, NENF, MRPL50, FNIP2, MCL1, PHPT1, USP54, CCT3, EIF2S1, UCHL5, PLXDC1, NCAPH2, NUDT2, LRRC28, TEX30, NDFIP2, RBKS, STAU1, CMC2, LUZP1, LMF2, MUL1, UBE4B, COX6A1, SELENOF, EIF2AK1, ABHD2, USMG5, PSMD6, COPS6, PPAT, TP11, HMOX2, COX5A, ATP5H, B4GALT1, ALDOA, DYNC2L11, KIF1B, 4933434E20RIK, SKA2, GRAMD3, IPP, ATAD1, IGHG2C, ACOT13, MTX2, NDUFA2, DTD2, NAA38, SOD1, TMEM106B, MTERF4, ENKD1, MON2, COPS4, TMED6, PSENEN, PEX11A, MYO5A, POLH, PSMB4, GALK1, PSMD13, NDUFA11, CHAC2, FABP3, ITFG2, MDH2, MRPL33, NAB1, NUP35, HSD17B4.

**Table 6-11** | Novel components uniquely **upregulated** in ASCs identified by proteogenomic analysis. (Cont.)

<p>MINOS1, CCT7, PTDSS2, UBE2K, SACM1L, RNF14, NTHL1, UEVLD, GSTM5, MRPL34, DDX54, JAG1, DUSP19, PYCRL, MBD3, CPD, DNAJC12, ICAM1, RARS, METTL26, ATP5J2, ACACB, TIMM21, CISD1, ST6GALNAC4, RABL3, PSMD8, RAD54L, PSMD14, EMC1, GLRX2, UBF1, PSPH, TUBB4B, MIA3, MRPL22, GCLM, FKBP1A, RPAP3, MRPL35, TIMMDC1, PRMT1, RAB6A, DHRS7B, P3H1, ZCRB1, RABEPK, MRPS14, UBL4A, TSR1, KIF18A, ELP5, ACSL4, NEK4, MRPS15, SMYD2, EIF4EBP1, ZFP706, MRPL53, CDK4, CIAO1, POMP, ELP4, CARM1, LSS, CNOT9, MRPL57, TIMM50, ACSL5, AIMP2, LARP1, HIGD1A, MRPS18A, NDUFA9, ARMC1, FAM120A, NSD2, FAM96A, AVEN, PEX19, GLS, MRPL48, PSMG1, ATP1A1, TTL12, WDR92, MRPL28, MRPL16, DCTD, BYSL, ISYNA1, MRPL27, PSMD1, PSMC4, PEX14, CTH, ELP3, PSMD4, MECR, PEX3, SLC52A2, TOX4, DDX19A, ZCCHC9, TARBP2, GMPS, EIF4E, KDM5C, NTMT1, OGT, NLN, UBOX5, EIF2B3, MRPS5, PSMC5, DDX3X, CHEK2, OGFOD1, POLR3D, PREPL, LACTB2, METTL15, NDUFAF1, FNTB, MRPL18, MRPS23, MRPL58, MRPL46, TULP3, SLC25A17, DNAJC15, SCO2, MRPL4, RNASEH2C, NME6, FTO, SIRT4, POLDIP2, LARP4, MRPS24, DBI, AURKAIP1, EMC2, FZR1, RARS2, MRPL15, QSER1, TMEM129, MRPS11, CKAP5, MRPS17, TTC1, EEF1AKMT1, MRPL21, HNRNPAB, IGHG2B, KCN18, KIAA1549L, SCN9A, SLC9A4, PAK5, OR10A4, CYP2F1, TOPAZ1, OR2D3, ESRP1, TMEM200B, HEATR9, OR5R1, RFPL4AL1, OR4F15, SORCS1, MOBP, CBLN1, ELF3, AC092821.1, KIAA1024, U1, CYP4F12, TMEM178A, CD34, DYTN, ZNF462, EPHA10, ZNF750, LY6G5C, SLC6A20, UGT2B15, OR5A51, TPRG1, ZNF804A, PNMA8A, DPCR1, CHST6, ZSCAN5A, CYP4F11, UNC79, KIAA1456, RGPDI, C1S, VWDE, AL161911.1, SIGLEC15, MYO15A, MSMP, NLRP2, ZNF185, KAZN, PLB1, WNK3, CA14, FAM209A, TNK1, AC118553.2, OR2AG1, SERPINA1, RGPD5, CA5A, RGPD8, FAM47E-STBD1, CAV1, NLRP7, FAM92B, MAGEC1, PTPRQ, CES1, RBMXL3, PCDHA6, AC083902.2, MUC22, CYP4F8, OR2T6, RTL1, OR10P1, CNTNAP5, C1QTNF1, CES3, LINGO4, SPPL2C, OR6C3, OR2J3, SPATA31D4, OR9Q1, TSHZ2, CCDC129, KRT6C, OR10D3, OR2AT4, DSG1, MAP3K15, OR13C3, FAM47A, TPTE2, OR52N2, DCAF8L2, OR4K5, SEPT5, BGLAP3, GFY, OR4K2, PRSS57, OR2AJ1, ADAMTS19, OR13C5, GSGI2, OR51Q1, FAM47B, OR4C13, OR2L8, NPY4R2, GPR119, OR6C74, OR52K1, PCDHA4, CISH, TTL8, SBK3, OR7G3, ZNF536, SLC6A1, UGT2B17, OR5AU1, OR13H1, ZDHHC11B, TEX13C, MUC17, OR8U1, OR10A2, OR5V1, OR4C15, OTOL1, KCNK18, OR9A4, PCDHGB1, PCDHA13, MICALCL, ZNF205, OR1M1, OR13F1, OR4L1, MAGEB6P1, OR9G1, NEU4, TRIM75P, DNAAF3, OR6K3, OR6K3, ADH1B, OR6Q1, PCDHGA6, TMC3, MAGEA10, SP9, SPATA31D1, OR4M1, OR4C46, OR13C2, OR10C1, PCDHGA5, PCDHGB4, OR6N1, OR9G4, PER3, OR1S1, OR51A4, OR6C6, VN1R5, TMEM151B, H2AFZ, HRNR, OR14J1, SPATA31D3, PCDHGB2, TP63, CYB5D1, OR4F6, TNR, OR1N2, ZNF334, MTCP1, OR1B1, OR52E4, FAM205A, TTC16, OR2AK2, KIAA1755, MYOC, CROCC2, AC135068.1, KRTAP29-1, KRT40, RUFY4, SERPINB4, OR5A1, LRRC30, GM867, OR9K2, RDH9, OR5A2, 1520401A03RIK, OR51D1, PELI3, NLRP9, CTNNBIP1, CFAP47, HOXB6, MFS2D2, HSD3B7, BORCS8, ZNF385B, OR51F2, DCAF12L2, TMEM253, PCDHGA12</p>
---

**Table 6-12** | Novel components uniquely **downregulated** in ASCs identified by proteogenomic analysis.

Gene Symbol (Total 1204)
<p>RIPOR2, STK26, TMEM131L, AIDA, PIK3CD, IFIT2, FCMR, ARHGAP45, SLC2A3, GMIP, DMXL1, NFATC1, ZMAT1, AMPD3, SP110, RASSF3, LRRFIP1, ZFP639, PRKCE, TRIM7, FOXO1, FAM208A, LGALS8, CD200, PTBP3, ANKRD13A, ZFP263, VEZF1, ZNF518A, CD2AP, INO80, NFATC3, ZBTB5, APIG2, DENND6A, CD79A, SLC25A24, MPRIP, STX7, BCL7A, TOP2B, ATG16L2, ZFP422, RXRB, MTM1, AKAP8, AP1S2, CEP135, KAT8, TRIO, PPP1R12A, AGO1, CEP68, PSD4, RABEP1, NR3C1, PDLIM2, MAPK14, CEP295, TCP1L2, BICRAL, ZBTB7A, PHF2, ARID4A, FMO5, PDP1, STRIP1, EZH1, ARID1A, PIBF1, NUMA1, ATXN3, PGM2L1, SETX, HMG1, ZFP638, ZFP740, ARID4B, SUGP2, ARPC2, JARID2, CCDC82, NONO, AKAP8L, PEAK1, AGPAT2, NFATC2, DOPEY2, ZC3H6, CAMK2D, TMEM2, THUMPDI, UTRN, KANSL3, PRKX, UBA7, RBM5, ATP2A3, MSL2, ANKRD11, MECP2, LNPEP, YWHAZ, RPRD2, NEK9, HECA, WASHC4, NR2C2, GANC, MCTP2, USPI2, CEP170, VPS13C, ARHGAP15, RP9, MAP4, PPTC7, SYF2, ACCS, PRR14, CNTRL, UVRAG, ILF3, IST1, TMEM63A, RBL2, DHX15, PDCD7, CPSF7, NIPBL, ANXA11, CREBBP, RCOR1, DENND5A, OTULIN, RYBP, ESCO1, MAP3K2, SLC23A2, NRF1, SNX6, ABCD4, DDX59, SPIN1, TAF3, NAFSYN1, WASHC2, RAPGEF6, ZFP746, SNRK, DIDO1, KBTBD2, EPC2, SFPQ, PDCD4, NCOA1, EP400, OSSEP, CCNY, TAPT1, NFATC2IP, ADO, BRWD3, HACE1, TERF2, NCOR1, STK24, CCNDBP1, PLEKHM1, PHF20, EIF4A2, ODF2, TERF1, ACIN1, EFR3A, BRD4, UIMC1, ACSL1, EFHD2, ARHGFE7, PGLS, KAT5, HMGXB4, PITPNM2, COPG2, ZBTB24, PPP3CA, RAB21, PP1L4, SH3BP2, ELF2, METTL3, AFF4, GABPB1, CAPG, TRIP12, SLTM, DNMBP, FBXO41, STMN3, KIAA1683, OSBPL10, H3F3B, CAPZA1, HIST1H4E, L3MBTL4, LAMB1, SIAH1A, 0610030E20RIK, GDF7, MARCH1, MRFAP1, ZFP273, VMA21, CFAP44, CALD1, SLC4A7, DTX1, ZBTB4, 1830077J02RIK, FAM129C, CRYBG1, SESN1, DENND4A, ABCA1, ZFP821, RETREG3, TTC9, SLC38A11, HMGB1, NLR3, TRIM58, SLC38A1, COLGALT1, GMFB, JAKMIP1, MARCKSL1, TIMP3, ACOT2, CYB561A3, PITPNM1, CHML, PLEKHA2, SLC43A2, MAP3K8, RUBCNL, TEX10, PHTF2, TRAK1, CDC14A, FYN, DOCK8, TNFRSF13C, STS, ZFP329, MNDAL, H2-Q6, TEP1, ZFP157, ZFP512, L3MBTL3, LGALS9, SLF1, 2810021J22RIK, CD55B, MGAT5, CRIP3, AKAP5, ZFP65, RUBCNL, UBE2R2, ADAM28, KLHL36, PIP4K2B, CCND3, R3HDM2, ZFP90, TRP53BP1, MYO9B, ZBTB14, CHKB, SEPT6, TMEM260, CNPY4, RETREG1, NSMF, GAB3, RELT, UHRF1BP1L, QK, CNRIP1, RNF145, BIN2, DIAPH1, DEPDC5, SIDT1, XPO1, DHX57, B3GNT7, USF3, CPNE1, RFX7, ZFP148, GPSM1, TES, CASP2, GATAD2A, DDX60, ANGEL2, CBFB, 9130401M01RIK, ATR, ITPR1, SMARCA2, CARMIL2, SMURF2, SLC39A10, CEPT1, ZMYM6, ZMYM5, ERP29, STX17, 2610507B11RIK, DUSP10, RUFY1, PARP6, MBTD1, GM14698, HEATR1, MLLT11, TCERG1, USP34, FAM107B, MED4, CEP85L, KLF16, CHD1L, JAK2, LPP, SYNE3, CHD1, TAF7, ZEB2, DDX6, PACS2, DTX4, OTUD4, RNMT, CYLD, CCDC88C, LBR, ATM, UBAC2, EGLN2, INVS, ZFP729B, RCC2, RAP2C, ATP8A1, TIGD2, RAB37, CAPRIN2, MYBPC2, ICE1, CYP4F18, DDX31, BICRA, SLC12A6, GLUD1, NFKBIA, NUP153, PDS5A, ITSN2, TARBP1, FANCM, IKZF1, SKAP2, PARP11, SNX10, TAF1A, MTERF1A, GABPB2, MCM9, SLC16A7, SIPA1L3, KDM5B, UNC93B1, CAP1, FAM102B, ERP27, ZCCHC7, INAFM1, HIP1, APPL1, ZFP729A, PRKDC, CASD1, UBA2, TWISTNB, PTK2, RAE1, POLI, CHMP2B, TAZ, ZC3H4, PTEN, BEGAIN, GNA13, DYRK2, SURF6, LPCAT2, CAST, PIK3R1, SLC6A6, ZFP143, NOP53, SENP1, ABHD17B, PDS5B, TMEM131, GPALPP1, BLOC1S2, STX6, PHF21A, ZFP407, TSGA10, DHX9, ZBTB37, CTDSPL2, MICALL1, PCMI, PUM3, ATG16L1, BOD1L, YTHDC2, ICE2, MEPCE, BBS4, WDR37, CCDC191, NUDT3, KANSL2, SIN3A, CHD6, GPM6B, KLHL42, ZCCHC8, SIPA1L1, PRRC2C, SLC2A1, TUBA1A, FBXO31, ZDHHC17, ZDHHC23, FBXO11, RAB4B, RBM27, FBXL3, EXOC2, PHF23, CHD8, POLG2, IREB2, ZFP28, KDM2B, 2610008E11RIK, PRDM4, AKAP7, STRADA, ARF6, FASTKD5, RASA1, FBF1, CCDC137, ZFP335, FAM117B, CC2D1A, CLASRP, ADRB2, TACC1, MTF2, DET1, DNAJC2, ABL2, OXSR1, MANBA, R3HCC1L, TGS1, ZBTB11, PHF3, ZFP35, TMEM222, DSTYK, RNF220, CYP2D22, IER5, SEC24B, TTF1, TAGAP, PPP4R3B, TAF4, GON4L, PLCG2, RAB32, PCMTD2, CHD4, WARS2, AHNAK, FIP1L1, ARHGEF6, CHMP1B, FAM126A, KLHL5, HARS2, MGEA5, PDSS1, ABRAXAS1, SMCHD1, SCIMP, GMCL1, NUCKS1, ARID5B, UCKL1, DDB2, B230219D22RIK, LASP1, SETD1B, TAOK3, ASH1L, CSEIL, TSC22D2, FRAT1, TOB2, FBXL8, PADI2, KDM6B, BRPF1, ATP2B1, HNRNPR, RPUSD2, MINDY2, UBXN1, USE1, IWS1, THOC1, PPP4R3A, CUX1, CNOT6L, MAF1, LEMD3, RECQL,</p>

**Table 6-12** | Novel components uniquely **downregulated** in ASCs identified by proteogenomic analysis. (cont.)

---

ZBTB10, ZFP317, IFT74, TNKS2, EPB41, TMEM55B, KLHL20, ALOX5AP, SVBP, GGPS1, KLF7, ASTE1, MAML1, INPP5E, MAP2K1, TERF2IP, TBC1D32, MAU2, MFAP1B, PAK2, BEND5, PHIP, MFSD14A, RNF31, VPS13B, MSL1, SPI, SP TLC2, KAT2A, SNX3, TESK2, RAD52, KLF13, SPOPL, ATXN7L3B, BCL10, MAML2, FNTA, ZBED4, LETM1, GNA12, TIA1, THOC2, GPR137B, CEP120, ZFP207, JMJD1C, HNRNPU, TRIM21, TAF5L, COQ2, CLK2, PIK3R4, RPL18, CEBPZ, ABCB1A, ACTR5, AP5Z1, TAF11, MAP3K3, ARL6IP6, PPIG, SDCBP, USP3, PCF11, OSBP8, XIAP, TRPS1, SELP, NRDE2, RLF, MAPK1, TMEM43, NOL11, RALGAPB, PTS, USP7, SPPL2B, MBNL1, CELF1, SART3, ERMP1, FAM168A, SACS, RALGAPA2, VPS18, CAR2, CREBRF, DCAF4, RBBP6, ZYG11B, IFI208, SNX25, RCN2, KCTD2, AQR, LMBRIL, HECTD1, RDH14, PUM2, RPGR, PHF14, ZFP326, GPATCH11, DNAJC27, MKL2, PYROXD2, KANSL1, ATXN2L, SZT2, CD1D1, CTCF, GPS2, MEF2C, MYNN, PTBP2, TBL1X, MCRIP1, AP3B1, UBE2W, MYD88, DDX55, SMC5, HNRNPC, SF3A1, ELF1, SNRNP200, ARRDC2, EIF3F, CMPK2, LTA4H, SFN, CCDC22, VHL, GPBP1L1, MKLN1, WIPF2, U2SURP, DDX39B, PHF20L1, YTHDC1, WDR11, POLR1C, KMT5B, CRTCI, BNIP3L, RWDD3, TRA2A, AGFG1, MCM3AP, POLR2D, FBR5, TECPR1, CHTOP, FIG4, MTERF3, BTBD1, RPL8, PDP, ABHD10, NCBP3, BICD2, BNIP2, 4921507P07RIK, NUP205, UBL3, FBRSL1, TGIF1, GCC2, ZFP292, WSB1, PSMD5, KDM3A, RANBP6, SLC12A7, HUWE1, HSPBAP1, SECISBP2L, RNF216, SIRT1, TPK1, CAB39L, ZFP830, FBXO21, PLCG1, SDCCAG3, ATE1, FTH1, NPAT, DIAPH2, OVGP1, LONP2, ZHX2, CUL5, MEAF6, BRMS1, BMT2, UTP15, SRSF1, SF3B2, XRN2, GPATCH8, DYRK1A, ZFP865, SMARCA5, MCMBP, KMT2A, AP3M2, PDZD8, RLM, SERPINE2, ZFP131, SMARCD1, RPL37, RPS11, SPATA13, UBT, TPR, MLXIP, RALGAPA1, RERE, CIPC, TOR4A, API5, ZFYVE27, RPL37A, NOP16, ING1, TSR2, YPEL3, CERS4, ITPR3, BCOR, DDX23, SIRT5, PHKA2, FILIP1L, GIGYF1, TRAPPC8, CLK1, PTPN18, AMDHD2, SPG11, SGSH, MOB3B, SLC25A32, TMEM134, THUMP2D, LRBA, INO80D, EIF1B, DDX27, PRDM2, TNPO1, STX4A, TNIP1, DHX36, PAN3, ZFP619, BRD8, FAM193A, RBM34, DPF2, ZBTB3, SPRNT, RICTOR, DGKD, SF1, EXOC6B, XPA, ENSA, ZFP524, TMPO, BDP1, MYO9A, GATD1, CDK5RAP1, IFI205, DNMT3A, HNRNPDL, AMZ1, ILF2, ZFP58, SH3GLB1, ALDH16A1, 1700037H04RIK, TSNAX, RPS25, CHIC2, SKAP1, IARS2, MAVS, SPG21, ZC3HAV1, IRF9, HNRNPUL2, HNRNPM, CBL1, TRIM11, NCL, LUC7L3, RBM6, KCTD13, NSD3, RACK1, RPL23A, SBNO1, AHI1, SAMHD1, CPSF6, PPCS, VPS16, ZFP346, POGZ, QDPR, ERBIN, FAM160A2, FAM20B, DENND4C, RASSF1, ZHX1, RASA3, IKZF3, ANKRD12, AFTPH, CUL3, KDM5A, MTMR3, MYL6, MDM1, NFRKB, PPM1A, TGFBRAPI, SFSWAP, ITPR2, AGL, NSUN6, UCK1, KHNYN, METAP2, ZFP953, STRN, MNT, WDR47, CABIN1, SDE2, HEXA, CAND1, EPC1, ZFP800, VPS26A, RSRC1, LEMD2, AKT2, PHF1, RPS6KA3, GMEB2, ZC3H12A, NDRG1, TLN1, PATJ, RAB24, LAMTOR1, FOXN3, FBXO38, HPS4, FCHSD1, UBE2I, KCTD18, VPS11, ATP9B, USP6NL, IKBKG, APPL2, ATP2C1, RELA, GDII, ING3, GNL3L, CCDC93, SPICE1, KDM2A, RNF114, SART1, TTC14, HIST1H1D, SELENOO, MAPK8IP3, ZFP654, TUBGCP6, TMEM71, TAB2, SNX12, CLCN7, ZFP429, CAPZA2, MTRF1L, SMARCE1, WASHC1, ZFP809, SMC3, AKAP11, RUFY2, COQ8A, FAM32A, PIAS1, AP5M1, 2410004B18RIK, ZFP455, MED17, ZFP383, LRRC37A2, CCDC154, ZNF578, C1QTNF3-AMACR, TTC34, LINC01125, ZNF611, SMIM18, ZNF28, ZNF813, SBSPON, CCDC7, SPDYE2, UGT8, CYP3A5, ZNF808, SPIN2B, STEAP1B, ZNF354A, ZNF816, SPDYE6, RLN2, SPDYE1, OR13A1, ZNF525, HEATR4, SPDYE3, SLC9C1, ZNF320, SPDYE16, UPK3BL1, ZNFX1-AS1\_2, ZNF888, RGPD6, ZNF816-ZNF321P, MICB, LRRC37A3, PRKN, FAXDC2, ZNF853, LRRC37B, HIST1H4K, DLEU1\_2, KIAA1324L, AC073264.3, RNU6-118P, RGPD2, FPGT-TNNI3K, TAS2R14, SOX4, TCL1A, TCL1B, LCN10, COL19A1, OR2A7, AVPR2, HTR3A, HIF1A, ZNF711, ZNF860, FFAR3, FAM222A, CEACAM16, QRF, NRARP, TCRG-C2, RIMKLB, WDR38, PPP1R3F, RDH5, ZFP82, RNASE12, CLEC9A, MANSC4, SCARNA2, SLC38A8, GML, SLC36A3, TIAM2, HOXB3, DYX1C1, LINC00854

---

**Table 6-13** | Upregulated Transcription factors in ASCs uniquely identified by our multi-omics analysis. Genes soft validated by proteomics are given in bold

---

Gene Symbol (Total 41)

**BHLHA15**, NACC2, BHLHE41, GFII1, ZBTB42, SUB1, CDC6, TOX2, EPAS1, PCBD1, **BRIP1**, **MYBL2**, **ATF6**, **WFOX**, **E2F8**, BMPR1A, NFIL3, CREB3, NFXL1, CEBPB, TOX3, **ETV6**, MTF, **DDBI**, E2F7, **MAFG**, **TDFP1**, LZTFL1, E2F6, FOXO3, PTCH1, HINFP, **HIVEP3**, **ATF6B**, **SUV39H2**, MLX, **HMGN3**, RNF14, **PSMC5**, **DDX3X**, **HNRNPAB**

---

**Table 6-14** | Downregulated Transcription factors in ASCs uniquely identified by our multi-omics analysis. Genes soft validated by proteomics are given in bold

---

Gene Symbol (Total 124)

**AFF3**, **CITA**, **SCML4**, **SP100**, **DTX1**, **ZBTB4**, DENND4A, **HMG1**, **ARID1B**, **SP110**, **LRRFIP1**, **SP4**, **SMARCD2**, **FOXO1**, **RUNX1**, **FOXK1**, **FLI1**, **ZFP263**, **VEZF1**, ZFP90, **BPTF**, **RARA**, **NFATC3**, **ZBTB5**, **REST**, **SKI**, **RFX5**, **RFX7**, **ZFP148**, **ATR**, **SMARCA2**, **MAFK**, **MLLT11**, **STAG2**, **KLF16**, **ZFP422**, **RXR**, **LPP**, **TAF7**, **ZEB2**, **ATM**, **NR3C1**, **NFKBIA**, **IKZF1**, **TAF1A**, **ZBTB7A**, **ARID4A**, **FUBP1**, **ATF7IP**, **EZH1**, **ARID1A**, **ZMYM2**, **HMG1**, **ARID4B**, **MLLT6**, **JARID2**, **SIN3A**, **NFATC2**, **PRDM4**, **CC2D1A**, **MTF2**, **TGS1**, **TFEB**, **TTF1**, **MECP2**, **NR2C2**, **ARID5B**, **DDB2**, **CUX1**, **ZBTB10**, **TFAM**, **KLF7**, **MAML1**, **SPI**, **KLF13**, **PDCD7**, **HNRNPU**, **CREBBP**, **RCOR1**, **CREB1**, **CEBPZ**, **RYBP**, **TAF11**, **TCEA2**, **TRPS1**, **FOXP4**, **NRF1**, **TAF3**, **CREBRF**, **APEX1**, **DIDO1**, **MKL2**, **CTCF**, **MEF2C**, **TBL1X**, **NCOA1**, **ELF1**, **TGIF1**, **NCOR1**, **PHF20**, **ZHX2**, **BRD4**, **SMARCA5**, **KMT2A**, **SMARCD1**, **UBTF**, **MLXIP**, **BCOR**, **ELF2**, **PRDM2**, **BDP1**, **IRF9**, **SRF**, **TAF8**, **ZHX1**, **IKZF3**, **STRN**, **MNT**, **GMEB2**, **FOXN3**, **RELA**, **SMARCE1**, **PIAS1**, **MED17**

---

**Table 6-15** | Contradictory regulation. Genes **upregulated** in mice ASCs but **downregulated** in humans ASCs across both RNA-Seq and microarray data

Gene Symbol (Total 258)
TMEM154, RILPL2, CORO2B, ADK, IL13RA1, SMIM10L1, AGPAT4, SLC12A2, TMBIM1, HTATIP2, GPR65, SNX22, PTPN22, CERS6, PSIP1, WEE1, TSPAN13, NCK2, FADS3, ABCB1B, MAP1LC3B, HK2, HS3ST3B1, FDFT1, QSOX2, NEIL1, RNF125, MTHFR, PSME4, CNST, GPM6A, METTL21A, PHF10, PPF1BP1, MTMR10, FAM135A, AKIRIN1, CEP44, KIF16B, TPST1, MCOLN2, PAWR, NVL, AGPAT5, UXS1, LIG4, USP47, SLC46A3, CKAP2, IPMK, ITGB2, GNB5, PHKB, JAZF1, P2RY14, CMPK1, CENPK, DYNLT3, MAPRE2, CCSER2, ZFP9, SMAD3, NRIP1, KIF5B, SLC36A4, ZRANB3, ABR, TLR2, SIAH2, APP, SERINC5, FAM91A1, LMO7, ZCCHC18, GNB4, YPEL2, CENPC1, CENPV, YBX3, MEF2A, RCBTB1, CLN8, UBAC2, HGSNAT, PIGB, ZFP386, TBC1D22B, HINT3, LPIN2, INTS6L, THAP2, PON2, APOLD1, ANKLE1, RIOK3, KCTD9, STX3, KAT2B, NKTR, CCS, DESI2, CCDC71L, ZDHHC2, CNOT8, ERCC6L2, GNAZ, SIK3, LHPP, GCA, GALNT7, FNBP1L, ATL2, LYRM2, ATP6V1H, ASCC3, EIF3B, GKAP1, TMEM203, CLHC1, SIDT2, SENP6, CEP192, CMS1, FAM216A, TGFBR1, RBSN, ISCA1, 201011101RIK, SFT2D2, NEK7, SETD7, AMFR, HAUS3, TTC32, AKAP9, SWSAP1, SKIL, ALKBH2, HACD2, WDR59, TXLNG, ZHX3, IFFO2, MED23, HIVEP3, TBC1D22B, CDC123, MAN2A2, KCNMB4, IRS2, MDM4, IRAK1BP1, TMUB2, CCDC6, WHAMM, ZBTB20, UBE2Z, TTC13, DNAJC21, REPS1, CENPL, SC5D, LARS, MED30, NNT, MOB1B, PCED1A, RIPK1, MTMR4, ZC3H7A, SIN3B, ZC3H11A, ESF1, SLC25A33, TCF12, PBDC1, ZFP330, CFAP20, DEGS1, MAP2K3, RANBP2, MFN1, CHMP4B, OTUD3, MRPS25, USP22, NCOA2, ITFG2, 311004302IRIK, JAM3, POLE3, KAT7, TASP1, INTS3, CAR52, DCLRE1B, BRAF, RAPGEF5, A230050P20RIK, DNAJC13, 4930453N24RIK, PRNP, PNPLA8, ARL14EP, ERO1LB, TSPAN3, IL1RAP, RABGGTB, CCDC88B, ACER3, ADAT2, PPP3CC, DYNLL2, METAP2, PBRM1, BIRC2, MED22, MAN2B2, COX19, GRK5, GABARAPL2, USP9X, CBX4, PJA1, TM2D3, COX7C, XPNPEP1, SLC25A13, ARHGAP12, MCUB, ARL5A, AKAP1, ZFP281, PLEKHM2, TSSC4, TTC39B, TUBE1, UBR3, OGFRL1, MXD4, IGF1R, HIPK1, CHUK, IFRD1, RBM7, SEPSECS, TOP1MT, SYTL1, SLC15A2, RNASEH2B, AHS2, AGO3, MEGF8, NPLOC4, ZIK1, SNRNP27, MZT2, DIRC2

**Table 6-16** | Contradictory regulation. Genes **downregulated** in mice ASCs but **upregulated** in humans ASCs across both RNA-Seq and microarray data

Gene Symbol (Total 244)
LDLR, KLHL6, ZBTB32, SLFN8, CD38, RAPGEF4, TLE3, RNF144A, STAT4, AIM2, OPTN, SAMSN1, GNS, BCL9, RFLNB, PMVK, THEMIS2, NABP1, VOPPI, MTHFD1L, CNP, PSEN2, CHST15, LTK, CTSC, DUSP7, PGLYRP2, TOR3A, HEXB, USB1, MTERF2, DEX1, SMARCB1, NETO2, CTSA, SLC39A3, ZCCHC24, PTPRS, TTBK1, NUDT19, PARP2, AACS, ZDHHC18, WW2, DENND6B, UNC13D, PRR12, CD276, SMPD3, TTC38, SGK1, CPNE5, B4GALT3, ELAC1, PYM1, PQBP1, CISD2, PQLC2, TRADD, FBXO22, SLC25A11, DCPS, AAAS, RECQL5, UGDH, TUFT1, MAPKAPK3, GADD45A, CASP6, OTUD7B, MAPK12, S100A11, SFXN2, ERG28, BUD23, ABCD2, CDT1, SMIM20, KCTD12, HSCB, RGL1, CPT1A, SLAMF1, CKLF, DENND1B, DGLUCY, S100A6, LRRCC1, RGS13, SLC35D2, NPM3, MAGI3, 3110009E18RIK, PRMT5, TSEN34, PPCDC, BATF, 5031439G07RIK, ZFP574, IFI35, NCLN, FBXW5, SMUG1, SWT1, MAP2K6, SLC9A3R1, CLCC1, DTNB, ZFP593, SLC37A4, SCAMP3, CETN3, VAMP5, FDXR, VPS72, CDC42EP4, TMEM206, CEP70, MAF, GCLC, PLBD1, ALKBH7, SH2B2, POP7, ARMC9, GPD2, TMX2, EYA2, SUOX, FAM69A, GSTK1, MDP1, STN1, KIF1BP, CABLES1, DNAJB5, MEF2D, UTP11, WRB, MRPS6, PDP2, MCCC1, HYI, INSR, RAD17, CD226, FGD6, EMC6, PNPT1, MTUS1, FAM50A, BCAR3, CYFIP1, KIF1C, CCDC32, CENPS, ICAM2, ALAS1, LRRK1, AKR7A5, PRPSAP1, B4GALT7, COX11, ARMCX5, MRPL49, UBE2CBP, LIN52, HSPB11, MRPS16, TMEM8, SNX15, TMEM173, CNOT3, LRRCC42, NDST1, NUDT6, SRD5A3, POLR2H, PAM, URM1, 2310022A10RIK, NLK, NCKIPSD, QDPR, ACTG1, EBNA1BP2, NDUFAF6, ENOX2, TOMM5, RPL22L1, UCHL4, PRKAG1, PPA, MCOLN1, ZFP598, PDIK1L, FAM118B, JMJD4, ZFP664, AAK1, PPP1R7, ACP6, DGKZ, SFXN4, FAIM, ERLIN1, PDCD2L, COX18, P2RX7, PPP1R13B, FAM206A, PIP4K2C, MPG, ADM, ATG4B, MAST1, SLC25A19, PRDX5, CTSH, IRF5, TRMT2A, ELK3, MBOAT7, TXNL4A, LSM1, PUS3, MRPS31, PIPNA, CALHM2, PAK1, ERCC2, FHOD1, RAB35, ZFP827, HSPA8, SLC26A6, MRPL3, ELP6, MCC, PUS7, TEX9, HCFC1R1, MAPK7, CHST7

**Table 6-17** | Cluster Differentiation markers upregulated in ASCs vs. NBCs across species in RNA-Seq and/or microarray data. Ordered by FDR adjusted *p*-value.

Gene Symbol (Total 42)
SLAMF7, SLC44A1, ENTPD1, LAMP2, CD93, SPN, BST2, SELPLG, CD2BP2, HMMR, EPCAM, BSG, CD274, FAS, FLT3, ITGB1, TNFRSF10B, TNFRSF13B, ITGAL, SLC3A2, LY75, IGF2R, ICAM1, ALCAM, SDC1, TNFRSF17, CD28, CSF2RB, LAG3, PVR, CD59A, NRPI, IL15RA, CD68, BMPRIA, ITGA3, IL6RA, IL12RB1, JAG1, THBD, IL2RB, MUC1,

**Table 6-18** | Upregulated proteins in the proteome of CD138+ plasmablast CD93+ plasmablast or both

Gene Symbol (Total 2679)
KIF23, HMMR, LSS, DIAPH3, CYP51, KIF4, GPT2, ZWILCH, LIG1, CKAP2, SHCBP1, TIMELESS, BCAT1, FANCD2, CIP2A, WDHD1, KIF15, HMGCS1, PSAT1, LARS, MKI67, HELLS, CHAF1B, MTBP, CDC45, SLC7A3, PLXND1, SMC4, UHRF1, SLC3A2, ALDH18A1, PBK, ALPL, CDK1, BHLHA15, NCAPG, SLC7A1, KNTC1, ESCO2, SMC2, KNL1, KIF2C, NCAPD2, ERCC6L, TOP2A, DERL3, NCAPH, GTSE1, KIFC1, TFRC, SLC38A2, BLM, MTHFD2, MCM4, RRM2, ACACA, ORC1, CHEK1, PRC1, MANF, NUSAP1, MCM5, BARD1, DHCR24, SEC24D, THADA, SLC7A5, EIF2A, HK2, GARS, KIF22, RBL1, YARS, CTH, MCM2, TUBB6, ATF6, IMPDH2, INCENP, HEATR3, POLE, FAR1, SHMT2, PSMB5, ASNS, BRCA1, ACSL3, CAD, SQLE, GLT8D1, AURKA, PLK1, AURKB, CAR5, TTF2, NAA25, EZH2, SLC1A5, TPX2, DDX20, CEP55, MCM7, FAM98A, TYMS, DSN1, SPC24, HMGCR, ACSL4, AFG3L1, BRIP1, ARMCX3, KIF11, TACC3, FANCI, ATAD5, CTPS, KLHDC4, IRF4, MGAT4A, FASN, GEMIN5, POLA1, KIF20A, CLIC4, LARP1, STIL, AEN, KPNA2, IMPDH1, RAD51, MTRR, RRP1B, DLGAP5, SRPK1, EIF2B1, IGHV6-3, GINS1, ANKRD52, HBS1L, WDR62, ABCF2, HIVEP3, CD36, CMSS1, ARL6, MVK, SHMT1, GM48551, KTN1, HDLBP, TNFRSF13B, EIF2B3, HAT1, EIF4G1, IFIH1, PDCD11, OS9, SRPR, PITRM1, RRM1, INF2, KNOP1, YBX3, SND1, CHD7, IARS, RIF1, KIF20B, ALCAM, OASL1, TOPBP1, ZFP281, SGO1, NCAPG2, CCNA2, FLT3, DHDDS, FPGS, NFKBIZ, HERC1, IPO7, CDCA5, CSDE1, TK1, MRPS34, TIMM17A, LIAS, PITPNB, RFC4, TRIP13, BUB1B, NSD2, EPRS, PPSH, PRDM1, ER11, AAC5, TTK, UMPS, KSR1, RSL1D1, ECT2, MAN1A1, RACGAP1, AARS, MCM6, PASK, GART, NCAPH2, MARS, HIRIP3, IQGAP3, CDCA2, SPDL1, EXT2, PRIM1, SIL1, ARID3A, MSTO1, MCM3, TRIM32, RDH11, MIPEP, NOB1, GCAT, USP36, POLE2, ST6GAL1, UBE2S, HEATR1, IGHV10-1, FASTKD5, EDRF1, NUF2, ELP4, FOCAD, CD93, KIFC5B, MAP1LC3B, NARS, DUT, PHF19, CHTF18, AKAP1, PECK, NIFK, CD69, PAICS, DENND4A, EIF2B4, CD86, MDN1, HJURP, LRRC40, ANKRD17, IGKV13-85, ASCC2, ZW10, GCN1L1, DNMT1, HERPUD1, SPC25, ZWINT, GFMI1, CDCA8, LMAN1, MTHFD1, PSRC1, CLPB, KDELC1, ECM29, IPO5, IDI1, SLC16A1, EDEM3, CPOX, HID1, RILPL2, ERCC6L2, MMS22L, PRMT5, MPP7, LCMT2, IPO11, NOA1, FANCA, CLASP1, DCTD, CRELD2, ELL2, ABCB1B, BYSL, REXO2, NOC2L, CLUH, GFPT1, NLE1, RIOK1, PRIM2, URGCP, FADS2, SEC24A, UCK2, PIP5K1A, IFT80, FDP5, RPF2, FAM111A, UTP14A, ZFP330, TARS, CCNB2, RFC2, PIK3R6, HELZ, LDLR, CDC27, MCM10, MYBBP1A, FTSJ3, LSG1, ANKRD46, IGHM, AHS2, ABCC4, SLC33A1, TMEM126A, ABCE1, EIF3D, TRAP1, PHGDH, PHF10, EIF4A1, CCNB1, TTC37, PFKP, PMVK, GM17296, XPO5, RRBP1, ELP3, SRM, CLPTM1, NSA2, IGHV1-22, RPS3A1, ANKRD16, TICRR, SUV39H1, PDE3B, FEM1B, NDC80, UBE2J1, LARP4B, TRAF3, UPF1, ERO1L, RPAP3, MTHFR, CENPK, PRMT7, GNL3, NOP2, GTF2E2, GEMIN2, ARFGAP3, IMP3, TTC27, TRIP4, BDH1, BSG, SLAMF7, DNAJC15, FADS1, NCAPD3, SLC4A8, PSMC3IP, KIF18B, EEF2KMT, ZC3H7A, XPOT, UTP20, ST7, OAT, PDCD2L, DNA2, PIGG, UTP6, IGHV1-52, AMPD2, LAP3, CBX4, SAC3D1, TMEM214, GRPEL2, UEVLD, BIRC6, BIRC5, PA2G4, NAT10, HSD17B7, PDIA6, EAF2, DNAJC21, POLA2, MPHOSPH10, IFRD2, GTF2E1, RPL23, CKAP2L, KBTBD8, HSP90B1, DALRD3, BOLA2, ABCF1, DUS1L, TMLHE, RPL4, RPS7, WDR3, MCM9, SEC14L1, RPLP0, MPP6, CKAP4, WDR36, NKRF, LRRC59, POLR1A, NAA15, ATP13A3, DHFR, FNDC3B, PWP2, DICER1, MRPS31, GM29394, MAOA, FKBP11, DAP3, PDEADIP, RIOK2, MVD, RPS17, TM9SF4, CCDC47, DZIP3, IGHV14-2, OGT, NOL10, AMFR, USP33, IGKV9-120, EIF3A, IGHV1-15, RECQL4, RPS12, AEBP2, RPS5, IGHV14-1, GEMIN4, MRPS9, EIF2B5, SPAG5, IPO4, CLSPN, METAP1, OTUD6B, TOP1, ASCC3, RBMS2, KDM4C, NOL6, PI4K2B, GUSB, DHRS13, GLIPR2, IGKV6-13, DHODH, PPAT, CEBPZ, MRPL13, LARP4, SYVN1, NEK6, FKBP4, WDR6, MEF2B, LYAR, MRPS14, MINPP1, POP1, MSH6, MRPL14, ORC2, FOXK2, RBM47, HYOU1, PEX13, DDX3Y, NOMO1, DHX33, IGHV1-5, MRPL37, DHRSX, SLC12A2, SSR4, CCT5, PREP, RFC5, FCF1, PCK2, DUSP12, DDX3X, MRPS23, COLGALT1, SLC39A10, TIGAR, HSPA13, BMS1, DSCC1, SCPEP1, TBCE, HELQ, CDH17, TONSL, IGHV1-53, NMT2, NPC2, SEL1L, NAA20, DGKD, GOLIM4, MRPS7, SEC11C, RWDD4, TXNDC15, ZRANB3, EVI2A, ATP2A2, MLKL, COG4, SLC43A1, SRGAP2, DDX27, IGHV10-3, UBE2C, ZFP598, RPS2, ANKRD28, EEF2, IKBKAP, NEK4, NET1, UTP4, PMS1, MRPL47, NUP88, R3HDM4, NCLN, EIF2S3X, AIMP1, NSUN2, TBL3, TIPIN, ETNK1, NBR1, MTRF2, IGHV1-80, RIOK3, DCAF13, DONSON, PELO, NOP14, WDR74, USP45, IMPACT, NEL3, GSTCD, SLC1A4, P3H1, FAM83D, IGHV6-6, LRRC8D, SRP72, DNAJC3, POLRMT, MCRIP2, NUP98, TXNDC5, SGO2A, ZNHIT3, RABGAP1L, CHAF1A, ANAPC7, MRPS22, NDC1, HUWE1, IGKV3-2, PRDX4, B4GALT1, CDCA7, FAM96A, DHX30, ERO1LB, HNF1B, NFXL1, ZFP280B, UBR2, CDK5RAP1, IGHV1-76, EIF3G, FASTKD1, RAD54L, RPL7, DIEXF, NUP107, KIF14, HSPA14, IGHV5-16, BAG6, ALKBH8, NAA16, TXNDC11, PSMB7, ASB6, MGAT2, SLC25A16, EIF2B2, PIH1D1, ARHGAP11A, MRPL2, FIGNL1, IGKV8-27, MLH1, MRPL28, MRPS6, GM20425, RAD51B, GLIPR1, SKA3, PABPC1, CENPE, MRPL17, YME1L1, BOP1, ASPM, YARS2, RRP7A, GSTT2, RTEL1, GSPT1, TRMT2B, UTP11, RB1, PNPLA6, GINS4, E2F8, SREBF2, UBR5, HSPH1, MRPS18B, SASS6, MRPL3, NVL, H2-T23, EIF4G2, ERP44, COB1, TRAF6, PCX, EIF2S2, IGHV1-69, SLC39A14, MINDY3, LGALS9, PSMC2, NUFIP2, MTDH, SSR1, MYO19, EDEM1, MORF4L2, KDM4A, SDHAF2, PPFIBP1, TBL2, G3BP1, WDR43, SRP19, ATP8B2, COPA, TUBE1, USP16, TECPR2, HELZ2, ORF11, MYBL2, MIS18BP1, ATG2B, ACAD11, WDR46, GTPBP10, ASF1B, IMP4, SARS, PCNA, MDM4, RAB39B, CDCA3, NBAS, DNAJB11, EIF3B, CIT, IGHV1-55, NSDHL, SLC25A33, ORC3, PSMD1, MAK1B, PYCR2, GALNT7, POLR1B, PDSS2, ERAP1, USP10, TELO2, PTTG1, PRR11, PGPEP1, CORO2A, DNAJB12, AMIGO2, XPO4, MRPS27, PVR, DTL, TEX30, LRWD1, TMED1, CAR13, SCFD2, PIEZO1, DNAJC7, KDM5C, DHX37, GRSF1, URB1, CERS5, RAD51C, PISD, FAM162A, CDC20, EPM2AIP1, TPP2, BZW2, KIF18A, SLC29A1, UBE2O, CCT3, H2-K1, UTP18, FCRL5, BBC3, PIN1, LMO7, ANKMY2, DAPK2, EIF3J1, NOL11, IGHV1-4, SAPCD2, PSMC6, PDXDC1, HSPIN, PREPL, PSMB6, PLCD3, ALDH7A1, CUL2, ERGIC1, NUP214, TFDPI, RPL18A, PSM7, PLXNA1, PUM1, DROSHA, PER1, PSMD12, NUGGC, TRMT2A, RPS4X, ATG4A, DCUN1D5, EIF3C, CCT4, POLR1D, SIK3, TTI1, SEC23IP, GMDS, MAP3K20, AK6, NSUN4, STEAP3, NCDN, CASP8AP2, SLC35E1, RABL3, ESF1, INTS6L, ISYNA1, MARF1, EIF2AK2, CDK6, RACK1, IGHV1-74, MRPL9, NUFIP1, SLC35F2, DDX21, COX10, LACTB2, CPEB3, NUDCD1, WDR90, QPCTL, IGHV1-26, TRDMT1, DUS4L, DDX56, MIS12, EEF1A1, ATP6V0A1, EIF5B, PTCD3, PSMD6, BORA, MRPL20, NDC3L, UOCC1, IGFB2, FANCG, FAU, RPS3, LAMC1, RINT1, GNAS, IPO9, SRP68, USP37, PSMG2, RAD51AP1, FASTKD2, SLC25A19, GET4, IGHV5-17, WDR75, COG1, NIP7, MACO1, LNPCK, NMD3, ZFP280C, MRPS35, IGKV4-68, GIGYF2, RCL1, AMMECR1, SLC4A7, SEC61A1, DNAJC2, DDX51, TMCC3, SLC2A6, MTERF3, IGKV5-43, CHEK2, NOTCH1, MZB1, UFL1, EME1, FTSJ1, MGEA5, TBC1D15, GNG12, FZR1, NT5C2, TRAF1, IGHV1-18, DERL1, YTHDC2, CLUAP1, MRM3, USP28, STAMPB, MRPL44, EIF2AK3, TRIP11, CCT2, SDF2L1, XBP1, P4HB, LDB1, UBL4A, CCT8, DDX52, DDX18, IGHV1-42, TDPI, RRP12, SKIV2L, NEMF, SHQ1, IGHV1-84, ASB3, ELP2, MRPL46, UOCC2, IFIT1, ACAT2, RRS1, PINX1, PRRC2B, QARS, TRMT11, MRPL51, EXO1, MBNL2, SLC7A6, PPFIB2, PPA1, MEMO1, KNSTRN, ABT1, ATIC, FDFT1, GNPAT1, TRIM56, 1110065P20RIK, UXS1, SLC7A6OS, PNP2, EIF2S1, 4931406P16RIK, DNMT3B, MRPS11, MASTL, MMGT2, CD44, EPB41L5, NAF1, HSPA5, COG1, SMYD5, ADA, MRPL21, PSMG1, ERCC6, NSL1, HSPA9, PPFIA1, CLPX, NAP1L1, ATXN2L, SCAP, TANGO6, EMC8, MSMO1, FASTKD3, LIG3, JADE3, CNOT6, CCDC124, APEX2, IGKV19-93, FNDC3A, DRG1, ATRN, METTL16, PYCR1, CLN6, GGNBP2, TTC13, EMC1, AAAS, RARS, BCAR3, GLE1, RPL5, IGHV1-62-2, DRG2, ATRIP, TP53, TCF25, GMPS, SCFD1, FABP5, RPS14, KYAT3, YDJC, ALKBH4, SEMA7A, KTI12, GOLM1, IL21R, CHORDC1, JMJD6, ZCCHC7, PAIP2, CARM1, AIMP2, 4932438A13RIK, RAD18, SAMSNI,

**Table 6-18** | Upregulated proteins in the proteome of CD138+ plasmablast CD93+ plasmablast or both

Gene Symbol (Total 2679)
GRWD1, PDF, TUBA1A, MARS2, FAM129A, TMEM33, STARD4, NSUN5, HECTD1, GM9833, EIF3E, XRCC3, PLOD3, EDEM2, PSME4, IGHV1-47, COX11, BAG2, RFC3, RPS9, GPCPD1, TFB1M, SCD2, MIA3, VHL, USP34, ABCF3, BTAFA1, IL2RA, METTL13, HCCS, PUS10, MRPL16, SMARCAD1, IBTK, NOM1, UTP14B, CPPED1, IGKV4-57, BZW1, OSBPL8, RIC1, TRPC4AP, NME2, METTL6, LMAN2, CENPF, HARS, AVEN, RPS20, RETREG1, MRTO4, LRRC57, SMYD3, NAA35, DNAJC1, UBAP2, RPS19, ARCN1, SEC23B, RPL7A, GALE, GINS3, BRCA2, ERN1, IGHMBP2, YBX1, METTL15, MBNL3, RPL6, GM21987, OGFOD1, RPS6KB1, SSBP3, ACOT7, NUDC, NUCB1, GM17018, CINP, ZFP277, ARHGAP19, CRLS1, RPS6, MTFMT, TXNLI, DCAF10, ARHGAP21, GPN3, TWNK, LYRM7, HCFC2, BATF, HPRT, ALG11, COX15, TMEM154, REEP4, RPL10, DMD, NFKB2, CELF1, LTN1, SLC19A1, PDIA4, MRPL42, ABCC1, GLA, PSMD11, IGHV5-12, IRS2, DTYMK, MNS1, ORC5, PDIA3, GPHN, GBF1, MRPL4, ARMC6, MTHFD1L, RPL17, TM7SF2, GXYLT1, CACYBP, IQCB1, RPL21, CASC4, JAK3, PMPCA, ALG2, SDC4, PWP1, GM49333, UBE2K, ZCCHC4, COG5, IGKV17-127, LIG4, RALB, MRPL49, LPIN1, IGLC3, UTP15, RPF1, INSR, KARS, PIMI, AICDA, UBR7, MRPS16, PGS1, IGHV7-3, SURF1, ERLEC1, GOLGB1, SAR1B, CKAP5, LARP2, DDX6, EMG1, NAA50, DHX36, KPNA6, SRP54A, GPATCH4, ZBTB32, ATR, POLD3, IGHV1-81, DDX49, DDX54, TAOK1, R3HDM1, ERAL1, GTF3A, SCD1, COPB2, IGKV2-137, ANAPC1, SUZ12, METTL2, EEF1G, RPL3, BAZ1A, ANAPC5, AATF, DHRS1, ATXN7L3B, FOXRED1, SYNCIP2, CAND54, CRY1, ESPL1, NUP93, HSPD1, MTRM19, GAA, RPL9, NOL8, RPS10, PNPT1, IAP, MAP4K4, QSOX2, URB2, PRAG1, MRPS30, NAA40, DLG1, MRPS25, TSR1, KLHDC2, MCM8, DDX24, P4HA1, TPGS1, SEC63, SPATA5, CASP3, SELENOK, UBAP2L, APOBEC3, MRPL45, ITPK1, FAM221A, TMEM242, GALNT2, JUN, TJP2, RPL38, FOXM1, DNP1, MRPS15, PSMG3, DNAJC10, CCDC167, NMT1, XRCC1, MRPL23, MCTS1, MRPL1, PLPP5, FLVCR1, LAS1L, HIVEP1, ATAD1, SOAT1, PRKCI, SYNCIP, CAND2, SOWAHC, MRPL15, PSMC1, MTRM14, PUS7, TUBA1C, NUP160, STOML2, CCDC88A, TXLNA, MS4A6C, MIA2, PTAR1, MRPL32, TMEM131, ETF1, RPL36, SECISBP2, MRPS2, RFTN1, MECR, DYM, CLIP1, MKRN2, TXNRD1, IGHV9-1, NLN, GNE, MTAP, SRFBP1, COG2, RPA1, TRMT1, VEZT, MRPS24, RBM28, PSMD2, ADAP1, SENP3, FEN1, ABCB6, MRPL35, VCP, ORC6, FANCL, EYA3, WEE1, RNFT1, TUBB4B, AHS1, ATF6B, ZGRF1, FBXO3, CDC34, IGHV2-5, GBOX, PPP21B, ISG20, RABEPK, HEG1, MRPL22, ATP1A1, BCL2A1, CCDC88B, FAM117B, KDM4B, SETD6, DERL2, IDH3G, HMBOX1, NAMPT, GSTO1, EIF3K, RPL10A, HEXIM2, RPAP2, PSMD13, PMM1, UGDH, ENTPD1, TFB2M, ATP13A1, TULP3, CD81, WDR77, RPS13, TBCE, FANCB, STAU1, STRAP, FAM208B, NUP85, GEMIN8, SLC25A25, IGKV4-63, TGS1, TNRC6A, RPL24, TUBA1B, ACBD6, RPN1, RPL13A, GOT1, POLH, MRPL38, LRPPRC, VCPKMT, NQO1, SERBP1, NFX1, ATL2, TEFM, TROVE2, IGHV1-77, ITGB1, SIPA1L1, HERC2, SUMF2, MIB1, FAM129C, MRPL48, ALDH1L2, ZMAT3, TSR3, PIM2, MTF2, CD40, LONP2, POR, PSMD8, RANBP2, WDR4, NDUFAF1, ZBTB10, SLC38A10, PUM3, STHL2, RNF168, UBE2Q2, SLF1, WFS1, IGSF8, EIF5, PIK3C2A, SETD3, LIN7C, DHRS7B, QSER1, RPS11, PSMC3, NAA10, IGHV1-39, EIF3H, NT5DC2, OXA1L, RAB23, GEN1, EIF4ENIF1, RPL7L1, LAMP2, ZEB2, HAUS7, RPN2, GBP3, GOLGA2, ATG13, ANAPC4, TYW1, RPS15A, USP14, ALG5, NXN, MRPL24, MGAT1, RUVBL2, PDCC2, SIRT4, CNOT4, ATAD2, NUP188, ZC3HAV1, TRUB2, ARMC1, BC027231, WWOX, SLC35B2, IGHV13-2, GRPEL1, WDR92, DDX10, PON3, USE1, LATS1, JUNB, GLCE, EIF3L, RRP1, NSD1, SLC35A4, ULK1, SLC9A8, TMEM41B, FMR1, PPAN, RPL30, MRPL39, BET1, GLCCI1, RPL27A, UBXN8, CCDC69, LIN37, MUT, LMO2, TIMMDC1, SESN2, PRMT3, OPTN, GLMN, HSP90AA1, C1GALT1, MRPL30, UBE2T, HOMER1, TLL12, RPL32, FXR2, VPS25, MCMBP, NDUFAF7, ACACB, RRAS2, MILR1, ORC4, FTO, ZC4H2, COX18, CSNK2A1, ZDHHC20, COPE, MAN1A2, GNL2, MYO5A, ALG8, EHD2, ZNHIT2, TACO1, HSD17B12, EIF3I, PEX3, TBRG4, DNAJC11, EMC7, ICAM1, RPAIN, MARK3, SRSF6, IGHV14-4, PSMA2, HERC6, ZFP36L1, ENKD1, LSM14B, IDE, AGPS, AKAP9, SLC25A17, COA3, TCP1, ABCD3, HAUS5, NEU1, PSMG4, RMI2, STEAP4, OSTC, KLHL11, PPID, NOC4L, SLAMF6, FAM207A, HDAC6, RBBP7, ZFC3H1, PTPN21, UTP3, FAN1, APPBP2, PRDX1, DENND5B, TECR, ST8SIA4, ALAS1, PCGF5, CCDC6, STT3A, STX18, TOP3A, IGHV8-8, MRPL10, ODR4, SLC25A13, AKT1S1, NTMT1, MCLN1, SEH1L, HAUS3, RPL19, KIN, RANGAP1, ST8SIA6, TXNDC16, SDHAF4, RBM3, ZDHHC21, CENPH, PPP5C, CCDC97, KCTD10, ITM2B, GPBP1L1, CDC123, RANBP9, UTY, TNFRSF10B, BTF3, SLC25A10, MCPH1, UBN2, RPS16, UBXN4, IGLV1, DOCK9, HIGD1A, IGKV4-53, MAPRE2, TTI2, RGP1, MUTYH, DDX19A, ARMC5, ELOVL5, DAB2IP, CSNK1G1, EIF1AX, QTRT2, CAPN7, SUGT1, MPP1, RABGGTA, CENPN, IGKV1-110, IGHV9-3, STXBPI1, CDK2, CDCA7L, ANKHD1, MSH2, DDX31, RPL35A, SUV39H2, PROSER1, PKM, SEPSECS, NAPS, EIF2S3Y, STK3, APRT, RNF219, UBE3A, GOLPH3, ANKRD39, SLC25A4, ENTPD7, PDP2, MAN1B1, EMC10, CDV3, METTL18, GM38394, CALR, PABPC4, QRSL1, CCNT1, ATP7A, 2310035C23RIK, POLR3B, NUP155, CBX2, UBOX5, MRPS17, POMT1, PIDD1, RPSA, NOP53, IVNS1ABP, GPN1, MTERF4, TRMT1L, ETV6, TMEM9, SPECC1L, RPL31, EDC4, TM9SF1, ACLY, USP4, PHAX, FBXW4, UBE4A, PTC1, DNAJA3, DOT1L, C1GALT1C1, CNOT8, ZFYVE16, CASKIN2, RPL27, FLNB, SDF4, PSMC5, TMED9, NUP35, ST6GALNAC4, ST3GAL6, ECSIT, ATP2B1, ULBP1, ITFG2, ZDHHC5, AGPAT4, NKAP, PES1, RAE1, CHFR, CCT6A, GLS, MAN2A1, NMRAL1, SPCS2, H13, MRPS21, RXRA, PWWP2A, EEF1E1, ENY2, MRPL57, TARSL2, TNPO3, DNAJA1, DARS, OTUD5, KLHL25, RPS18, IGLV2, DUS2, OGFOD3, PSMD4, JCHAIN, CCND2, TCF12, IGKV9-124, YIF1B, PSMD3, 2700097009RIK, CHD1L, SEC22B, INTS11, LGMN, SLC31A1, NBEAL2, DLG3, ZCCHC9, CERS2, RTTN, ELP6, CAPRIN1, NOL9, EIF3M, RPAP1, EIF5A, TLR7, DDX28, SPPL2A, GNB1L, ADAR, TMEM97, DCP1A, TUBB2B, USP47, FUBP3, PTPMT1, UROD, LUZP1, TMEM129, BPGM, NUP205, COX7A2L, TM9SF2, FAF2, ING2, RHEB, IFT172, ZDHHC17, SDHA, MTG2, GEMIN7, SMARCC1, PHIP, BTB, NASP, TOP3B, EXD2, LINS1, SPOP, RBM34, ATAD3A, FKBP2, PIMREG, BEND3, MSANTD4, CDK8, MGME1, INTS5, CEP72, ADAM9, SURF2, GDE1, RAB9A, ORMDL2, AAMP, ADRM1, FAM193A, GNAI3, PSMD14, UBE4B, SAP30BP, IGKV4-80, CDC23, EXOC3, STIPI, ANAPC2, RICTOR, MPC2, CSNK1B, MRPL41, VPS37C, ZC3H8, UFP3B, METTL5, CBWD1, FARS2, GK, RPS28, MTX1, RPS8, ARL5C, DDHD2, ARL8B, DOPEY1, 38412, GM15800, RB1CC1, IGHV1-78, SNUPN, SLC30A6, NUP133, RCBTB2, BST2, RPL12, EXT1, SLC6A4, RRP9, CYP20A1, RCC1L, TDRD3, TMED6, STYX, LY9, MRPS10, CCT7, MRPS5, AP4B1, PHLPP1, IGHV5-6, SLC35A3, ACSL5, NEURL4, IGKV8-28, ZMYM1, ZDHHC13, PGAM1, LRR1, CSRP2, AARS2, NUP50, RNF26, AP4E1, CDC16, CTNNA1, RBFA, RPL35, CCN2L2, PARV6, FNTB, FAM120A, FUT8, PIGK, PUS1, TOE1, RPL8, METTL9, FCGR2B, USP42, STRADA, TIA1, CDK4, MND1, KPNB1, GTPBP1, H6PD, CAPN5, AURKAIP1, TSEN54, IGHV9-4, MANEA, LY75, MRPL27, BRAF, DNAAF5, MTA1, CNTNAP1, PMPCB, GPAT4, ZFP622, CYB5B, ZCRB1, MRPL11, EIF4E2, KCNK6, IGLC2, RNASEH2B, NAA11, IGKV6-29, IGHV1-82, IGKV12-38, ST14, ZNHIT6, KLHL7, CLCN6, PNO1, CEP85, SRP9, COG8, HYPK, CCNH, CNOT9, EB13, NFU1, MOSPD2, G3BP2, SMPD4, BRX1, ZBED3, DGCR8, RPS15, NME6, CDC7, USP15, TSPYL1, ITGAL, RWDD1, IGHV1-54, KCNAB2, RPS19BP1, EMC2, UBE2D3, ANAPC10, TMUB2, SELPLG, YIPF4, MBD3, SH3GL1, FARSA, PCE1B, AASDHPPT, PDK3, MELK, 2510039018RIK, GCLM, LTV1, CD3EAP, CSNK2A2, SIMC1, CNM4, RARS2, SPG20, ARMCX2, SSSCA1, DTX3L, PRKCSH, TPST1, CUL1, RAB3GAP1, ITGB6, DDX41, SRSF3, EEF1B2, KPTN, CANX, NME1, KLHL9, SEC61G, NAPA, BUB1, KPNA3, TSG101, USP54, ARL1, AUH, SKIV2L2, TMEM263, PRPS1, MAD2L1, MRPL53, GPS1, FARSB, RPLP1, NPM3, IMPAD1, NARF, GCC2, HAX1, NAA30, PTPRA, ATXN10, SUMF1, SLC30A7, PDE7A, UCHL5, EPHA2, SNRPB2, HSP90AB1, PEF1, EDC3, POFUT1, HDAC9, SLC35B1, EBNA1BP2, HAUS4, QTRT1, GTF3C5, TIMM50, ARFGEF3, EIF4EBP1, UFSP2, PDAP1, UBF1, DIS3, GRAMD3, TOMM40, TMEM165, IPP, NUP54, RAB3GAP2, ZFP825, NPAT, MT-ND4, KATNA1, PLXNC1, UAPI, NGDN, ABCC5, CENPM, SELENOS, ZBTB11, USO1, ACAP3, EHMT1, ZFAND1, ZFYVE21, PSMC4, PARP4, SDAD1, LZTR1, CHD1, GGTA1,

**Table 6-18** | Upregulated proteins in the proteome of CD138+ plasmablast CD93+ plasmablast or both

Gene Symbol (Total 2679)
TIMM10, GORASP1, TUBB2A, SLC25A32, EIF1AD, MRE11A, MAP3K5, CNBP, SRRD, PFKM, RPS25, POMP, PSME3, LSM12, PALB2, DNAJC25, MRPL52, TJAP1, MAP1A, ISG15, SLC04A1, RNF126, MED8, GRM1, HAUS2, PAM16, RPE, PTBP2, GFM2, OIP5, HSPBP1, GADD45GIP1, GNPTAB, RPL14, DPH1, COG7, TNPO1, RAB6A, FKBP3, MRPL43, PEX12, NUDT1, NUDCD2, UBR4, ENDOU, RASGRP3, OSGEPL1, EHHADH, CCDC25, IP6K1, E2F4, NXPE3, ERGIC2, ATPAF2, PEX10, CLASRP, ENTPD6, SLC39A6, MRPL58, SMYD2, MED24, ZBED4, NDUFB9, STX12, WARS, JOSD2, GAPVD1, ZFPL1, SLFN9, IGHV7-4, NDUFB10, CNIH4, RCBTB1, RRAD, NUDT9, PIGT, EXOSC10, POT1A, PLEKHF1, HSD17B4, GRN, RBM19, GCSH, PIP5K1C, SACS, FXR1, HMG20B, ARID5A, CHML, FAM91A1, PTER, DDX11, NLRC5, SNAP47, CCNL1, NUP62, HGH1, NUPL1, SELENOI, GBA, AMMECRIL, GMPPA, RABAC1, ELL, COQ8B, BID, TXNDC9, SCO2, TIMM44, MUL1, CYBB, PPP2R2D, WDR48, YTHDF1, MRPS33, SNX10, ECD, DDX39, NOTCH2, METTL22, CLCN3, CCDC18, SAR1A, DARS2, KRR1, FAH, SPCS1, TSEN15, XRN1, MAPKAP1, MAN2B2, POLR3A, CCDC84, APOO, VRK2, CCP110, POLQ, STUB1, UBE3C, PPIB, TLR12, TSTD2, NDUFAF2, HPDL, CASP4, ATG9A, AKIRIN2, TYSND1, TBC1D10A, IDH3B, LAMC2, HIST2H3C2, SMCR8, KLHL6, IPMK, NPHP4, C1QBP, SUCO, NUDT4, RNASEH1, RPA2, TMED10, ALKBH1, SPTLC2, BCKDK, IGHV1-64, ARF4, SMYD4, RAP1A, HM13, BLMH, TRIM44, GSTT1, MTG1, TTC4, SSR3, ELP5, SNX14, COG3, METTL4, PBX1, NEDD4, GCFC2, GLRX3, MRPS18A, MTOR, PRKCZ, SIK1, ADAM19, SLC38A9, TMC01, DUSP19, ARFGAP2, DCAF1, UBE2A, CEP290, DGKE, CNPY2, ARHGAP35, IGKC, RUVBL1, COPS5, AP4M1, TXNL4B, GTF2I, TENT5C, RSL24D1, LCMT1, MMS19, TBC1D10B, PRKCA, NOSIP, RPS27, USP30, GAS2L3, DPH6, TXN1, RELB, CDAN1, R3HDM2, NXT1, DDB1, SPN, PRRC1, RASA1, PCMI, MRPS12, PGK1, TMED7, CNOT1, HRAS, NPHP1, UHRF1BP1, LPCAT4, ATP6V1E1, AMBRA1, NT5DC3, CKS1B, SUPT16, NDUFB6, BCCIP, RRN3, CD19, MRPL18, ZC3H15, HSPA4, UBE2Z, PCLAF, CUL4B, HAUS6, UBE2E1, ALDOA, TTC1, GTPBP4, BCLAF1, OLA1, TPRN, HNRNPAB, ATP1B3, CIAO1, TLK2, APOBEC1, PLSCR1, PPP6R3, CLDND1, IPO8, PTPN9, CAMSAP1, POP4, NACA, PELP1, POLR1E, POLE4, PRMT1, RPL23A, DCAKD, ROMO1, YRDC, MYDGF, KPNA4, PARL, ADSL, PEX16, YOD1, PBDC1, FBXL15, EMD, RPS21, CCDC86, PLXNB2, PFDN1, RMND5A, ZNFX1, WDR12, GABARAPL2, CSE1L, DYNCL1L2, PEX14, EIF4E, HMGN3, GLRX2, PSMB8, GNL1, PAPP5, 2810428115RIK, KLHDC3, AGGF1, ZFAND6, MAT2A, PAK1IP1, INTS7, GYPC, FKBP1A, DENR, PGAM5, KIFAP3, TOX4, RPS24, COPZ1, DTWD2, RFC1, EMB, SLC25A28, KIF1BP, TTC9C, IL2RG, SLC6A9, DHX29, ADSS, TMEM39B, ZC3H12D, NUP37, DCTPP1, ARPP19, CBX5, HNRNPA1, CNOT10, ARL6IP1, RAD50, CFP, RPL15, GBP2, MRPL50, THUMP3, TARBP2, RABGGTB, SWAP70, TPT1, LACCI, ZPR1, SLIRP, PATL1, ATXN2, NUS1, MAPK8, FAM78A, MCL1, CDK12, PPP2R5D, PARP9, TM9SF3, DCAF4, TRAF4, TMEM161A, MRPL33, USP1, GEMIN6, RPS27L, GSTT3, URII, NPLOC4, MED28, CTNND1, RRP8, POLR3E, NOP16, ADNP, TRMT61A, SLC20A2, THG1L, STMN1, TEX10, MED30, MAFG, DHX58, TCF3, DTD2, ARID3B, PCGF6, GANAB, HMBS, TRMT10A, TAF15, DPH2, UXT, ZFP142, PEX19, MTERF1B, TRIM25, CLN3, SLC43A3, DNMT3A, H2-Q8, TNPO2, ELAC2, CYP4A32, DNMI1, SS18L2, SENP6, GTF2H2, ABHD13, ZCCHC3, FOXJ3, POM121, PRRC2C, GTF2H1, IMPA2, RTN3, AHCTF1, EXOC4, RASA4, DBI, DDT, PUSL1, DNAJA2, RC3H1, WNK1, EEF1AKM1, PARI14, PHLDB1, MRPL34, LDHA, ELOA, ECNE1, IGKV4-57-1, RPL2, P2PCA, LARP7, PPP2R1A, EFL1, GRCC10, FAF1, TMED5, TMX2, POLR2K, PYCRL, ATG7, TRAK1, ZFP644, PI4KB, RBM45, ZFP511, NBN, NRD1, OTUB2, PRRC2A, USP46, CLK3, TRMU, HINT1, SKA1, RWDD2B, GATM, 38777, RTCB, DAD1, PNKD, GOPC, ZFP53, CIAPIN1, POLR3D, RCOR3, ATM, ZFP444, CD48, MRS2, ADPRH, MRPL55, ZKSCAN1, SREK1, DDOST, DIS3L2, TOPORS, ZMIZ2, RPL34, FAS, TRUB1, TAF1C, CCNC, PRPF38A, NAA38, IGHV5-9-1, GBP5, DHCR7, CLASP2, ARFIP2, USP20, XPO6, PPP1R16B, LMCD1, HSBP1, REPIN1, ASCC1, PXN, EED, TSEN2, ICE2, PAN3, REXO4, EIF3F, EIF1, SURF4, H2-D1, RNPEP, NUB1, POLR3G, COIL, SLC38A1, RALBP1, GOT2, BICD2, DHPS, UBE2E3, ICE1, RSPRY1, USP38, PPP4C, SLC25A15, OVCA2, DESI2, PHF6, TWISTNB, LIMS1, RPUSD4, CSNK1G3, RAN, FNTA, EIF4B, ZFP64, METTL23, OSBP, LLPH, PPP4R3B, ARID5B, CD83, MED9, MED11, CHAC2, RUNX3, IGF2BP3, INTS3, RNF10, OXCT1, PPP1CA, TET2, SLC25A3, COPS3, SDHAF3, CLK2, GATB, HDAC4, XPO1, NGRN, GAPDH, NUP43, CRIL, EEF1D, TSFM, SLC30A9, NDRG3, TMEM199, KLHL20, WRNIP1, EXOSC1, PFDN4, RMDN3, GMPR2, AIM2, MED1, USP9X, D16ERTD472E, DYNLT1A, ATP5F1, CHMP7, RPUSD3, KAT14, CSNK2B, RRP15, CLNS1A, NUDT5, CNOT2, NCOA7, GM10881, NABP1, PFDN2, XPNPEP1, FAM193B, BHLHE41, YKT6, SPAST, ZFP160, MAG3, CALU, RPL13, UBE2L3, TMEM209, ARNT, PSMB2, NUDT7, PTGES3, YEATS2, SERHL, ADI1, DDX1, SASH3, MRPL40, RNASEH2C, NUP153, ITGAV, MTHFSD, TSSC4, ZMYM4, KCMF1, TANC1, MT-ATP6

**Table 6-19** | Downregulated proteins in the proteome of CD138+ PB, CD93+ PB or both

Gene Symbol (Total 2340)
ARHGEF18, ALB, SERPINB1A, ZFP318, CPT1A, MECPP2, CROCC, ATP8A1, FGB, SFXN3, ANXA1, SELENBP1, ELMSAN1, SIPA1, GANC, ALDH2, CDKN1B, ZYX, RIN3, EPHX1, NUDT16, A430078G23RIK, AIF1, PHF1, ELMO2, FGG, GSN, LMNB2, FCRLA, H2-OB, IGH, BLVRB, SIGIRR, GPIBB, DOPEY2, PDCD4, KCTD12, ITPKB, IFI209, FGA, LSP1, ALOX12, SRPK3, ARRB1, H2AFY, PGAP1, SMARCA2, FCMR, TTC38, EVL, LPCAT2, MYL4, KBTBD11, CFAP43, TRIM65, SORL1, HS1BP3, GBP9, RUFY2, PLD4, ZC3H12A, SENP7, MTM1, CCDC71, PLCB2, CAPG, CASP6, TGM2, SUN2, AGO1, PITPNM1, VWF, SSH1, KMO, RPS6KA4, HIDE1, PARP3, B4GALNT1, CMPK2, HAAO, NFKBIE, TRIM21, SFXN2, CEP120, ADD1, DFFB, ACP6, CHD3, NDRG1, CRYL1, LRRK2, UCKL1, PNKP, ITGA2B, PCIF1, ZFP512B, H2-EB2, PLEC, AS3MT, GSTM1, VCL, HIST1H1T, BIN1, ARHGAP4, EML3, SETD1B, PITPNM2, FBXO22, FSCN1, MYL9, LIPE, FHOD1, EXOC6B, RBL2, GP9, TMEM71, GPD2, P2RX4, CEP95, ACTN1, CGGBP1, COQ8A, TRAF5, HMGB1, CD200, UBA7, ALDH6A1, AP3M2, PLBD1, TUBB1, PLGRKT, AP1S2, C4B, PAXX, ARID1B, CIRBP, SPTBN1, ENGASE, PTPRJ, RPS6KA5, KLC4, ARHGAP45, CYP4F18, CDT1, GRIPAP1, ITGB7, VRK3, ZHX2, DFFA, VASP, GPD1L, CAST, UBLCP1, TLR9, PIK3R1, ITPR2, RCS1, PDPR, CEP135, TMX4, WDR44, SPTAN1, SH2D3C, IKKBK, TRP53I1, KAT6A, TNFSF13B, PPP1R12A, DGUOK, PRKAR1A, HMGAI, NFKBIB, SPICE1, WDR1, FYB, NAAA, CALHM6, THBS1, CTSH, ARHGEF1, NPEPL1, ANXA5, ABR, NUMA1, SH3BP1, ACAA2, PHF2, EPC1, SH3KBP1, BANK1, TSC22D4, ANKRD44, NAGK, HSPA1B, ASAP1, UCK1, CALHM2, LTA4H, NIN, CST3, STOM, ATP2A3, FGR, ADD3, NUDT14, LMNA, PARVB, NQO2, RGS14, STRN, LCPI, EIF4A2, NFATC1, GBE1, RIPOR2, PPBP, PRKAR2B, ARHGAP27, NR2C2, ZADH2, SERPINB9, TREML1, MAPK14, PACSIN1, GIMAP9, PGM2L1, HDAC10, HCK, CAPN1, PACS1, SELENOH, ZFP639, SP4, ANXA2, CBRI, RFX5, BLK, ARID4A, PRKD2, ACSS1, PRR12, CAPNS1, DNTTIP1, DLGAP4, AHDC1, RFX1, SMARCC2, HDGF, FILIPIL, SLC25A11, BIN2, FRY, IVD, TBC1D5, MLYCD, SYNE3, MDM1, MORC3, ARMC10, RFXAP, ABCG3, AMPD3, ING1, GAS7, RASSF2, DCXR, KHNYN, LRMP, NEK9, ALDH1B1, MAP2K6, AK3, PLEKHA1, PCTP, PPP1R9B, CARD6, GLYCTK, BRCC3, TRAF3IP3, DHDH, CREBRF, DAPK3, TLR3, CEP128, ZFP524, RFL1, HIVEP2, ZFP35, TESPA1, IDNK, IRAK1, LCK, ABLIM1, TFAM, PLEKHO1, VIM, LGALS3, NR3C1, SIGLECG, SELENOO, IRAK4, SH3GLB2, TBC1D9B, PPTC7, SCML4, HMOX1, THUMP2, RAB32, TAF4, NFATC2, ALDH4A1, CDK13, METTL14, UVSSA, TDRD7, BASP1, BEND5, PDE5A, SMARCD1, MBD2, MSRA, MCCC1, PAFAH1B3, SIN3A, LENG9,

**Table 6-19** | Downregulated proteins in the proteome of CD138+ PB, CD93+ PB or both (cont.)

Gene Symbol (Total 2340)
MNDAL, CNP, CNN2, TMEM63A, DIABLO, CIC, ADSSL1, RHOG, TMCC1, ANPEP, IBA57, ACINI, DENND1C, BICRAL, CAPI, AAMDC, HVCN1, GLUD1, ASPH, FBXO7, RELA, PHC1, SIRT7, NRF1, AKAP13, TBC1D14, LEMD2, DEK, STK17B, CHTF8, STK10, MAP4K2, MCCC2, MYO18A, CBX1, RAC2, DNAJC9, TCOF1, BRWD1, FIG4, ECI1, ACOT2, OTUD7B, API5, NLRX1, CBFB, FAM50A, ZGPAT, SDCCAG3, EZH1, GRAMD4, STAT6, ZMYND11, PF4, METTL3, GIMAP8, PKN1, TWF2, VWA5A, DCLK2, SYNJ1, UIMC1, GGCT, MLLT6, CORO1B, ZFP512, PGM1, FAM98C, RASGRP2, MFN2, IL16, GP5, STARD10, PTPN6, LACTB, OXSM, DDX58, FLNA, PHF23, EAR1, CD2AP, ITGB3, TLE3, DBNL, TNKS1BP1, EPSTI1, CCDC90B, HIFX, HBB-BT, LMNB1, CPM, ATF7IP, TAF7, SP2, PEAK1, MPO, CUTC, PRDX5, IFI203, ZFP518A, HCFC1, GPX1, RPRD1B, SAMHD1, MTSS1, MYOF, GIGYF1, TERF2IP, ATAD2B, STK11IP, KANSL1, PPOX, MAPK12, ADPRHL2, SORD, TRP53BP1, RIDA, GDPGP1, MDP1, ARHGEF3, ZMAT1, HSPA2, MARCKS, HIST1H1A, SETD1A, POLR2G, PPP1R21, ANXA6, HHEX, AFF3, ARHGAP25, ENTDP5, ETS1, RABEP2, RAP1GAP2, CCDC93, CDKN2AIP, MYO1F, ANXA11, VPS33A, HSPA4L, INPP5F, TRIM7, ARHGAP26, SLC9A7, BC017158, RAB11B, CRTCL1, TBCK, GMIP, UGP2, TMOD3, COBLL1, GMFG, ENO3, APOBR, EML2, GALK2, PARP1, IKZF3, CIZ1, XPC, CEP78, SERPINE2, PDXK, PANK2, FCER2A, BCL9, ARHGAP6, MEF2D, CD2, KDM7A, PPM1M, TMUB1, ARHGEF11, WDR91, LYZ2, BCR, GPAT3, IGHA, PDP1, CAR1, NFATC3, NFATC2IP, PFKFB4, GGA2, CPT2, A430005L14RIK, CLU, EFHD2, EZR, BNIP3L, PYGL, MDH1, SYNE2, SNX8, IQGAP2, CRAT, VPS33B, KYAT1, SPIN1, VPS8, VTI1B, ABHD14B, AGL, GNS, XPNPEP3, PML, CHKB, ABHD16A, DOK3, KDM1B, DNMT2, TAF6, INPP4A, LMBRD1, FLYWCH1, WDR81, NSF, CAMK1D, BRD3, FYCO1, MPRIP, ALYREF2, TNIK, CD22, TEC, SNX30, TOR4A, DENND2D, HBA-A1, UBAC2, ANP32A, CORO7, CAMP, GM340, APOE, PSMB10, SPAG9, ABI3, GRAP2, SHTN1, IFIT3, 9030617003RIK, FLI1, 2310033P09RIK, ZFP148, SERPINB6B, DNMBP, PTRGL, MDH1, MICAL1, ECH1, D6WSU163E, FAM32A, KYNU, MPEG1, SLC25A20, GGPS1, BACH1, CDS2, DPP7, ZBTB3, ACCS, PGGHG, CD84, SMARCE1, PIK3C2B, RUFY1, ZMIZ1, DPF2, IPCEF1, SLC15A4, GMEB1, EEA1, STK19, TAX1BP1, RUBCN, TIPRL, ZC3H6, AKR1B3, CCAR2, SKAP2, GSDMD, NAXE, IST1, BRMS1, RGL2, VAV2, EP400, PLAUR, RASA3, RNF169, FBXO38, DDX59, SFN, CPSF7, TOR1AIP1, ZFP263, FMO5, RDH12, NLRC4, HEXA, HSD17B10, TPD52, PPM1F, BAHD1, KLHL14, RASSF5, STS, MAST3, CNTRL, SCAF1, HSD1L, FAM49B, ARL6IP4, ACSL1, RABEP1, COMMD7, CNTRB, IQGAP1, APIG2, CARMIL2, ANKRD13A, MYH9, COMMD5, METTL7A1, TBC1D20, ELL3, PAG1, EPS15L1, MUM1, SGSH, TKT, AKAP2, SPTB, PAX5, RMND1, CEP170, NAT2, RNASEL, SERPINB6A, DHX34, TALDO1, LNPEP, BCL7A, MEF2C, DCPS, TGFB1, NFKBIA, SLC4A1, PLS1, TMPPE, FGD2, PSD4, ODF2, SCIMP, TCP11L2, ABRAXAS1, AHNAK, ECHS1, CDK5, RCOR1, KMT5C, IFI47, TIMP3, ZFP668, MAML1, AKAP5, ACAP2, KMT5B, NUDT3, ANKRD12, SP100, SNX29, CPNE1, GPALPP1, CEP250, EPX, PDK1, PMM2, HIST1H2BK, SRF, INPP5D, FCGRT, AKAP12, ETFDH, IRF8, GOLGA7, ACSF2, SCLY, GDI1, SQOR, SAFB2, OTUB1, SCRNB, BRD9, TRIM36, ERBIN, POLD4, ELF2, IRF9, NIT1, CDYL2, MTIF2, RAB4B, MAFK, TCEA1, ANXA4, BRD8, PBX2, TECPR1, EGR3, CBLB, PTK2B, IL4RA, HBPI, PAXBP1, SLC25A12, TNIP2, CD97, STK26, SMC5, TNFAIP8L2, RNMT, NCOR2, PIK3C3, NDUFV3, ALDH3B1, ACP5, AKAP8, PPM1A, DRAP1, HMGB2, MFF, BPHL, EPN1, ENDOG, MSN, OSGEP, DIDO1, SLC40A1, SMUG1, PPT1, RGS19, TRIM59, FCER1G, CXXC5, MICU2, SPG11, SP110, ZC3H4, H2-OA, TRAPPC5, HIST1H1B, PLAC8, BC026585, TMEM201, POT1B, ARID1A, TBXAS1, WBP11, ZBTB7A, CRK, SESN3, SSH3, SLAMF1, ARMT1, PRR14, FBXL8, PIK3API, UAP1L1, NMNAT1, HPCAL1, MRC1, UPF3A, RBM5, LRCH1, PIBF1, EIF4EBP2, HIP1, SCAF8, F13A1, SNX4, FKBP15, ASPSCR1, TAPT1, CNST, TM6SF1, STX7, PLEKHG2, SLC28A2, SAMD1, ZFP638, LCP2, TIFA, RREB1, FAM102A, CEP295, 2610507B11RIK, WAS, STRIP1, PFKFB2, OSTF1, ELK3, SNX18, FDXR, SLFN5, CR2, RENBP, SH3BP2, RGS18, DDX23, SNX2, CAB39L, MINDY2, ABCB10, CHIL3, CNRIP1, BOD1L, OARD1, CLCA3A1, H2-KE6, AKAP8L, ZFP422, NGP, TALPID3, PPP3CB, COPS7B, HK1, HECA, RNF6, REL, PRKX, YWHAZ, SHPK, TBC1D10C, MOB2, APEX1, SURF6, FAM208A, APIM1, ALDH16A1, PPCDC, EMSY, NEMPI, LRBA, SERPINA3K, RAB27B, SPIB, QDPR, AGPAT2, JAK1, PPP2R5A, KDM2A, STX6, HSF1, KDM6B, PIP4K2A, TARDBP, TMEM2, UBXN6, GM4737, REST, ERLIN2, QK, MGMT, N4BP3, 1600014C10RIK, GPM3, SAP130, PUF60, RFLNB, FAM213A, NFRKB, CBR3, NEDD4L, RSF1, ESCO1, TPCN1, WDR82, AMDHD2, PKNOX1, SMC3, PPP1R11, SDCCAG8, BDP1, PRCP, LASP1, PCCA, DDAH2, SNAPIN, RDH13, NT5C, HPS1, ABCD1, DENND6A, SFSWAP, DCTN1, MAP3K3, ITIH4, IKBKG, GPIBA, SRI, CUX1, RESF1, CD55, ARHGAP17, TLN1, ALDH3A2, TUT4, PGLS, GSTK1, SLC25A46, MON1A, WDR70, RIPK2, MTA3, AP1S3, RAB31L1, ZFP407, ZMYND8, MAST2, APEH, LBH, ACOX3, TRAPPC9, MAPK3, TAF3, SYNE1, ANXA3, EFR3A, TBC1D8, LANCL1, SDE2, RPS6KA3, DTD1, DGKZ, UNC119, NFIC, YWHAQ, ATXN3, DUSP28, ITGAD, PRDM2, ITPR3, APPL2, LRP1, WDR45B, FOXN3, ANKRD11, ZFP608, SIRPA, LTBP1, TRIM14, CAPZA2, PLA2G4A, ACAP1, RFTN2, CNND2, ZFP385A, CALD1, PDIA5, CREBBP, SART3, PCCB, CEP68, CREB1, IFNGR1, SLC23A2, DHX15, RAB28, TFE8, HAGH, PYGM, EHD3, ZFP592, SELENOT, ARHGAP18, SERPINA1B, TNIP1, HNRNPK, PLCG1, A1607873, ZC3H3, RARA, GNPDA1, PPP6R2, NCF2, PALM, NCOA6, KAT5, DAXX, GM20498, PTS, UPRT, CDK17, ERMPI, ASRGL1, WIPF1, PHF3, AH11, HERC4, VCAM1, ETV3, ACTB, ASAH1, AAK1, RYBP, HOOK3, FOXO1, CAPZA1, HACE1, LBR, GMEB2, CORO1A, M6PR, ZBP1, INPPL1, UBE2N, LSM6, IRGC1, TTC28, PSME1, SIGLEC1, TUBGCP6, COMMD1, MACROD1, HOMER3, RUNX1, IFIT2, ITGA6, BPNT1, FMNL1, ENSA, OXR1, KIAA0226L, ACYP1, SRSF4, BAG1, HMGB3, BLES03, PRKCE, EMC6, PCYOX1, TGTP1, NFIA, BUD13, USP25, PSCAD, CABIN1, ACADM, STX8, EVI5L, ARHGEF7, IQSEC1, SMARCB1, COX7A2, CTD18, CAVIN2, SMARCD2, HLCS, NCF4, CTSA, CARD9, BRMS1L, LRRFIP1, USP12, TBL1XR1, ZFP831, CHMP1A, CD38, YWHAB, FES, MNT, GM43302, MTFR1L, ROCK1, PDCD6, ITGAX, RALY, TOP2B, CAMKK2, AKAP11, RSU1, RASSF1, POLL, FAM20B, NMI, POLG2, H2-EB1, TRAPPC8, STK24, RNF114, THA1, MAPK8IP3, PTPA, CCNK, TRIM34A, LIPT1, NFS1, ATG5, NRIF1, MSL1, TERF1, IPO13, DENND1B, JAKMIP1, DUSP7, GPX4, PPM1E, NIPBL, PTGS1, TMED8, CASP1, ZFP691, NAPG, PDE2A, H2-DMA, ZFP335, ZFP143, FUS, STIM1, TRIM24, SMG6, CYLD, HDHD5, SUGP2, S1PR4, SETDB2, SLFN1, ARHGAP9, PPA2, MFN1, MAP3K8, FN1, LPP, HPS4, CYP2D22, CD79B, SETD2, GIT2, RIC8A, KMT2E, PPP1R13B, PXX, KDM5A, HP, BRD4, EPS8, VIPAS39, RNF170, DENND1A, TNK2, PPP3CA, BLOC1S1, MCUR1, ARHGAP12, TSPAN13, EC12, MSL3, CD74, CLEC12A, FAM49A, GIMAP1, RING1, ATP9B, RBM14, AGAP2, MYL6, IMPA1, ZFP629, FBXO28, NRBP1, TMEM243, LIMK2, RXRB, VPS13C, PPP1R18, CCDC82, ACS2, POC1B, MAP3K2, DDX17, PACS2, 3110001I22RIK, CTDPI1, LSM3, BCL9L, SMAP2, MFAP1B, NSD3, USP11, PTPN23, BACH2, TST, EPC2, MOB1A, APBB1IP, COX4I1, CCDC71L, GIMAP7, LTF, LCN2, LTB, STK4, DDX60, FAM160A2, TMEM131L, NHEJ1, TK2, NUCB2, AKR1B10, L3MBTL3, ERCC4, ZBTB24, KAT8, RPRD2, JARID2, FUK, TMPO, GGA1, SRSF1, CARD11, WARS2, SF3A3, FOXO3, CXCR5, PHACTR4, D10JHU81E, GIMAP5, MSANTD2, OSBPL5, STAG1, ZFP553, RP9, FMNL2, WASHC5, NFYA, ACAT1, GNPDA2, PTBP3, CEP41, VPS16, FBXL4, CWF19L1, ARHGDI1, PRKACA, PFN1, CISD3, RAB27A, FAM151B, MGTST1, THUMPDI, SSBP1, NADK2, DCAF11, VEZF1, NT5DC1, BECN1, ICOSL, CAMK2D, MTR, PLEKHF2, RAD21, MAPKAPK2, FAM126A, VPS35, CTPS2, TIGD2, SUDS3, PSMB9, NUDT13, PAK1, PSPC1, PIRB, 9930111J21RIK2, DCPIB, TUT7, FECH, CBX8, ZMYM6, CLEC16A, BTLA, BRPF1, SIKE1, PRCC, HNRNPUL2, BCL2L13, SH3GLB1, RALA, FN3KRP, ZBTB2, TOMM34, MOCS1, RAB11FIP1, FBXO4, PANK4, EMILIN1, ACAD10, KMT2D, RBCK1, LIMAI, RAP1GDS1, NCOA3, ACTL6A, KLF3, PADI2, MITD1, ULK3, SBF1, DEF8, LETM1, GRB2, 9030624J02RIK, CBX3, HNRNPUL1, TRIP12, IRF2BPL, ACAD12, 6330416G13RIK, RBM6, HPS3, S100A9, C3, RSRC1, TGM1, GRK6, FBXL17, CRLF3, SF3A1, CSTF3, ZFP740, TRIO, B3GNT8, H2-DMB2, SGPL1, CPQ, PDXP, H1F0, 5031439G07RIK, VAMP8, MTMR3, AHR, SIRT5, GLB1, FCHSD2, SLTM, ZBTB9, GM49405,

**Table 6-19** | Downregulated proteins in the proteome of CD138+ PB, CD93+ PB or both (cont.)

Gene Symbol (Total 2340)
MAP2K1, SEPHS2, RDH14, RPS6KB2, SMC6, SLC25A24, ZFP207, NOL12, SATB1, ELF4, PDCD7, TRAPPC12, DET1, PIP4K2C, METAP2, LRSAM1, SLC27A4, TMX3, CLIC1, EML4, VPS11, HPS6, USF2, CD180, MBP, GATD1, PRG2, DENND5A, SINHCAF, RAD23A, CCSAP, RAB6B, WASHC2, UHRF1BP1L, CRIP2, MAF1, STAG2, BTG1, GFER, ZBTB4, DOCK5, MEN1, DPP9, NUA2, GLRX, PPCS, AP5Z1, DENND4B, COTL1, PRKRIP1, PAPSS1, RASSF3, HK3, THNSL1, SLAIN2, EPB41, TMEM65, PRPF31, FAHD1, UBTF, UBXN7, BLOC1S5, PIK3CA, ARFIP1, RHBDF2, LSM7, RFXANK, 4933427D14RIK, ARPC2, FAM107B, BORCS5, ABCD4, PSME2, GPATCH11, SLC2A3, OSBPL2, NCBP2, ABHD17B, UNC93B1, ARHGAP15, UBE2R2, TAF11, PBXIP1, GTPBP3, NADSYN1, TUBA8, RBBP5, UBE2I, MAVS, PHKG2, CWC25, CASP7, DYNLL2, POGZ, SCIN, CMTR1, VPS4A, B2M, TXNRD2, FCHSD1, RBPMS, AKT3, PLEKHM1, ZCCHC17, POLR2A, MEAF6, ABHD10, AFTPH, NAGA, UVRAG, BC017643, TP53RK, WDR37, LPGAT1, CHMP2A, MSL2, UBL5, CLYBL, CCDC12, GSE1, KCTD14, OGG1, SART1, CELF2, STXB3, CXXC1, SYNRG, MYL12B, BRWD3, SNX6, NME3, SFPQ, MMP9, THY1, SLC25A45, NPEPPS, CCNY, CSTF1, EPB41L3, TRAPPC1, GRAMD1A, 37316, VAV3, FXYD5, SNX3, CAR2, PLAA, KBTBD2, DIS3L, JPT1, ACSF3, BCL6, GCC1, TGOLN2, DYNLT3, 1110004F10RIK, EXOSC9, SUMO1, POLR2C, DLD, NCOA1, ANKRD27, UBE2G1, MAN2C1, ACADVL, ACADS, SLC41A3, SMIM14, SF3B4, VPS26A, CPSF6, MKL2, ZFP292, AP1B1, MRL1, UNC13D, IDH2, CC2D1A, TAF8, TADA3, SMC1A, CEP57, EAR6, PCMT1, CNPY4, IDH1, CDK19, LENG1, RMDN1, HEXB, PIP4K2B, KIF1C, GIMAP4, NIPSNAP2, RAPGEF4, COQ5, IFI208, SCAF11, NFYC, MAP4, DPY30, BAG5, SNAP29, TAF5, EARS2, ZHX1, ECE1, POLR2J, ZFP746, GLYR1, OFD1, BICRA, SH3PXD2A, FADD, FTH1, DAPP1, MYO1C, HADH, TCMO4, MAPKAPK3, NUDT18, SETX, RANBP3, HIST1H1D, HSPA12A, CUTA, LYRM4, WASF2, CCDC9, TRIOBP, ATP6V0D1, SNRK, CSK, ACTR2, BPTF, ARAP1, PCBD2, CARMIL1, ISOC1, SAFB, MCTP2, APOOL, LAGE3, SRSF2, ILF2, ZFP830, CD79A, ILF3, XDH, BMP2K, TERF2, NSFL1C, ATP1B1, COPG2, RAI1, PDLIM5, DOCK11, ATG3, SPTA1, PHRF1, TOMM5, IGSF6, PDE1B, PHF20, SYPL, GABPB1, ARPC5, PPM1G, HEMGN, STX4A, ZNF326, SNRPA, MTMR12, ZFP280D, DDI2, PPIL4, CECR2, UQCRC2, PPM1K, NCK1, PRPH, STK38L, WDR47, LRCH4, RAB21, RALGPS2, HNRNPA3, USB1, YWHAG, SYF2, CAND1, POC5, ZFP865, PHC2, ABHD12, FKBP5, PTPN14, AKT2, SCAMP3, NFKBID, SH3BGR2, LSM2, ANKFFY1, PRDX2, BAK1, INO80, ABCB1A, NCKAP1L, CCAR1, RGS3, CHTOP, CBFA2T3, SETD5, YEATS4, NUCKS1, SERPINH1, KANDL3, PTPN18, MMAA, OSBPL7, HMGN1, RHOF, SNX12, FAM104A, CSAD, RETREG3, SARNP, SYNGR2, 1810043G02RIK, TAF1, KXD1, ZC3H18, DIP2A, KLF2, SRRT, SKI, BLOC1S3, ADO, MOB3A, TICAM1, ADCY7, MICALL1, ELMO1, PRKRA, FAM192A, MPPE1, IGHG2C, MPME1, UTRN, CYP4F14, PIK3CD, CYB5R3, EXOG, CISD2, BCL11A, COX6C, QRICH1, ARHGAP1, ZBTB14, HNRNPL, HP1BP3, CCDC22, GATAD2B, SLC9A9, ATF7, MAPK11, XRCC6, CLIP2, RNF146, TMEM43, BLOC1S2, WASHC1, ZBTB1, PQBP1, TMEM134, BRF1, HNRNPM, RCN1, NUMBL, DHX8, THYN1, CAPZB, STARD3, PPIA, RBM4B, FGD3, FAM213B, TAF9B, SRA1, TREX1, HMGXB3, RTCA, NIPSNAP1, HSDL2, DBN1, HBS1L, COA4, OGFRL1, RABIF, ZFP260, TOMM40L, NDUFS4, SUPT7L, TSNAX, HDAC8, CNN3, KALRN, PHC3, TCERG1, NT5C3, RNF31, PSMA3, MAGOH, DUSP3, PSMD5, LRRC45, STRN4, ZFP574, AQP1, TMEM175, SLC12A6, NONO, OPA3, IFI35, TGFBRAP1, DIAPH1, ARID4B, ZFP346, FIZ1, SLC29A3, EMC4, RBKS, CYC1, CLEC1B, PLP2, RAPIB, ACTR3, EAR2, STAM2, RCC2, FGD6, ANKIB1, ETHE1, NCKIPSD, AKR7A5, PSIP1, C2CD3, PITPNC1, DNAJC28, KDM3B, CD177, TSC22D2, EEFSEC, TRAPPC11, CENPV, TRRAP, PAK2, CDK5RAP2, PPP1R37, TES, APPL1, XIAP, PIN4, MYO6, HGSNAT, SPATA2, CUL4A, POLD1, ADRB2, XRN2, MSH3, GM608, BAP18, HMGXB4, HDGFL2, VARS2, GTF3C4, FRG1, MAP3K11, PLA2G15, SELL, LPXN, SLC25A35, MIOS, ZDHHC18, CFAP97, DNAJC16, SMAD5, CML1, AGO2, TEX9, CNPY3, PIAS2, PIAS1, NUMBL, DHX8, THYN1, CAPZB, STARD3, PPIA, RBM4B, FGD3, FAM213B, ATP5D, ANP32E, MYH10, MYADM, AFF4, WDR13, NUDT21, PDCD10, CEP44, NAIP2, RFK, FAM69A, MTMR6, PHYKPL, GBA2, RFX7, POLR2D, CTBP1, PTPRC, DGAT1, AK6, CDIPT, KPNA1, SMPD2, PLK4, FAM76B, FAHD2, PZP, EPS8L1, MAP2K4, PARN, XRCC5, MCAT, WASHC4, DCAF5, HTATSF1, GTF2A1, ARL6IP6, YTHDC1, PYM1, FLII, MYH14, LAMTOR1, CFAP20, ZMYM5, GATAD2A, SLA, ALOX5AP, XXYL1, TAB2, AIDA, SVIL, IRF2BP2, MLLT3, HSCB, SLC25A44, HIF1A, HIF1AN, CRY2, ESYT1, FBXO6, ABCA2, OSBPL11, ZER1, FBXO46, DNAJC17, MICAL3, KIAA1429, UBE3B, YJU2, TNFRSF13C, UBE2W, TBCA, PHOSPHO2, CYFIP2, HGS, TAF6L, NAXD, MTHFS, LRIF1, UCHL3, SLC12A7, SNX5, PPP6R1, TBC1D8B, FBXL20, LCLAT1, CACTIN, TBC1D17, ZFP41, ETFA, MATK, NEDD9, ZFP787, EXOSC8, POLB, NSMAF, POLR2E, SSFA2, TCIRG1, THOP1, NDUFA4, ZFP84, PRPF19, HEBP1, CDC73, TAPBPL, RAPGEF6, ZBTB18, CCHCR1, CUL3, LLGL1, GDA, LYPLA1, DOK1, TTC33, SYMPK, ZFX, ARL8A, UBN1, CRKL, PHF20L1, MCRIP1, VCIPI1, TUBA4A, RASSF4, VRK1, ZFP593, IGSF5, APMAP, PPIF, MYD88, DERA, HPS5, IGHV1-63, SFXN5, INIP, DMAP1, FLOT1, ZFP276, CYB5A, SNX25, PRKCB, PIP4P1, NDST1, SP1, ZFP953, TNRC6C, ATXN1L, THAP11, OPA1, LZTFL1, AGK, NFYB, 1700037H04RIK, BUB3, PLPBP, PEX5, ITS2, SMTN, MROH1, MFS16, MORF4L1, ALS2, CTDSP1, MIER3, ACO2, HNRNPH2, FYTDD1, AP3S1, 0610037L13RIK, NIT2, LZIC, BCOR, CEACAM1, PHF14, HNRNPF, N4BP1, ANKS1, IFI205, MAP2K2, AGPAT1, MBD4, GPAM, SRR, R3HCC1L, COQ9, SLC27A1, CCDC88C, LYPLAL1, RAB24, SCNMI, CCSER2, ANAPC11, SCAF4, CRNKL1, 4930523C07RIK, KHDC4, TMC8, RAC1, SMU1, EP300, DOCK4, PTPRE, PTGR2, ESYT2, HMCES, MYO9A, TMEM222, NCEH1, PPM1D, RGS10, UQCRC10, PTGES2, TTC14, ERP29, ANTXR2, HDAC5, VAT1, PURA, ACAD8, LGALS8, VPS4B, RCHY1, CKB, XYLT1, SMPD3, UNKL, COQ6, UROS, PIAS1, SPG21, DXO, CD1D1, UTP23, SGK3, DMXL1, ITGA2, CYTIP, HIST1H1C, DPY19L1, TRAPPC13, NDUFS3, PDLIM2, AFG1L, KDM5B, RAB3D, RUFY3, ATP2C1, TRIM8, MICU1, TRIM30D, CCDC15, MCU, PRKACB, MFS14A, GPANK1, STX17, GNAQ, KLF13, RAB1B, HNRNPLL, VWA8, ALG9, PPP3CC, CRT3, FAAH, GMCL1, ACADSB, JAK2, PLBD2, MTRFIL, SLC4A2, HIPK1, PRKCD, TMEM120A, ZFP131, TMEM260, SAYSD1, GPN2, TTC7, IRAK2, TRAPPC10, GPKOW, GOLGA4, KIF1B, UNC50, USP6NL, ACAA1A, TUT1, SUCLG2, NSUN6, VPS18, TXNIP, F5, DTX1, KDM6A, DNAJC14, CALCOGOL1, DOK2, SNX15, ZKSCAN3, RAB5B, DEAF1, COQ3, H3F3A, CHTA, CRLF2, DCUN1D2, NUDT6, OTULIN, CAPN2, BC037034, PIK3R2, PDZD8, ELMO3, CCNDBP1, ZBTB5, CAT, SUOX, ING3, ELF1, GZMA, PATJ, RNF216, BRD2, VPS72, COQ7, ATG16L2, AP5M1, HINT2, SRCAP, GNL3L, TCHP, DSP, EPPK1, FBXO18, NCOR1, PPIG, JUP, PLEKHG3, ZFP24, TMEM192, XPA, EVI5, PURB, MYLK, WDFY2, CEP350, DCTN2, GHDC, ARRB2, KCNA3, FHIT, ZNHIT1, SELP, ALDOC, LRRC8A, TRIM33, MAPK11P1L, ETA1, AZI2, RAB43, DSCR3, FCRL1, COL6A2, MTF1, ATG16L1, PNPLA7, TSPAN14, PDLIM7, SMARCA4, PIAS3, ARHGAP24, RPRD1A, TIMM22, AMACR, SHISA8, STAT5B, TOB2, LRRC1, DR1, NNT, GLTP, CPLX2, GALNT12, CMIP, S100A10, HMGN2, DCAF7, BCAS2, FTL1-PS1, DDHD1, SORT1, NEDD8, ASAP2, TBRG1, RASGRP1, DGKQ, ACOT8, TFDP2, TAF10, VAPB, MADD, TIMM17B, DPY19L3, UBE2V1, MIDN, HTT, SPATA6, ZBTB22, BRD1, TBCC, TINF2, TIRAP, MRPS36, PRSS34, SLC5A6, HIST1H2AA, IRGM1, 2210016F16RIK, CBR4, GSTZ1, A1413582, F8A, ITGA4, PARK7, NECAPI, MTMR10, PRKD3, NEIL1, UBL3, HTRA2, TNS3, CDYL, CID, 40057, CYP4F13, MYO7A, ALDH9A1, OCRL, ZAP70, NEURL3, ERLIN1, SH2B3, NDUFS8, PHLDB3, CHD2, STAMBPL1, RNH1, DTNB, IRF2BP1, INPP4B, FLOT2, VAMP5, SNX11, RMC1, NUBPL, IRGM2, GM49361, PRPF40A, VPS39, PROS1, SMAPI, PCBP3, ZDHHC8, STN1

**Table 6-20 | Genes/proteins uniquely upregulated in CD93+ isolated but not in CD138+ isolated cells**

Gene Symbol (Total 386)
RRN3, CD19, MRPL18, ZC3H15, HSPA4, UBE2Z, PCLAF, CUL4B, HAUS6, UBE2E1, ALDOA, TTC1, TTC39B, NOLC1, GTPBP4, BCLAF1, OLA1, TPRN, HNRNPAB, ATP1B3, CIAO1, TLK2, APOBEC1, PLSCR1, PPP6R3, CLDND1, IPO8, PTPN9, CAMSAP1, POP4, NACA, PELP1, POLR1E, POLE4, PRMT1, RPL23A, DCAKD, ROMO1, YRDC, MYDGF, KPNA4, PARL, ADSL, PEX16, YOD1, PBDC1, FBXL15, EMD, RPS21, CCDC86, PLXNB2, PFDN1, RMND5A, ZNFX1, WDR12, GABARAPL2, CSEIL, DYNC1L2, PEX14, EIF4E, HMGN3, GLRX2, PSMB8, GNL1, PAPD5, 281042815RIK, KLHDC3, AGGF1, ZFAND6, MAT2A, PAK11P1, INTS7, GYPC, FKBP1A, DENR, PGAM5, KIFAP3, TOX4, RPS24, COPZ1, DTWD2, RFC1, EMB, SLC25A28, KIF1BP, TTC9C, IL2RG, SLC6A9, DHX29, ADSS, TMEM39B, ZC3H12D, NUP37, DCTPP1, ARPP19, CBX5, HNRNPA1, CNOT10, ARL6IP1, RAD50, CFP, RPL15, GBP2, MRPL50, THUMPD3, TARBP2, RABGGTB, SWAP70, TPT1, LACC1, ZPR1, SLIRP, PATL1, ATXN2, NUS1, MAPK8, FAM78A, MCL1, CDK12, PPP2R5D, PARP9, TM9SF3, DCAF4, TRAF4, TMEM161A, MRPL33, USP1, GEMIN6, RPS27L, GSTT3, UR11, NPLOC4, MED28, CTNND1, RRP8, POLR3E, NOP16, ADNP, TRMT61A, SLC20A2, THG1L, STMN1, TEX10, MED30, MAFG, DHX58, TCF3, DTD2, ARID3B, PCGF6, GANAB, HMBS, TRMT10A, TAF15, DPH2, UXT, ZFP142, PEX19, MTERF1B, TRIM25, CLN3, SLC43A3, DNMT3A, H2-Q8, TNPO2, ELAC2, CYP4A32, DNM1L, SS18L2, SENP6, GTF2H2, ABHD13, ZCCHC3, FOXJ3, POM121, PRRC2C, GTF2H1, IMPA2, RTN3, AHCTF1, EXOC4, RASA4, DBI, DDT, PUSL1, DNAJA2, RC3H1, WNK1, EEFIKMT1, PARP14, PHLDB1, MRPL34, LDHA, ELOA, CCNE1, IGKV4-57-1, RPLP2, PPP2CA, LARP7, PPP2R1A, EFL1, GRCC10, FAF1, TMED5, TMX2, POLR2K, PYCRL, ATG7, TRAK1, ZFP644, PI4KB, RBM45, ZFP511, NBN, NRD1, OTUB2, PRRC2A, USP46, CLK3, TRMU, HINT1, SKA1, RWDD2B, GATM, 38777, RTCB, DAD1, PNKD, GOPC, ZFP53, CIAPIN1, POLR34, RCOR3, ATM, ZFP444, CD48, MRS2, ADPRH, MRPL55, ZKSCAN1, SREK1, DDOST, DIS3L2, TOPORS, ZMIZ2, RPL34, FAS, TRUB1, TAF1C, CCNC, PRPF38A, NAA38, IGHV5-9-1, GBP5, DHCR7, CLASP2, ARFIP2, USP20, XPO6, PPP1R16B, LMC1D, HSBP1, REPIN1, ASCC1, PXXN, EED, TSEN2, ICE2, PAN3, REXO4, EIF3F, EIF1, SURF4, H2-D1, RNPEP, NUB1, POLR3G, COIL, SLC38A1, RALBP1, GOT2, BICD2, DHPS, UBE2E3, ICE1, RSPRY1, USP38, PPP4C, SLC25A15, OVCA2, DESI2, PHF6, TWISTNB, LIMS1, RPUSD4, CSNK1G3, RAN, FNTA, EIF4B, ZFP64, METTL23, OSBP, LLPH, PPP4R3B, ARID5B, CD83, MED9, MED11, CHAC2, RUNX3, IGF2BP3, INTS3, RNF10, OXCT1, PPP1CA, TET2, SLC25A3, COPS3, SDHAF3, CLK2, GATB, HDAC4, XPO1, NGRN, GAPDH, NUP43, CR1L, EEFD, TSFM, SLC30A9, NDRG3, TMEM199, KLHL20, WRNIP1, EXOSC1, PFDN4, RMDN3, GMPR2, AIM2, MED1, USP9X, D16ERTD472E, DYNLT1A, ATP5F1, CHMP7, RPUSD3, KAT14, CSNK2B, RRP15, CLNS1A, NUDT5, CNOT2, NCOA7, GM10881, NABP1, PFDN1, XPNPEP1, FAM193B, BHLHE41, YKT6, SPAST, ZFP160, MAGI3, CALU, RPL13, UBE2L3, TMEM209, ARNT, PSMB2, NUDT7, PTGES3, YEATS2, SERHL, ADI1, DDX1, SASH3, MRPL40, RNASEH2C, NUP153, ITGAV, MTHFSD, TSSC4, ZMYM4, KCMF1, TANC1, MT-ATP6

**Table 6-21 | Genes/proteins uniquely downregulated in CD93+ isolated but not in CD138+ isolated cells**

Gene Symbol (Total 267)
DOCK4, PTPRE, PTGR2, ESYT2, HMCES, MYO9A, TMEM222, NCEH1, PPM1D, RGS10, UQCR10, PTGES2, TTC14, ERP29, ANTXR2, HDAC5, VAT1, PURA, ACAD8, LGALS8, VPS4B, RCHY1, CKB, XYLT1, SMPD3, UNKL, COQ6, UROS, PIAS1, SPG21, DXO, CD1D1, UTP23, SGK3, DMXL1, ITGA2, CYTIP, HIST1H1C, DPY19L1, TRAPPC13, NDUFS3, PDLIM2, AFG1L, KDM5B, RAB3D, RUFY3, ATP2C1, TRIM8, MICU1, TRIM30D, CCDC15, MCU, PRKACB, MFSD14A, GPANK1, STX17, GNAQ, KLF13, RAB1B, HNRNPLL, VWA8, ALG9, PPP3CC, CRTC3, FAAH, GMCL1, ACADSB, JAK2, PLBD2, MTRF1L, SLC4A2, HIPK1, PRKCD, TMEM120A, ZFP131, TMEM260, SAYSD1, GPN2, TTC7, IRAK2, TRAPPC10, GPKOW, GOLGA4, KIF1B, UNC50, USP6NL, ACAA1A, TUT1, SUCLG2, NSUN6, VPS18, TXNIP, F5, DTX1, KDM6A, DNAJC14, CALCOCO1, DOK2, SNX15, ZKSCAN3, RAB5B, DEAF1, COQ3, H3F3A, CIITA, CRLF2, DCUN1D2, NUDT6, OTULIN, CAPN2, BC037034, PIK3R2, PDZD8, ELMO3, CCNDBP1, ZBTB5, CAT, SUOX, ING3, ELF1, GZMA, PATJ, RNF216, BRD2, VPS72, COQ7, ATG16L2, AP5M1, HINT2, SRCAP, GNL3L, TCHP, DSP, EPPK1, FBXO18, NCOR1, PPIG, JUP, PLEKHG3, ZFP24, TMEM192, XPA, EVI5, PURB, MYLK, WDFY2, CEP350, DCTN2, GHDC, ARRB2, KCNA3, FHIT, ZNHIT1, SELP, ALDOC, LRRC8A, TRIM33, MAPK11P1L, ETAA1, AZI2, RAB43, DSCR3, FCRL1, COL6A2, MTF1, ATG16L1, PNPLA7, TSPAN14, PDLIM7, SMARCA4, PIAS3, ARHGAP24, RPRD1A, TIMM22, AMACR, SHISA8, STAT5B, TOB2, LRRCC1, DRI, NNT, GLTP, CPLX2, GALNT12, CMIP, S100A10, HMGN2, DCAF7, BCAS2, FTL1-PS1, DDHD1, SORT1, NEDD8, ASAP2, TBRG1, RASGRP1, DGKQ, ACOT8, TFDP2, TAF10, VAPB, MADD, TIMM17B, DPY19L3, UBE2V1, MIDN, HTT, SPATA6, ZBTB22, BRD1, TBCC, TINF2, TIRAP, MRPS36, PRSS34, SLC5A6, HIST1H2AA, IRGM1, 2210016F16RIK, CBR4, GSTZ1, AI413582, F8A, ITGA4, PARK7, NECAP1, MTMR10, PRKD3, NEIL1, UBL3, HTRA2, TNS3, CDYL, CID, 40057, CYP4F13, MYO7A, ALDH9A1, OCRL, ZAP70, NEURL3, ERLIN1, SH2B3, NDUFS8, PHLDB3, CHD2, STAMBPL1, RNH1, DTNB, IRF2BP1, INPP4B, FLOT2, VAMP5, SNX11, RMC1, NUBPL, IRGM2, GM49361, PRPF40A, VPS39, PROS1, SMAP1, PCBP3, ZDHHC8, STN1, APLF, ZFP706

**Table 6-22** | Genes/proteins consistently regulated in all three platforms studied amongst RNA-Sequencing, Microarray, Label Free Proteomics**Upregulated Gene Symbol (Total 362)**

XBPI, KIF4, JCHAIN, EDEM2, LMAN1, SLAMF7, ELL2, PDIA6, ERN1, SDF2L1, FNDC3B, SEC24D, TXNDC5, CKAP4, FKBP2, EAF2, DNAJC3, MANF, SIL1, SEC11C, CRELD2, HSPA13, MZB1, DDOST, RPN2, REXO2, PRDM1, SPCS2, PRDX4, FNDC3A, TXNDC11, HSP90B1, SND1, OS9, HERPUD1, RPN1, STT3A, HDLBP, PYCR1, IRF4, KCNK6, FKBP11, RCBTB2, UBE2J1, SEL1L, HID1, CALR, SLC35B1, SEC63, SLC33A1, ARFGAP3, SPN, MTDH, ARF4, ARMCX3, SLC31A1, SSR4, LRRC59, HSPA5, HYOU1, DERL1, ALDH18A1, RPS27L, WFS1, ATG13, YARS, TMEM97, GLT8D1, CLPB, RRBP1, SEC24A, FUT8, GANAB, SELENOK, SRPR, PPIB, FOCAD, CNPY2, RRM2, SLC7A5, SEC22B, TJP2, ERP44, UCK2, ADA, TMEM214, LMAN2, GFPT1, AARS, RBM47, PDIA4, TMED9, ANKRD28, MANEA, ERLEC1, NCAPG, PVR, EVI2A, BST2, PA2G4, PPFIBP2, DNAJC1, ISG20, NPC2, SLC30A7, PI4K2B, DENND5B, CARS, CDC45, PHGDH, OSTC, PPA1, MARS, ARCNI, IDE, ALDH7A1, GLA, NUCB1, CLN6, FADS1, PCK2, SAR1B, PIGK, COPE, SEC23B, CCDC167, MAN1A2, COPG1, GMDS, SSR1, ARID3A, MKI67, CHEK1, CDC20, DTL, MYDGF, ERO1L, SDF4, SRP72, NME2, EXT1, TPX2, ERGIC2, FAH, MGAT1, TECR, CAD, GOT1, PSAT1, TUBA1C, NT5DC2, CDCA3, RTN3, ASNS, SLC39A14, EPRS, MEF2B, TUBB2B, KIF20A, ALG5, TMC01, EMC7, HMMR, GSTO1, CKAP2L, IMPA2, FAM129A, AURKA, NOMO1, SLC1A4, MRPL37, TECPR2, CASP3, SPC25, TK1, CDCA2, IQGAP3, TYMS, BRIP1, BSG, GOLGB1, ST7, LAP3, PFKM, NAPA, EIF2S2, MTHFD2, ARFGEF3, GLCCII, ATF6, WWOX, KIF22, HAX1, FLT3, IARS, ESCO2, UAP1, GUSB, QPCTL, ITGB1, COPB2, ENTPD7, GGCX, SRP19, ATP8B2, SAPCD2, ALDH1L2, E2F8, TBL2, KIF11, UBE2C, UFSF2, TARS, NEIL3, TMEM263, SARS, RAB39B, PDIA3, PIK3R6, FTSJ1, KIF18B, MELK, DHDDS, OAT, NAA20, GALE, CLIC4, FDPS, PLOD3, DIAPH3, PSMD7, TNFRSF13B, MCM10, ARHGAP21, KIF2C, MLKL, GAS2L3, DNAJB11, PLXNB2, FAM98A, COPZ1, MGAT2, PIGT, SLC7A1, HJURP, PBK, C1GALT1C1, CANX, SRM, ARL1, CCDC47, GLCE, MUT, RAD54B, TXLNA, LRR1, TXNL1, SLC3A2, SLC43A1, SLC30A6, IBTK, ALG2, YIF1B, TMED5, MRPL51, CCNE1, DHCR24, CINP, ALG8, DNAJB12, PCLAF, ERAPI, TMEM33, SOWAHC, GLIPR2, NXN, HASPIN, MAN2A1, NSDHL, GNL3, IMPDH1, SRP54A, EIF2A, ETV6, P4HB, GALNT2, TXNDC15, MSANTD4, BET1, AMPD2, CYP20A1, PSMB2, RHEB, PRDX1, MRPS34, SPC24, PLXNA1, NME1, DBB1, EYA3, MRPL13, EDEM3, TRIM32, GOLIM4, ATG4A, LY75, SCFD1, BCKDK, XPOT, MINPP1, MRPL20, ANAPCS, GSPT1, BCAT1, NDUFAF2, CERS2, TXN1, HCFC2, HMBS, UQCC2, SSBP3, CCDC88A, PSMB6, HEXIM2, MCRIP2, MRPS2, NUS1, SEC23IP, KLHL9, RDH11, H13, COG4, AUH, GRN, SESN2, PSMB5, SYVN1, ATP2A2, TM9SF2, ORMDL2, MRPL17, PABPC4, SRP68, TMED1, EIF5B, PDE4DIP, NDUFB9, AAMP, CLPTM1, VPS25, EIF3C, SPG20, HMGN3, CLDND1, EIF2B4, PSMB7, TMEM41B, LIG3, DCUN1D5, SCFD2, ADRM1, LNP

**Downregulated Gene Symbol (Total 393)**

ARHGEF18, SORL1, RIPOR2, SERPINB1A, RCSD1, RASGRP2, ZFP318, AFF3, ABLIM1, ARHGAP17, CD55, DEK, PXX, EVL, ETS1, TRAF5, HHX, STK26, LBH, FMNL1, CIITA, SIPA1, SCML4, SP100, BTG1, BCL6, TMEM131L, CR2, BACH2, KMO, MORC3, PRKCB, TBC1D10C, CDK19, JAK1, CD22, SPIB, BTLA, AIDA, IFNGR1, ZBTB18, HCK, ACAP1, PIK3CD, FAM49B, SNX2, STAT6, IFIT2, ELL3, FCMR, DOCK11, SESN3, PACS1, ADD3, ARHGAP45, GGA2, REL, ARHGEF1, ELMSAN1, BLK, BMP2K, RIN3, EML4, ANKRD44, RGS18, ARHGDIB, CXCR5, PTPRJ, KYNU, USP25, NFKBID, LCK, SMIM14, TNRC6C, CAMK1D, ARHGAP25, DENND4B, MAST3, SNX29, SLC2A3, GMIP, FGR, BIN1, IRF8, ATAD2B, PAX5, MBP, ARHGAP4, DMXL1, HPS3, ARID1B, STK10, SMAP2, FRY, ELF4, MTSS1, AKT3, NFATC1, TMED8, XYLT1, ZMAT1, ACTR3, NCKAP1L, SETD2, AMPD3, TBC1D5, SP110, NIN, RASSF3, CARD11, RAP1GDS1, LRRFIP1, ACAP2, SNX5, ZFP639, RABEP2, CDK13, NUA2, FKBP15, LPGAT1, IKBKB, GIT2, SIGLEC3, KDM7A, SP4, ITPKB, SATB1, CHD3, AKAP13, PRKCE, TRIM7, CSK, FOXO1, VAV2, CDK5RAP2, RUNX1, MAP4K2, ADD1, FAM208A, LGALS8, NAAA, ARHGAP9, FCER2A, BRWD1, CD200, PHC1, PTBP3, ANKRD13A, TMOD3, IQGAP1, FLI1, ZFP263, PLEKHF2, VEZF1, CELF2, ACTR2, STK17B, FUS, ZFP518A, INPP5F, YEATS4, CRTCC3, KAT6A, BPTF, CD2AP, INO80, SCAF11, SLC9A7, NFATC3, ZBTB5, REST, ADCY7, AP1G2, SKI, CD84, DENND6A, CD79A, INPP5D, SLC25A24, MPRIP, PLEKHO1, STX7, PLCB2, BCL7A, MAFK, TOP2B, CBLB, UBLCP1, ATG16L2, MPPE1, RASSF5, ROCK1, TOMM34, IRAK4, ZFP422, RXRB, MTM1, WDR82, AKAP8, METTL14, AP1S2, CEP135, KAT8, TRIO, LRCH1, SPTAN1, PPP1R12A, STK4, AGO1, CEP68, PSD4, RABEP1, NR3C1, SENP7, PDLIM2, MAPK14, DENND1C, MUM1, CEP295, TCP11L2, BICRAL, ZBTB7A, IL4RA, PPP1R21, PHF2, ARID4A, FMO5, PDP1, ARAP1, SCAF8, ATF7IP, STRIP1, EZH1, ARID1A, PIBF1, NUMA1, ATXN3, PGM2L1, SETX, HMGN1, ZFP638, IQSEC1, ZFP740, ARID4B, SUGP2, ARPC2, JARID2, CCDC82, NONO, AKAP8L, PEAK1, AGPAT2, NFATC2, DOPEY2, ZC3H6, OARD1, CAMK2D, TMEM2, THUMPD1, UTRN, KANSL3, PRKX, H2-EB1, CDKN2AIP, ZFP592, UBA7, RBM5, ATP2A3, TFEB, MSL2, ANKRD11, MECP2, SMC6, LNPEP, YWHAZ, RPRD2, NEK9, HECA, WASHC4, NR2C2, GANC, MCTP2, USP12, CEP170, H2-DMA, VPS13C, DAPP1, ARHGAP15, RP9, MAP4, KLF3, PPTC7, SYF2, ACCS, PRR14, CNTRL, TFAM, UVRAG, ILF3, IST1, ICOSL, TMEM63A, RBL2, DHX15, PDCD7, CPSF7, NIPBL, ANXA11, CREBBP, RCOR1, DENND5A, OTULIN, CREB1, UNC119, KLHL14, RYBP, IRF2BPL, ESCO1, XPC, MAP3K2, SLC23A2, NRF1, SNX6, ABCD4, DDX59, SPIN1, STK11IP, TAF3, NADSYN1, WASHC2, RAPGEF6, ZFP746, SNRK, APEX1, DIDO1, PIP4K2A, KBTBD2, EPC2, SLAIN2, SFPO, RHBDF2, PDCD4, NCOA1, EP400, OSGEP, CCNY, TAPT1, NFATC2IP, ADO, BRWD3, SYPL, HACE1, TERF2, NCOR1, STK24, RSF1, CCNDBP1, PLEKHM1, PHF20, EIF4A2, ODF2, TERF1, PHC2, ACIN1, EFR3A, BRD4, CCDC12, UIMC1, ACSL1, EFHD2, ARHGEF7, TTC7, PGLS, KAT5, HMGXB4, PITPNM2, COPG2, ZBTB24, PPP3CA, RAB21, PPL4, BRD3, SNX30, SH3BP2, ELF2, METTL3, AFF4, GABPB1, CAPG, WASF2, TRIP12, SLTM, DNMBP,

**Table 6-23** | Genes/proteins consistently regulated in two out of three platforms studied amongst RNA-Sequencing, Microarray, Label Free Proteomics**Upregulated Gene Symbol (Total 1136)**

SLC6A4, CALML4, DNAAF1, SCIN, CFAP46, RGS5, HIST1H2BG, TMEM167, CD59A, DERL3, SDC1, TMEM176B, MCTS2, MT1, TRAM2, BC004004, CREB3L2, PON3, TMEM176A, PLPP5, UBL5, SSR3, TRIB1, SLC44A1, ENTPD1, BCL2L1, WIPI1, CHID1, DNAJB9, BHLHA15, CCPG1, SEC61A1, GMPPA, NACC2, SLC39A7, YIPF2, KDELR2, TMED10, LAMP2, MAGED1, TNFRSF17, SLC04A1, 1700017B05RIK, PISD, SURF4, SPCS1, NANS, SLC5A2, CD93, YIPF6, GPR55, AP3S1, ALPL, SSR2, SELENOS, USO1, H1F0, LAX1, ZBP1, CALU, RAB3D, SLC39A11, TRP73, ITM2C, MGLL, MCFD2, PLBD2, TCEAL9, SSPN, CREG1, LARPIB, CHPF, FBXW7, FICD, TBC1D24, CITED2, CD28, KDELR1, AGA, SLC17A5, ZBTB38, GPR155, BHLHE41, MAGT1, CSF2RB, UBXN4, COP22, B3GNT9, BUB1B, SLC48A1, LAG3, TMED2, CAPN5, PLS1, TXLNB, CLPTM1L, GMPPB, UBA5, PIM2, ST8SIA6, BTD, EEF1AKMT3, FAM129B, SEC31A, GFII, GTSE1, TMED3, EXO1, CEP170B, VEGFA, TMEM205, BSCL2, OSBPL3, ATF5, TACC2, NCAPH, GGH, GCAT, SLC25A23, CCNE2, THBD, MVB12B, PECC, KRTPCAP2, NUGGC, POU2AF1, TAPBPL, FEN1, CIB2, CSNK1E, TMEM39A, CST3, UFL1, LMF1, GM20388, CDC25A, PRRG4, SELPLG, PLCD3, SKA3, NAGA, GORASP2, TWSG1, LAMC1, DPAGT1, KCNMA1, LGALS1, SLC31A2, SEC13, GLMP, CPTP, ALDH3B1, TMEM184B, SOCS2, MARVELD1, RCN3, SMO4, PCDH15, CCNB1, LY96, ATG9A, EDEM1, ITM2A, MAPRE3, LMNA, ANKRD55, PKM, ITM2B, BMY1, SLC1A5, GBA, CPOX, MBNL2, PTPN21, CTSL, PLEKHF1, STARD5, PGM3, PEX11G, NCAPG2, CLINT1, RAB2A, CLIC5, CSGALNACT1, TFG, YES1, CDK1, TMEM147, DLGAP5, DAP, ZBTB42, COG7, CTSB, CD2BP2, MAP1LC3B, SGO1, RAB26, ACAD11, PRSS16, KIFC1, NRP1, SERP1, GOLT1B, SUB1, MYO6, SEC61B, LAPTM4B, DAD1, ERGIC3, MANSC1, CFLAR, RCN1, PRR11, GSTT2, MYCBP, PLK1, BOLA2, NOL3, SLC43A3, NEK2, OGFOD3, CDC6, CORO2A, ASF1B, BRCA1, CCNA2, STK3, UNC13B, CKB, ENTPD4, RAD51, CAPN2, SGO2A, NMRAL1, MAGEH1, PSRC1, MAGED2, GPRC5D, TOX2, CEP55, GPX3, PAFAH1B3, NCKAP1, SPRED3, GOLPH3, EPCAM, EME1, AURKB, EPAS1, TTC7B, CDIP1, CST6, CDR2, GNG12, HARS, ABI2, PLA2G16, FAM114A1, STAC2, YIPF5, STAMPB, CHAC1, ASS1, CCNB2, MAOA, IL15RA, ABHD4, RECQL4, CDCA8, H2-T23, LRP8, YIPF3, TM9SF3, COPB1, ESPL1, ITGB6, BIRC5, GPR19, ABLIM2, HNRNPLL, COG6, FKBP14, GPT2, ANLN, NCAPD2, GNPAT1, PCBDF1, FUOM, METTL1, KCTD21, ABHD14A, GINS2, CACNG6, OST4, KLC3, ZFAND6, MAD2L1, TOP1, TM9SF4, EHHADH, MELTF, MFSD11, EIF3J1, LDLRAP1, CHST12, EBP, TMEM208, UBE2T, SETD3, MGST2, TTK, DHFR, CD274, ARHGAP11A, SCAMP2, FAS, ETTFDH, SDF2, ARFP2, SLC7A3, ORC1, SLC6A9, PDIA5, GPNMB, PKP2, ASRGL1, ZWINT, MYBL2, DERL2, ETL4, TICRR, AIG1, CYB561, BMP6, BUB1, CNKSR1, BIK, IFT122, SEC61G, ATXN1, YWHAE, CD68, TRIP13, CSRPI, SDC3, ARHGFE12, DCLRE1A, THYN1, MAN1B1, LRPAP1, H2-K1, EIF4E3, GOLGA2, MT2, LTBP3, CDPF1, NDUFA1, BMPR1A, SEC14L2, CDKN3, TEX2, TCTN3, TMEM141, CDKN2C, STT3B, NEK6, WDR45, SHCBP1, HIRIP3, ENPP6, MLEC, SLC35E1, ATP2B4, XK, CERCAM, FAM83D, NFIL3, DCAF12, CDK2AP2, IQCB1, RETSAT, SPOP, DTYMK, SQSTM1, LTBR, FAM162B, GINS1, ALG14, LGMN, CDC25C, PYCARD, EPN2, MBOAT2, MORF4L2, GOLGA3, EMC4, KDELC1, ULBP1, MAN1C1, SLC25A10, CLN3, ECT2, DUSP26, PCGF5, GOLGA5, CCR10, COMMD3, ACP2, 5730409E04RIK, LGALS3BP, PHLDB1, MGAT3, MPI, AZIN1, ASPM, PCX, ZNHIT1, CDV3, NCALD, DEPTOR, TMEM258, PNPO, EMP1, TMEM106C, SLC39A8, 1810055G02RIK, CREB3, SCD2, HSPB6, SYNDIG1L, ZC2HC1A, RAB4A, COPA, UBE2S, KPNA2, SLC7A11, YIF1A, TMEM256, SERINC3, SHB, SMOX, NDUFS5, PCNA, SWI5, LEPROT, DSTN, DARS, COX17, ITPRIPL2, ATRAID, AUNIP, SUMF2, ATP13A1, CCND2, GM17018, EIF3A, MTFR2, MOGS, DGCR6, SLC39A9, TMEM9, PTGR1, KAZALD1, HIBCH, MGAT4B, NFXL1, SLC38A10, TMEM255A, ARMCX2, NOD2, WDR62, PGAM2, ANG, 1110032A03RIK, ABCB6, MCPH1, ST14, SLC22A15, GIGYF2, PCYOX1, CAMSAP1, ARL3, ITGAL, MOSPD1, JUN, EDF1, ITFG1, PIPOX, PIF1, SLC11A2, INPP4A, KIF14, KNSTRN, RNF181, GALK2, FTL1, LCA5, ZDBHHC14, TMED7, 1110004E09RIK, B9D1, GSTP1, SEPT2, SMPDL3B, PTER, GOSR2, NMRK1, GAS8, H6PD, STXBPI, CEPB, MS12, UNC50, RPS6KA2, AKR1E1, MVB12A, CENPI, Rragd, WARS, ANXA5, ABHD5, CXADR, TOX3, RHPN2, GM20425, CRYL1, DSCC1, HBS1L, UGGT1, KIF23, SLC29A3, SPAG5, TEX35, POLE2, MORN2, STARD3NL, CTSD, NARS, MS4A3, TOR2A, DHCR7, SAR1A, RSPH1, TG, PRADC1, RHBDD1, KNTC1, ALAD, LIMS1, EHD4, ITGA3, MAP2K2, RINT1, CRELD1, 6430548M08RIK, DOLK, GLRX, CLIP1, DLG3, CD63, ALG3, QTRT1, CMC1, 1700047117RIK2, MARC2, ST6GAL1, TARSL2, CSNK2A1, ETFA, ACS2, RIMKLA, VMP1, RNASE4, TESC, HELB, RER1, SLC39A4, NDUFA4, AIFM2, MDH1, HYKK, LOXL3, SEPT8, XRCC3, TMEM107, UNG, GOLGA1, PLXND1, CUTA, FECH, MCEE, FADS2, YTHDF1, SHISA4, IPO4, PIMREG, ALG9, NDEL1, ETNK1, SHQ1, UBE2L3, GOT2, FAF2, GTF2A1, CARD10, DEPDC1B, NODAL, IER3IP1, PRR5, ZFP64, HMGB3, MSRB1, ZMPSTE24, MITF, PVT1, PRKCSH, SEC14L1, TMED4, TRAPPC2L, FAM221A, NUDT9, IFRD2, NEU3, AVPI1, CENPE, FOXRED1, CYB5D2, METTL6, ASNA1, ZFPL1, NUDT22, PSMC2, ECH1, DCAF10, IPO5, TM7SF2, GART, GOLPH3L, GEMIN5, LPIN1, GPR15, BLVRA, TRMT10A, RABGGTA, GDE1, IL2RB, PLOD1, EXTL2, NADK2, CARHSP1, HEATR5A, ADAP1, COMT, PIGG, HAGH, PSMA5, ACO2, IER3, PPM1H, AU040320, LPCAT3, E2F7, PRC1, NDUFV3, ATG5, CEP19, BCL2L2, TMBIM4, EB3, KCNAB2, FAHD2A, ANAPC13, RHBDD3, MAP3K20, VSIR, NEO1, MAFG, NUDCD2, GARS, TFDP1, SHPK, S100BP, AKR1B10, LZTF1L, ST8SIA4, ACADVL, NELFE, C1QBP, COL7A1, SSX2IP, TBC1D7, TBCD, E2F6, SLC25A39, LRRC8D, SMIM7, HSD17B14, CSM51, TOR1AIP2, NDCA5, CENPF, PUSL1, ATP5G3, VAT1, PSMD12, ACOXL, PSME3, FOXO3, MICALL2, BET1L, UMPS, PDXDC1, LAS1L, MTHFSD, PDK3, CLTB, HIST1H2BC, MUC1, DAAM1, IFNA1, LMAN2L, UBXN8, BLMH, CNTNAP1, SELENOM, THNSL2, INPP1, GRM1, PIEZO1, NT5C2, MIS12, PTCH1, RFT1, NDUFS3, GNAS, UFD1, SAMM50, TMTC2, DES11, CDC34, PRDX2, HIBADH, FAM45A, PICK1, ETFB, COX7A2, DNAJB13, PSMC1, CASQ4, FAM173A, IL12RB1, ATP1F1, RAB23, ISOC1, ISCA2, PSMD11, MBD1, GPR108, CHPF2, SLC35F2, GEMIN6, FITM2, UQCRCQ, LYAR, ATP6V1D, RNF5, DRG2, GBF1, IL6RA, TMEM219, SLC45A4, TIMP2, EIF4G1, SDHAF4, PLCD1, TXNRD1, GSTM4, HINFP, SLC25A25, ZFP428, CLPX, PDZD11, IDH2, ALG10B, DNAJC28, LRRC41, TMEM70, POP1, MRPS7, HIVEP3, ANKRD46, PLD3, ACO1, DNAJA3, SURF1, RFC3, MIB1, MRPS12, SAP30, CSTB, INF2, WBP1, RND2, TFB1M, MND1, ZFP825, ATF6B, P4HA1, RSPH3A, UROD, CSF2RB2, KDELC2, FLNB, SNAP47, NGLY1, PSMC3, CTAGE5, MR1, PAPSS1, TNFRSF10B, XCR1, ANKRD6, GLB1, HACD3, XKR8, ATP6AP2, SVIL, MTX1, TM9SF1, CBWD1, MKRN2, PRMT7, KLHDC8B, DOCK4, SLC35B2, PNKD, GCDH, IFT43, EMC9, DNPH1, BFAR, OPA3, ISG15, EIF2B11, PRKCI, ELOC, TRPT1, SLFN9, CASKIN2, TUBG1, VSIG10, BRCA2, SUV39H2, REEP5, BCAT2, DCTN3, NDUFB6, LNPI, IGF2R, SLC17A9, PRDX6, PRIM2, TKTL1, NDUFB11, TIMM17A, RAB13, MLX, DHRS3, ABCF2, NIFK, EIF2AK3, STIL, ATP5J, UBE2G1, AHS1, NDUFA13, NUPR1, FAAH, IKBIP, MAT2A, CRYZ, SIGMARI1, GINS3, NUDT5, GUK1, NENF, MRPL50, FNIP2, MCL1, PHPT1, USP54, CCT3, EIF2S1, UCHL5, PLXDC1, NCAPH2, NUDT2, LRRC28, TEX30, NDFIP2, RBKS, STAU1, CMC2, LUZP1, LMF2, MULL1, UBE4B, COX6A1, SELENOF, RGS10, EIF2AK1, ABHD2, USMG5, OXR1, PSMD6, COPS6, PPAT, TPI1, METTL9, HMOX2, COX5A, ATP5H, B4GALT1, ALDOA, DYNCL2L1, KIF1B, 4933434E20RIK, SKA2, GRAMD3, IPP, ATAD1, IGHG2C, ACOT13, MTX2, NDUFA2, DTD2, NAA38, SOD1, TMEM106B, SCYL1, MTERF4, ENK1, MON2, COPS4, TMED6, PSENEN, TBC1D15, PEX11A, MYO5A, POLH, PSMB4, GALK1, PSMD13, NDUFA11, FLOT1, CHAC2, FABP3, ITFG2, MDH2, MRPL33, NAB1, NUP35, HSD17B4, MINOS1, CCT7, PTSS2, UBE2K, SACM1L, RNF14, NTHL1, UEVLD, GSTM5, MMADHC, MRPL34, DDX54, JAG1, DUSP19, PYCRL, MBD3, CPD, DNAJC12, ICAM1, RARS, METTL26, ATP5J2, ACACB, TIMM21, CISD1, ST6GALNAC4, RABL3, PSMD8, RAD54L, PSMD14, BPGM, EMC1, GLRX2, UBF1, PSPH, TUBB4B, MIA3,

**Table 6-23** | Genes/proteins consistently regulated in two out of three platforms studied amongst RNA-Sequencing, Microarray, Label Free Proteomics (cont.)

MRPL22, GCLM, FKBP1A, RPAP3, MRPL35, TIMMDC1, PRMT1, RAB6A, VCP, DHRS7B, P3H1, ZCRB1, RABEPK, MRPS14, UBL4A, TSR1, TMEM165, CRLS1, KIF18A, ELP5, ACSL4, NEK4, MRPS15, SMYD2, EIF4EBP1, ZFP706, D16ERTD472E, MRPL53, CDK4, CIAO1, POMP, ELP4, CARM1, LSS, CNOT9, VPS37C, MRPL57, TVP23B, TIMM50, ACSL5, AIMP2, LARPI, HIGD1A, MRPS18A, NDUFA9, ARMC1, FAM120A, NSD2, FAM96A, NBR1, AVEN, PEX19, GLS, MRPL48, PSMG1, ATP1A1, TLL12, WDR92, MRPL28, MRPL16, GRSF1, DCTD, BYSL, ISYNA1, MRPL27, PYCR2, PSMD1, PSMC4, PEX14, CTH, ELP3, PSMD4, MECR, PEX3, SLC52A2, TOX4, DDX19A, ZCCHC9, TARBP2, GMPS, EIF4E, DHRS1, KDM5C, NTMT1, OGT, NLRN, UBOX5, EIF2B3, MRPS5, PSMC5, DDX3X, CHEK2, OGFOD1, POLR3D, PREPL, LACTB2, METTL15, NDUFAF1, FNTB, MRPL18, MRPS23, CKS1B, MRPL58, MRPL46, GSTT1, TULP3, SLC25A17, DNAJC15, NUDCD1, SCO2, MRPL4, RNASEH2C, NME6, FTO, SIRT4, POLDIP2, ALCAM, LARP4, MRPS24, DBI, AURKAIP1, EMC2, FZR1, RARS2, MRPL15, QSER1, TMEM129, MRPS11, CKAP5, ATP5K, MRPS17, TTC1, EEF1AKMT1, MRPL21, HNRNPAB, IGHG2B

**Downregulated Gene Symbol (Total 1093)**

FBXO41, STMN3, KIAA1683, OSBPL10, H3F3B, CAPZA1, HIST1H4E, L3MBTL4, LAMB1, SIAH1A, SLC44A2, MS4A1, BANK1, 0610030E20RIK, GDF7, MARCH1, MRFAP1, EBF1, CD19, ZFP273, VMA21, COTL1, CD37, FCRL1, NCF1, NOTCH2, A430078G23RIK, CFAP44, BCL11A, S1PR1, CALD1, MYLIP, CCR7, TM6SF1, GPR174, LCP1, H2-OB, GRAP2, SLC4A7, RASSF2, CORO1A, MAPK11, KLF2, CD83, SLC28A2, PIK3API, DTX1, GCNT1, STAP1, GALNT10, ZBTB4, I830077J02RIK, CRLF3, FFAR1, NAP1L1, MYO1G, H2-EB2, FAM129C, CRYBG1, RASGRP1, RASGEF1B, APBB1IP, TLR1, GUCD1, SESN1, KIF21B, PTPN6, SLC25A37, HERC4, LY86, FGD3, DGKA, FMNL3, TMC8, DENND4A, TESPA1, TMEM243, PSTPIP1, MACF1, PIK3C2B, LTB, CD40, PTPRC, UBE2D1, ABCA1, CCR6, WDFY2, ZFP821, RETREG3, HVCN1, TTC9, RHOF, PPP1R18, CD72, FOXP1, S1PR4, NOD1, LMO2, SLC38A11, HMGB1, GRB2, EGR3, H2AFY, FAM172A, NLRC3, IRAK3, WDFY4, PARP8, TRIM58, SLC38A1, ARHGAP27, RPS6KA5, RGS19, SH3BGL3, COLGALT1, CDC42SE1, AKNA, MAP4K4, TGFB2, GMFB, PRKAB2, PDE7A, LYN, SUN2, ARRB2, DOCK8, TRAF3IP3, CDC40, TNFRSF13C, STS, ZFP329, ZCCHC11, MYCBP2, MALTI, HNRNPAO, CD180, GAPT, LRRK2, SMARCD2, MNDAL, H2-Q6, CMTM6, TEP1, RHOH, EPST1, ZFP157, SYK, ZFP512, L3MBTL3, LGALS9, SLF1, RHOG, FAM46A, 2810021J22RIK, FCHO1, CD55B, MGAT5, PFN1, DPH5, CRIP3, RRAS2, AKAP5, ZFP65, RUBCN, MTPN, ID3, FOXP1, SSH2, CIB1, DENND2D, COBLL1, UBE2R2, TET3, ADAM28, VASP, FAM26F, SIRT7, KLHL36, CBL, PIP4K2B, CCND3, PTPRCAP, R3HDM2, ZFP90, DOCK2, SMG1, TRP53BP1, MYO9B, ZBTB14, IFIT3, SOCS5, CD79B, AP1S3, CHKB, SEPT6, RNF44, CWF19L1, SH3BP5, TMEM260, CAT, CNPY4, EXOC6, RETREG1, TMUB1, PLAC8, NSMF, ARHGAP24, XPO6, RARA, PAPOLG, GAB3, RELT, P2RY10, UHRF1BP1L, QK, TCTN1, HEATR6, CNRIP1, LRMP, RNF145, SH2B3, CCAR2, CHD7, RFX5, RALGPS2, CNN3, BIN2, DIAPH1, DEPDC5, LENG9, SIDT1, XPO1, DHX57, B3GNT7, AKAP10, USF3, ARHGEF3, CPNE1, WDR1, PHACTR4, RFX7, ZFP148, GPSM1, MDN1, TES, SASH3, CASP2, TRIM14, GATAD2A, DDX60, ANGEL2, CBFB, NUP160, 9130401M01RIK, ATR, ITPR1, SMARCA2, CARMIL2, SMURF2, SLC39A10, GLMN, IFNGR2, CEPT1, ZMYM6, ZMYM5, SNX8, ERP29, GRIPAP1, STX17, 2610507B11RIK, MKL1, DUSP10, INTS4, CYTH1, RUFY1, HSDL1, PARP6, MBTD1, GM14698, NFKB1, ARRB1, LPXN, HEATR1, MLLT11, STAG2, TCERG1, USP34, FAM107B, MED4, CEP85L, KLF16, CHD1L, JAK2, LPP, MAP4K5, SYNE3, CHD1, TAF7, ZEB2, DDX6, LYL1, PACS2, DTX4, OTUD4, FBXW4, RNMT, H2-OA, CYLD, HTT, CCDC88C, LBR, ATM, UBAC2, EGLN2, INVS, ATP11C, ZFP729B, RCC2, RAP2C, ATP8A1, MOB1A, TIGD2, RAB37, CAPRN2, MYBPC2, ICE1, CYP4F18, DDX31, NUDT16L1, ZBTB25, BICRA, SLC12A6, GLUD1, NFKBIA, NUP153, PDS5A, ITSN2, TARBP1, FANCM, IKZF1, SKAP2, PARP11, WDR44, CLEC12A, CAPZB, SNX10, TAF1A, ABI3, MTERF1A, GABPB2, MCM9, SLC16A7, CD82, SIPA1L3, KDM5B, UNC93B1, CAP1, FAM102B, ERP27, ZCCHC7, INAFM1, UHMK1, HIP1, APPL1, ZFP385A, ZFP729A, PRKDC, CASD1, UBA2, TWISTNB, PTK2, RAE1, POLI, POLD4, CHMP2B, FUBP1, SIK1, ARAP2, TAZ, CDKN2D, N4BP3, ZC3H4, PTEN, GNG2, SSH1, RNP133, BEGAIN, GNA13, KXD1, DYRK2, SURF6, LPCAT2, CAST, ZFP831, CDK17, ZMYM2, PIK3R1, CTPS2, SLC6A6, ZFP143, NOP53, SENP1, ABHD17B, PDS5B, TMEM131, GPALPP1, MICAL1, BLOC1S2, STX6, PHF21A, ZFP407, TSGA10, DHX9, PDCD10, ZBTB37, CTDSPL2, MICALL1, PCM1, PUM3, ATG16L1, BOD1L, MLLT6, YTHDC2, LDB1, ICE2, MEPCE, CD74, IRGC1, BBS4, WDR37, CCDC191, NUDT3, KANSL2, SIN3A, CHD6, OSTF1, GPM6B, KLHL42, CARNS1, ZCCHC8, SIPA1L1, PRRC2C, CPM, ULK3, SLC2A1, TUBA1A, FBXO31, ZDHHC17, ZDHHC23, FBXO11, RAB4B, RBM27, FBXL3, EXOC2, VIM, PHF23, CHD8, POLG2, IREB2, ZFP28, KDM2B, 2610008E11RIK, PRDM4, AKAP7, RRM2B, CCDC50, STRADA, WAS, ARF6, FASTKD5, RASA1, FBF1, CCDC137, ZFP335, FAM117B, RNF38, CC2D1A, CLASRP, ADRB2, TACC1, MTF2, DET1, DNAJC2, ABL2, OXSR1, MANBA, R3HCC1L, CD53, TADA3, TGS1, ZBTB11, PHF3, ZFP35, TMEM222, DSTYK, RNF220, DLGAP4, 0610010K14RIK, CYP2D22, IER5, SEC24B, TTF1, TAGAP, PPP4R3B, TAF4, GON4L, PLCG2, KLHL24, RAB32, PCMTD2, CHD4, WARS2, AHNAK, SLC22A5, FIP1L1, ARHGEF6, CHMP1B, FAM126A, KLHL5, HARS2, SLC25A30, MGEA5, PDSS1, ABRAXAS1, SMCHD1, SCIMP, GMCL1, NUCKS1, ARID5B, CHD2, EXO5, NT5C, UCKL1, DDB2, B230219D22RIK, LASP1, SETD1B, TAOK3, ASH1L, CSE1L, TSC22D2, FRAT1, EHMT1, TOB2, FBXL8, PADI2, DHX34, KDM6B, BRPF1, ATP2B1, HNRNPR, RPUSD2, MINDY2, SDCCAG8, UBXN1, USE1, IWS1, ACTR6, THOC1, PPP4R3A, CUX1, TRIM34A, CNOT6L, MAF1, LEMD3, RECQL, FAM102A, MKNK2, ZBTB10, ZFP317, CDS2, IFT74, TNKS2, EPB41, TMEM55B, KLHL20, ALOX5AP, SVBP, ZADH2, GGPS1, KLF7, ASTE1, MAML1, INPP5E, MAP2K1, TERF2IP, TBC1D32, MAU2, MFAP1B, PCIF1, PAK2, BEND5, PHIP, MFSD14A, RNF31, VPS13B, MSL1, SP1, SPTLC2, KAT2A, SNX3, TESK2, RAD52, KLF13, MAP3K14, SPOPL, ATXN7L3B, BCL10, MAML2, FNTA, ZBED4, LETM1, GNA12, TIA1, THOC2, GPR137B, CEP120, FAM49A, ZFP207, JMJD1C, MTMR14, HNRNPU, TRIM21, TAF5L, COQ2, CLK2, PIK3R4, 9130011E15RIK, RPL18, CEBPZ, ABCB1A, ACTR5, AP5Z1, TAF11, MAP3K3, ARL6IP6, PPIG, SDCBP, USP3, PCF11, TCEA2, OSBPL8, XIAP, TRPS1, KHDRBS1, SELP, NRDE2, FOXF4, RLF, GBA2, DCP2, MAPK1, H2-DMB2, TMEM43, ARMC10, FES, NOL11, RALGAPB, TMEM245, PTS, USP7, SPPL2B, MBNL1, TLL3, CELF1, SART3, ERMP1, FAM168A, SACS, RALGAPA2, VPS18, CAR2, CREBRF, DCAF4, RBBP6, ZYG11B, IFI208, SNX25, RCN2, KCTD2, AQR, LMBR1L, HECTD1, RDH14, PRKD2, PUM2, RPRG, PHF14, ZFP326, GPATCH11, DNAJC27, MKL2, PYROXD2, KANSL1, CCDC94, ATXN2L, SZT2, CD1D1, CTCF, GPS2, MEF2C, MYNN, PTBP2, PKN1, TBL1X, MCRIP1, AP3B1, UBE2W, MYD88, DDX55, SMC5, HNRNPC, SF3A1, ELF1, SNRNP200, ARRDC2, EIF3F, PNKP, CMPK2, LTA4H, SFN, CCDC22, CMTM7, VHL, GPBP1L1, MKLN1, WIPF2, U2SURP, DDX39B, KLHL18, PHF20L1, YTHDC1, NUBP1, WDR11, POLR1C, KMT5B, CRTCL, BNIP3L, RWDD3, TRA2A, PEX6, AGFG1, MCM3AP, POLR2D, FBR5, TCEPR1, CHTOP, FIG4, MTERF3, BTBD1, RPL8, PDP, ABHD10, NCBP3, BICD2, BNIP2, 4921507P07RIK, NUP205, UBL3, TCOF1, FBRSL1, TGIF1, GCC2, ZFP292, WSB1, PSMD5, KDM3A, RANBP6, SLC12A7, HUWE1, HSPBAP1, SECISBP2L, RNF216, SIRT1, TPK1, CAB39L, ZFP830, FBXO21, PLCG1, SDCCAG3, ATE1, FTH1, NPAT, DIAPH2, LIMK2, OVGP1, LONP2, ZHX2, CUL5, MEAF6, BRMS1, BMT2, UTP15, SRSF1, SF3B2, TRIM65, XRN2, GPATCH8, DYRK1A,

**Table 6-23** | Genes/proteins consistently regulated in two out of three platforms studied amongst RNA-Sequencing, Microarray, Label Free Proteomics (cont.)

---

ZFP865, DTNBP1, SMARCA5, MCMBP, KMT2A, SCAPER, AP3M2, PDZD8, RLIM, SERPINE2, ZFP131, SMARCD1, RPL37, RPS11, SPATA13, UBTX, TPR, MLXIP, RALGAPA1, RERE, CIPC, TOR4A, API5, ZFYVE27, RPL37A, NOP16, ING1, TSR2, YPEL3, CERS4, ITPR3, BCOR, DDX23, SIRT5, PHKA2, FILIP1L, GIGYF1, TRAPPC8, CLK1, PTPN18, AMDHD2, SFG11, SGSH, MOB3B, SLC25A32, RNF41, SPAG9, TMEM134, ENO3, THUMP2, LRBA, DNAJC9, INO80D, EIF1B, DDX27, PRDM2, TNPO1, STX4A, TNIP1, SDHAF2, GRK6, DHX36, PAN3, ZFP619, BRD8, FAM193A, RBM34, DPF2, ZBTB3, SPRTN, H2-AA, RICTOR, DGKD, LSM3, SF1, EXOC6B, XPA, ENSA, ZFP524, TMPO, BDP1, MYO9A, GATD1, CDK5RAP1, LACTB, IFI205, DNMT3A, HNRNPDL, AMZ1, ILF2, ZFP58, SH3GLB1, ALDH16A1, 1700037H04RIK, TSNAX, RPS25, CHIC2, SKAP1, FBXL12, IARS2, MAVS, SPG21, ZC3HAV1, IRF9, HNRNPUL2, HNRNPM, CBLL1, TRIM11, NCL, GSN, LUC7L3, RBM6, KCTD13, NSD3, RACK1, RPL23A, SBNO1, SRF, TAF8, LMBRD1, SPTBN1, AH11, SAMHD1, CPSF6, PPCS, VPS16, ZFP346, POGZ, QDPR, ERBIN, FAM160A2, LSM6, FAM20B, DENND4C, RASSF1, ZHX1, RASA3, IKZF3, ANKRD12, AFTPH, CUL3, KDM5A, MTMR3, MYL6, EZR, MDM1, NFRKB, PPM1A, TGFBRAP1, SFSWAP, ITPR2, AGL, NSUN6, UCK1, BAHD1, KHNYN, METAP2, ZFP953, STRN, MNT, WDR47, CABIN1, SDE2, HEXA, CXXC5, CAND1, EPC1, ZFP800, VPS26A, RSRC1, LEMD2, AKT2, PHF1, GRAP, RPS6KA3, GMEB2, ZC3H12A, NDRG1, TLN1, PATJ, RAB24, LAMTOR1, FOXN3, FBXO38, GRAMD1A, HPS4, FCHSD1, UBE2I, KCTD18, VPS11, ATP9B, USP6NL, IKBK, APPL2, ATP2C1, RELA, ADPRHL2, GDI1, ING3, H2-AB1, GNL3L, CCDC93, SPICE1, KDM2A, RNF114, SART1, TTC14, HIST1H1D, SELENOO, MAPK8IP3, ZFP654, TUBGCP6, TMEM71, TAB2, SNX12, CLCN7, ANP32A, ZFP429, CAPZA2, NFKBIE, MTRF1L, SMARCE1, WASHC1, RIPK2, ZFP809, SMC3, AKAP11, RUFY2, COQ8A, FAM32A, PIAS1, AP5M1, 2410004B18RIK, ZFP455, MED17, ZFP383, CXORF57

---

**Table 6-24** | Genes upregulated in proteome data, but show inconsistencies in transcriptome.

---

CYP51, ZWILCH, CKAP2, TIMELESS, WDHD1, LIG1, FANCD2, HMGCS1, CIP2A, KIF15, LARS, SMC4, HELLS, UHRF1, CHAF1B, MTBP, SMC2, KNL1, ERCC6L, TFRC, ACACA, TOP2A, SLC38A2, BARD1, BLM, MCM4, NUSAP1, HK2, MCM5, IMPDH2, SHMT2, THADA, RBL1, NAA25, POLE, HEATR3, MCM2, FASN, DDX20, SQLE, FAR1, INCENP, TTF2, ACSL3, EZH2, CTPS, HMGR, FANCI, DSN1, YBX3, MCM7, TUBB6, POLA1, RRP1B, KLHDC4, TACC3, ATAD5, RRM1, AEN, SRPK1, ANKRD52, MTRR, HAT1, AACS, IFIH1, CD36, AFG3L1, ARL6, ERI1, PITRM1, PDCD11, KNOPI, NFKBIZ, KSR1, MGAT4A, MVK, PITPNB, TOPBP1, SHMT1, GM48551, CHD7, IPO7, MCM6, ZFP281, NSA2, PASK, KTN1, RIF1, HERC1, OASL1, PRIM1, MCM3, CSDE1, DNMT1, PAICS, PRMT5, CD86, KIF20B, CLASPI, RFC4, DUT, LIAS, SPDL1, RSL1D1, IPO11, DENND4A, FPGS, PHF19, RACGAP1, CD69, HEATR1, ABCE1, AKAP1, UTP6, MDN1, CLUH, PFKP, USP36, ERCC6L2, EIF4A1, MIPEP, NOB1, NAA15, MTHFD1, LRRC42, MYBBP1A, IDI1, SLC16A1, EXT2, KIFC5B, NUF2, MSTO1, FOXK2, MAN1A1, MPP6, CHTF18, GFM1, ANKRD17, LSG1, GCN1L1, MPP7, EDRF1, RILPL2, NOA1, FASTKD5, PDE3B, RPF2, UTP14A, ZNHIT3, NLE1, FEM1B, FAM111A, AEBP2, MCM9, MMS22L, RIOK1, PIP5K1A, GMI7296, LCMT2, FANCA, ECM29, TTC37, ABCB1B, UTP20, NAT10, GTF2E1, CDC27, PMVK, NOP2, DUS1L, YBX1, LDLR, NOL2L, IFT80, RFC2, ABCC4, GEMIN4, TRAP1, ZW10, FTSJ3, PSMC3IP, ASC2, PDCD2L, DZIP3, POLR1A, RPS3A1, IMP3, NOL10, NCAPD3, BDH1, BIRC6, TTC27, TRAF3, ANKRD16, GTF2E2, DDX3Y, MTHFR, MRPS31, POLA2, TIGAR, URGCP, LARP4B, ZFP330, HELZ, DNA2, RPL7, RPLP0, TBCE, HELQ, MPHOSPH10, CDCA7, PHF10, WDR3, DDX27, FKBP4, SUV39H1, PMS1, WDR36, EIF3D, NOP14, G3BP1, KBTBD8, OTUD6B, RIOK2, MYO19, XPO5, EEF2KMT, RWDD4, ABCF1, UPF1, AHSAA, RPS5, NDC80, DAP3, NKRF, ZFP598, PWP2, NOL6, CDH17, PABPC1, TMEM126A, RPL23, EIF2B2, EIF3B, DICER1, DHRS13, DALRD3, BMS1, RPS2, DHX33, IKBKAP, EEF2, GEMIN2, CBX4, CCT5, USP33, DUSP12, MSH6, HSD17B7, GLIPR1, DNAJC21, RPS12, SLC39A10, RBMS2, NSUN2, GALNT7, RPL4, MVD, WDR6, EIF2B5, SAC3D1, NAF1, RPS17, ASCC3, CUL2, SLC29A1, LGALS9, BAG6, TRIP4, BZW2, ZC3H7A, ZFP280B, NDC1, MLH1, ATP13A3, UBR5, SLC12A2, CLSPN, NMT2, DIEXF, UTP4, COLGALT1, HUWE1, RPS7, MRPL2, TCP1, GM29394, CCT4, ALKBH8, CD44, KDM4C, EIF2S3X, SRGAP2, CENPK, SCPEP1, CEBPD, POLR1B, EIF3G, TMLHE, NUFIP2, EIF4G2, PREP, SLC4A8, TONSL, PDSS2, RFC5, PELO, TBL3, RABGAP1L, DDX18, RTEL1, USP45, MRPS9, AMFR, BBC3, WDR74, DHRSX, METAP1, GRPEL2, R3HDM4, MRM3, NOL11, FABP5, NAA16, PEX13, FCF1, HNF1B, MRPL14, CDK5RAP1, MIS18BP1, UBR2, CCT8, WDR43, ESF1, XPO4, BORA, RPS19, CAR13, MRPL3, SLC25A16, YARS2, RRN3, NOC3L, TRDMT1, DHX30, ORC2, CHAF1A, BOP1, POLRMT, PPP5C, UQCC1, NASP, FAM207A, HSPA14, RAD51B, PER1, NUFIP1, RRP7A, DGK, DROSHA, FIGNL1, NCLN, DUS4L, NUP107, YMEIL1, ANAPC7, DONSON, NUDC, UBE2O, MRTO4, NVL, TRAF1, MRPS35, DHODH, TIPIN, NIP7, RIOK3, MAK16, AMMECR1, SLC25A33, SREBF2, FASTKD1, EIF2AK2, POLR1D, BAG2, KDM4A, PIH1D1, HSPH1, TTP2, DHX37, TIMM44, DCAF13, EIF3E, ZRANB3, GSTCD, DDX21, ASB6, URB1, SAMS1, HSPA9, EMG1, AIMP1, NUP88, RRP12, HELZ2, RPS3, NUP98, ERO1LB, MJMD6, WDR75, MDM4, DDX52, LDB1, RPL24, TRMT2B, NUP93, TUBE1, PIN1, LRWD1, CIT, ATG2B, WDR46, PUM1, DNAJC7, MRPL47, EEF1G, EEF1A1, PRMT3, DNAJC2, NSUN4, KARS, PSMG2, STEAP3, PNPLA6, UTP11, SDHAF2, RPS4X, TRMT2A, NEMF, NET1, MRPS22, RBBP7, WDR90, RAD51AP1, ORC3, MRPS6, MTERF3, IMPACT, SASS6, TDPI, NUP214, IPO9, FCRL5, BATF, USP16, CCT2, MRPL9, CD19, SCD1, RACK1, MSMO1, USP10, NMD3, GET4, MRPS27, BAZ1A, RAD51C, UBR7, RPL18A, MS4A6C, ACAT2, SMARCAD1, NAA40, TRAF6, DDX51, PNO1, ZFP280C, GINS4, ATXN2L, UTP15, YTHDC2, MAP4K4, MINDY3, GBP3, LMO7, IL2RA, SERBP1, ORF11, IFIT1, RB1, ELP2, GTPBP10, RCL1, YDJC, PPFIBP1, MRPS18B, IMP4, DHX36, DAPK2, ERAL1, CASP8AP2, RANGAP1, PRRC2B, JADE3, SLC2A6, MILR1, COX10, NOL8, PINX1, DDX6, PGPEP1, SKIV2L, PTCD3, MASTL, TMCC3, STARD4, BCAR3, TEO2, EPM2AIP1, WDR77, SENP3, PTTG1, MGMT2, ZCCHC7, EPB41L5, BTAF1, PPP2R1B, DDX24, MARF1, SMYD5, ATIC, RPL7A, ATRIP, COX15, SUGT1, UTP18, AICDA, TTI1, PWP1, CDK6, RPAP2, TDRD3, EIF4ENIF1, ERGIC1, ANKMY2, NBAS, FAM162A, QARS, RPS9, PMM1, ABT1, DDX10, SIK3, GRWD1, FLVCR1, EIF3I, GEMIN8, SLC25A19, DDX56, ETF1, NMT1, CERS5, RPL6, NSUN5, NOTCH1, RPL5, PUS10, GPHN, GXYL1, AK6, FAU, PNP2, FASTKD2, RPS14, FDFT1, MRPS30, GPCPD1, AMIGO2, TRMT11, NGDN, NCDN, HECTD1, SIMC1, NEDD4, TMEM154, PSMG3, INTS6L, PSMC6, RPF1, PSME4, 4931406P16RIK, KTI12, EIF1AX, CLUAP1, UTP14B, PPIA1, SLC7A6OS, EIF5, SMYD3, CAPRIN1, SECISBP2, TTC13, SEMA7A, SLAMF6, MRPL44, ZC3H15, TROVE2, KPNA6, RPS20, RPL3, RRS1, ATXN7L3B, SYNCRIP, METTL13, NAA50, TUBA1B, NAMPT, NFKB2, USP28, ARMC6, DNMT3B, RANBP2, LTN1, CCDC124, MRPS16, ORC6, CACYBP, MEMO1, ATP6V0A1, CNOT6, LIN7C, IL21R, ATR, UBAP2, NAP1L1, SLC4A7, PSM2D, EIF3K, PGAM1, STRAP, BCL2A1, USP37, ARHGAP19, JAK3, RETREG1, 1110065P20RIK, CDC123, UBE2Z, APEX2, NAA35, UBAP2L, QTRT2, ERCC6, RPL17, FANCG, USP34, HPRT, SLC7A6, NSL1, GLE1, CPEB3, NOC4L, LSM14B, ASB3, MRPL45, HEG1, RPL27A, CHORDC1, ABCCI1, VHL, COG1, CD40, RPL21, ZGRF1, PPID, CELF1, MGEA5, TOP3B, NOM1, REEP4, RPL10A, EIF3L, SPATA5, SOAT1, LSM12, MRPL1, MTMR14, GGNBP2, TPGS1, MARS2, RPL9, SUZ12, AAAS, ACOT7, APOBEC3, MACO1, HSPA4, PTAR1, FBXO3, GM9833, LARP2, GPATCH4, ZBTB32, HCCS, TCF25, TRIM56, PUS7, CUL1, FMR1, ATL2, RPS11, KYAT3, HSPBP1, PHLPP1, MUTYH, METTL16, FAM129C, TRIP11, EMC8, SCAP, MRPL32, GPN3, WEE1, TUBA1A, UXS1, TBRG4, MRPL30, TEFM, 4932438A13RIK, ATRN, ABCF3, PDF, ZEB2, LRPPRC, TNRC6A, PIMI,

---

**Table 6-24** | Genes **upregulated** in proteome data, but show inconsistencies in transcriptome. (cont.)

---

ANAPC1, USP14, ADAM9, DRG1, MRPS10, EEF1B2, RPS16, RPL10, APPBP2, TMEM131, RPL38, KLHDC2, PAIP2, FXR2, GM21987, CUL4B, MBNL3, FASTKD3, RPL36, HSPD1, MKLN1, RAD18, MTHFD1L, FAM117B, ANAPC10, AATF, TANGO6, PGS1, EIF4E2, GOLM1, ZBTB10, PDCD2, COG5, URB2, STAU2, PES1, LONP2, GM49333, NUP160, TRPC4AP, GAA, BZW1, LRRC57, ALKBH4, PMPCA, BTF3, ORC5, DLG1, KPNB1, HSP90AA1, NUP188, COX11, OSBPL8, RPS6KB1, EIF3H, IAP, TP53, RIC1, MNS1, FOXM1, CRY1, ABCD3, PRAG1, PUM3, SRFBP1, JUNB, RPS10, LY9, KPNA3, DNAJC10, ELP6, MTFMT, RRAD, INSR, UBE2E1, PHIP, ZCCHC4, ALG11, NUP155, RWDD1, MRPL49, SLC19A1, RPS6, NSD1, RPL30, MTA9, DMD, R3HDM1, NFX1, ZMAT3, CPPED1, VCPKMT, RMI2, RBM28, KDM4B, MCM8, MARK3, XRCC1, VEZT, NUP85, RPL32, TFB2M, ARL5C, DYM, POLR3B, PPAH, GTF3A, HMBX1, HAUS6, TWNK, ZFP622, RFTN1, NOP53, TNPO3, WDR4, ANKHD1, TRMT1L, ITPK1, TTC39B, ZFP277, HERC2, IRS2, SDC4, TRUB2, COG2, MSH2, LYRM7, NQO1, GTF2I, SMYD4, TSEN15, UTP3, QSOX2, HERC6, CYB5B, MTF2, PTBP2, MRPL42, NOLC1, TAOK1, CDCA7L, RPA1, METTL22, LIG4, HSP90AB1, ORC4, GPN1, SLC9A8, NAA10, RALB, EIF5A, MRPL23, TCF12, COX18, APOBEC1, CCDC97, TMEM242, FARSA, RPSA, ANAPC4, MCMBP, CCNH, CNPT1, CENPN, GAPVD1, MRPS25, ZNHIT6, MCTS1, DDX49, TYW1, SIPA1L1, POLD3, USP47, METTL2, TRMT1, GNE, HIVEP1, GEN1, RPL27, ZC3HAV1, PTC1, RTTN, RPS8, EIF3M, MRPL19, BCLAF1, MRPL10, RPL12, MIA2, HAUS7, GTPBP4, DOT1L, PROSER1, ATAD2, NUP205, SLC25A28, STOML2, SMARCC1, RPL13A, CD81, LTV1, STIP1, GLMN, E2F4, CSNK1D, OLA1, TLR7, CAND2, FANCL, ATAD3A, RBM3, TPRN, RNFT1, OXAIL, FCGR2B, TXNDC16, ZDHHC5, CCDC88B, NBEAL2, ATP1B3, ZDHHC20, BC027231, IVNS1ABP, PELP1, TSR3, SETD6, RPL31, DOCK9, ACBD6, KPNA4, TLK2, RPS13, SLF1, MRPL38, CNOT4, UBE3A, KIN, POLR3A, PIK3C2A, ATPAF2, CSNK1G1, NUP133, CCDC69, SURF2, POR, NACA, EEF1E1, NEU1, HAUS5, IGSF8, CCDC6, RPL14, GNL2, ANAPC2, QRSL1, PDP2, ZFC3H1, UGDH, SEH1L, PPP6R3, USP15, KLHL11, POLE4, AGPS, PTPN9, NAA11, IP08, FANCB, UBN2, RNF219, POP4, GNB1L, SPECC1L, PSMD3, IDH3G, RUVBL1, PHAX, HOMER1, FAM193A, FAM208B, ARMC5, POLR1E, ANKRD39, AKAP9, NUP50, HSD17B12, ACLY, MRPL41, FARSA, EHD2, MTA1, ELL, CLCN6, G3BP2, TGS1, MRPL39, GRPEL1, AGGF1, FAN1, RPS27, RRP1, CEP72, RRAS2, ZMYM1, CNOT1, RPL23A, RPL7L1, DCAKD, FBXL15, CENPH, NKAP, CD3EAP, ATP7A, CCT6A, LATS1, NXPE3, CSE1L, SLC39A6, AGPAT4, 2700097009RIK, RPAIN, UBE2E1, ULK1, SDAD1, DENR, WDR48, EDC4, GNM1L, UBE2Q2, RNASEH2B, NDUFAF7, GNAI3, PHAX, RNF168, RNF126, TTI2, FARSA, USP4, NUP62, DNAJC11, INTS5, PARL, HYPK, ENY2, DDX31, RPS15A, OPTN, MRPL24, PEX16, DTWD2, CDC16, TACO1, ADSL, SLC35A4, GOPC, AARS2, PFDN1, DCAF1, CTNNA1, PSMA2, PBDC1, DDX41, RANBP9, SMPD4, EMD, UBE4A, ALAS1, RPL8, HGH1, PSMG4, UTY, YOD1, USE1, APRT, BEND3, CCNT1, TNPO1, ING2, CNOT8, DDHD2, COX7A2L, PUS1, DAB2IP, LMO2, SRSF6, GADD45GIP1, EBNA1BP2, RPS21, CCDC86, VRK2, DYNC1L2, C1GALT1, NUDT1, CDK8, LIN37, GM15800, ZFP36L1, RAE1, RMND5A, RPS24, SNUPN, HDAC6, TTC4, PSMB8, RPL35A, ZNFX1, UPF3B, WDR12, GABARAPL2, PWWP2A, KLHDC3, ATXN10, RSL24D1, ZNHIT2, NAPSA, COA3, OTUD5, RPAP1, ARID5A, CASP4, DHX29, STEAP4, LINS1, GEMIN7, STX18, ODR4, DNAJA1, DNAAF5, POMT1, XRN1, ELOVL5, TOP3A, CCNL2, MAPRE2, AKT1S1, CDK2, RPS15, EIF2S3Y, RGP1, UBR4, PTPMT1, PAPD5, SDHA, KCTD10, RBFA, SKIV2L2, IMPAD1, 28104281I5RIK, PAK1IP1, GNL1, PIDD1, TAF15, INTS7, DNMT3A, KPTN, GBP2, KRR1, EMB, RBM19, SLC25A13, CSRP2, SEPSECS, FBXW4, GTPBP1, METTL18, GYPC, MPP1, COQ8B, INTS11, TPT1, SH3GL1, HAUS3, ECSIT, RPL19, SLC25A4, ZDHHC21, CYP4A32, NOL9, PGAM5, RB1CC1, TIA1, KIFAP3, KLHL25, GTPBP1L1, CSNK2A2, RABGGTB, KIF1BP, TTC9C, SLC35A3, OSGEPL1, MED28, ADSS, CBX2, CDC23, BRAF, ATXN2, EXD2, ADAM19, RFC1, CAPN7, IL2RG, BRX1, ZPR1, STX12, 2310035C23RIK, PPP2R5D, MRPS21, GM38394, FUBP3, TMEM39B, ST3GAL6, ZDHHC13, RNF26, ZDHHC17, EMC10, NFU1, EXOC3, NUP37, CHD1L, SPPL2A, DUS2, ZC3H12D, METTL5, DCTPP1, RPS18, USP42, ZFYVE16, COPS5, DGCR8, CENPM, ARPP19, HNRNPA1, TMUB2, NARF, NLRC5, PALB2, ATP2B1, RCC1L, TOMM40, RBM34, CCDC18, FKBP3, SUMF1, UBE2A, ARL6IP1, DDX28, CBX5, AASDHPT, CNOT10, CDK12, RAD50, CHFR, STRADA, CFP, SAP30BP, PRRC2C, ADAR, SRP9, SWAP70, GORASP1, RPL15, SLIRP, THUMPD3, MAN2B2, COG8, CYBB, PATL1, MRPL43, PEX10, GGTA1, H2-Q8, IFT172, PARG9, NPAT, LACCI1, DOPEY1, TSTD2, FNTA, GK, FAM78A, DCP1A, SRRD, MAPK8, RRP9, MGME1, MPC2, MOSPD2, MTG2, USP1, RPS28, TRAF4, GTF3C5, PARVG, DCAF4, POLR3E, TSPYL1, TMX2, URII, LZTR1, USP20, STUB1, TMEM161A, 38412, NEURL4, TOPORS, NPLOC4, RAB9A, GSTT3, CLK2, RPS19BP1, KLHL7, ZC3H8, PARP4, RRP8, ARL8B, GPAT4, GPS1, TRMT61A, TCF3, CTNND1, RICTOR, STMN1, MRPL11, ALKBH1, NT5DC3, SLC20A2, THG1L, MED30, H2-D1, RAB3GAP2, NOP16, DPH2, RAN, GRCC10, ADNP, ZFAND5, EXOSC10, TEX10, AP4B1, CCDC84, ZBED3, STYX, PRPS1, CNOT2, DHX58, MLLT10, TNPO2, SSSCA1, EED, ELAC2, PGK1, UBA3, DARS2, PCMI, KLHL20, CD83, NBN, ARID3B, DNAJC25, TOE1, TRIM25, POM121, PCGF6, PDE7A, RPLP1, WNK1, MTERF1B, FXR1, NPM3, EXOC4, GTF2H2, RPL35, CDC7, UXT, NAA30, AP4E1, MAP3K5, ABCC5, PEF1, SENP6, MTOR, LDHA, TSEN54, ZFP644, LARP7, FOXJ3, CEP85, ZBTB11, ZFP142, OIP5, GLRX3, GTF2H1, UBE2D3, DNMI1, TPST1, PMPCB, SS18L2, PCED1B, MRE11A, POMGN1, COIL, PTPRA, ABHD13, NOLCH2, ZCCH3, RASGRP3, CLK3, IP6K1, EFL1, H2-AA, POFUT1, TSG101, AHCTF1, AMMECR1L, TIMM10, ENTPD6, DPH1, PARP14, RASA4, RBM33, DDT, CNM4, YIPF4, SREK1, ECD, UHRF1BP1, DTX3L, HAUS4, RAB3GAP1, RAB3IP, DNAJA2, SRSF3, ELOA, RC3H1, 2510039O18RIK, PPP2R1A, RPLP2, HM13, RSPRY1, SPTLC2, DDX11, SNRPB2, EDC3, EPHA2, FAF1, HDAC9, ATM, ARFGAP2, SUCO, PPP2CA, BID, KATNA1, HMG20B, CHD1, RCOR3

---

**Table 6-25** | Genes **downregulated** in proteome data, but show inconsistencies in transcriptome. (cont.)

---

ALB, CPT1A, SFXN3, FGB, CROCC, EPHX1, ELMO2, SELENBP1, ANXA1, CDKN1B, ALDH2, ZYX, KCTD12, ALOX12, KBTBD11, SRPK3, NUDT16, FGG, LMNB2, AIF1, PGAP1, FCRLA, CFAP43, FGA, BLVRB, IFI209, SIGIRR, HIDE1, GP1BB, TTC38, HS1BP3, VWF, TGM2, MYL4, PLD4, DFFB, B4GALNT1, LSP1, GBP9, ALDH6A1, TUBB1, PARP3, CASP6, SFXN2, PLGRKT, GPD2, HAAO, LMNA, FSCN1, HIST1H1T, CCDC71, FBXO22, VCL, ACTN1, ITGA2B, PLEC, GP9, CRYL1, P2RX4, ACP6, ANXA5, HDAC10, KLC4, ASPH, GSTM1, ABR, C4B, GPD1L, FYB, AS3MT, CALHM2, GBE1, ZFP512B, LIPE, PARVB, PAXX, PLBD1, FHOD1, EML3, TMX4, TRP53I11, SH3KBP1, CDT1, DGUOK, MYL9, HMGA1, ACAA2, DFFA, NAGK, PLEKHA1, CEP95, CST3, ITGB7, DNTTIP1, VRK3, CGGBP1, THBS1, TLR9, MLYCD, CALHM6, ENGASE, CIRBP, MAP2K6, ARHGAP6, GIMAP9, MYOF, PGGHG, SH2D3C, ALDH1B1, TREML1, STOM, CBR1, SH3BP1, NFKBIB, DHDH, NPEPL1, PRKAR1A, PACSIN1, TNFSF13B, TMCC1, ANXA2, RGS14, ACSS1, ASAP1, HSPA1B, AHDC1, IRAK1, PPBP, PRKAR2B, DCXR, UVSSA, IVD, TLR3, GLYCTK, TSC22D4, DDX58, SMARCC2, AK3, CTSH, ECI1, STARD10, CAPN1, NQO2, SLC25A11, SERPINB9, RFX1, SELENOH, SORD, DCLK2, PLS1, ZMYND11, NUDT14, CHTF8, ARHGEF11, PRR12, DIABLO, PDE5A, GP5, PCTP, ABCG3, SYNE2, GALK2, DAPK3, PPOX, PPP1R9B, HMOX1, HSPA4L, HIVEP2, GAS7, MFN2, ANPEP, PFKFB4, MCCC2, TDRD7, SMAD5, RAPIGAP2, CAPNS1, CUTC, BRCC3, HDGF, LALS3, ALDH4A1, IBA57, ADSSL1, CNN2, MCCC1, PAFAH1B3, CARD6, MYO18A, TEC, CRAT, CEP128, CBX1, TBC1D9B, RFXAP, HBB-BT, XPNPEP3, ZGPAT, RTF1, AAMDC, UGP2, DOK3, ENTPD5, ABHD16A, FBXO7, EAR1, GMFG, CIC, GIMAP8, MDPI, FAM98C, MYO1F, GGCT, LMNB1, OXSM, SYNJ1, CNP,

---

**Table 6-25** | Genes **downregulated** in proteome data, but show inconsistencies in transcriptome. (cont.)

---

PLAUR, NLRX1, TBXAS1, IDNK, FAM50A, HIFX, TNKS1BP1, IL16, BASP1, SH3GLB2, NDUFV3, TNIK, PYGL, MBD2, RPRD1B, ITGB3, PF4, ANXA6, LYZZ, APOBR, CEP78, TBCK, MICU2, GPAT3, MSRA, GRAMD4, NAXE, SP2, PML, TBC1D14, PGM1, DOCK4, GPX1, TWF2, LRSAM1, SHTN1, METTL7A1, CLU, HBA-A1, ABHD14B, PDXK, TLE3, CAR1, CORO1B, PTPRE, PRDX5, MPO, SHPK, KDM1B, KYAT1, OTUD7B, VWA5A, PPM1M, INPP4A, RMND1, OSBPL5, VPS8, MAPK12, CLCA3A1, HSPA2, BC017158, IFI203, CCDC90B, HIST1H1A, CIZ1, TAX1BP1, POT1B, WDR91, MEF2D, CNST, EML2, NLRC4, HCF1, WDR81, KMT5C, RNF169, AGPAT1, RIDA, GDPGP1, DBNL, HSD17B10, CDK5, TBC1D20, SETD1A, STK19, POLR2G, IPCEF1, VTI1B, SERPINB6B, PARP1, PPT1, FCGRT, ALYREF2, NFIC, CBX8, FYCO1, APOE, IFI47, AKR1B3, BCR, BACH1, ECHS1, EEA1, AKAP2, RREB1, SERPINB6A, CDYL2, SYTL4, CAMP, CPT2, FGD2, CEP250, SPTB, NSF, PANK2, MARCKS, EMSY, UAP1L1, VPS33A, TMPPE, SLC4A1, DNM2, UPF3A, RAB11B, UBXN6, 9030617003RIK, ALDH3A2, MON1A, DPP7, SQOR, TPD52, MTF2, ITIH4, GNS, BPHL, NFIA, TRIM36, PFKFB2, BCL9, VPS33B, IQGAP2, TIPRL, ALDH3B1, PCCA, A430005L14RIK, ENDOG, OXR1, ECH1, TAF6, PSMB10, MPEG1, GIMAP1, SNAPIN, ACP5, SMUG1, FLYWCH1, ANXA4, EPN1, NAT2, ACSF2, SAFB2, FCER1G, TNK2, RAB6B, EPB41L3, TNIP2, CD97, LRP1, MDH1, PDK1, DCPS, D6WSU163E, H2-KE6, MFN1, GSE1, PTK2B, GSDMD, TPRGL, BUD13, PBX2, GM340, TOR1AIP1, GMEB1, TRIOBP, TALDO1, 2310033P09RIK, MACROD1, MSN, SLC40A1, SCLY, ZFP668, NEDD4L, VCAM1, F13A1, HADH, SLC25A20, BRMS1L, FDXR, TCEA1, ZMIZ1, SIRPA, WBP11, TUT4, RAB31L1, CNTROB, PAG1, GSTK1, SCRN3, ABCB10, MTMR12, CARD9, CENPV, RAB27B, RNASEL, PPM1F, PDIA5, PCCB, DGKZ, DENND1B, COMMD7, HMGB2, TPCN1, EPS15L1, CHIL3, FAM213A, SCAF1, HK1, NMNAT1, ARL6IP4, BRD9, SLC25A12, EMILIN1, COMMD5, LCP2, PAXBP1, TKT, PIK3C3, TGFB1, ETFDH, MAPK3, ELK3, EPX, RALA, THA1, GP1BA, YWHAQ, TMEM201, ASAH1, TRIM59, NCF4, PTGR2, SLFN5, TBC1D8, AKAP12, NIT1, PLEKHG2, NCOR2, MRC1, MSANTD2, MGMT, HBP1, PANK4, HIST1H2BK, PMM2, OTUB1, SERPINA3K, MFF, GOLGA7, CBR3, BC026585, SGPL1, TEX9, PLA2G4A, AAK1, PKNOX1, HMGB3, TIFA, SLC25A46, HPCAL1, NGP, RFLNB, EIF4EBP2, SYNE1, ITGAD, CISD3, HIST1H1B, DRAP1, PPP2R5A, ASPSCR1, AGAP2, PTGS1, STAG1, TALPID3, ACOX3, ACAT1, SLAMF1, SSH3, CAMKK2, SH3BGR2, TRAPCC5, DAXX, CRK, MSL3, ERLIN2, USP11, ABCD1, FUK, SLC41A3, RTCA, APEH, SLC9A9, ARMT1, PSME1, MTA3, SAP130, TGTP1, KCTD14, RESF1, SAMD1, ACAD10, SNX4, ANXA3, HAGH, RING1, HOOK3, SNX18, HOMER3, AI607873, 5031439G07RIK, HSF1, ZBP1, FECH, RDH13, PCYOX1, KIAA0226L, DDAH2, NRIF1, RENBP, 1600014C10RIK, PHRF1, CPQ, APIM1, ZMYND8, CD38, ECI2, EARS2, COP57B, LANCL1, RNF170, NHEJ1, NCF2, DTD1, DUSP28, RNF6, EHD3, PPP3CB, ZFP629, DENND1A, MTFR1L, RAB28, FOXO3, MOB2, STIM1, PDCC6, WDR70, PPCDC, NEMPI1, PPA2, TMEM175, MTR, GM4737, TARDBP, PRCP, DCP1B, PUF60, DUSP7, SETDB2, POLL, PPP1R11, MCUR1, DCTN1, HLCS, EIV5L, HPS1, ACYP1, SELENOT, NAGP, ACADM, 9930111J21RIK2, POC1B, ZC3H3, GPX4, SRI, ACAD12, NCOA6, ARHGDA, PIP4K2C, BAK1, FAHD2, NFYA, HMCES, UBE2N, PALM, SLFN1, LTBP1, UPRT, NME3, MAST2, HP, ESYT2, TRAPCC9, RAPGEF4, CNDP2, HNRNPK, CASP1, SBF1, SUDS3, IFI35, DYNLT3, WDR45B, SIGLEC1, GNPDA1, CCSAP, ANKRD27, VIPAS39, SERPINA1B, HDHD5, HNRNPUL1, BAG5, ZFP608, PPP1R13B, TBC1D8B, TTC28, RFTN2, PPP6R2, PDE2A, CBX3, ASRGL1, HIF0, NCOA3, CCDC71L, CYC1, TRIM24, AKR1B10, IGSF5, MGST1, ETV3, TUBA4A, PYGM, ARHGAP18, PIRB, GIMAP7, BPNT1, EPS8, FN1, INPPL1, COMMD1, NAGA, GM20498, RPS6KB2, SMARCB1, GGA1, HK3, BAG1, KMT2E, RAB11FIP1, ACTB, USF2, M6PR, LTF, GM43302, TGM1, IPO13, SRSF4, ITGA6, PURA, TUT7, RGS10, YWHAB, NCEH1, D10JHU81E, BLES03, LK2, CAVIN2, ARPC5, PPM1D, BORCS5, POC5, CTSA, STX8, SLC27A4, COX7A2, PRCC, STXBP3, MYH14, SF3A3, TCN2, CKB, CTDP1, PTPA, TBL1XR1, RALY, CHMP1A, OFD1, DOCK5, LIPT1, COQ5, UQCRI10, ITGAX, SMPD3, PPM1E, CLEC16A, RSU1, B3GNT8, BECN1, SMG6, NFS1, NMI, PTGES2, HEMGN, ACSS2, CCNK, ATG3, ATG5, BLOC1S1, TST, DBN1, ZFP691, VPS35, PAK1, DDX17, FBXO4, RAB27A, TSPAN13, CASP7, RBCK1, ZFP553, RIC8A, 37316, RBM14, ZBTB2, CRIP2, WASHC5, ANTXR2, HDAC5, NUDT18, GIMAP5, NADK2, VAT1, THY1, OSBPL7, B2M, SCNM1, NRBP1, IMPA1, TUBA8, CLIC1, HEXB, FBXO28, ACAD8, TMEM65, PTPN23, NFYC, PSMB9, BCL9L, BCL2L13, KMT2D, 3110001I22RIK, COX411, DCAF11, FAM151B, SCIN, GTPBP3, ZBTB9, STARD3, PLA2G15, GNPDA2, ERCC4, NUDT13, ABHD12, TAF1, 4933427D14RIK, PIK3CA, NUCB2, TXNRD2, VPS4A, SNRPA, FMNL2, FBXL4, C3, MOCS1, VPS4B, CEP57, MAPKAPK2, RFXANK, SSBP1, CEP41, SLC25A45, 6330416G13RIK, PRG2, RCHY1, PRKACA, S100A9, NT5DC1, VAMP8, SMC1A, MEN1, AHR, EXOG, RAD21, SEPHS2, SF3B4, DXO, ACADS, ANKS1, DEF8, DD12, NPEPS1, TMX3, WDR13, 1810043G02RIK, ACTL6A, RABIF, VAV3, UNKL, NDUFS3, UROS, PSPC1, HMGXB3, COQ6, PDLIM5, PSME2, GLB1, SINHCAF, FBXL17, CSAD, SCAMP3, SNAP29, HNRNPL, MMAA, NDST1, OSBPL2, DUSP3, TP53RK, OGG1, CYTIP, MYO6, SRSF2, MYADM, FN3KRP, PBXIP1, HPS6, SUMO1, THNSL1, AKR7A5, HSCB, EPS8L1, HIF1AN, LIMA1, 9030624J02RIK, TAF6L, GLRX, HSPA12A, RASSF4, CSTF3, 1110004F10RIK, RUFY3, MRII, ISOC1, BC017643, MYO1C, NCBP2, GCC1, PDXP, GM49405, DGAT1, ZFP84, ZBTB1, GPANK1, ITGA2, TRAPCC12, SLC29A3, RAB1B, PCMT1, TCIRG1, VARS2, MTHFS, SAFB, CMTR1, FAAH, CRLF2, UBL5, LAGE3, EXOSC9, CLYBL, PARN, NIPSNAP2, FAM213B, ZCCHC17, NDUFA4, LYRM4, EAR6, SYNRG, COX6C, SRR, POLR2C, ETHE1, LENG1, PLAA, MAN2C1, CSTF1, APOOL, GLYR1, ACADSB, QRICH1, MYL12B, PLBD2, CXXC1, MYH10, FXYS5, UNC13D, HIPK1, ACSF3, PPIF, BRP1, PRKCD, ANKIB1, TRAPPC1, KIF1C, JPT1, TRAPPC11, SAYSD1, ATF7, N4BP1, MAPKAPK3, TMEM120A, SNX15, IGSF6, DLD, NAIP2, SETD5, CUTA, ZNF326, ACADVL, SLC4A2, UBE3B, RANBP3, NCK1, XDH, TICAM1, CFAP20, STRN4, GPKOW, PSMA3, RAI1, IDH2, LRCH4, MICAL3, HP1BP3, UQCRC2, DPY30, IRAK2, TRAPPC10, ATP6V0D1, FLOT1, SH3PXD2A, TAF5, PIK3R2, TXNIP, SLC27A1, DNAJC14, PRKRA, FADD, GOLGA4, TOMM5, MOB3A, PSIP1, TUT1, UNC50, KIF1B, IGHG2C, SMTN, COA4, ACAA1A, CCDC9, KDM6A, THYN1, PRPH, TMCO4, FBXO18, SVIL, ZKSCAN3, PQBP1, F5, NUDT6, GTF2A1, SARNP, PCBD2, PPM1K, CEP44, FIZ1, CCAR1, CARMIL1, SUCLG2, SLX4, NCKIPSD, MCAT, CAPN2, AI413582, DOK2, CALCOCO1, ZC3H18, FBXO6, RAB5B, NSFL1C, SUOX, AP3S1, DEAF1, COQ3, MAP2K4, CD177, SPTA1, PDE1B, PPM1G, DIP2A, ANAPC11, PRPF19, DCUN1D2, CACTIN, SRRT, BC037034, LRRC45, GZMA, LSM2, YWHAG, ZFP280D, CECR2, ELMO3, STK38L, BRD2, TAPBPL, HBS1L, RGS3, DNAJC28, HINT2, DSP, CYP4F13, MAGOH, IRF2BP2, PRDX2, ZFP41, VPS72, FAM104A, COQ7, SRA1, TCHP, MAP3K11, PLP2, NUMBL, ELMO1, BLOC1S3, DNAJC16, XXYL1, NSMAF, EPPK1, CYB5R3, SFXN5, TOMM40L, STAM2, PLRG1, SSFA2, CLIP2, BAP18, DTNB, LMBRD2, NDUFS4, OGFRL1, RNF146, FGD6, LLLGL1, HSDL2, PPP1R37, HDAC8, PIN4, C2CD3, FAM192A, THEM4, PPME1

---

**Table 6-26** | Genes predicted to be co-regulated with CREB3 and CREB3L2

Transcription Factor	Microarray	Additional genes identified by Multi-omics Analysis
CREB3L2	ITGB1; CLIC4; RAPH1; HDLBP; LAMC1; TRAM2; ALDH1L2; AFF1; TNS3; C11ORF24; ARHGEF12; ZBTB38; SSR3; FNDC3B; SLC39A14; SND1; PCYOX1; H6PD; EPAS1; ABHD2; SLC7A1; TM9SF3; ITPRIPL2; LIMA1; GANAB; PRRC1; PGRMC2; TMEM184B; ATP2B4; ELL2; DAP; CPD; ASXL2; PLXNB2; ALDH18A1; NFE2L1	KDM5C, MCFD2, B4GALT1, CHPF, KDELR1, PLOD3, PLOD1, GOLGA3, SEC61A1, TIMP2, CAPN2, SLC39A7, NOMO1, PDIA3, CKAP4, PDIA4, TMEM214, GORASP2, CANX, KDELR2, PLBD2, COPA, FAM114A1, SEL1L, RRBP1, ATP1A1, TM9SF4, FAM129B, HSP90B1, LMAN1, LARPI, FAM120A, MLEC, SSR1, FLNB, UGGT1, SEC31A, SEC24A, HSPA5, WFS1, GALNT2, EDEM1, SURF4, GBF1, SYVN1, GNG12, ARCN1, EXT1, GOLGB1, CALU, HYOU1, CALR, COPG1, P4HB
CREB3	ARF4, PSMD8, MYDGF, SEC61A1, NSDHL, LGALS1, TMEM147, GUK1, TMED3, ARFIP2, TMED2, NANS, PSENE1, RAB2A, SEC13, ANXA5, SARS, FTSJ1, COPZ1, BSCL2, GORASP2, DAD1, KDELR2, AURKAIP1, ERGIC3, PHPT1, PAFAH1B3, CUTA, MAGEH1, CD63, COPB2, FKBP2, COPA, FAM114A1, MAGED1, GOSR2, DCTN3, ARL1, TULP3, SLC3A2, PLD3, PSMB6, PSMB7, ASNA1, TMEM205, EMC4, EMC7, ATP6V1D, SEC23B, AAMP, SEC31A, MDH1, YIPF3, SLC31A1, YIPF5, DSTN, YIF1A, TMCO1, PSMC4, MARS, CALU, MAGED2, COPG1, COPE, SRP54, CLTA, DUSP14, RABGEF1, TRMT112, MRPL40, PPME1, PELO, C11ORF24, ANXA2, GLRX3, NDUFS8, TMX2, AKIP1, CSNK2B, KDELR3, HM13, MFF, RABAC1, POLR2L, PRELID1, RMDN3, ATAT1, CYCS	TIMM17A, ZFPL1, TPI1, STARD3NL, DDOST, IFT43, MORF4L2, MMADHC, SAR1A, PSMD13, COPB1, RNF5, EIF2B4, ADRM1, MRPL28, MRPL21, TM9SF1, PSMC3, RHEB

**Table 6-27** | Top 500 Differentially regulated Genes/Proteins across species and platforms. Sorted by median adjusted  $p$ -value for multiple group comparison.

Rank	Symbol	mSpIPC .array	mBMPC .array	hPB .array	hPC .array	hBMPC .array	mPB .rSeq	mSpIPC .rSeq	mBMPC .rSeq	hPB. rSeq	hBMPC .rSeq	mPB .prot	mPB93 .prot	ANOVA Microarray Adj $p$ -value	ANOVA RNA-Seq Adj $p$ -value	ANOVA MS/MS Adj $p$ -value
1	ARHGEF18	-5.47	-2.85	-2.29	-2.49	-2.78	-4.91	-4.63	-4.47	-3.01	-6.2	-4.75	-4.14	1.90E-08	8.30E-08	2.60E-08
2	XBP1	4.33	4.56	4.28	4.63	4.57	5.45	5.46	5.67	4.17	7.33	1.38	1.8	1.90E-08	2.60E-08	9.60E-05
3	JCHAIN	10.19	3.08	3.37	3.26	2.18	11.23	12.18	12.62	7.56	8.17	1.8	1.58	3.40E-08	3.40E-08	1.30E-02
4	EDEM2	5.47	4.17	3.48	3.02	2.51	3.35	4.21	4.48	1.34	2.7	1.73	1.21	2.00E-08	3.50E-08	1.40E-03
5	LMAN1	3.92	4.04	3.4	3.04	2.12	3.97	3.91	4.07	2.55	3.73	2.43	2.16	1.90E-08	4.10E-08	7.00E-06
6	SLAMF7	4.21	4.86	9.3	10.36	11.63	3.04	4.46	4.54	7.05	9.06	1.68	1.43	1.90E-08	4.20E-08	2.40E-05
7	ELL2	3.55	3.04	2.27	2.02	1.89	4.51	4.25	4.28	4.27	4.73	2.83	2.18	3.70E-08	4.70E-08	1.30E-05
8	ALDH18A1	2.15	2.47	3.82	2.99	2.7	2.73	2.84	2.82	2.5	1.96	2.68	2.63	2.70E-08	9.90E-07	5.00E-08
9	NCAPG	8.06	3.44	6.61	2.98	3.42	5.78	3.15	2.83	7.16	0.23	2.86	3.16	1.90E-08	3.40E-06	5.00E-08
10	SLC44A2	-5.01	-4.41	-3.74	-3.12	-3.78	-1.92	-4.09	-4.99	-1.41	-3.26	-0.83	0	1.90E-08	5.00E-08	2.80E-01
11	SORL1	-9.41	-8.72	-6.75	-6.88	-6.47	-7.32	-6.76	-8.63	-3.31	-6.56	-4.38	-4.86	1.90E-08	5.10E-08	3.10E-07
12	PDIA6	3.1	2.79	2.74	2.19	1.93	3.76	2.63	2.44	3.89	4.66	1.25	1.05	1.90E-08	5.20E-08	3.20E-05
13	ERN1	4.98	5.24	2.97	3.99	4.47	4.31	5.15	4.71	5.13	5.23	2.63	2.41	3.10E-08	5.20E-08	1.00E-03
14	SDF2L1	6.24	5.36	6.95	6.18	6.5	4.22	4.16	4.14	3.36	5.52	1.56	1.79	2.00E-08	5.30E-08	1.80E-04
15	SEC24D	8.03	5.77	5.64	5.17	6.31	4.5	4.45	4.45	3.23	3.48	4.54	3.74	1.90E-08	5.50E-08	2.60E-07
16	SIGLECG	-3.1	-3.51	-6.05	-6.04	-5.27	-4.48	-4.42	-4.3	-2.5	-6.7	-1.37	-1.05	2.50E-08	5.50E-08	4.80E-05
17	FNDC3B	5.5	6.1	8.46	8.37	8.85	5.52	6.86	6.99	4.2	7.02	3.01	2.19	1.90E-08	5.50E-08	5.50E-05
18	IGHA	10.34	4.61	7.14	6.19	6.69	4.44	13.22	14.82	11.77	12.84	-5.44	-5.35	2.30E-08	5.50E-08	1.10E-04
19	TXNDC5	3.16	5.05	3.2	3.12	2.9	4.9	5.23	5.16	5.52	6.76	2.29	1.57	2.00E-08	5.90E-08	1.60E-04
20	CKAP4	5.06	4.55	5.78	5.63	5.11	3.78	4.06	4.12	4.43	5.83	1.59	0.37	5.50E-08	6.00E-08	4.70E-05
21	RIPOR2	-3.46	-3.26	-0.85	-4.95	-6.24	-3.25	-3.45	-3.19	-0.81	-8.89	-2.47	-2.12	2.00E-08	6.10E-08	1.50E-05
22	MS4A1	-3.91	-2.86	-9.56	-9.22	-8.59	-1.15	-3.56	-4.08	-7.28	-9.23	-0.32	0.18	1.90E-08	6.10E-08	6.60E-02
23	SERPINB1A	-8.98	-6.65	-2.99	-2.8	-2.49	-4.11	-6.41	-6.05	-4.09	-0.68	-3.5	-3.72	3.10E-08	3.90E-06	6.30E-08
24	FKBP2	3.51	3.26	3.87	3.71	4.17	3.71	4.04	4.07	2.28	5.14	1.17	0.88	2.10E-08	6.40E-08	2.00E-02
25	RCSL1	-4.53	-3.43	-5.51	-5.72	-8.4	-4.42	-4.19	-4.11	-1.97	-3.16	-2.41	-2.11	1.90E-08	7.20E-08	6.10E-06
26	EAF2	8.58	5.94	4.29	5.16	6.04	6.38	6.8	6.48	1.78	2.64	2.06	2.05	1.90E-08	7.20E-08	1.90E-05
27	IGHG2B	10.17	9.56	0.57	0.91	2.99	7.33	12	13.47	7.08	11.82	0.01	-0.08	2.00E-08	8.20E-08	9.50E-01
28	DNAJC3	3.86	3.89	2.99	1.4	1.28	3.08	3.34	3.38	2.76	3.79	1.66	1.53	2.00E-08	9.50E-08	8.20E-05

**Table 6-27** | Top 500 Differentially regulated Genes/Proteins across species and platforms. Sorted by median adjusted *p*-value for multiple group comparison. (cont.)

Rank	Symbol	mSpIPC .array	mBMPC .array	hPB .array	hPC .array	hBMPC .array	mPB .rSeq	mSpIPC .rSeq	mBMPC .rSeq	hPB. rSeq	hBMPC .rSeq	mPB .prot	mPB93 .prot	ANOVA Microarray Adj <i>p</i> -value	ANOVA RNA-Seq Adj <i>p</i> -value	ANOVA MS/MS Adj <i>p</i> -value
29	MANF	3.95	3.41	4.37	3.94	4.16	4.04	3.45	3.45	5.21	4.84	2.62	2.07	1.90E-08	9.90E-08	2.30E-07
30	ZFP318	-7.68	-6.6	-5.41	-5.49	-5.79	-5.4	-4.59	-5.56	-4.33	-3.89	-3.52	-3.27	1.90E-08	3.60E-07	1.00E-07
31	SIL1	3.2	3.85	4.17	4.12	3.31	3.51	4.02	4.02	2.8	4.04	1.5	1.65	2.70E-08	1.00E-07	1.60E-06
32	RASGRP2	-2.43	-1.69	-6.28	-6.93	-6.49	-4.44	-3.01	-2.87	-1.76	-6.47	-1.86	-1.78	1.00E-07	1.00E-07	5.40E-05
33	GLIPR1	3.83	3.52	1.95	2.02	-0.69	4.48	3.74	4.64	-2.96	-1.79	1.79	2.28	3.20E-08	1.10E-07	2.70E-05
34	DNAJB11	2.68	1.87	3.3	2.41	2.55	3.26	2.68	2.64	5.49	5.74	1.06	0.7	2.40E-08	1.10E-07	3.10E-04
35	ENTPD1	7.74	6.69	8.47	8	0.09	1.87	3.76	3.77	1.94	-4.35	1.35	0.28	1.90E-08	1.10E-07	3.80E-03
36	SLC44A1	3.73	3.98	9.84	9.76	8.3	3.2	4.5	4.49	2.3	2.7	1.48	0.5	1.90E-08	1.10E-07	1.30E-01
37	TOP2A	6.53	3.51	6.06	1.66	-1.27	4.6	1.25	1.27	1.92	-3.69	3.72	3.51	2.30E-08	7.60E-06	1.20E-07
38	CRELD2	4.48	4.07	3.89	3.25	3.3	4.93	5.09	4.9	1.74	3.74	2.23	1.73	1.90E-08	1.20E-07	1.30E-05
39	AFF3	-3.25	-2.2	-5.4	-5.4	-5.4	-4.6	-4.28	-3.99	-3.75	-9.68	-3.13	-2.43	1.00E-07	1.30E-07	1.50E-04
40	CHID1	4.12	5.05	4.85	4.71	5.48	2.94	3.55	3.56	1.68	3.27	0.28	-0.43	2.20E-08	1.30E-07	1.60E-01
41	EBF1	-6.61	-2.52	-4.9	-4.9	-4.9	-2.7	-4.23	-4.08	-5.86	-4.96	-0.14	-0.11	1.90E-08	1.30E-07	7.40E-01
42	SIPA1	-4.09	-3.36	-2.94	-2.74	-2.49	-3.02	-3.1	-3.15	-1.78	-3.56	-1.98	-2.01	6.70E-08	5.30E-07	1.40E-07
43	IGHD	-11.35	-6.3	4.85	5.12	6.04	-7.67	-5.33	-5.81	0.05	-4.83	-3.87	-4.61	1.90E-08	1.30E-06	1.40E-07
44	RRM2	8.53	5.31	10.07	7.19	3.35	4.65	2.25	2.09	9.95	0.7	4.33	4.22	1.90E-08	1.70E-06	1.40E-07
45	ABLIM1	-4.04	-1.74	-2.66	-3.86	-7.3	-1.87	-3.2	-2.89	-3.18	-3.32	-1.59	-1.31	2.10E-08	1.40E-07	4.00E-05
46	CD19	-4.68	-3.77	-2.16	-3.84	-2.83	-2.31	-3.88	-4.37	-2.5	-6.27	0.13	1.14	2.00E-08	1.40E-07	1.50E-04
47	RPN2	3.44	2.56	3.64	3.48	2.79	2.91	2.66	2.67	2.3	3.49	0.78	0.58	3.50E-08	1.40E-07	5.90E-03
48	DDOST	3.46	3.01	2.4	1.72	1.86	3.57	2.7	2.71	2.18	3.26	0.78	0.95	2.60E-08	1.40E-07	4.70E-02
49	REXO2	4.11	3.6	2.91	2.06	1.97	3.19	4.18	4.19	1.35	1.41	1.52	1.68	2.00E-08	1.50E-07	3.90E-06
50	ARHGAP17	-3.5	-2.52	-3.24	-3.35	-2.94	-3.16	-2.04	-2.11	-5.18	-3.84	-1.44	-0.85	2.00E-08	1.50E-07	2.80E-03
51	CYP51	1.67	0.32	1.74	0.9	-1.31	3.01	-0.27	-0.1	1.58	2.19	3.89	4.19	3.80E-07	1.60E-07	2.50E-08
52	PRDM1	7.9	7.97	10.36	10.24	9.78	8.97	8.59	8.39	8.14	7.82	5.17	4.64	1.90E-08	1.60E-07	2.50E-06
53	RIN3	-4.17	-3.86	-3.76	-3.8	-3.33	-4.32	-3.7	-5.05	0.19	-3.59	-2.96	-2.98	1.00E-07	3.80E-06	1.60E-07
54	CD55	-7.23	-5.21	-4.84	-5.41	-3.37	-4.24	-3.26	-3.14	-1.93	-1.47	-2.55	-2.23	1.90E-08	1.60E-07	1.60E-03
55	SPCS2	3.39	3.41	4.16	4.17	3.49	2.75	3.22	3.31	1.8	3.04	0.75	0.3	1.90E-08	1.60E-07	1.90E-02
56	CCPG1	2.73	3.19	4.79	5.1	6.01	2.97	3.31	3.7	1.58	4.52	0.01	0.62	2.00E-08	1.60E-07	3.90E-01

**Table 6-27** | Top 500 Differentially regulated Genes/Proteins across species and platforms. Sorted by median adjusted *p*-value for multiple group comparison. (cont.)

Rank	Symbol	mSpIPC .array	mBMPC .array	hPB .array	hPC .array	hBMPC .array	mPB .rSeq	mSpIPC .rSeq	mBMPC .rSeq	hPB .rSeq	hBMPC .rSeq	mPB .prot	mPB93 .prot	ANOVA Microarray Adj <i>p</i> -value	ANOVA RNA-Seq Adj <i>p</i> -value	ANOVA MS/MS Adj <i>p</i> -value
57	ELMSAN1	-3.55	-2.05	-2.36	-4.49	-0.75	-3.2	-2.19	-2.22	-0.43	-1.75	-1.94	-1.83	2.40E-08	3.40E-06	1.70E-07
58	PRDX4	3.87	4.39	5.66	5.65	6.09	5.54	5.25	5.38	3.96	6.1	1.69	1.35	1.90E-08	1.80E-07	1.30E-04
59	TXNDC11	4.09	3.29	2.58	2.31	1.34	3.5	3.75	3.77	3.91	3.96	3.09	2.65	2.00E-08	1.80E-07	1.30E-04
60	PXK	-3.88	-3.43	-2.3	-2.21	-5.35	-2	-3.22	-3.47	-2.97	-3.77	-0.97	-0.73	1.90E-08	1.80E-07	5.20E-03
61	SND1	2.8	2.97	2.52	2.27	1.15	2.62	2.53	2.49	2.6	2.27	1.46	1.43	3.30E-08	1.90E-07	1.20E-06
62	HSP90B1	3.13	3.42	2.51	2.3	2.31	4.08	4.03	3.94	3.64	4.53	1.45	1.26	1.90E-08	1.90E-07	3.20E-05
63	GMPPA	3.58	3.06	3.56	3.28	4.08	2.84	2.72	2.61	3	2.59	0.61	0.47	2.00E-08	1.90E-07	6.40E-02
64	HERPUD1	2.5	2.13	1.75	1.7	1.54	2.6	3.26	3.47	4.2	5.76	2.31	2.09	2.30E-08	2.00E-07	6.30E-06
65	COTL1	-3.66	-4.31	-0.88	-2.99	-2.38	-3.49	-4.72	-6.12	0.59	-4.77	-1.37	-1.27	3.50E-08	2.00E-07	9.60E-03
66	CD37	-7	-4.84	-8.74	-8.57	-8.2	-3.14	-2.69	-2.77	-3.52	-3.31	-0.53	-0.1	2.00E-08	2.00E-07	1.30E-01
67	EVL	-8.24	-4.66	-4.07	-3.33	-2.43	-5	-3.81	-4.84	-2.54	-4.03	-2.42	-1.76	2.10E-08	2.10E-07	8.40E-07
68	HDLBP	3.83	3.38	3.66	3.43	3.52	3.4	3.12	3.14	2.5	4.34	2.24	2	1.90E-08	2.10E-07	1.00E-06
69	SEC11C	2.25	1.43	5.26	4.58	5.76	3.58	3.5	3.72	4.11	5.46	2.96	2.27	2.10E-07	1.10E-07	9.70E-05
70	ETS1	-5.58	-3.29	-2.65	-3.97	-7.24	-2.2	-3.39	-3.33	-3.86	-5.2	-1.92	-1.16	3.80E-08	2.10E-07	2.10E-04
71	MZB1	3.17	1.9	4.27	4.25	4.77	5.17	3.85	3.89	4.08	5.45	2.91	2.87	2.10E-07	1.30E-07	3.20E-04
72	RPN1	3.5	2.91	3.55	2.84	1.21	3.31	2.48	2.48	2.35	2.93	1.1	1	1.90E-08	2.10E-07	3.10E-03
73	STT3A	3.12	2.09	8.1	7.53	4.11	2.91	2.44	2.42	3.72	3.28	1.34	1.5	4.90E-08	2.10E-07	3.50E-03
74	FCRL1	-4.68	-4.27	-9.77	-9.77	-5.38	-2.95	-3.55	-3.63	-3.36	-8.21	-0.19	-0.68	1.90E-08	2.10E-07	5.80E-02
75	SSR3	2.38	1.7	5.92	5.87	5.73	3.28	2.59	2.61	3.43	4.25	1.26	0.16	2.10E-07	7.20E-08	7.00E-02
76	SLC39A7	2.64	2.04	3.85	3.12	1.33	2.5	2.29	2.27	2.72	2.74	0.3	-0.68	2.10E-08	2.10E-07	9.40E-02
77	YARS	3.26	3.44	2.16	1.85	1.42	2.87	2.55	2.43	1.94	0.34	2.17	2.1	6.70E-08	1.00E-06	2.20E-07
78	PYCR1	8.3	10.2	1.9	1.48	1.07	7.1	7.73	7.85	9.57	9.71	1.99	1.57	7.40E-08	2.20E-07	9.50E-04
79	KCNK6	8.07	6.07	6.75	5.5	5.71	4.65	4.38	4.3	3.31	3.59	0.83	0.71	1.90E-08	2.40E-07	2.50E-02
80	FKBP11	6.95	7.06	3.88	3.67	3.79	5.45	4.66	4.48	6.62	8.01	2.97	1.57	1.90E-08	2.50E-07	7.30E-05
81	NOTCH2	-4.03	-2.28	-4.67	-4.67	-4.67	-4.29	-4.06	-4.53	-2.01	-2.45	0.97	1.03	3.70E-08	2.50E-07	3.80E-02
82	TMED10	3.02	2.59	2.87	2.63	2.66	2.9	2.73	2.8	2.33	4.06	0.7	0.5	2.10E-08	2.60E-07	8.50E-02
83	DEK	-2.18	-1.95	-6.26	-8.05	-4.38	-3.58	-5.03	-4.34	-5.46	-5.09	-1.48	-1.01	2.80E-07	1.70E-07	8.60E-05
84	LAMP2	2.48	3.69	6.72	6.68	5.49	3.26	1.93	2.13	1.28	3.19	1.67	-0.33	2.00E-08	2.80E-07	1.40E-03

**Table 6-27** | Top 500 Differentially regulated Genes/Proteins across species and platforms. Sorted by median adjusted *p*-value for multiple group comparison. (cont.)

Rank	Symbol	mSpIPC .array	mBMPC .array	hPB .array	hPC .array	hBMPC .array	mPB .rSeq	mSpIPC .rSeq	mBMPC .rSeq	hPB. rSeq	hBMPC .rSeq	mPB .prot	mPB93 .prot	ANOVA Microarray Adj <i>p</i> -value	ANOVA RNA-Seq Adj <i>p</i> -value	ANOVA MS/MS Adj <i>p</i> -value
85	RCBTB2	4.64	4.6	6.48	5.08	7.1	3.48	3.9	4.06	5.09	5.82	1.44	0.87	1.90E-08	2.80E-07	2.80E-02
86	TRAF5	-3.39	-3.62	-9.65	-9.65	-7.5	-4.96	-3.04	-3.57	-5.4	-4.93	-1.93	-1.65	1.90E-08	2.90E-07	3.70E-06
87	BCL11A	-6.01	-4.71	-6.68	-8.34	-5.7	-4.6	-6.12	-7.48	-5.25	-4.35	-0.47	-0.23	1.90E-08	3.00E-07	5.00E-02
88	UBE2J1	3.23	3.16	4.25	3.79	3.65	2.43	3.16	3.26	2.57	3.02	1.59	1.43	1.90E-08	3.10E-07	1.80E-05
89	SEL1L	4.21	5.23	6.84	7.13	6.47	3.94	4.6	4.72	1.91	3.63	1.88	1.52	1.90E-08	3.20E-07	8.20E-05
90	HHEX	-10.33	-5.62	-9.73	-9.88	-9.27	-5.34	-6.53	-6.45	-6.39	-7.36	-3.24	-1.76	1.90E-08	3.30E-07	2.20E-04
91	SURF4	4.32	3.39	3.48	2.88	2.15	2.51	2.06	2.05	3	3.28	1.13	2.28	1.90E-08	3.40E-07	7.20E-02
92	SLC35B1	3.5	2.45	2.83	2.3	2.3	2.93	2.06	2.19	3.51	3.33	2.29	1.89	6.20E-08	3.50E-07	3.90E-02
93	STK26	-2.91	-1.44	-3.24	-3.33	-3.31	-2.99	-2.76	-2.52	-4.95	-2.59	-1.1	-0.63	2.20E-08	3.60E-07	1.20E-03
94	SEC63	2.64	1.66	3.69	3.18	0.7	2.43	2.12	2.17	1.85	1.68	1.43	1.41	8.10E-08	3.60E-07	1.40E-03
95	SPCS1	2.18	2.81	1.95	1.99	1.68	3.27	3.97	4.09	1.97	3.36	0.49	0.04	2.60E-08	3.60E-07	5.30E-02
96	NANS	3.53	3.24	4.67	4.1	3.97	2.85	2.43	2.57	3.06	4	0.18	-0.19	1.90E-08	3.60E-07	7.30E-02
97	SLC33A1	3.44	3.77	4.08	3.85	1.64	3.4	3.6	3.61	2.6	2.79	1.9	1.33	1.90E-08	3.80E-07	2.20E-05
98	LBH	-4.05	-2.57	-6.62	-7.48	-7.39	-2.36	-3.47	-3.57	-4.74	-6.22	-1.83	-0.99	1.90E-08	4.30E-07	3.10E-03
99	FCMR	-9.1	-4.03	-3.52	-5.35	-2.53	-4.9	-3.7	-4.13	-2.59	-5.64	-5.17	-4.94	2.10E-08	2.70E-06	4.40E-07
100	KMO	-4.31	-2.05	-4.49	-3.72	-3.49	-5.47	-2.31	-1.41	-3.59	-3.33	-2.54	-2.84	2.10E-08	1.10E-06	4.70E-07
101	ARFGAP3	4.91	4.07	3.1	2.77	2.7	3.72	4.26	4.5	1.95	2.64	2.07	1.47	5.10E-08	4.90E-07	3.40E-05
102	SLC4A7	-2.88	-2.37	-0.93	-3.01	-5.08	-0.6	-2.07	-2.41	-3.42	-3.92	1.47	1.16	3.80E-08	5.00E-07	5.70E-04
103	CIITA	-7.45	-9.93	-7.51	-6.74	-4.84	-5.67	-6.28	-7.07	-5.36	-8.28	-1.01	-1.7	5.50E-08	5.00E-07	3.60E-02
104	SELENOS	2.39	2.27	6.66	6.53	7.33	3.18	3.13	3.16	4.68	5.39	0.67	0.36	1.90E-08	5.00E-07	5.70E-02
105	USO1	2.13	2.21	2.88	2.7	1.93	2.08	2.44	2.53	1.52	2.02	0.52	0.33	2.00E-08	5.10E-07	5.40E-02
106	MTDH	2.51	1.99	1.83	1.06	1.44	2.75	2.69	2.71	1.56	1.94	0.92	0.9	3.10E-08	5.20E-07	1.10E-04
107	ZBP1	2.89	2.8	4.73	5.4	1.58	2.75	2.35	2.23	2.4	2.08	-1.2	-1.18	1.80E-07	5.20E-07	1.70E-03
108	ARF4	4.05	2.65	4.84	4.1	4.28	1.96	2.12	2.28	2.33	3.29	1.11	-0.12	4.60E-08	5.30E-07	4.10E-02
109	CALU	2.96	2.44	7.35	6.46	4.01	2.44	2.36	2.33	2.25	3.22	0.62	0.75	2.10E-08	5.40E-07	8.00E-02
110	ARMCX3	4.88	4.44	2.16	2.01	1.48	3.75	3.43	3.37	1.64	2.61	2.94	2.46	2.00E-08	5.60E-07	7.10E-07
111	SLC31A1	3.81	2.96	6.89	6.3	4.74	2.52	1.64	1.74	2.13	3.28	2.2	3.04	2.20E-08	5.70E-07	2.40E-03
112	SSR4	3.44	3.03	2.72	3.28	3.38	4.67	4.1	4.38	3.85	6.19	1.42	0.54	6.60E-08	5.90E-07	1.10E-04

**Table 6-27** | Top 500 Differentially regulated Genes/Proteins across species and platforms. Sorted by median adjusted *p*-value for multiple group comparison. (cont.)

Rank	Symbol	mSplPC .array	mBMPC .array	hPB .array	hPC .array	hBMPC .array	mPB .rSeq	mSplPC .rSeq	mBMPC .rSeq	hPB. rSeq	hBMPC .rSeq	mPB .prot	mPB93 .prot	ANOVA Microarray Adj <i>p</i> -value	ANOVA RNA-Seq Adj <i>p</i> -value	ANOVA MS/MS Adj <i>p</i> -value
113	SELL	-8.14	-3.45	1.02	1.1	-3.71	-5.73	-3.59	-5.39	-1.78	-8.71	-1.43	-0.69	3.00E-07	6.40E-07	7.10E-02
114	ITM2C	5.75	5.52	3.45	3.91	4.38	4.23	5.37	5.37	3.44	4.33	0.06	0.79	4.80E-08	6.60E-07	1.80E-01
115	IRF4	4.01	4.83	3.75	3.57	0.95	3.32	2.86	2.84	4.08	1.47	2.67	2.35	2.20E-06	2.30E-07	6.90E-07
116	GLT8D1	2.5	2.54	2.71	2.92	4.14	2.14	2.58	2.65	0.86	1.75	3.28	2.1	3.10E-08	1.10E-06	7.00E-07
117	SP100	-2.59	-2.41	-2.9	-1.49	-3.42	-1.28	-1.96	-1.92	-1.8	-2.03	-1.56	-1.38	4.20E-07	7.40E-07	4.90E-04
118	SRPK3	-10.58	-4.71	0	0.13	0.15	-3.85	-4.23	-4.14	4.61	0.58	-3.74	-4.27	1.90E-01	7.40E-07	2.20E-07
119	B4GALNT1	-5.04	-4.89	0	0	0	-3.64	-4.49	-4.19	5.9	2.28	-1.81	-1.86	4.90E-01	1.80E-07	7.50E-07
120	PECAM1	-2.6	-2.48	1.04	2.78	2.75	-3.64	-2.06	-2.32	-3.34	2.76	-0.31	-0.43	3.50E-07	7.60E-07	1.90E-01
121	HSPA5	1.82	0.96	3.61	3.56	3.31	3.43	2.76	2.65	3.83	4.43	1.03	0.85	7.70E-08	7.70E-07	7.70E-04
122	DTX1	-7.03	-5.3	-2.97	-3.32	-3.23	0.28	-2.72	-3.42	-0.9	-5.19	-0.46	-1.05	3.60E-08	7.70E-07	3.30E-02
123	BCL6	-6.76	-5.48	-1.26	-2.49	-5.37	-2.23	-5.41	-7.46	-2.54	-4.83	-1.66	-1.5	1.60E-07	7.90E-07	1.40E-02
124	SCARB2	-3.45	0.19	2.76	3.1	2.49	-1.56	1.11	1.41	1.41	3.85	-0.43	-0.39	2.50E-08	7.90E-07	2.50E-01
125	HYOU1	2.81	2.94	6.6	5.85	5.24	3.56	2.73	2.63	3.69	4.11	1.8	1.53	3.20E-07	8.00E-07	6.40E-05
126	TMEM131L	-4.18	-2.4	-6.33	-6.27	-5.57	-2.5	-2.77	-2.78	-0.67	-2.55	-0.83	-0.89	2.00E-08	8.00E-07	2.60E-03
127	SLC7A5	2.95	4.31	3.62	2.69	0.73	4.34	4.07	3.65	7.24	7.22	3.7	4.1	8.30E-07	1.80E-06	1.40E-07
128	HELLS	3.11	0.78	3.61	1.13	-0.54	1.78	-0.76	-0.7	-0.07	-3.28	2.79	2.7	8.40E-07	4.40E-06	5.00E-08
129	CR2	-8.93	-8.77	-3.12	-1.33	-3.33	-3.77	-5.69	-8.09	-4.78	-3.85	-1.03	-1.29	5.60E-07	8.50E-07	2.50E-04
130	DERL1	2.83	2.35	2.11	1.54	0.99	2.14	2.19	2.16	2.71	3.32	1.17	0.82	1.90E-08	8.60E-07	7.50E-04
131	SAMD9L	-7.8	-5.38	5.64	8.5	-0.06	-5.3	-5.31	-5.67	-0.1	-2.27	-0.23	0.06	1.90E-08	8.70E-07	2.10E-01
132	BACH2	-8.12	-3.46	-9.34	-9.73	-9.17	-3.13	-4.33	-4.25	-5.44	-5.98	-0.79	-0.66	1.90E-08	8.90E-07	5.20E-03
133	CREG1	4.2	4.11	4.02	4.15	5.38	3.31	2.61	2.62	0.26	2.64	2.02	2.2	9.00E-07	9.00E-07	2.40E-01
134	OS9	2.59	2.29	2.19	1.28	2.33	2.83	2.24	2.25	1.52	2.93	2.03	1.94	1.30E-06	1.90E-07	9.80E-07
135	CD93	3.56	4.92	0	0	1.07	3.82	3.3	4.02	5.72	1.81	2.98	3.84	1.80E-01	4.00E-07	9.80E-07
136	IL4RA	-4.61	-6.69	-8.3	-8.44	-7.44	-4	-5.65	-6.81	-5.84	-7.41	-1.46	-1.02	1.90E-08	9.90E-07	9.50E-04
137	WFS1	4.05	5.5	5.99	5.37	6.07	2.82	3.94	4.08	2.87	5.52	1.7	0.3	3.40E-08	1.00E-06	4.80E-03
138	EPHX1	-8.51	-5.21	0	0	0.81	-0.66	-3.06	-4.15	1.74	3.88	-3.08	-3.46	1.10E-02	1.00E-06	1.20E-07
139	RPS27L	2.22	2.11	3.16	3.12	3.41	3.01	2.27	2.51	1.57	2.4	1.18	1.98	1.90E-08	1.00E-06	2.40E-02
140	PAG1	-2.55	-2.13	1.47	1.6	-2.56	-2.69	-2.76	-2.86	0.49	-1.96	-1.36	-0.99	2.70E-08	1.10E-06	5.90E-04

**Table 6-27** | Top 500 Differentially regulated Genes/Proteins across species and platforms. Sorted by median adjusted *p*-value for multiple group comparison. (cont.)

Rank	Symbol	mSplIPC .array	mBMPC .array	hPB .array	hPC .array	hBMPC .array	mPB .rSeq	mSplIPC .rSeq	mBMPC .rSeq	hPB. rSeq	hBMPC .rSeq	mPB .prot	mPB93 .prot	ANOVA Microarray Adj <i>p</i> -value	ANOVA RNA-Seq Adj <i>p</i> -value	ANOVA MS/MS Adj <i>p</i> -value
141	BTG1	-4.88	-2.77	-10.96	-11.23	-8.29	-3.47	-4.66	-5.04	-2.54	-2.04	-2.08	-1.51	1.10E-06	7.60E-07	1.40E-02
142	GALNT10	-4.16	-4.59	-2.37	-3.25	-5.16	-2.84	-3.59	-3.98	-1.77	-2.64	0.38	0.37	3.30E-07	1.10E-06	6.30E-01
143	STAP1	-4.42	-4.08	-2.83	-0.51	-0.39	-1.6	-3.5	-3.76	-2.37	-1.44	-0.15	-0.03	2.00E-07	1.10E-06	8.30E-01
144	ARHGAP4	-2.72	-3.5	-3.22	-3.22	-3.31	-1.85	-3.01	-3.37	-0.47	-1.92	-2.29	-2.43	2.50E-08	1.10E-05	1.20E-06
145	HMGCS1	-0.16	-0.62	-7.55	-7.55	-7.55	2.33	-1.3	-1.12	1.51	0.08	3.74	4.11	1.90E-02	1.20E-06	3.50E-08
146	CEACAM1	-5.77	-1.78	0.01	0	2.4	-4.46	-2.24	-1.48	5.49	1.86	-0.94	-0.34	1.90E-08	1.20E-06	1.10E-01
147	OASL1	-3.59	-4	0.72	8.1	0.41	0.38	-3.12	-3.78	-2.66	2.37	2.22	2.17	7.00E-07	1.30E-06	1.40E-06
148	SEC24A	2.84	2.06	3.41	3.27	2.98	3.16	2.55	2.31	2.58	3.11	2.56	1.68	2.80E-08	1.30E-06	1.90E-05
149	MORC3	-3.56	-1.74	-1.79	-2.18	-3.47	-2.28	-1.98	-2.06	-4.77	-2.77	-1.24	-1.11	5.70E-08	1.30E-06	2.20E-05
150	RRBP1	5.23	4.07	4.97	4.72	4.22	2.92	2.51	2.64	2.52	4.1	1.75	1.3	1.90E-08	1.30E-06	2.40E-05
151	TBC1D10C	-3.15	-2.43	-2.87	-2.53	-3.47	-2.13	-2.52	-2.49	-2.44	-3.15	-1.06	-1.2	2.40E-08	1.30E-06	4.60E-04
152	NAP1L1	-3.24	-1.51	-4.83	-4.53	-4.49	-1.37	-2.29	-2.37	-1.52	-1.02	1.74	1.61	1.90E-08	1.30E-06	5.70E-04
153	JAK1	-3.41	-2.16	-3.01	-2.85	-8.02	-2.1	-2.24	-2.26	-2.61	-1.38	-0.76	-0.53	1.90E-08	1.30E-06	1.90E-03
154	PRKCB	-4.39	-2.81	-4.36	-5.02	-4.94	-2.65	-3.17	-2.72	-1.92	-4.6	-0.42	-0.56	1.90E-08	1.30E-06	2.20E-02
155	CDK19	-4.4	-2.52	-2.19	-2.89	-6.69	-2.37	-2.64	-2.7	-2.59	-4.98	-0.56	-0.43	1.90E-08	1.30E-06	2.20E-02
156	KDELRL1	2.79	2.4	4.54	3.26	0.64	2.06	2.27	2.34	2.53	4.01	0.78	0.66	2.40E-08	1.30E-06	4.70E-01
157	CD22	-10.05	-7.83	-7.71	-7.65	-7.26	-4.75	-6.21	-5.9	-6.82	-9.5	-1.18	-0.8	1.90E-08	1.40E-06	3.60E-04
158	FAM129C	-7.94	-2.44	-8.78	-8.74	-8.57	-0.46	-4.36	-3.83	-7.46	-9.73	0.96	1.21	1.90E-08	1.40E-06	9.40E-04
159	FUT8	6.43	5.6	3.03	2.73	-1.53	3.7	5.52	5.79	1.21	1.33	1.11	1.44	2.00E-08	1.40E-06	6.30E-03
160	ADK	2.92	4.27	-1.66	-1.66	-1.66	2.97	4.22	4.37	-0.72	-2.78	0.26	0.36	1.90E-08	1.40E-06	2.80E-01
161	AGA	1.6	2.67	4.77	4.98	5	1.44	2.37	2.73	0.62	2.73	0.22	-0.42	3.10E-08	1.40E-06	3.10E-01
162	MYO1G	-1.88	-2.1	-1.76	-1.42	-5.46	-1.38	-2.07	-1.91	-0.01	-3.86	-0.08	0.26	2.10E-07	1.40E-06	5.10E-01
163	SRPR	2.47	2.45	1	0.78	0.27	2.7	2.69	2.74	1.87	2.48	2.46	2.08	6.10E-04	1.30E-07	1.50E-06
164	SELENOK	2.65	2.43	1	0.88	1.47	2.7	2.68	2.89	2.5	4.25	3.51	0.21	7.80E-08	1.50E-06	1.70E-03
165	CRLF3	-2.06	-2.16	0.48	-0.25	-1.54	-1.94	-2.14	-2.25	-1.01	-2.31	-0.84	-0.34	1.50E-06	1.20E-06	1.30E-02
166	PPIB	1.6	1.52	3.14	2.52	2.17	3.27	1.72	1.62	4.1	3.55	1.03	1.39	1.90E-08	1.50E-06	1.80E-02
167	CRYBG1	-3.97	-2.15	-4.05	-7.14	-6.62	-1.7	-2.04	-2.17	-1.61	-5.27	-0.31	-0.24	2.70E-08	1.50E-06	3.00E-01
168	AARS	1.45	1.93	5.08	5.18	3.79	1.82	1.79	1.73	2.76	2.85	2.53	2.65	1.30E-07	3.00E-06	1.60E-06

**Table 6-27** | Top 500 Differentially regulated Genes/Proteins across species and platforms. Sorted by median adjusted  $p$ -value for multiple group comparison. (cont.)

Rank	Symbol	mSplPC .array	mBMPC .array	hPB .array	hPC .array	hBMPC .array	mPB .rSeq	mSplPC .rSeq	mBMPC .rSeq	hPB. .rSeq	hBMPC .rSeq	mPB .prot	mPB93 .prot	ANOVA Microarray Adj $p$ -value	ANOVA RNA-Seq Adj $p$ -value	ANOVA MS/MS Adj $p$ -value
169	MARS	1.33	1.79	1.71	1.92	0.67	1.8	1.87	1.84	1.42	1.42	1.49	1.61	6.60E-07	7.40E-06	1.60E-06
170	EZH2	1.66	0.97	0.53	-0.44	-5.73	2.84	0.21	0.3	2.5	-2.46	2.03	2.04	1.10E-05	1.60E-06	3.10E-07
171	H2-DMA	-2.41	-3.27	-1.75	-2.67	-2.9	-4.72	-3.35	-3.91	-2.32	-3.51	-2.99	-2.32	1.90E-08	1.60E-06	4.60E-03
172	BTLA	-8.22	-2.1	-1.72	-5.27	-1.52	-2.9	-1.36	-1.67	-0.51	-2.22	-1.49	-1.5	2.60E-07	1.60E-06	4.60E-03
173	KLHL6	-4.47	-2.98	5.31	5.4	5.9	-1.5	-3.51	-4.12	1.69	2.44	0.53	0.44	1.90E-08	1.60E-06	7.10E-02
174	UBXN4	1.68	1.43	1.67	1.07	-1.82	2.15	1.98	2.17	0.75	2.18	0.76	1.3	2.50E-08	1.70E-06	3.70E-04
175	IFNGR1	-3.11	-1.93	-6.89	-7.09	-6.73	-2.43	-1.99	-2.03	-3	-1.71	-1.12	-0.45	2.10E-08	1.70E-06	3.50E-03
176	CNPY2	3.02	3.38	5.54	5.21	5.49	3.29	3.64	3.66	1.22	2.15	0.62	1.11	1.90E-08	1.70E-06	5.20E-03
177	BUB1B	4.86	3.6	7.27	4.89	0.19	4.01	3.15	3.1	2.01	-3.74	4.8	4.71	2.00E-08	2.00E-06	1.80E-06
178	HCK	-6.33	-6.65	-0.07	-5.13	-0.68	-4.37	-3.64	-5.35	-4.86	-5	-2.16	-2.19	1.90E-08	1.80E-06	9.90E-06
179	ZBTB18	-4.13	-2.06	-3.82	-3.82	-3.82	-3.22	-2.51	-2.88	-4.47	-3.69	-1.72	-1.98	2.60E-08	1.80E-06	3.60E-02
180	RPS6KA5	-5.68	-3.8	-0.3	0.83	-3.8	-3.06	-2.73	-2.79	-0.68	-2.36	-3.01	-3.38	2.00E-08	1.10E-05	1.90E-06
181	AIDA	-3.83	-1.97	-0.86	-0.45	-1.05	-1.77	-2	-2.01	-2.75	-1.58	-0.9	-1.12	1.90E-06	1.70E-06	2.20E-02
182	SLFN8	-5.28	-3.63	3.74	3.4	2.64	-2.89	-1.79	-2.34	0.39	1.71	-0.04	0.01	1.70E-07	1.90E-06	9.70E-01
183	SEC22B	2.58	2	3.61	3.15	3.37	1.75	1.72	1.78	0.59	1.26	0.53	0.34	1.90E-08	2.00E-06	1.90E-02
184	TLR1	-10.18	-10.52	-5.19	-6.57	-2.5	-5.54	-5.9	-8.52	-1.87	-0.16	-0.62	-0.54	7.10E-08	2.00E-06	1.50E-01
185	FOCAD	3.46	2.98	4.02	3.52	2.42	1.17	2.3	2.63	3.48	2.04	1.44	1.61	6.50E-06	1.60E-06	2.10E-06
186	CLPB	4.36	5.09	1.37	1.09	0.81	2.32	2.97	2.83	5.27	3.42	1.3	1.56	2.00E-05	1.20E-06	2.10E-06
187	ERP44	2.59	1.96	3.59	3.08	3.18	2.66	1.76	1.77	-0.22	0.74	1.3	0.93	1.90E-08	2.10E-06	2.30E-04
188	ACAP1	-3.29	-1.99	-0.72	-0.88	-2.71	-1.6	-2.77	-2.35	-1.83	-2.5	-1.13	-0.97	6.80E-08	2.10E-06	2.10E-03
189	UCK2	5.76	5.04	4.85	2.27	0.94	2.77	2.44	2.86	2.9	1.06	2.62	2.74	3.80E-08	2.20E-06	5.70E-06
190	FAM49B	-4.05	-3.36	-3.62	-3.44	-4.72	-2.22	-2.89	-2.81	-1.36	-1.18	-1.15	-1.03	2.70E-08	2.20E-06	3.50E-04
191	TMED2	4.36	3.27	1.63	1.33	0.54	2.91	2.26	2.47	3.03	3.46	1.29	2.59	2.10E-08	2.20E-06	1.30E-01
192	DENND4A	-2.89	-2.08	-3.89	-3.83	-6.05	-0.46	-2.07	-1.84	-2.85	-2.57	1.81	2.04	3.00E-08	4.00E-06	2.30E-06
193	PLS1	2.13	10.04	4.37	1.68	0.05	5.72	8.75	9.04	0.37	3.3	-2.73	-4.45	2.20E-07	2.30E-06	2.10E-05
194	PTPN6	-4.78	-1.77	-0.82	-2.98	0.44	-2.88	-1.49	-1.64	-1.25	-1.36	-1.35	-1.01	2.00E-08	2.30E-06	1.10E-04
195	ADA	8.14	6.25	6.2	8.04	4.58	4.34	3.55	3.82	5.1	4.65	1.37	1.39	1.90E-08	2.30E-06	4.20E-04
196	TMEM214	3.02	2.74	4.21	3.7	4.6	3.17	2.59	2.63	1.73	2.1	2.62	2.24	2.30E-08	2.40E-06	2.90E-05

**Table 6-27** | Top 500 Differentially regulated Genes/Proteins across species and platforms. Sorted by median adjusted *p*-value for multiple group comparison. (cont.)

Rank	Symbol	mSplPC .array	mBMPC .array	hPB .array	hPC .array	hBMPC .array	mPB .rSeq	mSplPC .rSeq	mBMPC .rSeq	hPB. rSeq	hBMPC .rSeq	mPB .prot	mPB93 .prot	ANOVA Microarray Adj <i>p</i> -value	ANOVA RNA-Seq Adj <i>p</i> -value	ANOVA MS/MS Adj <i>p</i> -value
197	SNX2	6.75	3.31	6.41	4.37	1.41	-2.12	-2.65	-2.65	-2.5	-2.84	-1.19	-0.94	2.80E-08	2.40E-06	1.30E-03
198	PLPP5	7.23	5.88	0.97	0.7	1.65	6.25	6.42	6.36	-1.46	0.39	2.43	2.29	2.40E-06	6.00E-08	1.80E-03
199	STAT6	-3.88	-2.28	-2.67	-2.68	-2.48	-1.96	-1.89	-1.9	-2.71	-3.22	-1.57	-0.89	7.20E-08	2.50E-06	1.20E-04
200	LMAN2	2.27	2.53	5.05	4.38	4.11	1.63	1.49	1.45	2.6	3.49	0.94	1.08	4.00E-07	2.60E-06	4.10E-04
201	ZDHHC13	-5.41	-5.5	1.26	0.63	-2.56	-2.22	-2.98	-3.44	0.95	-0.75	0.69	0.68	2.30E-08	2.60E-06	1.60E-02
202	GMPPB	3.62	3.5	4.86	4.19	3.76	2.51	2.08	2.46	2.9	3.8	0.22	0.1	1.90E-08	2.60E-06	3.10E-01
203	CLPTM1L	1.83	2.06	2.65	2.47	2.09	3.23	2.73	2.74	2.64	3.77	-0.03	0.08	1.90E-08	2.60E-06	9.10E-01
204	ELL3	-6.06	-5.43	-3.78	-3.78	-3.78	-6.75	-6.35	-7.78	-3.82	-7.42	-3.57	-3.95	8.40E-08	2.70E-06	1.70E-04
205	PIM2	2.17	2.4	1.75	2.07	-0.67	2.15	2.62	2.52	2.95	3.39	1.7	1.18	9.60E-08	2.70E-06	5.70E-03
206	UBA5	3.36	2.58	4.71	4.07	3.01	2.71	2.56	2.62	0.22	1.58	0.44	-0.15	1.90E-08	2.70E-06	2.00E-01
207	GIMAP3	-5.55	-3.44	0	0.1	1.47	-1.53	-4.3	-3.53	2.72	6.12	0.31	-0.02	2.70E-06	1.20E-06	5.60E-01
208	LY86	-4.07	-4.41	-6.93	-6.84	-2.53	-2.79	-4.11	-4.94	-5.21	-2.67	0.05	0.15	1.90E-08	2.70E-06	9.70E-01
209	PFKP	1.23	0.71	0.47	-0.2	-4.46	2.32	1.97	1.39	1.67	-3.76	1	1.27	2.00E-07	2.80E-06	3.00E-06
210	PACS1	-4.31	-1.73	-1.53	-1.49	-1.28	-1.88	-1.38	-1.2	-3.4	-3.72	-1.29	-1.14	1.20E-06	2.80E-06	1.70E-05
211	SESN3	-5.65	-4.13	-8.23	-8.23	-8.23	-2.11	-3.95	-4.35	-3.88	-5.92	-1.99	-1.73	1.90E-08	2.80E-06	8.50E-04
212	DOCK11	-2.03	-1.45	-1.8	-2.41	-4.65	-2.28	-2.16	-2.17	-1.54	-3.55	-0.49	-0.19	2.40E-08	2.80E-06	3.40E-02
213	GFPT1	6.68	4.75	6.71	6.29	7.15	2.17	2.47	2.36	3.05	3.11	1.81	1.88	1.90E-08	2.90E-06	5.50E-06
214	MTA3	-5.36	-3.29	2.32	2.5	1.06	-2.83	-2.5	-2.12	-0.86	0.94	-1.09	-0.98	5.10E-08	2.90E-06	1.60E-03
215	RASGRP1	-7.93	-3.82	0.01	-1.94	-3.12	-3.27	-4.64	-6.04	0.16	-3.25	-0.44	-0.69	2.90E-06	1.50E-06	7.60E-02
216	ADD3	-4.79	-3.07	-1.2	-1.79	-3.74	-3.15	-1.94	-1.97	-0.79	-1.73	-2.08	-1.45	2.40E-07	3.00E-06	2.00E-05
217	RBM47	7.61	8.06	9.22	9.12	8.13	4.86	5.71	5.4	6.52	7.18	3.47	3.11	1.90E-08	3.00E-06	5.50E-05
218	GGA2	-3.03	-2.32	-3.51	-6.24	-7.64	-2.53	-2.66	-2.57	-2.19	-2.53	-1.41	-1.42	1.90E-08	3.00E-06	1.00E-04
219	TLR7	-3.15	-0.01	6.06	8.11	0.46	1.89	0.08	-1.45	-1.37	-4.69	1.36	1.73	3.00E-06	9.10E-07	3.80E-03
220	GTSE1	2.45	1.88	6.09	2.24	1.58	3.67	-0.19	-0.35	3.13	-1.78	5.03	5.54	1.80E-03	3.10E-06	9.10E-08
221	MANEA	2.25	1.85	7.61	8.4	7.94	1.97	2.56	2.39	4.6	4.19	0.93	1.51	4.30E-07	3.10E-06	2.00E-03
222	MCM3	1.5	-0.06	2.03	1.09	-3.47	0.77	-1.97	-1.95	1.34	-1.29	1.88	2.16	2.10E-03	3.10E-06	1.50E-06
223	TMED9	1.05	0.66	2.75	2.54	2.38	2.27	1.76	1.82	2.12	4.26	1.2	1.17	1.70E-07	3.10E-06	9.00E-03
224	BLK	-5.77	-3.12	-3.9	-6.44	-8.11	-1.73	-2.44	-2.9	-2.8	-4.62	-1.75	-2.25	1.90E-08	3.50E-06	3.20E-06

**Table 6-27** | Top 500 Differentially regulated Genes/Proteins across species and platforms. Sorted by median adjusted *p*-value for multiple group comparison. (cont.)

Rank	Symbol	mSpIPC .array	mBMPC .array	hPB .array	hPC .array	hBMPC .array	mPB .rSeq	mSpIPC .rSeq	mBMPC .rSeq	hPB. .rSeq	hBMPC .rSeq	mPB .prot	mPB93 .prot	ANOVA Microarray Adj <i>p</i> -value	ANOVA RNA-Seq Adj <i>p</i> -value	ANOVA MS/MS Adj <i>p</i> -value
225	DGKA	-1.46	-1.85	-2.65	-2.27	-2.5	-0.91	-1.51	-1.61	-3.1	-2.06	-0.61	-0.41	6.50E-07	3.20E-06	2.00E-01
226	SLC12A2	2.89	3.09	1.1	-0.39	-2.75	1.83	3.38	3.85	-1.65	-1.18	1.43	1.47	3.80E-08	3.30E-06	3.60E-05
227	ERLEC1	2.65	3.23	3.31	2.76	2.73	2.64	2.45	2.89	1.62	3.42	1.47	0.94	1.90E-08	3.30E-06	2.60E-03
228	EVI2A	2.3	2.49	5.14	4.68	4.29	0.97	3.13	2.59	0.65	1.91	2.26	2.1	2.00E-08	3.40E-06	6.10E-05
229	SLCO4A1	3.32	0.51	6.87	6.05	2.76	4.18	-0.15	-0.24	0.74	3.24	1.31	1	3.40E-06	3.30E-07	5.10E-02
230	IDII	1.63	-0.86	-2.09	-2.47	-4.68	2.77	-1.89	-1.92	-2.48	-0.97	2.86	3.12	2.90E-05	1.30E-06	3.60E-06
231	BMP2K	-2.99	-2.25	-5.58	-5.53	-5.58	-3.41	-2.46	-2.55	-3.59	-2.52	-0.41	-0.57	8.40E-07	3.60E-06	4.10E-03
232	RASSF2	-4.1	-2.35	0.36	-1.17	-3.41	-2.24	-2.57	-2.55	-1.57	-3.24	-1.15	-0.93	3.70E-06	5.30E-07	3.00E-05
233	SLC16A1	1.69	1.26	4.69	3.87	-0.97	1.8	-0.66	-0.08	3.46	1.27	2.27	2.45	3.40E-05	1.00E-06	3.70E-06
234	BST2	3.03	3.17	2.41	3.75	1.6	2.05	2.26	2.71	-0.02	1.84	3.39	1.56	2.70E-07	3.70E-06	2.90E-02
235	OSBPL3	5.82	7.6	3.1	2.27	1	2.01	6.51	7.65	1.64	3.98	0.45	-0.29	2.40E-07	3.80E-06	2.80E-01
236	PGRMC1	0.74	1.39	2.85	2.17	0.23	-0.23	1.46	1.61	-4.43	0.08	-0.21	-0.11	3.80E-06	3.30E-06	6.00E-01
237	EML4	-2.43	-1.22	-5.69	-5.69	-5.69	-1.8	-1.67	-1.72	-0.86	-2.19	-0.84	-0.46	5.00E-08	3.90E-06	1.50E-02
238	PDIA4	5.45	2.73	9.04	8.27	8.12	3.61	3.23	3.05	4.43	3.6	0.77	0.51	4.00E-06	3.10E-06	2.20E-03
239	TK2	-6.55	-2.68	0.5	0.9	2.36	-3.23	-1.56	-1.55	-0.04	2.44	-3.12	-2.7	1.60E-07	4.00E-06	5.20E-03
240	NCAPH	5.53	2.13	3.77	1.06	0.17	3.94	0.06	-0.31	7.84	4.55	2.8	2.85	5.50E-05	4.10E-06	1.10E-07
241	REL	-4.45	-3.79	-7.08	-7.43	-7.39	-1.93	-4.1	-4.71	-6.4	-6.55	-1.3	-0.93	4.10E-06	3.20E-06	1.70E-03
242	TESPA1	-5.24	-6.32	-0.09	-0.09	-0.09	-5.56	-6.7	-6.39	-3.12	-5.51	-2.2	-3.28	1.10E-01	4.10E-06	2.50E-06
243	RAB1A	1.67	1	-2.73	-2.35	-2.73	1.86	1.52	1.67	1.55	2.1	-0.19	-0.42	4.20E-06	2.00E-06	2.50E-01
244	ANKRD44	-1.73	-1.09	-4.7	-4.52	-4.86	-2.34	-1.38	-1.17	-2.69	-2.3	-1.54	-1.27	5.80E-07	4.30E-06	1.20E-05
245	ZBTB4	-3.25	-0.78	-1.35	-0.92	-1.39	-1.91	-0.22	0.08	-2.67	0	-1.48	-0.97	4.30E-06	1.10E-06	1.60E-02
246	BCL2L1	1.93	2.11	2.86	3.29	1.56	5.01	1.54	1.71	1.7	3.74	0.1	0.93	4.30E-06	1.30E-07	3.40E-01
247	PECR	4.66	4.07	2.83	1.41	-0.18	3.43	3.72	3.99	0.37	-1.53	1.89	1.8	5.50E-08	5.00E-06	4.40E-06
248	SIRPA	1.88	5.46	0	0.01	1.9	-1.01	3.61	3.28	4.68	2.63	-2.63	-3.2	4.40E-06	3.30E-06	5.60E-04
249	DNAJC1	2.98	2.57	4.61	4.88	4.49	1.4	1.84	1.84	1.69	2.95	0.96	0.42	5.00E-08	4.40E-06	2.00E-03
250	FGR	-3.57	-1.52	-6.76	-6.85	-4.93	-5.32	-1.32	-1.02	-0.27	-8.28	-5.1	-5.72	4.40E-06	8.80E-06	4.50E-06
251	CD38	-3.57	-2.27	8.45	9.71	9.64	-0.92	-2.73	-3.51	3.69	5.17	-1.97	-2.03	2.50E-07	4.50E-06	1.90E-03
252	GCAT	3.52	4.16	0.55	0	0.59	4.41	4.04	4.46	3.24	5.73	3.12	2.79	2.50E-01	4.50E-06	4.20E-06

**Table 6-27** | Top 500 Differentially regulated Genes/Proteins across species and platforms. Sorted by median adjusted  $p$ -value for multiple group comparison. (cont.)

Rank	Symbol	mSplPC .array	mBMPC .array	hPB .array	hPC .array	hBMPC .array	mPB .rSeq	mSplPC .rSeq	mBMPC .rSeq	hPB. .rSeq	hBMPC .rSeq	mPB .prot	mPB93 .prot	ANOVA Microarray Adj $p$ -value	ANOVA RNA-Seq Adj $p$ -value	ANOVA MS/MS Adj $p$ -value
253	MACF1	-2.44	-1.28	-2.28	-2	-4.11	-2.24	-1.86	-1.95	-1.54	-2.48	0.1	0.13	6.80E-07	4.50E-06	5.20E-01
254	CCDC125	-2.03	-0.69	-0.56	-0.88	-3.78	-1.29	-3.5	-3.98	-1.96	0.77	-0.59	-0.57	6.10E-07	4.50E-06	7.30E-01
255	HSPA13	4.07	3.58	2.4	1.71	1.77	3.65	3.54	3.55	2.25	4.05	2.45	2.15	4.60E-06	1.30E-07	6.40E-05
256	CCNE2	4.17	0.94	5.23	1.11	0.12	3.69	1.81	1.44	3.42	1.89	0.13	-0.12	5.70E-07	4.60E-06	6.90E-01
257	NPC2	1.27	1.14	1.3	1.6	2.44	2.45	1.73	1.51	0.16	3.66	1.78	2.12	2.00E-08	4.80E-06	2.00E-05
258	CD40	-7.25	-2.96	-5.2	-4.71	-4.52	-4.24	-3.43	-3.19	-2.76	-0.82	0.8	1.08	1.90E-08	4.90E-06	6.90E-04
259	RAPGEF4	-7.44	-7.72	0	0.94	3.34	-6.28	-5.06	-4.91	6.61	6.67	-1.48	-2.11	4.80E-08	4.90E-06	3.10E-03
260	PARP3	-3.57	-1.94	0.03	0.33	1.43	-3.07	-1.48	-1.33	-0.27	1.31	-3.17	-3.02	8.10E-03	4.90E-06	9.00E-07
261	PTPRC	-1.38	-1.51	-6.27	-6.37	-6.34	-0.34	-1.82	-2.12	-1.45	-5.54	-0.48	-0.2	1.90E-08	4.90E-06	7.30E-02
262	CXCR5	-9.7	-10.79	-8.45	-8.42	-8.12	-4.08	-6.36	-9.1	-8.71	-8.03	-1.34	-1.9	1.90E-08	5.00E-06	7.10E-04
263	KRTCAP2	2.39	3.45	2.08	1.9	3.11	2.72	2.63	2.68	1.06	3.44	-0.2	0.69	2.00E-08	5.00E-06	2.30E-01
264	SPIB	-8.14	-3.74	-3.78	-3.86	-3.24	-5.77	-7.47	-6.06	-5.25	-8.06	-1.56	-1.23	5.10E-06	1.50E-06	1.60E-03
265	HTATIP2	3.43	3.9	-2.73	-3.24	-5.66	1.21	3.03	3.24	-2.82	-0.11	0.44	0.76	1.90E-08	5.10E-06	1.80E-01
266	PTPRJ	-3.81	-3.19	-0.77	-1.9	-5.3	-2.17	-2.52	-3.17	-1.69	-2.86	-2.37	-1.62	5.30E-07	5.30E-06	8.30E-06
267	ARHGAP6	1.75	2.99	3.82	3.01	-2.05	-0.5	2.12	1.84	5.36	2.51	-1.92	-3.12	4.90E-06	5.30E-06	6.30E-06
268	SLC30A7	4.31	3.52	2.9	1.79	0.15	2.52	2.55	2.75	1.4	0.82	1.08	1.05	1.70E-06	5.30E-06	2.70E-02
269	TAPBPL	1.96	2.25	3.13	2.85	3.84	1.2	2.06	2.18	0.29	1.89	-0.57	-0.64	1.10E-07	5.30E-06	3.90E-02
270	PLAUR	-2.65	-6.64	0	0	4.51	-7.99	-5.8	-5.95	2.33	6.34	-2.33	-3.04	3.40E-08	5.40E-06	6.40E-05
271	SESN1	-2.59	-0.97	-2.21	-1.9	-2.98	-2.28	-0.87	-0.86	-3.84	-1.21	-0.18	-0.42	5.40E-06	2.10E-06	4.00E-01
272	KYNU	-5.91	-7.73	-3.19	-3.77	-2.59	-5.39	-5.46	-6.79	-0.13	-6.97	-1.18	-0.67	2.70E-08	5.50E-06	5.00E-04
273	RILPL2	3.34	2.77	-0.35	-1.01	-0.4	3.81	2.35	2.02	-4.37	-0.01	3.49	3.48	7.60E-03	5.60E-08	5.50E-06
274	GALNT1	-1.3	-1.04	1.06	0.55	-3.81	-1.57	-1.73	-1.81	1.66	-0.92	-0.63	-0.15	4.80E-07	5.50E-06	1.10E-01
275	ABCA1	-2.44	-0.37	-5.18	-6.91	-6.38	-3.92	-0.12	-0.79	-2.48	-5.17	-0.74	-0.23	7.10E-07	5.50E-06	1.70E-01
276	ARHGAP45	-1.57	-1.89	-1.87	-1.89	-3.66	-1.78	-1.57	-1.61	-2.94	-2.74	-1.89	-1.59	2.10E-05	3.00E-06	5.60E-06
277	BCL2	-7.48	-0.39	-5.98	-3.3	-2.07	-2.71	0.22	0.1	-3.28	1.88	0.27	0.77	1.90E-08	5.60E-06	3.50E-01
278	ALDH2	-3.23	-3.13	2.04	0.19	1.85	-3.34	-2.66	-2.11	-2.31	1.63	-2.13	-2.03	5.70E-06	1.10E-05	1.70E-07
279	USP25	-5.01	-2.48	-2.74	-0.88	-5.19	-2.58	-1.99	-1.98	-1.67	-1.55	-0.68	-0.24	2.00E-08	5.70E-06	4.40E-03
280	WDFY2	-2.97	-3.4	-7.5	-7.44	-7.46	-2.3	-2.94	-3.55	-3.94	-4.55	-0.87	-1.27	2.10E-08	5.70E-06	5.00E-02

**Table 6-27** | Top 500 Differentially regulated Genes/Proteins across species and platforms. Sorted by median adjusted *p*-value for multiple group comparison. (cont.)

Rank	Symbol	mSplPC .array	mBMPC .array	hPB .array	hPC .array	hBMPC .array	mPB .rSeq	mSplPC .rSeq	mBMPC .rSeq	hPB. rSeq	hBMPC .rSeq	mPB .prot	mPB93 .prot	ANOVA Microarray Adj <i>p</i> -value	ANOVA RNA-Seq Adj <i>p</i> -value	ANOVA MS/MS Adj <i>p</i> -value
281	IL6ST	4.15	3.71	-2.15	-1.49	-0.68	3.74	3.28	2.94	0.6	3.35	0.61	-0.26	5.70E-06	1.60E-06	1.50E-01
282	CSNK1E	0.66	3.86	4.24	4.51	2.86	1.78	1.83	1.87	-0.03	2.35	0.5	-0.02	7.50E-07	5.70E-06	5.30E-01
283	DENND5B	3.63	3.73	1.68	1.61	2.09	1.78	2.16	1.99	0.04	1.65	0.68	0.03	1.10E-06	5.80E-06	5.70E-03
284	TMEM39A	3.61	2.64	4.71	4.33	2.92	2.3	2.34	2.25	3.23	2.8	0.44	0.54	2.00E-08	5.80E-06	2.70E-01
285	CST3	0.57	1.45	0	0	4.19	1.51	2.57	2.73	0.65	9.05	-2.52	-2.75	7.10E-04	5.90E-06	4.80E-06
286	CARS	1.54	2.49	1.2	1.36	0.2	2.11	2.19	2.08	1.9	1.24	2.5	2.82	3.40E-03	6.10E-06	1.90E-07
287	NFKBID	-3.03	-4.53	-4.01	-4.01	-4.01	-1.89	-3.77	-3.63	-1.69	-3.75	-1.76	-2.99	2.00E-07	6.20E-06	1.80E-03
288	UFL1	1.73	1.78	1.18	1.05	-2.11	1.82	1.71	1.73	-0.11	1.27	0.79	0.53	3.90E-07	6.30E-06	7.60E-04
289	CDC45	5.99	4.27	2.07	0	0	3.97	0.74	0.65	8.67	1.7	5.84	5.94	4.40E-04	6.40E-06	5.00E-08
290	NCF1	-6.93	-5.29	-0.53	-0.86	-2.85	-2.13	-4.47	-4.66	0.18	-3.27	0.13	0.56	6.50E-06	2.40E-07	2.00E-01
291	IRF2BP2	-2.11	0.48	-2.27	-1.99	-4.87	-2.16	-0.47	-0.67	-0.78	0.72	-0.61	0.02	5.70E-06	6.70E-06	4.10E-02
292	PHGDH	1.88	2.24	5.32	4.31	3.82	2.13	2.44	2.41	4.51	4.7	1.77	1.68	1.90E-08	6.90E-06	1.10E-05
293	FOXP1	-3.14	-2.06	-7.2	-7.19	-7.08	-1.4	-2.11	-2.56	-4.95	-4.95	-0.02	-0.18	4.30E-08	6.90E-06	5.40E-01
294	TTPAL	-2.46	-2.53	0.04	-0.23	-3.17	-1.45	-2.28	-2.09	0.3	-1.45	-0.3	-0.2	5.90E-08	6.90E-06	7.00E-01
295	CCNB1	8	4.64	8.72	4.99	-0.01	4.98	1.57	0.94	2.28	-1	5.22	5.57	1.90E-08	1.00E-05	7.00E-06
296	TNRC6C	-2.96	-1.19	-4.7	-4.49	-4.61	-1.9	-1.43	-1.69	-1.23	-2	-1.13	-1.54	2.20E-08	7.00E-06	2.20E-02
297	CAMK1D	-1.92	-1.01	-3.35	-3.1	-3.82	-2.01	-1.44	-1.48	-2.85	-1.92	-1.53	-1.14	2.20E-07	7.10E-06	3.00E-04
298	NOD1	-3.67	-2.25	-5.19	-4.09	-3.76	-3.07	-1.41	-1.89	-0.85	-3.01	-0.9	-1.17	2.30E-08	7.10E-06	2.70E-01
299	SELPLG	2.22	2.36	7.52	6.99	6.82	2.22	2.32	2.53	2.97	2.21	1.78	-1.01	1.90E-08	7.20E-06	3.00E-03
300	PPA1	4.29	3.34	1.69	1.77	0.94	2.5	2.09	2.01	1.72	0.74	1.44	1.7	9.90E-07	7.30E-06	1.90E-04
301	OSTC	2.68	2.17	3.42	2.84	2.89	2.73	1.97	1.65	2.81	2.72	1.46	-0.07	2.00E-08	7.30E-06	4.10E-03
302	GORASP2	3.93	2.32	3.33	2.97	2.6	1.35	1.36	1.39	2.83	3.17	0.05	0.07	1.90E-08	7.40E-06	9.60E-01
303	ARCN1	2.67	1.45	1.75	1.36	0.31	1.46	1.29	1.35	1.44	1.5	0.91	0.86	4.60E-06	7.70E-06	9.40E-04
304	P2RX4	-1.28	1.53	2.13	1.96	1.61	-0.41	1.47	1.52	3.72	3.39	-2.87	-3.01	1.70E-03	7.70E-06	1.70E-06
305	ARHGAP25	-2.47	-2.2	-2.01	-1.03	-2.68	-1.22	-1.77	-1.79	-0.57	-0.96	-1.02	-0.76	2.00E-08	7.80E-06	1.60E-04
306	HID1	0.15	5.38	1.15	1.03	2.27	9.44	8.95	9.32	6.95	9.43	6.92	6.27	1.90E-04	3.50E-07	7.80E-06
307	IDE	2.83	2.97	4.17	3.53	2.31	2.29	2.13	2.47	3.61	2.11	1.18	1.51	2.00E-08	7.80E-06	1.70E-03
308	ALDH7A1	3.93	4.01	0.38	0.09	2.46	1.53	2.54	2.95	-0.09	3.3	0.9	0.64	3.60E-08	7.90E-06	3.90E-04

**Table 6-27** | Top 500 Differentially regulated Genes/Proteins across species and platforms. Sorted by median adjusted *p*-value for multiple group comparison. (cont.)

Rank	Symbol	mSplPC .array	mBMPC .array	hPB .array	hPC .array	hBMPC .array	mPB .rSeq	mSplPC .rSeq	mBMPC .rSeq	hPB. .rSeq	hBMPC .rSeq	mPB .prot	mPB93 .prot	ANOVA Microarray Adj <i>p</i> -value	ANOVA RNA-Seq Adj <i>p</i> -value	ANOVA MS/MS Adj <i>p</i> -value
309	DENND4B	-2.56	-1.89	-1.17	-1.32	-2.38	-3.96	-2.35	-2.83	-0.01	-2.95	-0.63	-0.52	5.80E-08	7.90E-06	1.20E-02
310	PLD4	0.38	-1.15	-0.01	-0.01	-0.01	-2.1	-1.11	-1.32	0.6	-3.5	-2.73	-2.46	1.90E-01	7.90E-06	6.80E-07
311	FAM172A	-1.83	-0.78	-1.91	-1.68	-3.57	-1.67	-1.25	-1	-2.19	-1.41	0.06	-0.42	1.60E-07	8.00E-06	1.40E-01
312	STOM	0	1.53	2.13	1.55	0.11	-0.74	4.85	4.73	2.45	3.56	-4.01	-3.92	7.90E-04	3.60E-06	8.20E-06
313	TMEM65	1.81	2.26	3.86	3.15	-0.14	0.45	1.89	1.77	-6.71	-1.29	-1.67	-1.71	1.30E-07	8.20E-06	7.40E-03
314	SNX29	-3.71	-1.84	-3.76	-3.82	-3.61	-3.37	-1.43	-1.46	-2.11	-2.05	-3.07	-3.47	2.10E-08	8.30E-06	2.00E-04
315	GLA	6.8	5.49	2.69	1.88	-0.11	3.86	1.63	2.06	1.15	2.01	1.49	1.63	1.30E-07	8.40E-06	6.50E-04
316	SLC2A3	-2.85	-2.42	-4.08	-6.2	-7.75	-0.86	-2.8	-2.87	-0.64	-3.59	-1.04	-0.88	2.00E-08	8.40E-06	1.30E-02
317	LGALS1	4.09	3.29	6.87	5.04	7.86	3.02	3.04	3.15	5.38	4.71	0.48	1.07	2.20E-08	8.40E-06	2.50E-01
318	STAT4	-1.77	-2.39	4.66	3.66	0.98	-0.98	-1.99	-2.09	5.31	5.59	-0.11	0.55	2.20E-06	8.50E-06	1.30E-01
319	MCM5	1.54	-0.37	-0.39	-1.82	-3.8	0.68	-1.92	-2.2	0.48	-1.06	2.4	2.31	3.00E-05	8.80E-06	1.60E-07
320	MCM7	1.98	0.03	-0.8	-4.11	-6.29	0.67	-1.12	-0.99	0.94	-1.19	1.88	1.81	1.80E-04	8.80E-06	3.80E-07
321	LDLR	-0.28	-1.26	7.62	4.12	6.06	2.1	-2.22	-1.13	5.27	4.45	4.33	4.19	4.10E-02	1.40E-06	9.00E-06
322	RHOB	4.2	3.94	-3.16	-4.32	1.14	1.49	3.91	3.47	3.44	6.31	0.65	0.99	6.90E-07	9.10E-06	2.10E-01
323	PCLAF	6.09	2.52	8.4	4.11	4	4.32	0.76	0.52	7.43	1.24	1.02	3.26	1.20E-07	9.30E-06	1.20E-03
324	WDFY4	-1.8	-1.6	-4.5	-3.69	-4.2	-1.97	-1.98	-2.34	-1.72	-4.37	-0.18	0.19	6.60E-07	9.30E-06	1.40E-01
325	FADS1	4.5	4.36	4.93	2.7	1.54	5.19	2.94	2.76	2.24	4.13	3.56	3.89	9.50E-08	9.40E-06	9.90E-06
326	SEC13	2.13	1.58	3.28	2.73	2.03	1.65	1.07	1.34	2.12	2.66	0.26	0.54	2.00E-08	9.40E-06	2.50E-01
327	BIN1	-5.47	-5.54	-5.84	-5.95	-2.74	-5.27	-5.42	-5.93	-2.1	-2.24	-2.11	-1.63	1.90E-05	9.50E-06	3.20E-06
328	ALDH3B1	3.67	3.27	0.01	0	2.5	0.45	2.62	2.69	4.19	3.67	-1.02	-1.13	4.50E-07	9.60E-06	3.00E-04
329	IRF8	-6	-2.59	-5.73	-7.14	-5.75	-1.29	-1.73	-2.01	-4.52	-8.44	-0.86	-0.44	1.90E-08	9.70E-06	1.00E-03
330	IFIT2	-5.99	-5.56	-0.77	-0.71	-0.63	-5.08	-3.85	-5.02	-5.28	-5.32	-2.34	-2.45	9.70E-06	2.60E-06	1.50E-03
331	TOPBP1	0.55	-0.8	-0.88	-1.38	-3.09	0.21	-1.44	-1.26	-0.09	-1.52	1.68	1.85	9.10E-03	9.70E-06	8.70E-07
332	ATAD2B	-0.07	-2.64	-4.01	-5.03	-4.98	-1.77	-1.57	-1.96	-1.61	-3.91	-1.18	-0.97	1.70E-07	1.00E-05	1.20E-04
333	TXN1	2.48	1.58	5.74	5.88	5.84	2	2.28	2.55	3.54	3.3	0.73	1.26	2.40E-08	1.00E-05	7.70E-03
334	SHMT1	0	0.08	3.17	1.93	1.78	1.14	-1.1	-1.4	3.54	1.32	1.78	1.68	1.00E-01	1.00E-05	8.70E-07
335	GLMP	1.65	1.55	6.72	6.19	5	2.1	1.11	0.97	1.13	2.32	0.71	0.44	1.00E-05	9.40E-06	6.70E-01
336	ARHGEF1	-2.16	-2.34	-1.39	-1.33	-1.83	-2.39	-2.15	-2.37	-0.85	-2.36	-1.34	-1.06	1.20E-05	3.30E-06	1.10E-05

**Table 6-27** | Top 500 Differentially regulated Genes/Proteins across species and platforms. Sorted by median adjusted *p*-value for multiple group comparison. (cont.)

Rank	Symbol	mSpIPC .array	mBMPC .array	hPB .array	hPC .array	hBMPC .array	mPB .rSeq	mSpIPC .rSeq	mBMPC .rSeq	hPB. .rSeq	hBMPC .rSeq	mPB .prot	mPB93 .prot	ANOVA Microarray Adj <i>p</i> -value	ANOVA RNA-Seq Adj <i>p</i> -value	ANOVA MS/MS Adj <i>p</i> -value
337	SCML4	-5.52	-6.12	-3.2	-3.2	-3.2	-3.85	-4.85	-6.16	-4.05	-6.45	-3.53	-4.24	5.60E-05	6.90E-07	1.10E-05
338	SLC7A6	1.1	0.88	-0.33	-0.84	-4.01	1.6	0.41	0.31	1.49	-0.83	1.95	1.68	3.50E-06	1.10E-05	6.30E-04
339	PAX5	-5.65	-2.97	-5.18	-4.02	-4.02	-4.52	-6.84	-6.84	-3.3	-5.76	-1.21	-0.76	2.00E-08	1.10E-05	7.20E-04
340	PRAG1	-2.92	-2	-3.38	-4.79	-0.97	1.21	-1.03	-1.28	1.62	-1.78	0.87	0.83	1.10E-05	6.40E-06	1.50E-03
341	MBP	-2.44	-1.33	-0.83	-2.18	-6.24	-3.07	-1.91	-1.76	-0.74	-1.62	-1.51	-1.02	1.20E-06	1.10E-05	1.40E-02
342	COLGALT1	-1.83	-1.34	-3.3	-3.07	-2.72	-0.43	-1.46	-1.53	-0.34	-1.87	1.11	1.15	1.70E-06	1.20E-05	3.60E-05
343	ARID1B	-1.64	-0.68	-3.93	-4.12	-5.39	-1.53	-1.3	-1.38	-1.84	-0.91	-1.89	-1.5	3.20E-05	1.20E-05	5.80E-06
344	MAP4K4	-2.45	-1.79	-3.68	-4.6	-8	-1.06	-1.66	-1.68	-1.67	-3.05	0.67	0.99	8.00E-07	1.20E-05	1.80E-04
345	EDEM1	5.84	4.65	0.88	0.19	-1.23	5.03	4.71	4.83	2.22	0.9	2.03	1.17	4.70E-08	1.20E-05	3.00E-04
346	SAR1B	2.12	1.79	5.55	5.71	5.18	2.02	1.38	1.49	2.98	3.14	0.94	0.88	2.00E-08	1.20E-05	1.30E-03
347	TGFBR2	-4.58	-1.9	-2.47	-2.47	-2.44	-1.54	-1.93	-2.07	-5.11	-3.49	-0.26	0.03	1.90E-08	1.20E-05	3.00E-01
348	GMFB	-1.86	-1.27	-1.46	-1.64	-2.95	-1.21	-1.36	-1.23	-3.12	-1.72	0.08	0.72	2.50E-08	1.20E-05	3.30E-01
349	STK10	-3.56	-3.48	-1.23	-0.96	-4.84	-2.86	-3.59	-3.88	-2.84	-1.72	-1.21	-1.24	2.60E-08	1.30E-05	3.10E-05
350	LMNA	0.4	2.74	0	0	0.46	0.71	2.53	3.38	7.36	6.79	-2.54	-3.72	2.40E-04	1.30E-05	1.00E-06
351	AKAP12	-5.3	-0.88	0	0	3.1	-1.45	-4.94	-3.31	1.68	5.84	-1.46	-0.96	6.00E-07	1.30E-05	9.10E-04
352	FNDC3A	2.58	2.65	1.64	1.92	2.56	2.55	3.13	3.22	0.87	3.05	1.51	0.75	1.30E-05	1.80E-07	1.40E-03
353	PIGK	3.1	2.75	8.01	7.8	7.31	1.72	1.25	1.44	1.68	2.57	0.56	0.09	1.90E-08	1.30E-05	2.50E-02
354	PSIP1	1.86	1.98	-1.82	-1.83	-2.69	1.8	1.45	1.51	-1.49	-0.91	-0.48	-0.5	1.00E-07	1.30E-05	3.00E-02
355	PDE7A	-2.66	-1.94	-2.15	-2.22	-2.05	-1.13	-1.95	-2.11	-2.68	-3.11	1.31	1.19	1.10E-07	1.30E-05	3.20E-02
356	MCM4	0.54	-0.85	3.85	0.12	0	0.37	-1.83	-1.75	3.1	-0.6	2.14	2.01	3.80E-02	1.30E-05	1.40E-07
357	APMAP	2.17	2.28	3.12	2.65	1.05	0.89	1.68	2.14	-0.48	1.85	-0.39	-0.38	1.90E-08	1.30E-05	6.00E-02
358	CPSF2	-2	-2.06	3.18	2.61	-0.25	-0.17	-1.4	-1.28	-2.44	0.08	-0.37	-0.08	9.00E-06	1.30E-05	9.90E-02
359	USP24	-2.21	-1.63	-5.36	-5.34	-5.59	-1.48	-1.48	-1.61	-2.05	-3.22	-0.34	-0.16	2.30E-07	1.30E-05	1.40E-01
360	LYN	-2.93	-1.98	-3.25	-3.99	-3.67	-2.04	-2	-2.19	-1.21	-2.75	-0.42	-0.31	2.00E-08	1.30E-05	1.70E-01
361	SEC62	0.83	0.71	5.53	5.46	6.71	1.48	1.32	1.47	-1.95	1.22	-1.03	-0.04	1.70E-06	1.30E-05	3.80E-01
362	WEE1	2.3	3.38	-0.34	-1.28	-7.15	1.95	2.17	2.45	0.18	-6.37	1.37	1.6	1.70E-07	1.40E-05	9.60E-04
363	SMAP2	-1.89	-1.18	-2.29	-2.78	-1.95	-1.37	-1.38	-1.27	-3.07	-0.31	-0.79	-0.88	3.30E-07	1.40E-05	2.10E-03
364	ITM2B	1.18	1.19	1.53	1.37	1.57	1.93	1.51	1.71	-2.37	2.08	1.06	0.83	1.80E-06	1.40E-05	1.00E-02

**Table 6-27** | Top 500 Differentially regulated Genes/Proteins across species and platforms. Sorted by median adjusted *p*-value for multiple group comparison. (cont.)

Rank	Symbol	mSplPC .array	mBMPC .array	hPB .array	hPC .array	hBMPC .array	mPB .rSeq	mSplPC .rSeq	mBMPC .rSeq	hPB. .rSeq	hBMPC .rSeq	mPB .prot	mPB93 .prot	ANOVA Microarray Adj <i>p</i> -value	ANOVA RNA-Seq Adj <i>p</i> -value	ANOVA MS/MS Adj <i>p</i> -value
365	PRKACB	-2.66	-0.72	0.37	0.01	-3.81	-1.27	-1.26	-1.13	-2.02	-1.83	-0.38	-0.73	3.90E-08	1.40E-05	1.80E-02
366	LCK	-6.16	-4.28	-3.19	-3.94	-3.26	-1.2	-3.83	-4.44	-1.84	-3.46	-2.36	-2.61	2.00E-04	6.40E-06	1.50E-05
367	SEC23B	2.04	1.39	4.84	4.73	4.78	1.92	1.58	1.58	1.78	2.24	1.33	1.16	2.10E-08	1.50E-05	1.20E-03
368	TPM4	-2.16	-2.01	1.66	-1.14	0.96	-1.21	-2.13	-2.19	-0.21	1.87	-0.52	-0.17	2.20E-08	1.50E-05	1.50E-01
369	TMEM154	0.79	3.25	-0.09	-0.44	-1.47	3.97	4.38	4.32	-0.72	-2.92	2.67	3.19	1.60E-05	5.10E-08	4.10E-04
370	ARHGAP27	-1.57	-1.72	0	0	0	-1.36	-1.64	-1.7	-0.96	-1.8	-1.12	-0.97	4.90E-01	1.00E-05	1.60E-05
371	ELF4	-2.73	-2.7	-1.22	-1.81	-4.76	-1.73	-2.06	-1.74	-5.08	-4.48	-0.95	-1.57	1.80E-05	1.50E-05	5.50E-04
372	KLF2	-3.99	-3.95	-3.47	-7.02	-0.01	-7.41	-4.26	-4.79	-2.46	1.19	-1.54	-1.61	1.80E-05	5.70E-07	2.00E-02
373	ISG20	5.48	4.35	0.74	1.64	1.24	3.92	3.66	3.54	1.7	1.79	1.59	1.69	1.90E-05	4.50E-06	1.60E-03
374	CPQ	0	7.4	0.45	0.3	3.13	0.75	6.22	6.39	-0.82	4	-1.72	-2.26	2.30E-05	3.10E-06	1.90E-03
375	COPE	2.34	2.3	2.41	1.7	3.16	2.16	2.13	2.24	2.41	1.42	0.81	0.75	2.30E-05	1.30E-05	5.30E-03
376	PCK2	0.49	1.71	2.38	2.29	0.8	1.97	2.23	2.1	1.86	0.01	1.4	1.58	4.30E-05	9.90E-06	2.40E-05
377	HERC4	-1.44	-1.37	0.97	1.16	-0.76	-1.88	-1.8	-1.98	-0.78	-2.29	-0.81	-0.56	2.40E-05	2.40E-06	3.60E-03
378	LCP1	-2.37	-2.36	1.23	-0.93	-0.69	-3.03	-3.08	-2.95	-0.23	-3.35	-1.49	-0.84	1.70E-04	4.20E-07	2.60E-05
379	FGD2	-3.2	-2.46	2.29	2.12	0.18	-2.37	-3.55	-3.4	0.27	-2.16	-1.21	-1.34	2.60E-05	8.50E-07	1.90E-04
380	KIF21B	-2.15	-2.36	-0.6	-1.75	-5.2	-1.81	-2.13	-2.12	-0.51	-2.49	0.25	0.16	2.60E-05	2.10E-06	2.80E-01
381	TWF2	-3.5	-2.96	2.47	1.97	1.4	-2.81	-2.57	-2.55	0.64	-1.3	-1.2	-0.95	2.70E-05	1.10E-06	8.40E-05
382	FRY	-3.74	-2.44	-1.5	-1.5	-1.36	-2.9	-2.32	-2.03	-0.56	-5.51	-5.31	-4.12	2.50E-02	1.50E-05	2.90E-05
383	MCFD2	2.43	2.65	0.77	0.55	1.45	1.92	2.25	2.11	0.73	1.98	1.51	0.98	2.90E-05	8.50E-07	3.00E-01
384	SKA3	3.72	1.2	3.69	0.31	0	5.03	1.23	0.58	3.6	-1.67	3.98	3.78	3.00E-05	7.20E-06	1.00E-04
385	SCCPDH	0.8	4.22	2.07	1.04	1.46	1.91	2.47	2.74	-3.26	1.41	-0.35	-0.84	3.00E-05	1.30E-05	1.50E-01
386	SLAMF6	-2.03	-2.99	2.84	2.58	-0.42	0.57	-1.7	-1.58	-4.27	-2.76	1	1.6	3.10E-05	4.20E-06	4.60E-04
387	AP1B1	-1.27	-0.92	-1.38	-1.43	-1.9	-1.48	-1.14	-1.1	0.32	-1.92	-0.55	-0.43	3.20E-05	1.30E-05	2.10E-02
388	CLSPN	5.8	3.86	0.05	0.12	0.49	4.36	1.32	0.67	3.98	-1.2	5.82	5.57	1.10E-01	7.80E-06	3.60E-05
389	MSN	-1.27	-0.9	0.78	0.23	-1.34	-1.6	-2.55	-2.48	0.3	-1.72	-1.01	-1.01	4.10E-05	4.30E-07	4.60E-04
390	GMIP	-4.68	-3.4	-1.2	-1.32	-1.25	-2.36	-3.03	-2.93	-0.61	-2.2	-1.1	-0.96	4.30E-05	8.50E-06	1.30E-04
391	LRRC59	3.58	2.82	1.65	0.72	0.51	2.66	2.4	2.39	3.58	3.49	1.72	1.39	9.90E-05	6.70E-07	4.40E-05
392	SLFN5	-0.4	-2.17	3.42	7.11	1.04	-4.63	-3.6	-4.66	2.89	0.16	-4.57	-4.28	4.40E-05	4.90E-06	8.50E-04

**Table 6-27** | Top 500 Differentially regulated Genes/Proteins across species and platforms. Sorted by median adjusted *p*-value for multiple group comparison. (cont.)

Rank	Symbol	mSplPC .array	mBMPC .array	hPB .array	hPC .array	hBMPC .array	mPB .rSeq	mSplPC .rSeq	mBMPC .rSeq	hPB. .rSeq	hBMPC .rSeq	mPB .prot	mPB93 .prot	ANOVA Microarray Adj <i>p</i> -value	ANOVA RNA-Seq Adj <i>p</i> -value	ANOVA MS/MS Adj <i>p</i> -value
393	MAGT1	1.87	2.37	1.76	1.39	0.64	3.02	2.42	2.32	0.82	2.11	0.6	0.58	4.90E-05	1.60E-06	1.90E-01
394	CLN6	2.9	3.79	1.65	1.07	1.76	1.91	2.66	2.75	3.17	4.22	1.65	1.48	5.00E-05	9.00E-06	7.10E-04
395	GIMAP8	-5.7	-2.96	0	0	3.84	-1.7	-2.97	-2.8	5.38	5.59	-1.2	-1.1	1.50E-04	3.60E-07	5.60E-05
396	RHOF	1.76	3.29	-2.43	-2.49	-1.25	-1.89	-3.11	-2.53	-1.65	-2.19	-0.94	-1.05	5.60E-05	6.40E-06	1.50E-02
397	DHRS7	1.91	2.77	2.58	2.36	2.61	1.76	2.52	2.55	-0.45	1.81	0.45	-0.03	5.60E-05	6.70E-07	1.60E-01
398	DNAJC25	1.65	2.16	0.98	0.23	-1.48	2.59	2.16	2.4	-1.4	0.77	1.16	1.16	5.70E-05	1.50E-05	3.10E-02
399	TMED3	1.93	3.22	1.89	0.68	1.89	1.86	2.84	2.71	2.53	2.36	0.15	0.19	6.30E-05	3.10E-06	8.70E-01
400	ANKRD28	1.11	2.08	4.09	3.73	2.82	1	2.18	2.08	1.3	4.57	1.17	1.09	8.40E-05	3.10E-06	6.50E-05
401	GRB2	-1.23	-1.23	1.33	1.23	-0.09	-0.41	-1.34	-1.47	-0.96	-1.38	-0.74	0.01	6.70E-05	7.40E-06	5.10E-03
402	DMXL1	-1.88	-1.64	-1.58	-1.08	-2.18	-0.15	-1.51	-1.78	-2.16	-1.38	-0.27	-0.55	6.80E-05	1.20E-05	1.50E-02
403	TNIK	-3.92	-6.88	0.05	0.92	0.01	-3.58	-5.71	-9.07	3	-2.65	-2.18	-2.58	3.50E-02	4.70E-06	6.80E-05
404	SMIM14	-5.7	-2.5	-0.77	-1.53	-2.62	-3.47	-2.13	-2.07	-3.09	-1.77	-1.37	-1.19	7.10E-05	6.80E-06	1.70E-02
405	LMF1	2.01	3.9	2.03	2.65	3.97	2.05	2.14	2.27	0.62	4.48	-0.29	-0.65	7.20E-05	6.60E-06	1.30E-01
406	CD9	4.65	2.53	1.19	-2.19	3.33	3.09	2.5	2.1	-1.89	4.49	-0.34	-0.55	7.20E-05	1.90E-06	2.80E-01
407	RGS19	-1.96	-1.15	0.61	0.71	1.05	-2.17	-1.59	-1.47	-0.47	-2.65	-1.28	-1.83	2.30E-03	1.10E-05	7.50E-05
408	METTL7A1	2.18	2.63	3.3	2.85	3.71	-2.97	2.63	3.11	0.44	1.11	-3.34	-4.23	5.40E-04	7.00E-07	8.80E-05
409	TLE3	-0.98	-0.44	2.47	2.33	1.66	-3.07	-1.78	-1.72	1.4	1.76	-1.17	-0.96	6.90E-04	7.50E-06	9.10E-05
410	ATG13	1.72	2.72	1.31	1.13	0.76	1.76	1.67	1.62	2.15	1.31	0.74	0.38	9.30E-05	1.00E-06	7.60E-03
411	BTD	0.78	3.61	1.97	2.44	4.21	2.1	2.57	2.63	-0.83	2.51	1.09	0.71	9.30E-05	2.90E-06	2.20E-02
412	ABCG1	-3.08	-1.38	-1.72	-8.88	-6.05	-2.48	0.34	-0.19	-2.33	-6.25	-0.09	-0.32	9.30E-05	8.10E-06	8.00E-01
413	ZBTB38	1.64	2.45	4.38	4.72	3.33	2.17	2.34	2.16	3.03	3.87	-0.21	-0.01	9.50E-05	1.50E-06	4.80E-01
414	CALR	1.03	1.24	2.07	0.55	2.68	2.76	1.99	1.79	3.02	2.52	0.53	0.54	9.90E-05	3.50E-07	6.70E-03
415	PIM1	2.13	2.71	-0.37	-1.66	-1.88	4.55	3.89	4.15	-2.39	2.97	2.22	2.19	1.10E-04	4.30E-08	1.10E-03
416	SAMSN1	-0.65	-0.59	4.51	4.12	2.37	1.43	-0.49	-0.65	1.87	2.45	1.12	1.57	9.50E-03	1.40E-05	1.10E-04
417	CEP170B	0.65	2.74	0.25	0.03	0.86	2.39	1.54	2.16	6.69	4.41	-0.19	-0.01	1.20E-04	3.10E-06	6.20E-01
418	NUCB1	3.12	3.93	2.34	2.1	1.46	3.69	3.55	3.66	2.47	2.2	0.99	0.96	1.60E-04	9.00E-06	8.90E-04
419	APBB1IP	-2.07	-3.28	-2.56	-2.69	-2.67	-3.04	-2.89	-2.8	0.27	-1.89	-0.77	-0.98	1.60E-04	1.90E-06	1.20E-03
420	CORO1A	-2.47	-3.6	-0.39	-1.47	-1.39	-2.99	-3.7	-3.63	1.33	-3.6	-1.82	-1.22	1.60E-04	5.50E-07	3.90E-03

**Table 6-27** | Top 500 Differentially regulated Genes/Proteins across species and platforms. Sorted by median adjusted *p*-value for multiple group comparison. (cont.)

Rank	Symbol	mSpIPC .array	mBMPC .array	hPB .array	hPC .array	hBMPC .array	mPB .rSeq	mSpIPC .rSeq	mBMPC .rSeq	hPB. .rSeq	hBMPC .rSeq	mPB .prot	mPB93 .prot	ANOVA Microarray Adj <i>p</i> -value	ANOVA RNA-Seq Adj <i>p</i> -value	ANOVA MS/MS Adj <i>p</i> -value
421	AGPAT4	1.81	1.52	-3.57	-3.66	-3.57	2.65	1.17	1.08	0.07	-3.86	1.07	1.15	1.70E-04	3.20E-06	6.40E-03
422	HPS3	-3.23	-1.54	-1.06	-1.12	-3.29	-2.4	-1.85	-1.76	-3.9	-2	-0.65	-0.46	1.70E-04	1.20E-05	1.10E-02
423	IL2RA	-0.03	1.47	-0.44	-0.49	-0.43	5.37	-0.18	0.71	-2.84	-5.11	2.54	3.35	3.70E-01	1.70E-07	1.80E-04
424	GRAP2	-0.78	-2.47	0	0	0.25	-4.59	-1.62	-1.72	-3.27	-5.72	-1.53	-1.42	1.00E-01	4.70E-07	2.00E-04
425	AP3S1	2.56	3	1.16	0.56	0.88	2.91	2.79	2.82	0.47	2.52	-0.85	-1.04	2.10E-04	4.30E-07	3.50E-02
426	RAB3D	3.29	6.74	0.07	0	1.31	2.34	4.49	4.62	1.79	3.14	-0.47	-1.32	2.20E-04	5.50E-07	1.50E-02
427	PKM	1.8	0.09	2.84	2.06	1.86	1.87	-0.18	0.03	1.3	-0.57	0.54	0.69	2.30E-04	1.40E-05	2.60E-03
428	PTPN22	1.58	1.43	-0.46	-1.66	-1.42	1.94	1.78	1.58	-0.25	-1.46	0.14	-0.14	2.30E-04	1.20E-05	2.20E-01
429	MAST3	-2.79	-1.96	-0.41	-0.51	-1.14	-2.56	-1.91	-2.03	-0.14	-1.39	-0.99	-1	2.40E-02	8.20E-06	2.40E-04
430	GANAB	1.95	1.33	2.44	1.74	0.96	1.91	1.83	1.77	1.15	2.06	0.78	1.12	2.40E-04	1.40E-06	2.90E-02
431	GPR183	-3.2	-2.86	1.14	0.22	-1.84	-2.48	-3.75	-3.96	-1.14	-3.45	-0.66	-0.71	2.40E-04	1.80E-07	4.80E-01
432	STARD4	2.38	-0.84	-0.01	0.04	-2.52	1.57	-1.56	-1.45	-1.53	1.01	2.41	2.93	1.30E-03	5.50E-06	2.50E-04
433	SEC31A	2.19	1.46	2.09	1.85	2.36	1.68	1.37	1.47	1.22	1.27	0.35	0.05	2.50E-04	3.10E-06	9.30E-02
434	FGD3	-2.21	-2.1	-2.26	-2.01	-2.8	-1.16	-1.57	-1.73	0.79	-2.44	-0.5	-0.54	2.60E-04	3.10E-06	2.20E-02
435	PLCD3	3.81	4.78	0.15	0.14	0	4.11	5.37	6.11	5.76	1.45	2.66	2.29	1.30E-01	7.20E-06	2.70E-04
436	PPFIBP2	0.93	2.68	1.28	0.74	-0.08	0.26	2.3	2.23	2.46	1.39	0.96	1.02	3.80E-03	4.40E-06	3.10E-04
437	TJP2	3.02	5.66	0.01	-0.09	1.47	3.15	3.2	3.34	4.01	1.74	0.91	1.23	2.30E-02	2.10E-06	3.50E-04
438	CD83	-3.98	-2.04	-11.42	-10.96	-5.24	-2.37	-2.98	-3.3	-6.43	-7.7	-0.24	1.61	3.60E-04	6.10E-07	3.10E-02
439	ACSF2	-2	-2.88	0.32	-0.13	-0.06	-1.97	-2.54	-2.49	2.18	0.09	-1.78	-1.79	2.60E-01	8.90E-07	3.60E-04
440	EXO1	5.22	1.69	1.8	0.1	0.09	2.45	-1.05	-3.53	4.41	-1.84	2.35	2.34	5.10E-04	3.10E-06	4.00E-04
441	DGKE	2.47	1.08	0.44	0.53	1.77	-0.58	1.47	1.61	2.16	0.14	0.85	0.04	4.80E-04	4.20E-06	6.40E-02
442	C2CD5	0.48	1.37	-0.13	-0.22	-3.06	0.4	1.4	1.34	0.38	-1.45	-0.23	0.12	4.80E-04	1.10E-05	9.10E-02
443	BHLHE41	2.49	3.99	10.23	9.95	8.98	1.91	3.36	3.67	3.06	5.26	1.06	1.19	5.00E-04	1.60E-06	7.10E-02
444	LAMC1	1.7	2.61	3.18	1.61	0.88	2.15	1.46	1.25	-0.94	2.02	2.46	1.93	3.80E-03	8.10E-06	5.10E-04
445	UBE2E3	1.98	2	-2.46	-2.02	-2.3	1.66	1.51	1.78	-0.51	1.59	0.66	1.19	5.10E-04	2.90E-06	7.60E-02
446	ZFP260	0.8	1	4.29	5.22	1.99	1.69	1.57	1.48	-2.76	-2.34	-1.11	-0.62	5.30E-04	8.70E-06	5.60E-02
447	GNE	3.25	2.96	0.05	-0.69	-1.17	1.73	2.52	2.5	-1.06	1.04	0.75	0.15	5.80E-04	1.70E-06	3.00E-03
448	HSH2D	-3.01	-3.04	-1.13	0.15	1.05	-2.18	-2.66	-3.16	2.03	2.05	-0.59	0.15	5.80E-04	1.50E-05	8.20E-02

**Table 6-27** | Top 500 Differentially regulated Genes/Proteins across species and platforms. Sorted by median adjusted  $p$ -value for multiple group comparison. (cont.)

Rank	Symbol	mSpIPC .array	mBMPC .array	hPB .array	hPC .array	hBMPC .array	mPB .rSeq	mSpIPC .rSeq	mBMPC .rSeq	hPB. .rSeq	hBMPC .rSeq	mPB .prot	mPB93 .prot	ANOVA Microarray Adj $p$ -value	ANOVA RNA-Seq Adj $p$ -value	ANOVA MS/MS Adj $p$ -value
449	ZBTB32	-0.04	-1.07	3.09	1.56	0.81	3.37	-0.96	-0.34	10.15	2.64	2.36	2.51	8.70E-03	1.80E-06	8.50E-04
450	IRAK3	-5.9	-3.01	-1.86	-1.86	-1.8	-1.24	-2.53	-2.04	-0.16	-3.13	-0.04	0.03	8.70E-04	8.50E-06	9.50E-01
451	PLEKHG2	-5.78	-4.59	0.84	0.32	-0.1	-1.92	-3.23	-3.99	1.46	-3	-1.27	-1.13	2.00E-01	1.40E-06	9.30E-04
452	TVP23B	1.51	1.97	0.89	0.67	1.32	1.38	1.68	1.81	1.25	3.7	0.8	-0.14	1.10E-03	3.90E-06	1.20E-01
453	TNFAIP8L2	-8.16	-5.76	0.07	0.16	0.15	-4.59	-5.08	-5.17	-3.14	-5.77	-2.38	-0.92	4.90E-01	1.40E-05	1.10E-03
454	TNS3	4.39	4	2.18	2.37	1.63	3.05	3.77	3.81	-1.9	1.59	0.31	-0.57	1.40E-03	1.70E-07	1.20E-02
455	FMNL1	-1.76	-2.79	-4.5	-2.95	-3.93	-2.07	-2.61	-2.3	0.08	-1.89	-1.09	-0.98	1.50E-03	4.30E-07	2.50E-03
456	LMO2	-5.91	-2.46	-3.07	-2.79	-1.52	-2.63	-3.26	-3.31	-2.44	-3.71	1.71	1.08	1.60E-03	7.20E-06	8.80E-03
457	SBNO2	-1.49	-0.83	1.58	0.67	0.45	2.34	-0.89	-1.14	3.04	-0.2	-0.01	-0.09	1.60E-03	5.10E-06	8.80E-01
458	TMEM97	3.22	2.63	5.06	3.08	1.44	3.15	-0.21	0.02	2.4	0.65	1.54	2.27	4.00E-03	1.10E-06	1.80E-03
459	OPTN	-2.15	-0.96	2.25	2.67	1.77	1.48	-0.71	-1.29	2.25	2.31	1.21	0.9	2.90E-03	9.70E-06	7.60E-03
460	SH3BGRL3	-2.13	-1.84	-1.2	-1.9	-0.63	-2.99	-2.95	-2.78	-0.37	-1.04	-1.02	-0.6	2.90E-03	1.10E-05	2.70E-01
461	B2M	-1.46	-0.49	0.83	1.13	0.78	2.22	2.32	2.4	0.28	2.24	-2.51	0.23	3.00E-03	1.30E-06	6.90E-03
462	RPIA	-0.14	1.45	9.84	10.48	10.9	0.04	1.47	1.7	-5.2	-2.31	-0.36	-0.76	3.10E-03	6.50E-06	1.30E-01
463	HIF0	5.36	8.06	-0.09	0.69	0.45	6.01	6.25	6.28	5.07	5.82	-1.23	-1.42	3.70E-02	5.10E-07	3.60E-03
464	SLC38A1	-0.83	-0.67	-0.36	-1.87	-5.89	-0.13	-1.37	-1.45	-1.37	-3	0.51	2.48	3.60E-03	9.80E-06	5.90E-02
465	NAGA	-0.22	1.17	6.81	5.46	4.81	0.95	1.87	1.95	3.1	2.27	-1.58	-1.99	1.60E-02	7.30E-06	3.90E-03
466	NBEAL2	-0.43	-2.64	-1.14	-0.67	-0.28	-3.36	-3.44	-3.16	6.01	2.87	0.54	0.68	1.70E-02	8.50E-06	4.10E-03
467	IGKC	0.65	0.79	2.87	3.08	2.92	6.36	7.15	7.49	4.05	7.51	1.68	1.85	4.10E-03	3.40E-08	4.70E-02
468	AKNA	-0.83	-2.76	-0.51	-0.88	-2.37	-2.17	-2.55	-2.5	-1.95	-2.63	-0.35	-0.07	4.10E-03	1.20E-05	1.30E-01
469	TMEM243	-4.16	-2.74	-0.31	-0.8	0.29	-2.77	-3.11	-3.15	-1.47	-0.02	-1.12	-0.68	5.40E-03	4.20E-06	7.30E-03
470	SPN	4.33	2.62	1.21	1.01	1.01	3.68	2.55	2.23	6.78	4.61	1.47	1.84	5.60E-03	5.10E-07	3.50E-02
471	S1PR4	-2.69	-2.9	0.53	-0.22	0.21	-2.64	-2.89	-2.86	0.24	-3.6	-1.26	-0.42	1.50E-02	7.00E-06	5.70E-03
472	AIM2	-1.21	-1.23	5.69	6.44	3.65	-2.82	-1.26	-1.15	3.4	1.45	0.27	0.56	5.70E-03	9.30E-06	9.50E-02
473	STX5A	1.22	1.05	-0.75	-1.05	-1.02	1.88	1.58	1.55	1.24	1.42	0.3	0.47	5.70E-03	2.60E-06	4.30E-01
474	ARHGDI	-1.08	-0.94	-0.48	-2.11	-1.06	-1.67	-2.04	-1.6	-0.18	-2.45	-1.23	-1.03	7.60E-03	4.80E-06	6.10E-03
475	PSTPIP1	-6.3	-3.72	-1.67	-1.59	-0.95	-2.73	-3.57	-2.95	-2.75	-3.94	-0.3	-0.08	6.40E-03	4.40E-06	4.50E-01
476	ALDOC	-1.61	1.02	3.33	1.86	1.53	2.6	-0.44	0.3	-2.28	1.19	-0.33	-0.89	6.50E-03	1.30E-05	6.10E-02

**Table 6-27** | Top 500 Differentially regulated Genes/Proteins across species and platforms. Sorted by median adjusted *p*-value for multiple group comparison. (cont.)

Rank	Symbol	mSplPC .array	mBMPC .array	hPB .array	hPC .array	hBMPC .array	mPB .rSeq	mSplPC .rSeq	mBMPC .rSeq	hPB. rSeq	hBMPC .rSeq	mPB .prot	mPB93 .prot	ANOVA Microarray Adj <i>p</i> -value	ANOVA RNA-Seq Adj <i>p</i> -value	ANOVA MS/MS Adj <i>p</i> -value
477	LAX1	1.1	0.97	5.62	6.07	5.43	2.16	2.04	2.12	4.05	4.7	0.15	0.06	6.80E-03	5.10E-07	8.30E-01
478	TMC8	-1.3	-1.76	-2.28	-3.69	-3.95	-1.93	-2.14	-2.43	-2.46	-4.35	-0.55	-0.31	7.20E-03	3.90E-06	1.20E-01
479	PIK3CD	-0.81	-1.64	-3.37	-6.14	-5.37	0.17	-1.24	-1.6	-2.68	-5.4	-0.49	-0.26	7.70E-03	2.20E-06	4.90E-02
480	MAP7	-0.01	-0.2	2.75	1.67	3.21	-2.91	-3.74	-3.58	1.13	3.99	-0.41	-0.67	8.30E-03	8.60E-06	3.60E-01
481	TWSG1	-2.61	-0.3	2.35	1.62	1.1	3.1	0.83	0.46	0.71	3.15	-0.19	0.75	8.50E-03	7.40E-06	2.00E-01
482	CARMIL1	-1.06	-1.15	0	-0.34	3.13	-3	-0.51	-0.16	-1.38	2.27	-1.06	-0.66	1.20E-02	6.80E-06	3.20E-02
483	GALNS	2.96	2.27	-0.41	-0.53	-1.39	1.93	2.39	2.47	2.89	-0.01	0.65	0.42	1.30E-02	1.40E-06	2.80E-01
484	TM9SF1	0.8	0.51	5.72	5.63	6.05	1.4	1.07	1.14	0.01	2.43	0.94	0.67	2.50E-01	1.40E-05	1.50E-02
485	PLBD2	1.79	3.61	0.87	0.76	0.65	1.24	2.14	2.24	0.87	2.75	-0.39	-1.01	1.60E-02	8.90E-07	2.20E-02
486	CAPN5	0.5	3.56	0	0	0.08	1.05	2.54	2.65	6.07	3.2	1.29	0.16	1.90E-01	2.30E-06	2.30E-02
487	RETREG3	-1.27	-1.06	0.12	0.52	-0.09	-1.77	-0.9	-0.61	-2.26	-0.31	-0.8	-0.7	1.80E-01	5.90E-06	3.00E-02
488	MYO1E	-3.82	-4.28	0	0.02	0.47	-1.9	-2.15	-2.46	-1.77	-7.13	-0.25	-0.05	3.10E-02	6.70E-06	4.60E-01
489	GBP7	-0.63	0.79	-3.08	-3.07	-3.07	-1.88	1.14	0.78	-6.81	-3.11	-0.21	0.24	3.30E-02	8.90E-06	2.40E-01
490	FHIT	2.42	6.85	0	0	0.19	1.43	3.48	3.67	-4.85	1.5	-0.37	-0.66	1.60E-01	1.00E-05	6.10E-02
491	SPR	3.21	3.99	0.68	0.93	0.04	0.48	2.46	2.33	3.37	4.91	-0.52	-0.5	7.90E-02	4.30E-06	1.70E-01
492	ADPRM	-0.12	0.45	-3.64	-3.33	-2.24	1.8	0.51	0.61	-1.41	-2.8	0.06	1.18	8.80E-02	1.00E-05	2.40E-01
493	MFSD2A	3.04	2.77	0.07	-0.11	-0.4	5.01	1.64	1.36	5.78	4.82	0.96	1.43	1.90E-01	1.40E-05	1.20E-01
494	IFT20	3.7	2.92	-0.6	-0.48	-0.75	2.84	2.31	2.29	1.18	1.1	-0.36	0.06	1.20E-01	3.70E-06	6.70E-01
495	IGHV4-1	0.39	4.23	0	0	0.06	7.24	8.94	9.51	8.99	12.92	0.43	0.35	1.90E-01	4.70E-08	1.40E-01
496	LIMD2	-3.68	-1.48	-0.65	-0.53	-0.14	-2.24	-1.98	-1.63	0.2	-4.01	-0.51	0.45	1.40E-01	1.20E-05	4.20E-01
497	SFMBT2	-0.07	-0.32	1.7	3.07	0.3	5.86	-2.89	0.83	3.03	3.96	-1.7	-0.66	2.50E-01	6.30E-06	1.70E-01
498	CHST3	-0.01	0	0	0	0.16	-2.22	-3.62	-4.02	3.7	4.99	-0.45	-0.32	4.30E-01	9.20E-08	1.80E-01
499	JUND	0	0.05	-2.28	-1.93	-1.31	-1.64	0.48	0.32	-1.72	2.54	0.5	0.24	1.90E-01	3.60E-06	3.30E-01
500	SIPA1L2	-4.75	-3.04	0	0	0.02	-1.2	-2.7	-3.55	6.07	2.34	-0.62	-0.03	1.90E-01	1.30E-05	3.40E-01