

Arsenite methyltransferase—a unitasking arsenite methylator or a hidden protein moonlighter?

Andrew Little

Department of Pharmacology & Therapeutics

McGill University, Montréal

A thesis submitted to McGill University in partial fulfilment of the requirements of the degree of  
Master of Science

May 2024

© Andrew Little, 2024

## ACKNOWLEDGEMENTS

I wish to express my sincerest gratitude to my supervisor Dr. Koren Mann for providing the opportunity to work on this project and for her guidance during the entire time. I benefited from the learning environment in the lab and it has spurred my professional and personal growth greatly.

Cynthia Guilbert, our lab manager, deserves special acknowledgement for her work on this project, such as isolating and establishing the mouse embryonic fibroblast populations as well as performing the IP-MS experiment. I also express my sincere appreciation to Cynthia Guilbert for the valuable training and mentorship.

To the members of my thesis advisory committee, Dr. Paul Clarke, Dr. Jean-François Trempe, and Dr. Anne McKinney, I express gratitude for directing my progress during the development of this project.

I thank the members of the Lady Davis Institute Flow Cytometry Core and Cell Imaging Core, Christian Young and Matthew Duguay, for assistance with cell sorting and confocal microscopy. I thank Christian and Matthew for teaching me about these methods and, at the same time, for making the cell imaging and sorting experiences memorable with many laughs.

Thank you to Dr. Sonia del Rincón, Dr. Samuel Preston, and Madelyn Abraham for the valuable input and ideas shared during lab meetings.

Lastly, I give special thanks to my lab colleagues Ayse Nazlı Zengin, Braeden Giles, Rowa Bakadlag, and Roni Juran for their friendship, support, and encouragement along this journey and for the zest that they each brought to the lab environment.

## ABSTRACT / RÉSUMÉ SCIENTIFIQUE

### *English Abstract*

**Background:** Inorganic arsenic is metabolised by arsenic 3-methyltransferase (AS3MT) to form methylated arsenic species that are excreted more rapidly. AS3MT was suspected to have an additional function in view of its unexplained evolutionary conservation together with the stress response and slower proliferation exhibited basally by *As3mt* knockout cells.

**Aim:** To characterise *As3mt* knockout versus wild-type cells and evaluate AS3MT-dependent basal differences from a proteomic standpoint. The hypothesis is that AS3MT expression affects cell function via a basal activity independent of arsenic methylation.

**Methods:** Wild-type and *As3mt* knockout mouse embryonic fibroblasts (MEFs) were exposed to 1.0  $\mu$ M inorganic arsenic for 24 hours. Together with unexposed control cells, total protein extracts were collected for LC-MS from one wild-type and two *As3mt* knockout cell lines (N = 4 and N = 3 respectively). The spectral count data were analysed by unpaired Student's T-tests ( $\alpha = 0.05$ ) followed by pathway enrichment analysis to screen for pathways altered basally in an AS3MT-associated manner. Hits were validated by immunoblot to verify associations with AS3MT expression. Growth curve, scratch wound migration, and collagen contractility experiments were conducted to evaluate whether the *As3mt* knockout MEFs exhibit a phenotypic difference. Lastly, *As3mt* addback MEFs were harnessed to test the AS3MT dependence of protein expression differences and the growth phenotype. IP-MS was carried out separately to predict AS3MT interacting partners in AML12 hepatocytes.

**Results:** Pathway enrichment analysis by Enrichr revealed that the basal expression differences correspond with endocytic trafficking, cell adhesion, and cytoskeletal integrity, which aligned with the pathways represented among identified IP-MS hits. Immunoblots validated the expression differences detected by LC-MS, as they showed among the *As3mt* knockout MEFs a

decreased basal expression of  $\beta$ -catenin and vimentin and the increased basal expression of HO-1, VPS29, and myoferlin relative to the wild-type MEFs. Subsequent immunoblots and growth curves indicated that the *As3mt* addback MEFs neither phenocopied the protein expression of the wild-type control population nor showed a higher growth rate than the empty vector controls. In parallel, the growth curve, wound healing, and collagen contractility experiments showed wide phenotypic variability between wild-type MEF populations and disaffirmed the hypotheses that *As3mt* knockout is associated with slower basal proliferation, migration, and contractile phenotypes.

**Discussion:** The analysis of both the LC-MS data from MEFs and IP-MS data from AML12 cells pointed to the hypothesis that *As3mt* knockout MEF populations experience a disturbance in the basal expression of proteins implicated in cell adhesion and endocytic trafficking. Although individual targets were validated with respect to an expression difference between the initially used wild-type control and *As3mt* knockout populations, the AS3MT associations became inconsistent in view of the variability among subsequently introduced wild-type controls. Furthermore, the data from the addback MEFs did not support AS3MT dependence, as there was neither an AS3MT-dependent change in growth rate nor expression of LC-MS hits upon *As3mt* reconstitution.

## Résumé en Français

**Contexte:** L'arsenic inorganique est métabolisé par l'arsenic 3-méthyltransférase (AS3MT).

Cette réaction forme des espèces d'arsenic méthylées qui sont excrétées plus rapidement.

AS3MT est soupçonné de posséder une fonction supplémentaire en raison de sa conservation évolutive ainsi que la réponse au stress et le taux de prolifération plus lent manifesté à la base par les cellules *As3mt* knock-out.

**Objectif:** La caractérisation des cellules *As3mt* knock-out et l'évaluation des différences basales qui dépendent de l'expression d'AS3MT d'un point de vue protéomique. L'hypothèse est que l'expression d'AS3MT affecte la fonction cellulaire via une activité basale indépendante de la méthylation de l'arsenic.

**Méthodes:** Les fibroblastes embryonnaires de souris (MEFs) de type sauvage et *As3mt* knock-out ont été exposés à 1,0  $\mu\text{M}$  d'arsenic inorganique pendant 24 heures. Avec les cellules non exposées, des extraits de protéines ont été prélevés d'un type sauvage et de deux lignées cellulaires *As3mt* knock-out pour la CL-SM (N = 4 et N = 3 respectivement). Les données de comptage spectral ont été analysées par un test T de Student non apparié ( $\alpha = 0,05$ ) suivi d'une analyse d'enrichissement des voies pour dépister les voies altérées basalement de manière associée à AS3MT. Les résultats ont été validés par immunoblot pour vérifier les associations avec l'expression d'AS3MT. Des expériences de courbe de croissance, de migration et de contractilité du collagène ont été menées afin d'évaluer si les MEFs *As3mt* knock-out présentent une différence phénotypique. Enfin, des MEFs reconstitués avec *As3mt* ont été utilisés pour tester la dépendance AS3MT des différences d'expression des protéines et du phénotype de croissance. IP-SM a été réalisée séparément pour prédire les partenaires d'interaction d'AS3MT dans les hépatocytes AML12.

**Résultats:** L'analyse de l'enrichissement de la voie par Enrichr a révélé que les différences d'expression basale correspondent au trafic endocytaire, à l'adhésion cellulaire et à l'intégrité cytosquelettique, qui s'alignent avec les voies représentées parmi les résultats IP-SM. Les immunoblots ont validé les différences d'expression détectées par LC-MS, car ils ont montré chez les MEFs *As3mt* knock-out une diminution de l'expression basale de la  $\beta$ -caténine et de la vimentine et de l'augmentation de l'expression basale de HO-1, VPS29 et myoferlin par rapport aux MEFs de type sauvage. Les données produites à partir de la reconstitution d'*As3mt* (les immunoblots et une courbe de croissance) ont indiqué que les MEFs d'addback *As3mt* ne présentaient ni une expression de protéines similaire au MEF de type sauvage, ni un taux de croissance supérieur aux contrôles vectoriels vides. En parallèle, la courbe de croissance, la cicatrisation des plaies et les expériences de contractilité du collagène ont montré une grande variabilité phénotypique entre les populations de MEF de type sauvage et ont invalidé les hypothèses selon lesquelles *As3mt* knockout est associé à une prolifération basale, à une migration, et à des phénotypes contractiles plus lente.

**Discussion:** L'analyse des données CL-SM des MEFs et des données IP-MS des cellules AML12 a mis en évidence l'hypothèse que les populations de MEF *As3mt* knock-out subissent une perturbation de l'expression basale des protéines impliquées dans l'adhésion cellulaire et le trafic endocytaire. Bien que les cibles individuelles aient été validées par rapport à une différence d'expression entre les populations de contrôle de type sauvage initialement utilisées et les populations de type *As3mt* knock-out, les associations avec AS3MT sont devenues incohérentes compte tenu de la variabilité entre les témoins de type sauvage introduits par la suite. En outre, les données provenant que la reconstitution d'*As3mt* ne soutenaient pas la

dépendance à celle-ci AS3MT, car il n'y avait pas de changement dans le taux de croissance ni d'expression des protéines identifiées par CL-SM suite à la reconstitution d'*As3mt*.

## TABLE OF CONTENTS

<b>ACKNOWLEDGEMENTS.....</b>	<b>2</b>
<b>ABSTRACT / RÉSUMÉ SCIENTIFIQUE.....</b>	<b>3</b>
English Abstract.....	3
Résumé en Français.....	5
<b>TABLE OF CONTENTS.....</b>	<b>8</b>
<b>CONTRIBUTION OF AUTHORS.....</b>	<b>11</b>
<b>CONFLICTS OF INTEREST STATEMENT.....</b>	<b>11</b>
<b>LIST OF ABBREVIATIONS.....</b>	<b>12</b>
<b>LIST OF FIGURES.....</b>	<b>15</b>
<b>LIST OF TABLES.....</b>	<b>16</b>
<b>1. INTRODUCTION &amp; LITERATURE REVIEW.....</b>	<b>17</b>
<b><i>1.1 Introduction to Arsenic.....</i></b>	<b><i>17</i></b>
<b><i>1.2 Introduction to Arsenite (+3 oxidation state) Methyltransferase.....</i></b>	<b><i>17</i></b>
<b><i>1.2.1 Arsenic Metabolism.....</i></b>	<b><i>17</i></b>
<b><i>1.2.2 AS3MT Expression.....</i></b>	<b><i>21</i></b>
<b><i>1.2.3 The AS3MT Protein.....</i></b>	<b><i>21</i></b>
<b><i>1.2.4 The AS3MT Gene.....</i></b>	<b><i>22</i></b>
<b><i>1.2.4.1 GWAS related to the AS3MT locus.....</i></b>	<b><i>22</i></b>
<b><i>1.2.5 Evolutionary development of AS3MT.....</i></b>	<b><i>25</i></b>
<b><i>1.2.5.1 Conservation of AS3MT.....</i></b>	<b><i>25</i></b>
<b><i>1.2.6 Additional AS3MT Activity.....</i></b>	<b><i>30</i></b>
<b><i>1.2.6.1 AS3MT-dependent metabolic differences in mice.....</i></b>	<b><i>30</i></b>
<b><i>1.2.6.2 Slower proliferation associated with As3mt knockout.....</i></b>	<b><i>31</i></b>
<b><i>1.2.6.3 Elevated basal HO-1 expression associated with As3mt knockout.....</i></b>	<b><i>31</i></b>
<b>2. HYPOTHESIS, RATIONALE, &amp; RESEARCH OBJECTIVES.....</b>	<b>32</b>
<b><i>2.1 Hypothesis.....</i></b>	<b><i>32</i></b>
<b><i>2.2 Rationale.....</i></b>	<b><i>32</i></b>
<b><i>2.3 Research Objectives.....</i></b>	<b><i>33</i></b>
<b>3. MATERIALS &amp; METHODS.....</b>	<b>35</b>
<b><i>3.1 Chemicals and Reagents.....</i></b>	<b><i>35</i></b>
<b><i>3.2 Arsenite exposure.....</i></b>	<b><i>35</i></b>
<b><i>3.3 As3mt knockout mice and isolation of MEFs.....</i></b>	<b><i>35</i></b>
<b><i>3.4 Cell culture.....</i></b>	<b><i>36</i></b>
<b><i>3.4.1 Mouse Embryonic Fibroblasts.....</i></b>	<b><i>36</i></b>
<b><i>3.4.2 Murine AML12 Hepatocytes.....</i></b>	<b><i>36</i></b>
<b><i>3.4.3 Platinum-E 293T cells.....</i></b>	<b><i>37</i></b>
<b><i>3.5 Retroviral transduction.....</i></b>	<b><i>37</i></b>
<b><i>3.5.1 Plat-E transfection and virion production.....</i></b>	<b><i>37</i></b>
<b><i>3.5.2 MEF transduction.....</i></b>	<b><i>38</i></b>

<b>3.6 <i>As3mt</i> Reconstitution</b> .....	<b>38</b>
3.6.1 Cloning the <i>As3mt</i> insert into a pMIG construct.....	38
3.6.2 Site-directed mutagenesis of <i>As3mt</i> .....	39
3.6.3 Transduction of <i>As3mt</i> <sup>-/-</sup> MEFs for <i>As3mt</i> reconstitution.....	40
<b>3.7 Flow Cytometry: sorting GFP-positive MEFs</b> .....	<b>40</b>
<b>3.8 <i>As3mt</i> Genotype Analysis</b> .....	<b>41</b>
<b>3.9 Sry Genotype Analysis</b> .....	<b>41</b>
<b>3.10 RT-qPCR</b> .....	<b>42</b>
<b>3.11 Protein extraction, quantification, immunoblotting</b> .....	<b>42</b>
3.11.1 Protein extraction.....	42
3.11.2 Protein quantification by Bradford Assay.....	43
3.11.3 Immunoblotting.....	43
<b>3.12 Liquid Chromatography-Mass Spectrometry</b> .....	<b>44</b>
<b>3.13 Immunoprecipitation-Mass Spectrometry</b> .....	<b>47</b>
<b>3.14 Immunofluorescence</b> .....	<b>47</b>
<b>3.15 Confocal Microscopy</b> .....	<b>49</b>
<b>3.16 Growth Curves</b> .....	<b>50</b>
3.16.1 Growth curve seeding.....	50
3.16.2 Crystal Violet staining.....	50
<b>3.17 Scratch Wound Migration Assay</b> .....	<b>51</b>
<b>3.18 Collagen Contractility Assay</b> .....	<b>51</b>
<b>3.19 Statistical Analyses</b> .....	<b>52</b>
3.19.1 Analysis of LC-MS spectral count data.....	52
3.19.2 Image analysis of immunofluorescent VPS35 punctal staining.....	53
3.19.3 Analysis of phenotypic assay data.....	53
<b>4. RESULTS</b> .....	<b>54</b>
<b>4.1 LC-MS Screen of AS3MT-associated and arsenic-induced protein expression differences</b> .....	<b>54</b>
<b>4.2 IP-MS search for putative AS3MT interacting partners</b> .....	<b>59</b>
<b>4.3 Adherens Junctions and Cytoskeleton</b> .....	<b>60</b>
<b>4.4 Vimentin phenotype of wild-type and <i>As3mt</i><sup>-/-</sup> MEFs</b> .....	<b>64</b>
4.4.1 Immunofluorescent imaging of vimentin in MEFs.....	65
4.4.2 Functional comparison between wild-type and <i>As3mt</i> <sup>-/-</sup> MEFs.....	66
<b>4.5 Endocytic trafficking phenotype of wild-type and <i>As3mt</i><sup>-/-</sup> MEFs</b> .....	<b>68</b>
<b>4.6 Hypothesis testing with expanded model and <i>As3mt</i> addback MEFs</b> .....	<b>73</b>
4.6.1 Expanding MEF model with more wild-type controls.....	73
4.6.2 Hypothesis testing with <i>As3mt</i> addback MEFs.....	75
<b>5. DISCUSSION</b> .....	<b>78</b>
<b>5.1 Summary and interpretation of results</b> .....	<b>79</b>
5.1.1 IP-MS data.....	79

5.1.2 Comparison to existing AS3MT knockout screening.....	80
5.1.3 Adherens Junctions.....	81
5.1.4 Cytoskeletal Organisation and Vimentin.....	82
5.1.4.1 Vimentin staining pattern in MEFs.....	83
5.1.4.2 Scratch wound migration and collagen contractility assays.....	85
5.1.5 Endocytic Trafficking.....	85
5.1.5.1 Retromer complex phenotype of MEFs.....	87
<b>5.2 Limitations.....</b>	<b>87</b>
5.2.1 Narrow scope of the proteomics approach in MEFs.....	87
5.2.2 Limitations of the MEF model.....	88
5.2.3 LC-MS and data analysis.....	90
5.2.4 Immunoblot validation of LC-MS hits.....	91
5.2.5 Proliferation curves of As3mt addback MEFs.....	91
5.2.6 IP-MS and data analysis.....	92
<b>5.3 Future directions.....</b>	<b>93</b>
<b>5.4 Implications.....</b>	<b>93</b>
<b>6. CONCLUSION.....</b>	<b>95</b>
<b>7. REFERENCES.....</b>	<b>96</b>
<b>APPENDIX A: Supplementary Figures and Tables.....</b>	<b>110</b>
<b>APPENDIX B: Macro program for VPS35 punctal quantification.....</b>	<b>115</b>

## CONTRIBUTION OF AUTHORS

The following people contributed to the work presented in this thesis:

- Dr. Koren Mann designed the project and supervised the research.
- Cynthia Guilbert established the mouse embryonic fibroblasts by isolating the primary MEFs and immortalised four of the six populations that were used in the experiments. In addition, Cynthia Guilbert transfected the AML12 hepatocytes and carried out the IP-MS on these cells. Cynthia Guilbert also provided assistance with molecular cloning.
- The liquid chromatography-mass spectrometry was done through the RI-MUHC proteomics platform by Lorne Taylor, Amy Wong, and Jennifer Nedow.
- Dr. Jean-François Trempe ran the MaxQuant program for label-free quantification of mass spectrometry data.
- Braeden Giles wrote the Macro program for the quantification of puncta in confocal microscopy images.
- Huizhong Huang performed RT-qPCR for the quantification of *Vps35*, *Vps29*, and *Vps26a* mRNA.

Unless specified above, all work in this thesis was completed by Andrew Little. This thesis was written by Andrew Little and edited primarily by Dr. Koren Mann. The methods section and French abstract were also edited by Cynthia Guilbert.

## CONFLICTS OF INTEREST STATEMENT

No conflicts of interest, financial or otherwise, are declared by the authors.

## LIST OF ABBREVIATIONS

AF	Alexa Fluor™
AML12	Murine hepatocyte cells (Alpha Mouse Liver 12)
<i>AS3MT</i>	Arsenite (+3 oxidation state) methyltransferase (human gene at 10q24.32 locus)
<i>As3mt</i>	Arsenite methyltransferase (murine gene)
AS3MT	Arsenite methyltransferase protein
bp	Base pairs
BSA	Bovine serum albumin
CI-M6PR	Cation independent mannose 6-phosphate receptor
CIMT	Carotid intima-medial thickness
CSC	Cargo-selective complex (retromer recycling complex)
DAPI	4',6-diamidino-2-phenylindole
DEGs	Differentially expressed genes
DIC	Differential interference contrast
DMEM	Dulbecco's Modified Eagle's Medium
ECM	Extracellular matrix
EV	Empty Vector
FBS	Fœtal bovine serum
GO analysis	Gene Ontology pathway enrichment analysis
GWAS	Genome-wide association study
HO-1	Hæm oxygenase 1 (HMOX1) protein
ICD	Inclusion-Cell Disease
IP	Immunoprecipitation

IP-MS	Immunoprecipitation-Mass Spectrometry
IRES	Internal ribosome entry site
KEGG	Kyoto Encyclopædia of Genes and Genomes
KO	Knockout
LAMP	Lysosome-associated membrane protein
LC-MS	Liquid chromatography-mass spectrometry
LFQ	Label-free quantification
M6P	Mannose-6-phosphate
M6PR	Mannose-6-phosphate receptor
MEF	Mouse embryonic fibroblast
NCATS	National Center for Advancing Translational Sciences
NS	Not significant
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
PMI	Primary methylation index
PML	Promyelocytic leukæmia protein
ppb	Parts per billion
ppm	Parts per million
RNAi	RNA interference
ROI	Region of interest
RT-qPCR	Quantitative reverse transcription polymerase chain reaction
SEM	Standard error of the mean

siRNA	Small interfering RNA
SMI	Secondary methylation index
SNP	Single-nucleotide polymorphism
SRP	Signal recognition particle
TGN	Trans-Golgi network
TJP	Tight junction proteins TJP1 and TJP2 (also known as ZO-1 and ZO-2)
<i>Vim</i>	Murine vimentin gene
VPS	Vacuolar sorting protein
WT	Wild-type genotype
ZO	Zona occludens proteins ZO-1 and ZO-2 (also known as TJP1 and TJP2)

## LIST OF FIGURES

<b>Figure 1.</b> Two posited reaction schemes for AS3MT-catalysed arsenite methylation.....	18
<b>Figure 2.</b> Domains of the human AS3MT protein.....	22
<b>Figure 3.</b> Clustal Omega multiple sequence alignment of orthologous <i>AS3MT</i> and <i>ArsM</i> genes.....	27–28
<b>Figure 4.</b> Preliminary AS3MT-associated growth and HO-1 expression data in MEFs.....	31
<b>Figure 5.</b> LC-MS experimental design for screening of protein expression differences.....	54
<b>Figure 6.</b> Differentially expressed proteins detected in MEFs by LC-MS.....	55
<b>Figure 7.</b> Differential adherens junction-related protein expression between wild-type and <i>As3mt</i> knockout MEFs.....	61
<b>Figure 8.</b> <i>As3mt</i> knockout MEFs express lower levels of $\beta$ -catenin basally.....	64
<b>Figure 9.</b> <i>As3mt</i> knockout MEFs express lower levels of vimentin basally.....	65
<b>Figure 10.</b> Non-filamentous vimentin staining pattern in MEFs.....	66
<b>Figure 11.</b> Scratch wound migration of wild-type and <i>As3mt</i> knockout MEFs.....	67
<b>Figure 12.</b> Collagen contractility of wild-type and <i>As3mt</i> knockout MEFs.....	68
<b>Figure 13.</b> Differential endocytic trafficking-related protein expression between wild-type and <i>As3mt</i> knockout MEFs.....	69
<b>Figure 14.</b> Immunofluorescent VPS35 staining in wild-type and <i>As3mt</i> <sup>-/-</sup> MEFs.....	71
<b>Figure 15.</b> Variability in expression of mass spectrometry hits among wild-type MEF populations.....	74
<b>Figure 16.</b> Fibronectin, vimentin, and HO-1 expression do not show AS3MT dependence in MEFs.....	75
<b>Figure 17.</b> Growth rate does not show AS3MT dependence in MEFs.....	77
<b>Supplementary Figure 1.</b> Confirmation of <i>As3mt</i> Genotype.....	110
<b>Supplementary Figure 2.</b> mRNA expression of retromer complex subunits in wild-type and <i>As3mt</i> knockout MEFs.....	111
<b>Supplementary Figure 3.</b> Sex determination of MEFs by <i>Sry</i> genotype analysis.....	112
<b>Supplementary Figure 4.</b> Fibronectin expression in wild-type and <i>As3mt</i> <sup>-/-</sup> MEFs.....	112

## LIST OF TABLES

<b>Table 1.</b> Summary of genome-wide associations of the 10q24.32 locus with pathologies..	23
<b>Table 2.</b> Summary of primary antibodies used in immunoblots.....	44
<b>Table 3.</b> Secondary antibodies used in immunoblots.....	44
<b>Table 4.</b> Summary of antibodies used in immunofluorescence experiments.....	48
<b>Table 5.</b> Pathway enrichment analysis of AS3MT-associated basal proteomic differences.....	58
<b>Table 6.</b> Hits from the FLAG <sup>®</sup> -AS3MT IP-MS experiment in AML12 hepatocytes.....	60
<b>Table 7.</b> AS3MT-associated expression differences of adherens junction proteins.....	63
<b>Supplementary Table 1.</b> Primers used in PCR reactions.....	113
<b>Supplementary Table 2.</b> Primers used for RT-qPCR.....	113
<b>Supplementary Table 3.</b> List of adherens junction adapter proteins.....	114

## 1. INTRODUCTION & LITERATURE REVIEW

### *1.1 Introduction to Arsenic*

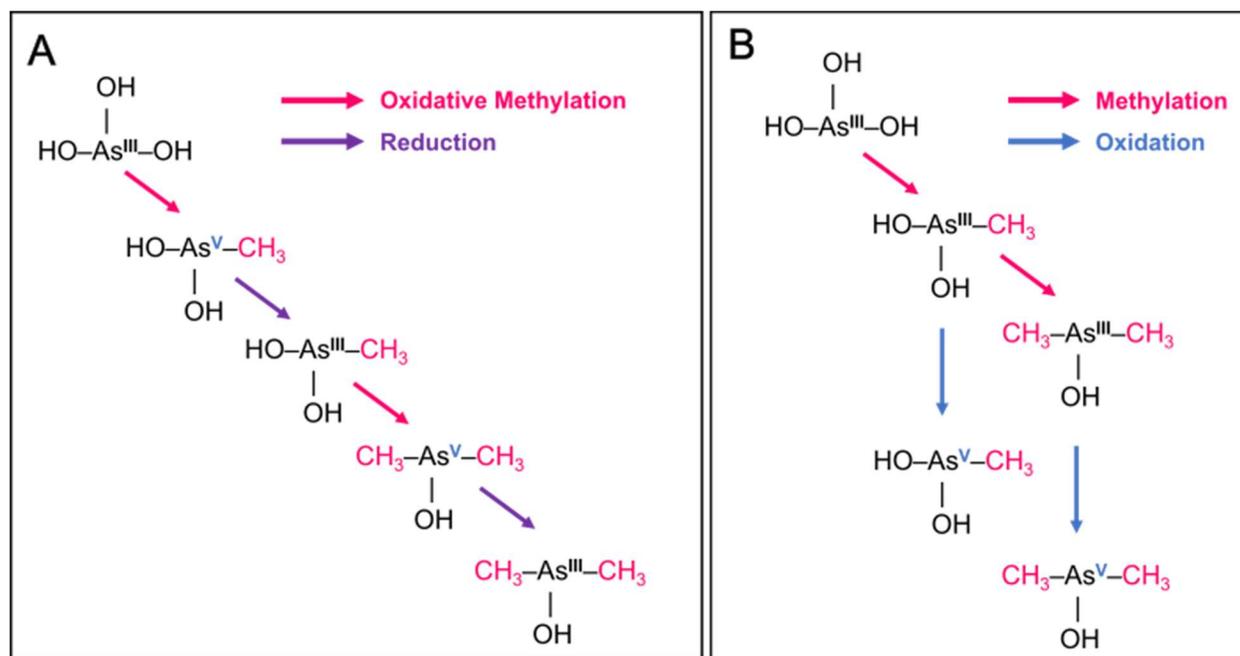
Arsenic is a toxic metalloid that poses a public health risk due to its presence in water and food as a contaminant and due to the fact that it is toxic at low exposure levels (Ghosh & Sil, 2015). Arsenic is naturally abundant in the Earth's surface and the most common exposure sources are contaminated groundwater, rice, seafood, and industrial dusts (Mantha et al., 2017; de Rosemond et al., 2008; Korte & Fernando, 1991; Bhumbla & Keefer, 1994). Acute arsenic toxicity presents with symptoms such as nausea, vomiting, abdominal pain, diarrhoea, and peripheral neuropathy (Ratnaike, 2003), whereas the outcomes of chronic exposure are increased risk of cancer (Speer et al., 2023), diabetes (Sung et al., 2015), cardiovascular disease (Simeonova & Luster, 2004; Nong et al., 2016; Chen et al., 2013; Negro Silva et al., 2017), and, in arsenicosis-endemic areas, neurodevelopmental disorders, and skin lesions (Tolins et al., 2014; Smith et al., 2000).

### *1.2 Introduction to Arsenite (+3 oxidation state) Methyltransferase*

#### *1.2.1 Arsenic metabolism*

Arsenite (+3 oxidation state) methyltransferase (AS3MT) is the enzyme that catalyses adenosylmethionine-dependent methylation of trivalent arsenicals (Stýblo et al., 2021). It and its orthologues are expressed among all domains of life (Chen et al., 2017a). The only established function of the AS3MT protein is the catalysis of arsenite methylation. The biomethylation of trivalent arsenic species generates mono-, di-, and trimethylated arsenic metabolites, and there are two putative pathways for the arsenic methylation pathway: the Challenger scheme and the alternate scheme (Stýblo et al., 2021) (**Figure 1**). These pathways differ because the Challenger scheme proposes alternation between oxidative methylation and reduction of arsenic species (Challenger, 1947), whereas the alternate scheme posits that the addition of methyl groups does

not entail oxidation-reduction cycling, instead suggests that the oxidation of methylated trivalent arsenicals occurs after their generation (Currier et al., 2011). Intracellular reductants such as glutathione, thioredoxin, thioredoxin reductase, and NADPH are essential for the methylation reaction (Hayakawa et al., 2005), as they support the reaction scheme either by providing reducing equivalents directly for arsenic reduction or by maintaining catalytic cysteines in the reduced state (Hayakawa et al., 2005; Mersaoui et al., 2022; Waters et al., 2004). It has been reported that the arsenite-glutathione conjugate—not arsenite—is the true substrate for the AS3MT-catalysed methylation (Hayakawa et al., 2005). It is important to note that the arsenite methylation reaction is involved in one-carbon metabolism because its activity is dependent on the consumption of S-adenosyl-L-methionine (SAM) as the methyl donor (Mersaoui et al., 2022; Buchet et al., 1985).



**Figure 1. Two posited reaction schemes for AS3MT-catalysed arsenite methylation. (A) Challenger scheme. (B) Alternate scheme.**

The physiological purpose of arsenite methylation is to increase the rate of arsenic clearance by urinary excretion, thereby lowering the whole-body arsenic burden (Stýblo et al.,

2000; Stýblo et al., 1995; Vahter, 1999; Thomas et al., 2001; Thomas et al., 2007). AS3MT is solely responsible for metabolising arsenic and producing methylated arsenic metabolites, and conversely, the only established function of AS3MT is arsenite biomethylation. Deficiency in AS3MT activity is in part deleterious because it permits accumulation of inorganic arsenic in tissues and raises susceptibility to arsenic-induced toxicity (Chen et al., 2011; Thomas et al., 2001; Drobná et al., 2009; Dodmane et al., 2013). On the other hand, ablation of AS3MT activity is also beneficial to an extent, as AS3MT is required for the promotion of arsenic-induced atherosclerosis and *As3mt* knockout in mice is protective against this pathogenesis (Negro Silva et al., 2017).

Arsenic methylation is considered both a detoxification process and an activation process at the same time (Stýblo et al., 2021), as some of the methylated arsenic intermediates in the pathway are more toxic and reactive than inorganic arsenic (Thomas et al., 2007; Stýblo et al., 2000). While pentavalent metabolites  $\text{MAs}^{\text{V}}$ ,  $\text{DMAs}^{\text{V}}$ , and  $\text{TMA}^{\text{VO}}$  are less toxic than inorganic arsenic (Vahter & Concha, 2001), trivalent methylated arsenicals are more cytotoxic and reactive than inorganic arsenic and their pentavalent counterparts (Stýblo et al., 2000). For example,  $\text{MAs}^{\text{III}}$  and  $\text{DMAs}^{\text{III}}$  are more highly deposited in tissues compared to  $\text{MAs}^{\text{V}}$  and  $\text{DMAs}^{\text{V}}$ , which are excreted relatively rapidly (Vahter & Concha, 2001).

The differences have not been entirely resolved between the biological activities of the downstream metabolites and how they relatively contribute to arsenic toxicity, yet methylation indices have been established as a framework assess urinary arsenic profiles (Wei et al., 2018; Yamauchi & Takata, 2021). These indices entail the calculation of ratios between methylated and unmethylated arsenicals: the primary methylation index (PMI) quantifies how efficiently the first methylation reaction occurs (production of monomethylated arsenicals), whereas the

secondary methylation index (SMI) expresses the rate of the second methylation (generation of dimethylated arsenicals) (Wei et al., 2018; Yamauchi & Takata, 2021). A high PMI indicates a higher proportion of monomethylated arsenic metabolites compared to inorganic arsenic molecules, which signifies incomplete methylation (Yamauchi & Takata, 2021). A high SMI indicates higher levels of demethylated arsenic metabolites relative to monomethylated arsenic species, and suggests higher methylation efficiency (Yamauchi & Takata, 2021). These methylation indices are applicable diagnostically because they are used to evaluate relationship between arsenite methylation capacity and susceptibility to disease outcomes such as peripheral neuropathy and skin lesions, as a high PMI and low SMI are indicators of a higher risk of more severe arsenic toxicity symptoms (Yamauchi & Takata, 2021).

This understanding of arsenic metabolism gives rise to another distinction: the difference between the effects of arsenic toxicity that originate from AS3MT activity indirectly—its impact on methionine metabolism or by competing with an additional AS3MT function—versus what unique effects that arise directly from the generation of downstream metabolites. Further questions entail how each of these relationships contributes to the pathogenesis and risks associated with arsenic exposure and how these relationships may interact.

Given that the arsenic methylation phenotype depends on *AS3MT* genotype (Thomas et al., 2007), the relevance of AS3MT concerning public health is underscored by its contribution to interindividual differences in sensitivity to arsenic toxicity among human populations. Single nucleotide polymorphisms (SNPs) can both increase and decrease AS3MT activity (Agusa et al., 2011; Wood et al., 2006), and they correlate with disease outcomes in arsenic-exposed human populations such as carotid intima-medial thickness (CIMT) (F. Wu et al., 2015)—a diagnostic parameter that is used to quantify the degree of atherosclerotic vascular disease. In view of these

observations AS3MT is clearly relevant in the context of arsenic epidemiology, as *AS3MT* genotype analysis has potential for diagnostic application given its nature as a risk factor for arsenic-enhanced atherosclerosis. Additionally, pharmacologic control of AS3MT-catalysed arsenite methylation may be harnessed to combat the risks of arsenic exposure and prevent the adverse outcomes of arsenic toxicity.

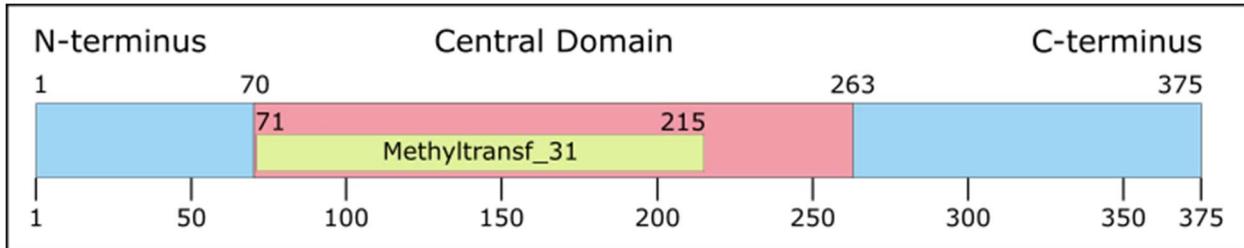
### 1.2.2 *AS3MT* Expression

AS3MT is a cytosolic protein that has a molecular weight of 43 kDa among mammalian homologues (Stýblo et al., 2021). AS3MT is most highly expressed in the liver and adrenal glands among mammals, yet it is also expressed in tissues throughout the body including the brain (Lin et al., 2002). In humans, the full AS3MT isoform is more highly expressed during prenatal neurodevelopment and exhibits a decline in postnatal expression, whereas an alternatively spliced isoform, AS3MT<sup>d2d3</sup>, maintains high expression in tissues that differentiate into neuronal fates (Li et al., 2016).

### 1.2.3 *The AS3MT Protein*

The AS3MT protein has three domains (**Figure 2**): an N-terminal domain that interacts with SAM, a methyltransferase domain, and a C-terminal domain with unknown function (Kabiraj et al., 2023). Mammalian AS3MT contains four conserved cysteine residues which are located in the active pocket for the arsenic methylation activity (Thomas et al., 2007; Chen et al., 2017b). These residues are required for arsenite methylation as the methylation function of AS3MT is abolished when point mutations are applied to these positions (Dheeman et al., 2014). Although only one active pocket is established—where the catalytic cysteine residues are located and where arsenite methylation occurs—based on the structure-function predictions of various AS3MT homologues using PrankWeb, a machine-based ligand binding prediction method,

AS3MT may have additional functional pockets where polar residues are overrepresented (Kabiraj et al., 2023).



**Figure 2. Domains of the human AS3MT protein.** The AS3MT primary sequence consists of 365 amino acids. The protein possesses an N-terminal domain (residues 1–70), a central domain (residues 71–263) that contains the methyltransferase domain (Methyltransf\_31; residues 71–215), and a C-terminal domain (residues 264–315).

#### 1.2.4 The AS3MT Gene

The *AS3MT* gene is located on chromosome 10q24.32 in humans, nearby *BORCS7*, *CYP17A1*, *CNNM2*, and *NT5C2*. The *AS3MT* locus is adjacent to *BORCS7*—a gene that is transcriptionally oriented in the same direction with an overlapping promoter. *BORCS7* encodes BLOC-1 related complex subunit 7, which is a protein that serves as an important factor for the anterograde axonal transport of lysosomes in neurons (Snouwaert et al., 2018).

Mammalian *AS3MT* has eleven exons (Thomas et al., 2007), and the methyltransferase domain spans exons 3–5 (Drobná et al., 2009). An alternatively-spliced variant, AS3MT<sup>d2d3</sup>, lacks exons 2–3 and is expressed highly in neuronal tissues, but it does not methylate arsenic (Li et al., 2016).

##### 1.2.4.1 GWAS related to the AS3MT locus

The *AS3MT* locus has been associated with numerous pathologies in genome-wide association studies (**Table 1**).

**Table 1. Summary of genome-wide associations of the 10q24.32 locus with pathologies.**

Blue: neurologic diseases; red: cardiovascular diseases, yellow: cognitive phenotype association, green: arsenic-related diseases.

<b>Associated Disease</b>	<b>Reference(s)</b>
<b>Schizophrenia</b>	Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium, 2011; Ripke et al., 2013; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Duarte et al., 2016; Hannon et al., 2019; Li et al., 2016; Gusev et al., 2018; Pardiñas et al., 2018; Yu et al., 2017; Washer et al., 2022
<b>Depression</b>	Li et al., 2016
<b>Psychosis</b>	DeMichele-Sweet et al., 2018
<b>Epilepsy</b>	Fan et al., 2021
<b>Attention-Deficit Hyperactivity Disorder</b>	Park et al., 2015
<b>Cross-disorder (neuropsychiatric):</b> Attention-Deficit Hyperactivity Disorder, Anxiety, Autism Spectrum Disorder, Bipolar Disorder, Major Depressive Disorder, Schizophrenia.	Lotan et al., 2014
<b>Alzheimer’s Disease</b>	DeMichele-Sweet et al., 2018; Pishva et al., 2020
<b>Migraine</b>	Winsvold et al., 2017
<b>Ischæmic Heart Disease</b>	Schooling et al., 2018
<b>Coronary Artery Disease</b>	Gong & O’Bryant, 2012; Winsvold et al., 2017
<b>Hyperlipidæmia</b>	Gong & O’Bryant, 2012
<b>Cerebrovascular Disease</b>	Liu et al., 2023
<b>Training-Induced Cognitive Plasticity</b>	Zhao et al., 2021
<b>Arsenic-Induced Skin Lesions</b>	Pierce et al., 2012; Das et al., 2016

10q24.32 was reported in numerous GWAS as one of the strongest and most consistent signals of GWAS schizophrenia association, and the *AS3MT* gene was identified as the lead protein-coding region in this locus for association with schizophrenia risk (Washer et al., 2022; Duarte et al., 2016). Furthermore, this association was corroborated in a transcriptome-wide association study (Gusev et al., 2018). It is noted at the same time that the *AS3MT*<sup>d2d3</sup> variant is expressed more abundantly in schizophrenia patients relative to controls (Li et al., 2016).

Lastly, it has recently been observed that in *AS3MT*<sup>-/-</sup> SH-SY5Y cells, a neuronal human cell line, there is dysregulation of genes implicated in schizophrenia (Washer et al., 2022). This conclusion was supported by RNA-seq data that showed enrichment of schizophrenia-related genes among the list of differentially expressed genes in *AS3MT*<sup>-/-</sup> SH-SY5Y cells (Washer et al., 2022). In conjunction with pathways involved in neuronal development, Washer *et al.* noted changes in gene expression at the pathway level spanning inflammation, ECM formation, and RNA processing (Washer et al., 2022).

Altogether, this information prompted Washer *et al.*'s hypothesis that an unknown, additional *AS3MT* function is related to neurodevelopment and differentiation, and this is why *AS3MT* belongs to a network of genes implicated in schizophrenia risk. It remains to be elucidated whether *AS3MT* SNPs and *AS3MT* expression are causally related to schizophrenia pathogenesis, and, if they indeed are, the mechanism responsible for the influence of *AS3MT* function on schizophrenia risk. The association of *AS3MT* knockout with a disturbance in expression of RNA processing-related genes connects with a report by Qiu *et al.*, who suggested a two-pronged role of *AS3MT* in m<sup>6</sup>A mRNA modification (Qiu et al., 2023). Qiu *et al.* concluded that, in L-02 cells (human foetal hepatocytes), *AS3MT* increases the METTL14-mediated m<sup>6</sup>A methylation of NLRP3 mRNA by strengthening its association with the

METTL14 methylase, thereby causing more mRNA stability (Qiu et al., 2023). It was argued based on *in silico* predictions and a co-immunoprecipitation experiment that, in addition to the m<sup>6</sup>A methylation modulation, AS3MT directly interacts with the NLRP3 protein to promote inflammasome activation (Qiu et al., 2023). These authors suggested that arsenic upregulates AS3MT protein expression and, therefore, gives rise to these two inflammasome-promoting AS3MT activities, and that this mechanism is responsible for arsenic-induced hepatic insulin resistance and inflammation (Qiu et al., 2023).

#### *1.2.5 Evolutionary development of AS3MT*

Regarding microorganisms, *AS3MT* corresponds with *ArsM* genes in bacterial orthologues (Qin et al., 2006). Between the *AS3MT* and *ArsM* gene families, the primary peptide sequences of nearly all arsenite methyltransferase orthologues and homologues range from 348 to 382 amino acids (Thomas et al., 2007). The AS3MT protein is remarkably conserved among phylogenetically distant species (**Figure 3**). The extent of the conservation is intriguing in view of the reduced selection pressure that has evolutionarily incentivised the integrity of AS3MT activity, considering that it has remained very much conserved despite the fact that oceanic arsenic concentrations have drastically decreased since the time of the primordial sea of the Archean and Proterozoic *Æons* (Zhu et al., 2014; Chen et al., 2017a). The conserved sequence motifs of the methyltransferase domain and the suite of catalytic cysteine residues of AS3MT are especially conserved (Song et al., 2009; Marapakala et al., 2012), and known to be shared in common with other methyltransferases of proteins and small molecules (Stýblo et al., 2021).

##### *1.2.5.1 AS3MT Conservation*

Upon comparing the peptide sequences of AS3MT homologues between phylogenetically distant species, a high extent of conservation is evident (Thomas et al., 2007; Chen et al., 2017a).

This conservation is exhibited in **Figure 3**, a Clustal Omega multiple sequence alignment between ten phylogenetically distant organisms. The arsenic selection pressure waned long before the phylogenetic divergence between these distantly-related organisms, as the concentration of oceanic arsenic has drastically lowered since the time of the primordial sea (Zhu et al., 2014; Chen et al., 2017a). It is unknown what kind of additional evolutionary selection pressure, if any, has incentivised such conservation of the AS3MT protein nevertheless. Regardless, it can be inferred that the AS3MT protein confers a fitness advantage in oceanic conditions of 1-2 ppb (13-27 nM) arsenic and even on land. There are three possible explanations for this conservation:

- A) the lower environmental concentration of the Cenozoic era, 1.7 ppb, is still high enough to exert a selection pressure that demands detoxification by arsenic methylation;
- B) a methylated arsenic metabolite is essential for normal cell function through an unknown mechanism (Uthus, 1992; Uthus & Poellot, 1991); or,
- C) there exists an additional function of AS3MT, independent from its catalysis of arsenite methylation—the hypothesis of this investigation.



K.alabastrina	ahdigfatprvlaishievh-deklkalvndtkffsityrlfktppqgdihprgnalvsvt	229
B.lanceolatum	akevgfctprlvlsaglvnlddhnpkmdlvgdikfvsityrlfklpqntske----aaqvv	288
P.proliflica	aeevgfsqprlvvtatvvpvd-nkelqdvlgdfkfsatyrlfkvpkg-kak----tcqvi	281
B.pectinirostris	aeevgfsrprlvvtatiitvd-nkelqdvlgdfrfvaatyrlfklpkg-nsk----pcrvi	281
E.macularius	aqdvqfwpvqlvtasrītis-nkelesivgdcrfvsatfrlfkislddsaq----rcevi	271
S.araneus	aqkvqfcprrlvvtasīitiq-nkelesivgdcrfvsatyrlfklpkteamq----rcqvi	274
H.sapiens	aqkigfcprrlvvtanlitiq-nkelervigdcfrfvsatfrlfkhsktgptk----rcqvi	274
M.musculus	aqkigfcprrlvvtadiitve-nkelegvlgdcrfvsatfrlfklpktepa-----rcrvv	275
C.yangmingshanensis	vaeagfrdvrllvsvgpvdvs-dpqlrklvdpdqfysctfrcfkvatleatr-edygqsat	295
C.polyspora	mqdlgcpdmrvvkenpitld-dtevaakigmvgfrsvtvrafkmp-ledrc-edfgqlav	272
	. * ::: : : : : * : * * *	
K.alabastrina	yngaiegheeyvldvnnvfkrnrvptvvgacetacvlsislwlekyftfydyfevk-qdga	288
B.lanceolatum	yngsvtgheaslkdqytfqegdivedpelsailqesrfaddftvrpaskkktkadel	348
P.proliflica	ynggitgaedifqfdlytftknevaqvdegevacilthsrfaedftfpppggscepc---	338
B.pectinirostris	ynggitgvnsfppfdqytfkaeevmevdgdldakilsssrfaedftfepsgqg--qc---	336
E.macularius	ynggiighekelefdvnykfkqgevevdeetaailqnsrfatklflinpirqklpitegc	331
S.araneus	ynggikghekelifdanftfkagevevdghtatilknsrfapefqsptgakpaacggs	334
H.sapiens	ynggitghekelmfndanftfkagevevdeetaailknsrfaqdflirpigeklptsggc	334
M.musculus	ynggikghekelifdanftfkageavavdeetaavlknsrfapdfllftpvdaslpapqgr	335
C.yangmingshanensis	ylggig---eefklrffftfprekpvrvdrntaeiirhsrlhqwfsvsa-eqqh--mglf	349
C.polyspora	ykgtltqphafdlddhhhhletgkplricgnsadmlsksrhyaehfqfwdgkthh--fglf	330
	* * : : * : : . : : * *	
K.alabastrina	ekmnvpgvastlvlmqkayeesiesepsagccppkakssgccppkakssdcccppkpasd	348
B.lanceolatum	ccpg-----akgvltnpfegpld-----kessccgpk--g--cc-----	378
P.proliflica	-cak-----pkttdvdpfdlasr-----s-gtrggccstk--sadcck-----	372
B.pectinirostris	-gvk-----akenlvdpfelvlqlekqnp---g-sttrggccstq--sascck-----	376
E.macularius	sh-----ekdkirdpflveq---mearvsa-qsvsqccgsk--g--cc-----	367
S.araneus	cgpq-----sedliadpflklaeesekskprcsa-sgargccask--t--sc-----	375
H.sapiens	sale-----lkdiitdpflklaeesdkmksrcvp-daaggccgk--k--sc-----	375
M.musculus	sele-----tkvlirdpflklaedsdkmkrhap-egtggccgkr--k--nc-----	376
C.yangmingshanensis	kand-----syallhaplsmqveqlv-----sgaaalehhh--hhh-----	383
C.polyspora	dcap-----pnkg-----qlta-----sgaacc-----	348
K.alabastrina	nasagsclpkikpsgccppkpitadtpkaestgkleilnasskd	392
B.lanceolatum	-----	378
P.proliflica	-----	372
B.pectinirostris	-----	376
E.macularius	-----	367
S.araneus	-----	375
H.sapiens	-----	375
M.musculus	-----	376
C.yangmingshanensis	-----	383
C.polyspora	-----	348

**Figure 3. Clustal Omega multiple sequence alignment of orthologous *AS3MT* and *ArsM* genes.** Alignment of ten organisms: *Homo sapiens*, *Mus musculus*, *Sorex araneus*, *Branchiostoma lanceolatum*, *Poeciliopsis prolifica*, *Boleophthalmus pectinirostris*, *Eublepharis macularius*, *Cyanidiococcus yangmingshanensis*, *Kickxella alabastrina*, *Crenothrix polyspora*.

The conservation is unlikely to be attributable solely to explanation A unless environmental concentrations ranging from 1-2 ppb arsenic are not high enough to cause detectable adverse events, but this only applies to aquatic environments, as it still would not explain why the AS3MT protein has still remained so conserved among terrestrial animals hundreds of millions of years after the transition from the sea.

Explanation B concerns the question of arsenic essentiality. It has been argued previously that arsenic is indeed an essential nutrient in trace amounts for chickens, goats, rats, and hamsters (National Research Council (US) Subcommittee on Arsenic in Drinking Water, 1999). Although there is no compelling evidence for essentiality in humans, a previous report proposed a recommended daily human requirement of 12  $\mu\text{g}$  arsenic per day (Uthus, 1992). This number aligned with the reported worldwide average dietary arsenic intake of 12-40  $\mu\text{g}$  (Uthus, 1992). Although the underlying physiological reason behind the purported arsenic essentiality was not known, it was speculated that arsenic, at low concentrations, acts as a micronutrient due to an effect on methionine metabolism that facilitates taurine and polyamine formation (Uthus, 1992), and by another report, that arsenic interacts with the micronutrient selenium in a beneficial way (Zeng et al., 2005). However, this hypothesis has been contradicted by following reports that organisms thrive in conditions that are as close to zero-arsenic as possible.

Explanation C is the basis of the hypothesis of this investigation. Perhaps AS3MT methylates another metalloid or an uncharacterised substrate belonging to an entirely different substrate class such as a protein or RNA species. Alternatively, AS3MT may be serving as a scaffold for a protein-protein or protein-RNA interaction.

### 1.2.6 Additional *AS3MT* Activity

The focus of the hypothesis is not the relevance of *AS3MT* regarding arsenic-related epidemiology, but the relevance of the *AS3MT* protein as it relates to a basal cell function. The hypothesis of an additional function served by *AS3MT* can be regarded as an inquiry about protein moonlighting—a gene sharing phenomenon whereby a protein can perform more than one physiologically relevant function, which in some cases are radically different from each other (Jeffery, 2003).

As a basis for this hypothesis, the strong conservation of *AS3MT* (section 1.2.5.1, **Figure 3**) and GWAS implicating the *AS3MT* locus (section 1.2.4.1) are noted in conjunction with *AS3MT*-dependent metabolic differences observed in mice (section 1.2.7.1) and preliminary data associating *AS3MT* expression in mouse embryonic fibroblasts with proliferation rate (**Figure 4A**) and h em oxygenase 1 (HO-1) expression (**Figure 4B**).

#### 1.2.6.1 *AS3MT*-dependent metabolic differences in mice

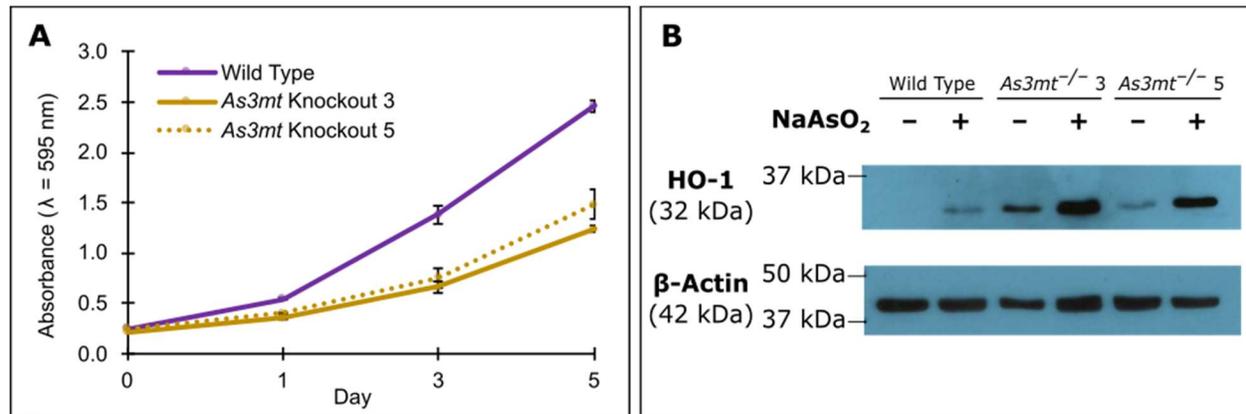
Although *As3mt* knockout in mice does not result in an overt phenotype besides severely impaired arsenic methylation, an *AS3MT*-associated difference in metabolism has been reported. It was noted by Huang *et al.* and Douillet *et al.* that upon gene deletion of *As3mt*, even in the absence of arsenic exposure, C57BL/6 mice present with higher body mass index and higher adiposity percentage (Huang *et al.*, 2016; Douillet *et al.*, 2017). In addition to obesity, the adverse metabolic phenotype of *As3mt*<sup>-/-</sup> mice extends to increased insulin resistance, which was dependent on sex and the level of arsenic exposure (Douillet *et al.*, 2017). It was proposed that, basally, the *As3mt* knockout is at least in part responsible for the adverse metabolic phenotype due to impaired phosphatidylcholine metabolism, as demonstrated by the hepatic phosphatidylcholine profiles of the *As3mt* knockout C57BL/6 mice (M. C. Huang *et al.*, 2016).

### 1.2.6.2 Slower proliferation associated with *As3mt* knockout

Based on preliminary growth curves, it has been noted that *As3mt*<sup>-/-</sup> mouse embryonic fibroblasts (MEFs) exhibit a slower growth rate in basal conditions (**Figure 4A**). An altered cell cycle, even in the absence of arsenic exposure, hints that AS3MT activity may influence cell proliferation through a mechanism unrelated to arsenic metabolism.

### 1.2.6.3 Elevated basal HO-1 expression associated with *As3mt* knockout

HO-1 is a member of the heat shock protein family and serves as a stress marker given its stress induction as part of the NFE2L2 pathway (He et al., 2020). Preliminary western blots demonstrated not only that arsenic exposure (1  $\mu$ M, 24 hours) induces HO-1 expression in MEFs, but also that HO-1 is expressed at a conspicuously higher level in *As3mt* knockout MEFs cultured without arsenic treatment (**Figure 4B**). Together, these data point to the hypothesis of an additional AS3MT activity, as the absence of this unknown activity in basal conditions gives rise to a stress response.



**Figure 4. Preliminary AS3MT-associated growth and HO-1 expression data in MEFs.** (A) Crystal violet growth curve displays relative growth rates of one wild-type MEF population compared with two *As3mt* knockout MEFs lines. Error bars represent the SEM of four technical replicates. (B) In MEFs, stress marker HO-1 is induced upon arsenic exposure (1  $\mu$ M, 24 hours) and more highly expressed basally between two *As3mt* knockout MEF lines relative to one wild-type population.

## 2. HYPOTHESIS, RATIONALE, & RESEARCH OBJECTIVES

### 2.1 Hypothesis

The hypothesis is that the AS3MT protein has an additional, basal function that is unrelated to its catalysis of arsenite methylation. This hypothesis addresses a knowledge gap because it is not yet known whether AS3MT has an additional activity beyond arsenite methylation, and, if it indeed exists, what this activity entails.

### 2.2 Rationale

Five evidence items prompted our investigation of a posited basal function of AS3MT:

1. the extensive conservation in the primary peptide sequence of AS3MT among homologues despite relief of evolutionary selection pressure (section 1.2.5.1, **Figure 3**);
2. the GWAS associations of the *AS3MT* locus with cardiovascular, neurological, and cognitive conditions observed independent of arsenic exposure (section 1.2.4.1, **Table 1**);
3. the observed sex-specific metabolic changes in *As3mt* knockout mice (section 1.2.6.1) (Huang et al., 2016; Douillet et al., 2017),
4. preliminary data indicating an association of *As3mt* knockout with slower proliferation of mouse embryonic fibroblasts (section 1.2.6.2, **Figure 4A**); and,
5. preliminary data suggesting that *As3mt* knockout is associated with elevated expression of stress marker HO-1 in mouse embryonic fibroblasts (section 1.2.6.3, **Figure 4B**).

### 2.3 Research Objectives

To investigate the question of an additional AS3MT activity, this research project committed to a proteomics purview. The experimental design was centred around mouse embryonic fibroblasts (MEFs) as a model. MEFs were chosen because they can be conveniently cultured and genetically manipulated. AS3MT-associated differences were studied by making comparisons between wild-type versus *As3mt* knockout MEF populations.

Aim 1) As a starting point, LC-MS (Liquid Chromatography-Mass Spectrometry) was chosen as a shotgun approach to screen for differentially expressed proteins between the wild-type and *As3mt*<sup>-/-</sup> mouse embryonic fibroblasts. In order to narrow the focus to specific pathways and proteins, the shotgun proteomics data was used as the input in pathway enrichment analysis in order to generate hypotheses and minimise bias. Aim 2) LC-MS hits were validated using immunoblots, also serving to validate the proteomics model as a technique screen for differentially expressed proteins among the MEFs.

An IP-MS (immunoprecipitation-mass spectrometry) dataset from a prior experiment on AML12 cells was retrospectively analysed in conjunction with the LC-MS data for the purpose of screening for potential interacting partners of AS3MT.

Aim 3) Phenotyping experiments were chosen in view of the differences associated with the pathways that were suspected to be disrupted in the *As3mt* knockout MEFs. The phenotypic profiling consisted of growth curves, immunofluorescent staining of proteins of interest, a scratch wound migration assay, and a collagen contraction experiment.

Altogether, these phenotyping experiments gauge whether the *As3mt* knockout is associated with a functional difference and whether the effect size is physiologically relevant.

Aim 4) Lastly, to explore the hypothesis beyond identifying associations, AS3MT dependence was tested using MEFs reconstituted with murine *As3mt*. These *As3mt* addback MEFs were used to determine whether the observed protein expression differences and proliferation phenotype are indeed AS3MT-dependent. In addition to reconstituting the knockout MEFs with wild-type *As3mt*, the MEFs were also reconstituted with a methylation-dead copy of *As3mt*. The goal of this addback was to provide a method to deduce whether an additional function of AS3MT is dependent on its methyltransferase activity or whether its mechanism is entirely different, for example, a role that is served by AS3MT via structural interactions.

### 3. MATERIALS & METHODS

#### 3.1 Chemicals and Reagents

The following chemicals and reagents were used in tissue culture: sodium arsenite (Sigma Aldrich, cat.: S7400); Polybrene<sup>®</sup> (10 mg/mL) (Santa Cruz Biotechnology, cat.: sc-134220); Geneticin<sup>®</sup> (G418) (Wisent, St Bruno, QC, Canada). Chemicals were diluted in sterile water and stored as aliquots at -20 °C.

#### 3.2 Arsenite exposure

Sodium arsenite was diluted serially to 2.0 mg/mL (by mass of arsenic atoms) then again, by a factor of 14, to the concentration 1.0 mM (83.33 ppm). The 1.0 mM sodium arsenite stock was added into tissue culture media at a dilution factor of 1000 (one µL per mL medium) to achieve the final 1.0 µM (83.33 ppb) concentration to which the MEFs were exposed for 24 hours. This exposure level was chosen because it is not cytotoxic, yet it is high enough to induce protein expression changes in response to arsenic-induced stress.

#### 3.3 *As3mt* knockout mice and isolation of MEFs

The *As3mt* knockout mice were obtained from Dr. David Thomas at the U. S. Environmental Protection Agency. The *As3mt* knockout in mice was generated using a deletion construct as described by Drobná *et al.* (Drobná *et al.*, 2009). The targeting construct was designed to delete exons 3–5 through homologous recombination. Exons 3–5 encode the methyltransferase domain. The mutant *As3mt* gene was transferred onto a C57BL/6 background by backcrossing as required (Drobná *et al.*, 2009).

Embryos from both wild-type and homozygous *As3mt*<sup>-/-</sup> C57BL/6 mice were trypsinised and mechanically disrupted to isolate primary MEF populations. Each MEF population originated from a single mouse embryo. Immortalisation of the MEFs was achieved via retroviral

transduction with a pBABE-neomycin construct (Addgene plasmid #1767) that encodes the SV40 large T antigen. Plat-E cells were transfected with the pBABE construct, virions were harvested, and MEFs were transduced according to the protocol described herein (section 3.5). After the MEFs recovered from the transduction process, MEFs were selected with Geneticin<sup>®</sup> (G418; 50 mg/mL stock in sterile deionised water). Resistance to the G418 antibiotic is conferred by the neomycin resistance gene encoded by the pBABE vector. G418 concentration was escalated to 300 µg/mL, and following selection, MEFs were maintained in culture as populations at low passage numbers for downstream experiments.

### ***3.4 Cell culture***

All cell cultures were maintained at 37 °C in a humidified incubator at 5 % CO<sub>2</sub>.

#### *3.4.1 Mouse Embryonic Fibroblasts*

MEFs were maintained in culture at low passage numbers as independent populations. The MEFs were cultured in Dulbecco's Modified Eagle Medium (DMEM; Wisent, cat.: 319-005-CL) supplemented with 1 M HEPES buffer (Wisent, cat.: 330-050-EL), 50X essential amino acids (Wisent, cat.: 321-010-EL), 50X non-essential amino acids (Wisent, cat.: 321-011-EL), 0.5X penicillin/streptomycin diluted to 0.5 % (Wisent, cat.: 450-201-EL), and 10 % foetal bovine serum (FBS) (Wisent). MEFs were cultured in 150-mm dishes (PROGENE<sup>®</sup>, cat.: 667651) and lifted with 1X trypsin (Wisent, cat.: 325-242-EL) every three to four days.

#### *3.4.2 AML12 Murine Hepatocytes*

AML12 cells were obtained from the American Type Culture Collection. The AML12 cells were cultured as populations in DMEM/F12 (Wisent) supplemented with 10 % foetal bovine serum (Wisent), insulin (Sigma Aldrich), holo-Transferrin (Sigma Aldrich), and dexamethasone (Sigma Aldrich).

### *3.4.3 Platinum-E 293T cells*

Platinum-E (Plat-E) cells were cultured in DMEM (Wisent) supplemented with 10 % foetal bovine serum (Wisent) with 0.5 % penicillin/streptomycin (Wisent). Plat-E cells were cultured in 100-mm dishes (PROGENE<sup>®</sup>, cat.: 229621) and lifted with 1X trypsin (Wisent) every two to three days.

## **3.5 Retroviral transduction**

### *3.5.1 Plat-E transfection and virion production*

The Plat-E retroviral ecotropic packaging cell line (Cell Biolabs, cat.: RV-101)—a 293T-based cell line—was used for the virion production that was outlined previously (Guilbert et al., 2020).

The Plat-E were seeded at  $2 \cdot 10^6$  cells per 100-mm dish in plat-E medium overnight. The following calculations are for one 100-mm dish. In this sequence, 15 µg pBABE construct, and sterile deionised water (for a total volume of 500 µL), and 62 µL of 2 M CaCl<sub>2</sub> were added to a polypropylene tube. In a separate tube, 500 µL 2X Hank's Balanced salt solution (HBSS), pH 7.1 was prepared. The tube with the HBSS was aerated gently with a 1-mL pipette for one minute as the contents of the construct-containing tube were slowly added. The mixture was incubated for 20 minutes at room temperature.

The medium on the Plat-E cells was replaced with 10 mL new medium before adding the one-millilitre construct mixture. After incubating at 37 °C for 24 hours post-transfection, the medium was replaced with 5 mL of Plat-E medium and returned to the incubator. After 24 hours, 5 mL of virion-containing supernatant was harvested before adding another 5 mL medium and returning the dish to the incubator. The 5 mL supernatant was harvested once more after another

24 hours. Both times the supernatant was harvested it was filtered through a 0.2  $\mu$  M PES filter (Fisher Scientific, cat.: 13-1001-06) and stored at -80 °C until transduction of the MEFs.

### 3.5.2 MEF transduction

MEFs ( $2 \cdot 10^5$ ) were seeded in a 100-mm dish. After incubation for 24 hours, the medium was replaced with 4 mL supplemented MEF medium with the addition of 2 mL of viral supernatant along with Polybrene<sup>®</sup> diluted to a concentration of 8  $\mu$ g/mL. After incubating the MEFs for 24 hours with the viral supernatant, the replacement of the medium and addition of more viral supernatant and Polybrene<sup>®</sup> was repeated once more. After the second 24-hour incubation, the viral supernatant was replaced with supplemented MEF medium.

### 3.6 *As3mt* Reconstitution

The *As3mt* knockout MEFs were reconstituted with FLAG<sup>®</sup>-tagged murine AS3MT by transduction with a pMIG construct containing a FLAG<sup>®</sup>-*As3mt* insert with three copies of FLAG<sup>®</sup> in tandem. Through the process detailed herein (section 3.7.1), the FLAG<sup>®</sup>-*As3mt* insert was cloned into the pMIG backbone—an IRES-GFP-containing retroviral backbone vector.

#### 3.6.1 Cloning the *As3mt* insert into a pMIG construct

The wild-type murine *As3mt* insert was amplified by PCR from the preëxisting p3X-WT*mAs3mt*-FLAG<sup>®</sup> construct that was used for the transfection of AML12 hepatocytes (section 3.7). This amplification of the insert was performed with the primers provided in

**Supplementary Table 1**, which introduced XhoI and ApaI restriction sites at the ends of the amplicon.

Following a column purification of the PCR product to remove contaminant p3X starting material, the purified FLAG<sup>®</sup>-*mAs3mt* insert was double-digested with XhoI and ApaI restriction

enzymes (New England Biolabs<sup>®</sup>, cat.: R0146S and cat.: R0114S respectively) in 10X CutSmart<sup>®</sup> Buffer (New England Biolabs<sup>®</sup>, cat.: B6004S) and column-purified once more.

In parallel, the pMIG backbone (Addgene plasmid #9044) was prepared by double restriction digestion with XhoI and ApaI in CutSmart<sup>®</sup> Buffer (New England Biolabs<sup>®</sup>). A gel extraction and column purification were carried out on the digested pMIG to remove the filler piece.

The digested FLAG<sup>®</sup>-*mAs3mt* insert and digested pMIG backbone were ligated together for transformation of One Shot<sup>™</sup> Stbl3<sup>™</sup> chemically competent *E. coli* (Invitrogen, cat.: C7373-03). The transformed Stbl3<sup>™</sup> *E. coli* was plated onto a carbenicillin dish for ampicillin selection before colonies were picked and grown in liquid culture. Plasmid DNA was isolated from the *E. coli* using the E. Z. N. A.<sup>®</sup> Plasmid DNA Mini Kit I (Omega Bio-Tek, cat.: D6943 – 02) according to the manufacturer's instructions. The plasmid was isolated and then sequenced by whole-plasmid sequencing to confirm the identity of the construct as pMIG-WT*mAs3mt*-FLAG<sup>®</sup>.

### 3.6.2 Site-directed mutagenesis of *As3mt*

Alongside the reconstitution with the wild-type *As3mt* gene, *As3mt* knockout MEFs were also reconstituted with a construct encoding a FLAG<sup>®</sup>-tagged copy of a methylation-dead mutant AS3MT. Site-directed mutagenesis was carried out on the previously-generated pMIG-WT*mAs3mt*-FLAG<sup>®</sup> insert in order to mutate DLG to HLA in exon 3 of the *As3mt* gene. The mutation of these residues within the SAM-binding pocket renders the AS3MT protein methyltransferase-dead, given that SAM binding necessitates contact with nine residues LDLGS<sup>G</sup>SGR (Martin, 2002; Petrossian & Clarke, 2009). This mutagenesis was carried out via PCR using the primers provided in **Supplementary Table 1**. Following the mutagenesis and

amplification, a digestion was carried out with DpnI (New England Biolabs<sup>®</sup>, cat.: R0176S) to cleave the methylated, p3X-WTm*As3mt*-FLAG<sup>®</sup> template construct. After transformation of Stb13<sup>™</sup> *E. coli* and plasmid isolation, Sanger sequencing was carried out on the plasmid mini preparations to confirm the mutation before proceeding with downstream transduction.

### *3.6.3 Transduction of *As3mt*<sup>-/-</sup> MEFs for *As3mt* reconstitution*

In parallel, the wild-type and mutant HLA-*As3mt* inserts and empty vector pMIG were used in the transduction of MEFs. Plat-E cells were transfected with either the wild-type FLAG<sup>®</sup>-*As3mt* or mutant FLAG<sup>®</sup>-*As3mt* construct, or with empty vector (EV) pMIG. The viral supernatant corresponding with each of these constructs was harvested and used to transduce the *As3mt*<sup>-/-</sup> MEFs. During this process, the success of transfection and transduction was confirmed by viewing the GFP signal by widefield microscopy. After two passages transduction the MEFs were sorted for GFP signal by flow cytometry (section 3.7).

### *3.7 Flow Cytometry: sorting GFP-positive MEFs*

After lifting MEFs with trypsin, the transduced cells were washed with PBS and passed through 70 μM nylon mesh cell strainer to filter out aggregates. After counting, cells were pelleted and resuspended in 200 μL FACS buffer (2.0 % FBS (Wisent) in 1X PBS (Wisent, cat.: 311-011 CL)) and transferred into 5-mL round-bottom polypropylene for sorting.

MEFs were selected by GFP-signal using the BD FACSAria<sup>™</sup> Fusion cell sorter. The cells were sorted twice, at times three weeks apart, to guarantee that the construct remained expressed in the MEFs. During the second sorting, the MEFs were split into three populations by GFP intensity level (low-, medium-, and high-GFP intensity). After expanding in culture, the medium GFP-intensity MEFs were studied in downstream experiments. This ensured that cells

with similar vector copy numbers were compared across experimental groups (wild-type *As3mt*, HLA methylation-dead mutant *As3mt*, and empty vector control).

### **3.8 *As3mt* Genotype Analysis**

The *As3mt* knockout genotype of the MEF populations was confirmed via PCR and RT-qPCR (**Supplementary Figure 1**). DNA extraction from MEFs was carried out with the E. Z. N. A.<sup>®</sup> MicroElute<sup>®</sup> DNA Kit (Omega Bio-Tek, cat.: D3096-00S) following the manufacturer's instructions.

PCR reactions on the genomic DNA extracted from the MEFs used the primers provided in **Supplementary Table 1**. These primers were designed and applied in genotype analyses to verify *As3mt* knockout as described by (Drobná et al., 2009). Primers 1 and 2 were used for the detection of wild-type *As3mt*, whereas primers 1 and 3 were used for the detection of the *As3mt* knockout genotype. Amplification of DNA from the wild-type *As3mt* gene using primers 1 and 2 yields a 330 bp band. Amplification of the modified (knockout) *As3mt* gene using primers 1 and 3 yields a 400 bp band (Drobná et al., 2009). An image of a gel upon electrophoresis of the reaction products is displayed in **Supplementary Figure 1A**.

### **3.9 *Sry* Genotype Analysis**

The sex of each MEF population was determined via PCR with primers that amplify the male-specific *Sry* gene alongside separate control reactions using primers that amplified *IL3* as an autosomal control (**Supplementary Table 1**). The *Sry* amplicon is 402 bp whereas the *IL3* amplicon is 544 bp. An image of the PCR result is displayed in **Supplementary Figure 3**.

### **3.10 RT-qPCR**

MEFs were seeded on day 0 at a density of  $5.0 \cdot 10^5$  cells per 150-mm dish and maintained at 37 °C in a humidified incubator until harvesting at day 2. RNA extraction from the MEFs was carried out using the E. Z. N. A.<sup>®</sup> Total RNA Kit (Omega Bio-Tek, cat.: 101319-242) following the manufacturer's instructions. RNA samples were stored at -80 °C. First-strand cDNA synthesis was performed on 1.0 µg RNA samples using the iScript<sup>™</sup> cDNA synthesis kit (Bio-Rad, cat.: 1708891) following the manufacturer's instructions, and using the Eppendorf Vapo.Protect<sup>™</sup> Mastercycler<sup>®</sup> Pro thermocycler. The 20 µL first-strand cDNA synthesis reaction products were diluted to a volume of 200 µL (dilution factor of 10).

RT-qPCR was done using the primer combinations in **Supplementary Table 2** on the cDNA from the MEF RNA. qPCR was performed in a 10 µL reaction mixture in triplicate for each target multiplied by each cell type. After quality controlling primers for efficiency and specificity, the primers were used in RT-qPCR to compare between cell types. The  $2^{-\Delta\Delta C_t}$  analysis method was employed to interpret the threshold cycle data. One reference gene transcript, 36B4 mRNA or 18S rRNA, was used to normalise qPCR data, which were analysed via the  $2^{-\Delta\Delta C_t}$  method. The means of the  $2^{-\Delta\Delta C_t}$  values were compared between MEF populations with reference to standard error of the mean (S. E. M).

### **3.11 Protein extraction, quantification, immunoblotting**

#### **3.11.1 Protein extraction**

MEFs were seeded in 150-mm dishes on day 0. The seeding density of MEFs used for the protein extracts for LC-MS was  $4.0 \cdot 10^5$  cells per dish, whereas up to  $1.0 \cdot 10^6$  cells were seeded per 150-mm for subsequent westerns (within each experiment, all cell types were seeded at the same density). After maintaining at 37 °C at 5 % CO<sub>2</sub> in a humidified incubator, on day 2

MEFs were harvested using rubber cell scrapers and resuspended in lysis buffer: 1X RIPA buffer (with Triton<sup>®</sup> X-100, pH 7.4) (BioBasic, cat.: RB4478) with 1X cOmplete™ EDTA-free Protease Inhibitor Cocktail (Roche, cat.: 11873580001) and 1X PhosSTOP™ (Roche, cat.: 04 906 837 001) Phosphatase Inhibitor. Lysis reactions were carried out 30 minutes on ice with periodic vortex mixing, before centrifuging at 13,200 rpm at 4 °C for 15 minutes to clear debris. The proteinaceous supernatants were collected and stored at 80 °C for downstream applications.

### *3.11.2 Protein quantification by Bradford Assay*

The Bradford Assay (Bradford, 1976) was performed for protein quantification of RIPA lysates. Two-microlitre volumes of protein lysates were added into cuvettes containing 798 µL Milli-Q<sup>®</sup>-purified water and 200 µL Bradford protein-binding dye (Bio-Rad, cat.: 5000006). In parallel, 2, 4, 8, 16, and 32 µg/µL standards were prepared with bovine serum albumin (BSA) with 200 µL Bradford reagent and water to a volume of 1.000 mL. Absorbance was measured by spectrophotometry and interpolated on the BSA standard curve to calculate protein concentration.

### *3.11.3 Immunoblotting*

Upon protein quantification, protein concentrations were equalised between aliquots of RIPA lysates before adding 5x Laemmli buffer. Samples were boiled at 95 °C for three minutes then loaded onto a SDS-polyacrylamide gel (between 10-12 % depending on the molecular weight of target(s)). After carrying out SDS-PAGE, the protein was transferred onto a nitrocellulose or PVDF membrane (depending on primary antibodies to be used). Membranes were stained with 1X Ponceau S Acid Red 112 (C. I. 27195) to evaluate loading evenness and transfer fidelity. The membranes were cut based on the molecular weight of intended targets. Membranes were blocked with 5 % milk (v/v) in 1X tris-buffered saline (TBS). After 1-2 hours

of blocking, primary antibody dilutions in 5 % milk (v/v)/TBS were applied to the membranes for incubation overnight at 4 °C on a rocker. The primary antibodies that were used in immunoblots are displayed in **Table 2**.

**Table 2. Summary of primary antibodies used in immunoblots**

Target	Producer, Catalogue Number	Host Species	Dilution
β-actin	Millipore <sup>®</sup> and Sigma Aldrich, cat.: A5441	Mouse	1 in 10,000
β-catenin	Cell Signaling, cat.: 8480s	Rabbit	1 in 1,000
Fibronectin	BD Transduction Laboratories <sup>™</sup> , cat.: 610077	Mouse	1 in 2,000
FLAG <sup>®</sup> M2	Sigma Aldrich, cat.: F1804	Mouse	1 in 2,000
HO-1	Enzo Life Sciences, cat.: ADI-SPA-896-D	Rabbit	1 in 1,000
Myoferlin	Novus Biologicals, cat.: NBP1-84694	Rabbit	1 in 1,500
Vimentin	BD Pharmingen <sup>™</sup> , cat.: 550513	Mouse	1 in 1,000
VPS29	Abcam, cat.: Ab236796	Rabbit	1 in 500

After the overnight incubation with the primary antibodies, the membranes were washed three times for ten-minute in TBS with 0.1 % v/v Tween (TBS-T). Membranes were subsequently treated for 1-2 hours at room temperature on a rocker with the required horseradish peroxidase-conjugated secondary antibody (**Table 3**), which was diluted 1:3000 in 5 % milk (v/v)/TBS.

**Table 3. Secondary antibodies used in immunoblots**

Target	Producer, Catalogue Number	Host Species	Dilution
Mouse Fc	GE Healthcare Life Sciences, cat.: NA931V	Sheep	1 in 3,000
Rabbit IgG	GE Healthcare Life Sciences, cat.: NA934V	Donkey	1 in 3,000

Following three more 10-minute TBS-T washes, membranes were revealed on x-ray films using the following developing kits: Cytiva Amersham<sup>™</sup> ECL<sup>™</sup> (cat.: RPN2106) and Millipore<sup>®</sup> Immobilon<sup>®</sup> chemiluminescent kit (cat.: WBKLS0500).

### ***3.12 Liquid Chromatography-Mass Spectrometry***

The mass spectrometry was carried out by the Proteomics and Molecular Analysis Platform at the Research Institute of the McGill University Health Centre (RI-MUHC). Protein

extracts were prepared from one wild-type and two *As3mt* knockout MEF populations, and from both control and arsenic-treated cultures corresponding with each population. The sample size was four biological replicates for the wild-type MEF population and three replicates for each of the two *As3mt* knockout MEF populations. Following quantification, the RIPA lysates were diluted in RIPA lysis buffer as required to concentrations between 1.6 – 4.2  $\mu\text{g}/\mu\text{L}$  and submitted as samples ranging from 60-80  $\mu\text{g}$  protein. Liquid chromatography-mass spectrometry was done with the EVOQ<sup>®</sup> Triple Quadrupole Mass Spectrometer system. Mass spectrometry data was analysed using the Scaffold 5 proteome software for spectral count analysis and through the MaxQuant software for the label-free quantification analysis.

For the spectral count analysis, peptides were grouped into clusters based on sequencing overlap between paralogous peptides, which aims to retain valid protein identification for proteins that share peptide sequences. This occurs when the same peptide is present within multiple proteins. A cluster is a group of proteins that have shared peptide evidence, and it can be treated as a proxy for the identification of a single protein. This protein grouping was done by the clustering algorithm of the Scaffold 5 software. An advantage of protein clustering is higher statistical power, as it retains valid protein identification when the peptide evidence would have otherwise been discarded on account of being insufficient from the perspective of another grouping strategy or analysis method.

For robustness, the list of detected proteins was filtered based on the following parameters: a minimum of four peptides detected, 99 % protein threshold, and 95 % peptide threshold. FBS contaminants were removed from the list of proteins. In order to make comparisons between experimental groups, the spectral count data were normalised with an imputed constant of 0.5 spectral counts. Imputation is the replacement of missing values with

estimates near the detection limit (Aguilan et al., 2020). This is needed in order to calculate ratios when there is data for a given protein from one sample or condition but not another (Aguilan et al., 2020). The spectral count data serve as a surrogate for protein quantity, because, for a given protein, the number of detected fragments reflects its abundance. For each peptide, three unpaired Student's T-tests ( $\alpha = 0.05$ ) were carried out in parallel, corresponding with AS3MT-associated and arsenic-induced expression comparisons. Of the significantly different protein clusters, all but four contained a single protein. For these four clusters, variant peptides were visually inspected and only those representing the main protein were retained and confirmed for statistical significance.

The label-free quantification (LFQ) data for each sample was obtained via MaxQuant. The MaxQuant data from mass spectrometry was analysed after the pathway enrichment analysis. The reason for using LFQ intensity data is that the spectral count fold differences are often distorted by dynamic exclusion filtration during tandem mass spectrometry (S. Zhang et al., 2016).

A minimum of eight observations was set for the LFQ data analysis. The LFQ intensities were normalised and imputed using a previously-described data processing method (Aguilan et al., 2020). The LFQ ratios were interpreted as a more accurate indicator of fold change provided that the dynamic exclusion filtration in tandem mass spectrometry can distort the magnitudes of fold differences if they are calculated solely based on spectral count data (S. Zhang et al., 2016). F-tests and T-tests were performed on these data based on the same previously-established analysis method (Aguilan et al., 2020).

### ***3.13 Immunoprecipitation-Mass Spectrometry***

IP-MS was conducted on AML12 hepatocyte populations overexpressing FLAG<sup>®</sup>-tagged AS3MT. AML12 cells were stably transfected with either empty vector (EV) p3X-FLAG<sup>®</sup> or p3X-WTm*As3mt*-FLAG<sup>®</sup> constructs (Mersaoui et al., 2022). Following transfection, the AML12 cells were selected with 900 ng/mL of G418 (Wisent).

FLAG<sup>®</sup>-tagged AS3MT was immunoprecipitated using an anti-FLAG<sup>®</sup> antibody (Sigma Aldrich, cat: F1804) and FLAG<sup>®</sup> M2 beads (Sigma Aldrich). The immunoprecipitation was performed according to the manufacturer's instructions. The FLAG<sup>®</sup> M2 beads were washed and sent to MS/MS for analysis at Université de Sherbrooke, where the immunoprecipitated proteins were trypsinised and analysed by mass spectrometry. The experimental N was two biological replicates for each the empty vector control and AS3MT bait immunoprecipitations.

To filter out non-specific interactions, the protein detection events in the empty vector control immunoprecipitations were subtracted from the spectral counts of the FLAG<sup>®</sup>-AS3MT samples to gauge which proteins were the most enriched due to possible AS3MT interactions. The spectral count difference (AS3MT bait IP – empty vector control IP) for each protein was averaged between the two biological replicates. Proteins with an average difference of five or more detection events were identified as the proteins enriched in the AS3MT bait IP.

### ***3.14 Immunofluorescence***

MEFs were seeded on 24-well plates (Fisherbrand™, cat.: FB012929) on glass coverslips (Cedarlane, cat.: 64-0715). The seeding density for the vimentin staining was 80,000 MEFs per coverslip, whereas the seeding density was sparser—20,000 MEFs per coverslip—in the immunofluorescent VPS35 staining experiment for the purpose of single-cell punctal staining quantification. After seeding, the MEFs were maintained overnight in the humidified incubator.

The cells were retrieved from the incubator the following day and washed twice with PBS as they remained on the cover slips within the wells. Next, cells were fixed with either 100 % methanol at -20 °C for ten minutes (vimentin staining) or 4 % paraformaldehyde (v/v in PBS) (Electron Microscopy Sciences) for 15 minutes at room temperature (VPS35 staining).

Following fixation, the wells were washed gently with PBS four times each. The cells were then permeabilised with Triton<sup>®</sup> X-100 (MP Biomedicals cat. 94854) for 15 minutes at room temperature. 0.2 % Triton<sup>®</sup> X-100 was used in the vimentin experiment, whereas 0.1 % Triton<sup>®</sup> X-100 was used in the VPS35 experiment. The coverslips were subsequently blocked using either 8 % donkey serum (Jackson ImmunoResearch Laboratories, cat.: 017000121) (v/v in PBS) for 40 minutes at room temperature (vimentin staining) or 4 % BSA (m/v in PBS) for two hours at room temperature (VPS35 staining).

After blocking, the coverslips were extracted from the wells. The required primary antibody dilution (**Table 4**) was applied to the coverslips for incubation overnight at 4 °C in a humidified water chamber.

**Table 4. Summary of antibodies used in immunofluorescence experiments**

	Target	Host species	Producer & catalogue no.	Dilution	Conjugated Fluorophore
Primary	Vimentin	Mouse	BD Pharmingen™, 550513	1:100 in 8 % donkey serum	N/A
	VPS35	Rabbit	Proteintech, 10236-1-AP	1:100 in 4 % BSA	N/A
Secondary	Anti-Mouse IgG (H+L)	Goat	Invitrogen™, A-11029	1:2,000 in 8 % donkey serum	Alexa Fluor <sup>®</sup> 488
	Anti-Rabbit IgG (H+L)	Chicken	Invitrogen™, A-21442	1:500 in 4 % BSA	Alexa Fluor <sup>®</sup> 594

After the overnight primary antibody treatment, the coverslips were washed four times with PBS for five minutes each time. The required secondary antibody dilution (**Table 4**) was applied to the coverslips and incubated at room temperature in the dark for one hour. The

coverslips were washed again for four times with PBS (five minutes each wash). The coverslips were then treated with a 1 in 500 dilution of 1 mg/mL 4',6-diamidino-2-phenylindole (DAPI) stock (Thermo Scientific, cat.: 62248) in PBS for ten minutes at room temperature. The DAPI solution was washed from the coverslips with PBS before mounting them onto microscope slides with ProLong™ Gold mounting medium (Invitrogen™, cat.: P10144).

### **3.15 Confocal Microscopy**

All coverslips were imaged with the Zeiss LSM 800 *AiryScan* confocal laser scanning microscope. The anti-vimentin-stained coverslips were imaged using the *AiryScan* SR mode with deconvolution at 63X magnification and 1.3X zoom for maximum resolution. The settings for all images for the vimentin (Alexa Fluor™ 488) track were laser power = 1.0 % and master gain = 856 V.

The laser settings were decided by optimising the image acquisition of the anti-VPS35-stained coverslips using *As3mt* knockout MEF population 3 as the reference population, given the expectation based on the mass spectrometry data that *As3mt* knockout MEFs will show a higher signal intensity relative to the wild-type populations. The settings decided upon for the VPS35 (Alexa Fluor™ 594 track) were laser power = 0.4 % and master gain = 670 V. For the purpose of punctal quantification only images at 63X with 1X zoom and these predefined laser settings were used in the analysis. The viable images consisted of N = 23 B2i wild-type cells, N = 30 B3b wild-type cells, N = 31 KO3 cells, and N = 21 KO5 cells (53 wild-type MEFs and 52 *As3mt* knockout MEFs in total).

Data were collected for the quantification of the number, size, intensity, and distance from the nuclear centrepoint of puncta by analysing the images using a Macro program

(**Appendix B**) on the *Fiji* distribution of *ImageJ*. The average punctal intensity, size, and distance from the nuclear centroid were averaged in order to make comparisons between cells.

### **3.16 Growth Curves**

#### *3.16.1 Growth curve seeding*

MEFs were seeded into 12-well plates (Fisherbrand™, cat.: FB012928) at 4,400 cells per well on day 0. Four wells (technical replicates) were allocated for each day of staining for each cell type. Cells were maintained in the humidified incubator until the plates were collected for fixing and staining from day 0 (once cells adhered to the dish) to day 5. Given that the crystal violet staining intensity served as a surrogate for confluence, proliferation of MEFs was gauged by increases in spectrophotometric absorbance from day 0 onward.

#### *3.16.2 Crystal Violet staining*

At each time point, plates were washed with PBS and fixed with 4 % paraformaldehyde (Electron Microscopy Sciences, cat.: 15710) for 15 minutes at room temperature. After fixation, cells were washed with purified Milli-Q® water and then stained with 0.1 % crystal violet solution (90.0 % water, 10.0 % ethanol, 0.1 % (w/v) crystal violet (JT Baker® Chemical Co., cat.: F907-03) for 20 minutes on a rocker. Blank wells were stained with crystal violet in parallel to be used to assess background in subsequent spectrophotometry. Following the staining, all wells were washed gently three times with water and left to dry overnight. The dried crystal violet residue in each well was resuspended in 200 µL 10 % acetic acid (v/v in water) as the plates were rocked for 20 minutes. Crystal violet suspension (150 µL) from each well was pipetted into separate wells on a 96-well plate (Fisherbrand™, cat.: FB012931). The crystal violet staining intensity for each well in the 96-well plate was measured by spectrophotometry ( $\lambda = 595$  nm). The absorbance values were blanked with the absorbances of the stained cell-free wells and

recorded with the MEF population identity and time point corresponding with well on the 12-well plate from which the crystal violet suspensions were transferred.

### ***3.17 Scratch Wound Migration Assay***

On day 0, MEFs were seeded in an IncuCyte® Imagelock 96-well plate (Sartorius, cat.: BA-04856). After counting the cells and preparing a master mix for each cell type, eight wells were seeded per MEF population. In each well 6,000 MEFs in a volume 200  $\mu$ L were seeded. This seeding density was chosen in order to create a confluent monolayer of MEFs in each well at the time of the scratch-making and subsequent migration. Cells were maintained overnight at 37 °C in the humidified incubator. On day 1, the scratches were made using the IncuCyte® 96-Well Woundmaker Tool (Sartorius, cat.: BA-04858) to ensure uniform scratches across all wells. Next, the medium was aspirated and replaced in the wells to remove floating cells. The cells were imaged and monitored by incubating the plate in the IncuCyte® Live-Cell Analysis System for 24 hours. Images were acquired every two hours with the scan type set to “Scratch Wound”.

Image analysis was carried out and the percentage wound confluence data was collected automatically by the IncuCyte® using software module cat.: 9600-0012. For each time point, the automatically-calculated percentage wound confluence values were averaged across the eight wells for each cell type and plotted as the function of average percent wound confluence over time.

### ***3.18 Collagen Contractility Assay***

MEFs were counted for the adjustment of cell suspensions to  $2.0 \cdot 10^5$  cells/mL. Two parts of the cell suspension were mixed with one part of Type I rat tail collagen solution (4.40 mg/mL) (Corning cat. 354236, Lot no. 24323001). For each collagen lattice 500  $\mu$ L of this mixture, which contained a concentration of  $1.33 \cdot 10^5$  cells/mL, was immediately transferred

into a well of a 24-well plate (Fisherbrand™, cat.: FB012929). Four technical replicate lattices were seeded per MEF population. The collagen-cell mixtures were incubated for 20 minutes at room temperature in order to solidify into collagen lattices. Cell culture medium (600 µL) was added by gently pipetting on the sides of each of the wells. Next, the collagen lattices were dislodged from the perimeter of the wells by tracing the edges with P200 pipette tips. The plate was placed in the humidified incubator and retrieved at the defined time points (days 0–3) for imaging and collagen lattice size measurement. Collagen lattice area was calculated by measuring the diameter of the lattices (in mm) on ImageJ. The lattice contractility was evaluated as a percentage of the initial (day 0) lattice area.

### ***3.19 Statistical Analyses***

#### *3.19.1 Analysis of LC-MS spectral count data*

Protein clusters were filtered based on the following parameters: a minimum of 4 peptides detected, 99.0 % protein threshold, and 95.0 % peptide threshold.

Each protein was treated as a separate hypothesis with respect to three two-tailed Student's T-tests ( $\alpha = 0.05$ ), in which the average spectral count of the representative protein was compared between the required experimental groups for these hypotheses: 1) a basal, AS3MT-associated difference, 2) an arsenic-induced expression change in wild-type MEFs, and 3) an arsenic-induced expression change in *As3mt* knockout MEFs. The missing values were imputed with a constant of 0.5 for normalisation and analysis.

The proteins meeting the criterion  $p \leq 0.05$  for each hypothesis or combination of hypotheses were recorded. For the sorting of hits into a Venn diagram of the three hypotheses, a cut-off for fold difference  $\geq 1.5$  was added to the  $p \leq 0.05$  criterion. An alternate analysis was considered: a multiple test correction according to the Benjamini-Hochberg procedure

(Benjamini & Hochberg, 1995). The subset of proteins that showed significant AS3MT-associated basal differences ( $p \leq 0.05$  for hypothesis 1) was used as the input for pathway enrichment analysis by Enrichr.

### *3.19.2 Image analysis of immunofluorescent VPS35 punctal staining*

For the hypothesis testing related to VPS35 punctal number, average punctal size, average punctal intensity, and average punctal distance from the nuclear centroid, 95 % confidence intervals ( $\alpha = 0.05$ , two-tailed) were used in the image analysis as the basis of comparison between cell types.

### *3.19.3 Analysis of phenotypic assay data*

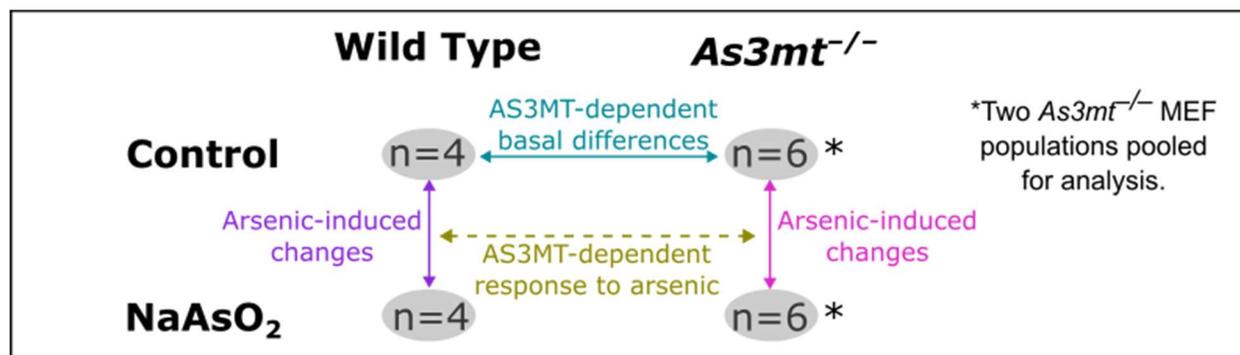
SEM was used to compare means between MEF populations with respect to the optical density data in crystal violet growth curves, the percentage wound confluence data in the scratch wound migration assay, and the percentage surface area data in the collagen contractility assay.

## 4. RESULTS

### 4.1 LC-MS Screen of AS3MT-associated and arsenic-induced protein expression differences

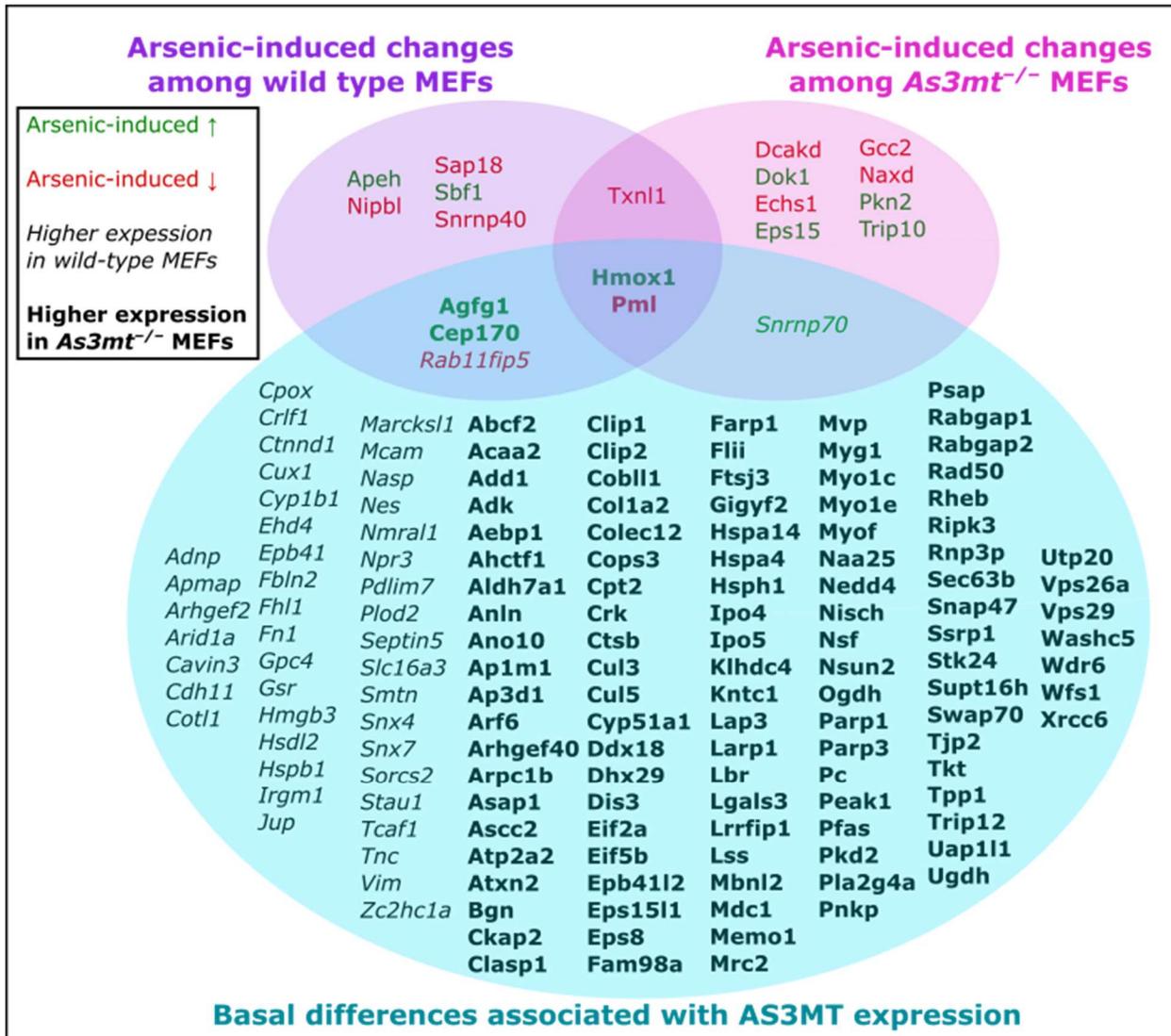
It was hypothesised that AS3MT may have a function in addition to arsenic methylation. Based on the *As3mt*<sup>-/-</sup> MEF phenotype of elevated HO-1 expression and slower proliferation in basal conditions, it was expected that the AS3MT-associated phenotypic differences are reflected in the basal proteome by protein expression differences. The proteomic perspective was justified for the initial investigation of this hypothesis because the proteome is extremely dynamic and relatively wide, the proteome has a preponderant influence on cell function given that proteins are the functional molecules that manage the cell's activities, the proteome considered the central link between the genome and the cell because is it the culmination of gene expression while it provides a foundation for biochemical pathways (Brown, 2002).

Shotgun proteomics was performed on *As3mt*<sup>+/+</sup> and *As3mt*<sup>-/-</sup> MEFs with and without exposure to 1  $\mu$ M arsenite for 24 hours. The mass spectrometry experiment was designed to screen for basal, AS3MT-associated protein expression differences between the wild-type versus *As3mt* knockout MEF populations, and, in parallel, to assess the changes induced by arsenic treatment (Figure 5).



**Figure 5. LC-MS experimental design for screening of protein expression differences.** The schematic displays the comparisons made. LC-MS served as a hypothesis-generating experiment regarding AS3MT-associated differences and arsenic-induced changes in protein expression.

Protein lysates were made from one wild-type and two *As3mt* knockout MEF populations, and from both control and arsenic-treated cultures corresponding with each population. Four biological replicates were tested for the wild-type MEF population whereas three biological replicates were tested for each of the two *As3mt* knockout populations.



**Figure 6. Differentially expressed proteins detected in MEFs by LC-MS.** The Venn diagram represents differentially expressed proteins (Student's T-test  $p \leq 0.05$ , fold difference  $\geq 1.5$ ). The Venn diagram displays arsenic-induced changes in wild-type MEFs (purple), arsenic-induced changes in *As3mt*<sup>-/-</sup> MEFs (pink), and basal AS3MT-associated differences (blue).

The data were filtered for robustness with the predefined thresholds (n = 4 peptides minimum, 99 % protein threshold, 95 % peptide threshold). The remaining data corresponded

with 2633 proteins in 2200 clusters. Significant protein differences were identified as a fold difference of  $\geq 1.5$  when comparing the average spectral counts between experimental groups with a  $p \leq 0.05$  in a Student's T-test. Comparisons were made with respect to 1) basal differences between wild-type and knockout MEFs; 2) arsenic-induced changes in wild-type MEFs; and 3) arsenic-induced changes in As3MT knockout MEFs (**Figure 6**). All but four of the significant protein differences were clusters containing only one peptide. For each of the four clusters containing multiple proteins (GPC4, MBNL2, STK24, VIM), the representative protein was retained and it was verified that its spectral count data met the fold change  $\geq 1.5$  and  $p \leq 0.05$  criteria.

Eleven proteins were identified that underwent a significant arsenic-induced change in wild-type MEFs, while twelve proteins showed significant arsenic-induced changes in the *As3mt* knockout MEFs. Of these, only three proteins were in common between the two MEF genotypes, two of which are known to be modulated by arsenic: HO-1 and PML (promyelocytic leukaemia protein). HO-1 is a stress protein induced by arsenic (Wang et al., 2012), while PML is degraded following arsenic exposure (Lallemand-Breitenbach et al., 2008). Interestingly, the basal expression of these proteins were also different between the *As3mt* genotypes. Both of these proteins are expressed basally at higher levels among the *As3mt* knockout MEF populations. HO-1 and PML served as controls and they both showed the expected expression differences with statistical significance. This suggested that the data and the analysis method are relatively robust. However, when a Benjamini-Hochberg multiple test correction (Benjamini & Hochberg, 1995) was applied with respect to the hypothesis of a basal AS3MT-associated difference, the spectral count data reached significance for only two proteins: HMGB3 (High Mobility Group protein B3) and PLOD2 (procollagen-lysine 2-oxoglutarate 5-dioxygenase 2). When a Benjamini-

Hochberg multiple test correction was applied to the arsenic-treatment comparisons, there were no proteins with significant corrected  $p$ -values for arsenic-induced expression differences in wild-type MEFs, while PML was the only protein whose arsenic-induced expression difference remained significant in the *As3mt* knockout MEFs. In order to have a larger pool of proteins for pathway analyses, we included those proteins significant by T-test.

The numbers of proteins that underwent significant arsenic-induced expression changes were unexpectedly low considering that there are multiple pathways that are known to be affected by arsenic toxicity, such as DNA damage repair, oxidative stress, and cell proliferation (Vergara-Gerónimo et al., 2021). In comparison, significantly more differences were observed between untreated wild-type and knockout MEFs, as 157 proteins had a significant basal difference. Thus, the focus was narrowed to these basal differences to address the question of an additional AS3MT function independent of arsenic.

These 157 proteins that were differentially expressed between the wild-type and *As3mt*<sup>-/-</sup> MEFs ( $p \leq 0.05$ , fold difference  $\geq \pm 1.5$ ) were used in pathway enrichment analysis by Enrichr. The NCATS BioPlanet 2019 and KEGG 2021 libraries were used in the pathway enrichment analysis because they are comprehensive, manually-curated databases that include both healthy and disease states. Additionally, the NCATS BioPlanet 2019 library is advantageous because it is designed using redundancy and consistency cross-evaluation (R. Huang et al., 2019).

**Table 5. Pathway enrichment analysis of AS3MT-associated basal proteomic differences.**

The analysis was done on the subset of differentially expressed proteins detected by LC-MS. *p*-value-based pathway rankings were provided by Enrichr from the analysis against the BioPlanet 2019 and KEGG 2021 libraries.

<b>BioPlanet 2019 Library</b>					
<b>Index</b>	<b>Pathway Name</b>	<b><i>p</i>-value</b>	<b>Adjusted <i>p</i>-value</b>	<b>Odds Ratio</b>	<b>Combined Score</b>
1	E-cadherin nascent adherens junction-like junctions pathway	$1.298 \cdot 10^{-5}$	$5.737 \cdot 10^{-3}$	19.17	215.65
2	Fcy receptor-mediated phagocytosis	$9.798 \cdot 10^{-5}$	$1.976 \cdot 10^{-2}$	8.92	82.34
3	TGF- $\beta$ regulation of extracellular matrix	$1.515 \cdot 10^{-4}$	$1.976 \cdot 10^{-2}$	3.43	30.15
4	Oncostatin M	$1.788 \cdot 10^{-4}$	$1.976 \cdot 10^{-2}$	4.42	38.11
5	Double-strand break repair	$6.549 \cdot 10^{-4}$	$5.373 \cdot 10^{-2}$	20.33	149.00
6	Apoptotic execution phase	$7.380 \cdot 10^{-4}$	$5.373 \cdot 10^{-2}$	10.78	77.75
7	Cholesterol biosynthesis	$8.509 \cdot 10^{-4}$	$5.373 \cdot 10^{-2}$	18.39	129.99
8	Adherens junction actin cytoskeletal organisation	$1.493 \cdot 10^{-3}$	$7.346 \cdot 10^{-2}$	14.85	96.62
<b>KEGG 2021 Library</b>					
<b>Index</b>	<b>Pathway Name</b>	<b><i>p</i>-value</b>	<b>Adjusted <i>p</i>-value</b>	<b>Odds Ratio</b>	<b>Combined Score</b>
1	Endocytosis	$7.465 \cdot 10^{-7}$	$1.067 \cdot 10^{-4}$	6.76	95.36
2	Fcy receptor-mediated phagocytosis	$1.166 \cdot 10^{-4}$	$8.340 \cdot 10^{-3}$	8.62	78.11
3	Steroid biosynthesis	$4.904 \cdot 10^{-4}$	$2.338 \cdot 10^{-2}$	22.72	173.12
4	Lysosome	$3.410 \cdot 10^{-3}$	$8.916 \cdot 10^{-2}$	5.27	29.96
5	Non-homologous end-joining	$4.512 \cdot 10^{-3}$	$8.916 \cdot 10^{-2}$	23.26	125.64
6	Fatty acid degradation	$4.653 \cdot 10^{-3}$	$8.916 \cdot 10^{-2}$	9.64	51.79
7	Pathogenic <i>Escherichia coli</i> infection	$4.681 \cdot 10^{-3}$	$8.916 \cdot 10^{-2}$	4.09	21.93
8	Ubiquitin-mediated proteolysis	$4.988 \cdot 10^{-3}$	$8.916 \cdot 10^{-2}$	4.80	25.45

According to this pathway enrichment analysis, proteins implicated in adherens junction structure, cytoskeletal organisation, and endocytic trafficking were overrepresented in the set of proteins that showed a significant basal difference between the wild-type and *As3mt* knockout MEFs (Table 5).

#### ***4.2 IP-MS search for putative AS3MT interacting partners***

Next, it was postulated that some of the differentially expressed proteins might bind to AS3MT and could narrow down important hits identified in the shotgun proteomics.

Accordingly, a previously-generated dataset of AS3MT-interacting partners was interrogated. In this prior experiment, immunoprecipitation with FLAG<sup>®</sup> beads was performed on AML12 hepatocytes overexpressing FLAG<sup>®</sup>-tagged AS3MT (N = 2 biological replicates) to identify potential protein AS3MT-interacting partners. A total of 321 proteins were detected in the FLAG<sup>®</sup> pull-downs from the AML12 hepatocyte lysates.

These prior IP-MS data were newly analysed based on the spectral counts. For each protein, the number of detection events in the empty vector control IP was subtracted from the number in the FLAG<sup>®</sup>-AS3MT bait IP (AS3MT bait – empty vector control). Proteins with an average difference of five or more detection events between the two biological replicates were considered IP-MS hits. The 29 hits that were defined by this method are displayed in **Table 6**.

The top hit was AS3MT, which had an average difference of 20 detection events. This signified that the most enriched protein was AS3MT itself and confirmed that the FLAG<sup>®</sup>-AS3MT immunoprecipitation was successful. Of note, the top interacting proteins included those involved in endocytic trafficking as well as adherens junctions and tight junctions. Protein including clathrin, dynein, myoferlin, and evelin-2 participate in the retrograde endocytic trafficking of receptors and growth factors between intracellular compartments (Johannes & Popoff, 2008; Popoff et al., 2009; Okazaki et al., 2012; Dong et al., 2019; Canty et al., 2021). Adherens and tight junctional proteins included desmoyokin, plectin, afadin, supervillin and tight junction proteins 1 and 2.

**Table 6. Hits from the FLAG<sup>®</sup>-AS3MT IP-MS experiment in AML12 hepatocytes.** The blue and green text denote cell adhesion- and endocytic trafficking-related hits respectively.

Identified Protein	Average $\Delta$ (As3MT Bait IP – control)	Identified Protein	Average $\Delta$ (As3MT Bait IP – control)
AS3MT	20.0	<b>Evectin-2</b>	6.5
Myosin 9	17.0	Dock7	6.5
Filamin A	13.0	Eif3a	6.0
<b>Tjp1</b>	12.5	Prrc2c	6.0
<b>Desmoyokin</b>	10.5	<b>Clasp1</b>	6.0
<b>Plectin</b>	10.0	<b>Supervillin</b>	6.0
<b>Clathrin</b>	9.0	Myosin 1e	6.0
<b>Tjp2</b>	8.0	Myosin 1b	6.0
Eif4b	8.0	Fxr1; Fxr2	6.0
Mthfd1l	7.5	Myosin 18a	6.0
Myosin 5a	7.5	Myosin 1c	6.0
<b>Afadin</b>	7.0	<b>Myoferlin</b>	5.5
Erc1	7.0	Prpf8	5.0
Senataxin	7.0	Hspa9	5.0
<b>Dynein</b>	6.5		

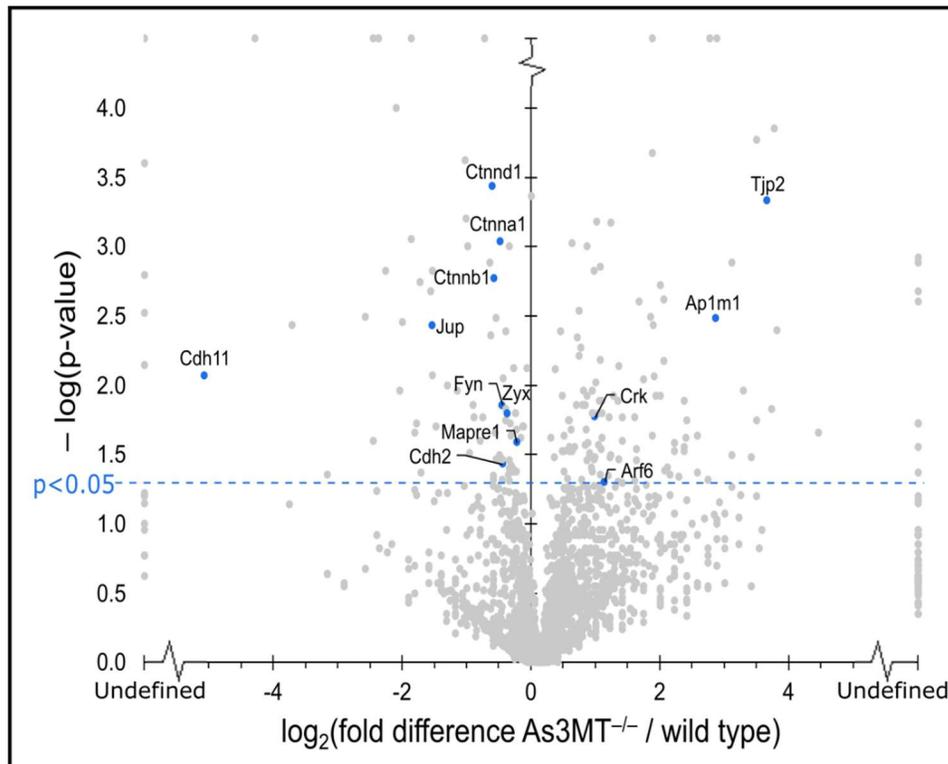
Adherens junctions are protein complexes that are anchored on their cytoplasmic face by the attachment between cadherins and actin filaments (Gates & Peifer, 2005; Hartsock & Nelson, 2008). These contacts are made through adapters such as catenins, spectrin, vinculin, afadin, and zona occludens-1 (ZO-1) (also known as tight junction protein 1 (TJP1)) (Gates & Peifer, 2005), both of which were detected as hits in the IP-MS experiment. Overall, the IP-MS data complement the LC-MS data, and both datasets overlap in two key pathways: endocytic trafficking and adherens junctions.

#### ***4.3 Adherens Junctions and Cytoskeleton***

Adherens junctions and actin cytoskeleton organisation were identified as enriched pathways (**Table 5**). According to the Enrichr search against the BioPlanet 2019 database, the

pathways ranked first, third, and eighth when listed by  $p$ -value were adherens junctions, ECM organisation, and actin cytoskeletal organisation (Benjamini-Hochberg-adjusted  $p$ -values =  $5.737 \cdot 10^{-3}$ ,  $1.976 \cdot 10^{-2}$ , and  $7.346 \cdot 10^{-2}$  respectively) (Table 5).

Several members of the catenin family showed a significant AS3MT-associated basal expression difference based on spectral count data. Proteins with a significant difference included  $\alpha$ -catenin (ctnna1),  $\beta$ -catenin (ctnnb1),  $\gamma$ -catenin (a. k. a. junction plakoglobin (Jup)), and  $\delta$ -catenin (a. k. a. p120-catenin (ctnnd1)). Catenins are present at the cell membrane as a component of adherens junctions, where they serve as adapters that connect actin filaments with cadherins (Hartsock & Nelson, 2008). According to the spectral count data, all four of these catenins were expressed at significantly lower levels in the *As3mt* knockout MEFs relative to the wild-type MEFs (Figure 7).



**Figure 7. Differential adherens junction-related protein expression between wild-type and *As3mt* knockout MEFs.** Volcano plot of LC-MS spectral count data displays fold differences of spectral count averages between wild-type versus *As3mt*<sup>-/-</sup> MEFs. The highlighted data represent the differentially proteins selected by Enrichr corresponding with adherens junctions.

To further interrogate the proteomics dataset, the data were reanalysed with a focus on the LFQ intensities to verify the expression difference observed in the spectral count analysis. LFQ intensity data was analysed in view of the wider dynamic range and because the fold differences conveyed by spectral counts are often distorted by dynamic exclusion filtration during mass spectrometry (S. Zhang et al., 2016). The objective was to evaluate whether the patterns hold true between the spectral count and LFQ intensity analyses. The LC-MS dataset was interrogated using an exhaustive list of adherens junction adapter proteins (Zaidel-Bar, 2013) (**Supplementary Table 3**). The group of proteins in **Table 7** represents all of the cadherins and adherens junction adapter proteins that were detected in the MEFs.

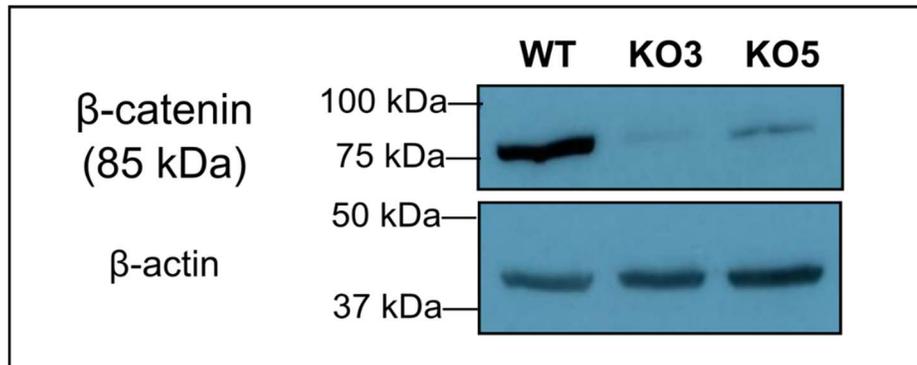
The LFQ data (**Table 7**) confirmed that there was a consistent downregulation of catenins among the *As3mt* knockout MEFs as well as the decreased expression of cadherins and many other adherens junction adapters—including tight junction protein 1. According to the LFQ intensity data, all three detected cadherins (cadherin-2, -3, and -11) were present at significantly lower levels in the *As3mt* knockout MEFs. Although not all adherens junction adapter proteins had lower expression in the *As3mt* knockout MEFs, all eight proteins with a significant expression difference were downregulated in the *As3mt* knockout MEFs. All five detected members of the catenin family were present at lower levels in the *As3mt* knockout MEFs with fold differences ranging from 0.1 to 0.5. Apart from the catenin proteins, the other significantly downregulated adherens junction adapters were protein 4.1, palladin, and moesin.

**Table 7. AS3MT-associated expression differences of adherens junction proteins.** Adapter proteins in blue font serve as adapters that link cadherins with the actomyosin cytoskeleton. Proteins in red font link cadherins with microtubules. Proteins in purple font serve as adapters between cadherins and both types of these filaments.

<b>Cadherin Protein</b>	<b>LFQ Intensity Fold Difference (<i>As3mt</i><sup>-/-</sup> / wild-type)</b>	<b><i>p</i>-value (two-tailed Student's T-test)</b>
Cadherin-3	0.051	0.000002***
Cadherin-11	0.073	0.000017***
Cadherin-2 (N-cadherin)	0.543	0.040122*
<b>Adherens Junction Adapter Protein</b>	<b>LFQ Intensity Fold Difference (<i>As3mt</i><sup>-/-</sup> / wild-type)</b>	<b><i>p</i>-value (two-tailed Student's T-test)</b>
Protein 4.1	0.060	0.000041***
$\gamma$ -catenin (Junction Plakoglobin)	0.106	0.021816*
$\alpha$ N-catenin (catenin $\alpha$ -2)	0.371	0.014943*
$\beta$ -catenin (catenin $\beta$ -1)	0.385	0.000095***
Palladin	0.424	0.003920**
$\alpha$ -actinin-1	0.458	0.051063 <sup>ns</sup>
p120-catenin ( $\delta$ -catenin)	0.481	0.010658*
$\alpha$ E-catenin (catenin $\alpha$ -1)	0.510	0.007123**
Vinculin	0.605	0.222598 <sup>ns</sup>
Moesin	0.638	0.015854*
Eplin	0.644	0.167144 <sup>ns</sup>
$\alpha$ -actinin-4	0.724	0.234428 <sup>ns</sup>
Tight junction protein 1 (ZO-1)	0.728	0.241332 <sup>ns</sup>
ILK	0.901	0.650069 <sup>ns</sup>
Radixin	0.940	0.692542 <sup>ns</sup>
Supervillin	1.108	0.791591 <sup>ns</sup>
IQGAP1	1.166	0.445750 <sup>ns</sup>
Dlg1	1.279	0.133385 <sup>ns</sup>
Ankyrin G	1.452	0.474953 <sup>ns</sup>
Afadin	1.537	0.462863 <sup>ns</sup>
$\beta$ 2 spectrin	1.599	0.235104 <sup>ns</sup>
$\alpha$ -adducin	1.947	0.088901 <sup>ns</sup>
Clip1	2.061	0.388655 <sup>ns</sup>
Ezrin	2.347	0.058794 <sup>ns</sup>

The AS3MT-associated difference in  $\beta$ -catenin expression was validated by immunoblotting (**Figure 8**).  $\beta$ -catenin was an attractive choice for hit validation because it is a protein moonlighter that serves multiple important functions: it contributes to the integrity of adherens junctions and also functions as a signalling component in the canonical Wnt signalling pathway (Brembeck et al., 2006; Valenta et al., 2012). This pathway influences differentiation

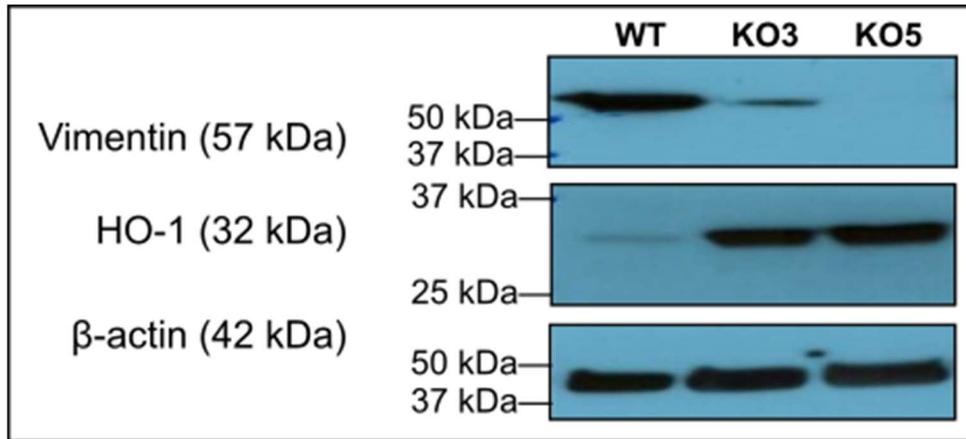
and proliferation (Brembeck et al., 2006). The direction of the expression difference matched the LFQ intensity data. Altogether, the LC-MS data and the  $\beta$ -catenin immunoblots provide more evidence for an *As3mt*<sup>-/-</sup>-associated disturbance in the expression of adherens junction proteins.



**Figure 8.** *As3mt* knockout MEFs express lower levels of  $\beta$ -catenin basally. The  $\beta$ -catenin blot is representative of two individual experiments.

#### 4.4 Vimentin phenotype of wild-type and *As3mt*<sup>-/-</sup> MEFs

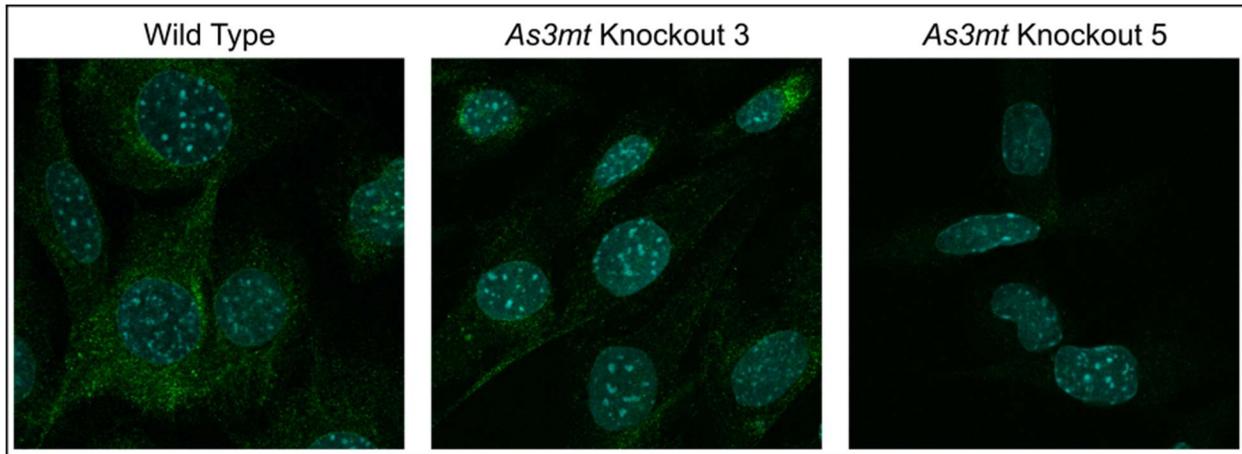
The type III intermediate filament, vimentin, was selected for immunoblot validation not only in view of its role in actin cytoskeletal organisation (Lopez-Menendez, 2022; Serres et al., 2020), but also given that its functions also relate to cytoskeletal stability, tensile stiffness, cytosolic organelle anchorage, growth, and migration (Arrindell & Desnues, 2023; Perlson et al., 2005; Ridge et al., 2022). The mass spectrometry data indicated that there was lower vimentin expression associated with *As3mt* knockout with an LFQ intensity ratio (KO/WT) of 0.117 ( $p = 2.13 \cdot 10^{-3}$ ). Differential vimentin expression was validated by immunoblotting (**Figure 9**), showing that vimentin was basally depleted in both *As3mt*<sup>-/-</sup> MEF populations relative to the wild-type control population.



**Figure 9. *As3mt* knockout MEFs express lower levels of vimentin basally.** The vimentin immunoblot is representative of seven individual experiments. Immunoblot demonstrates downregulation of vimentin in *As3mt* knockout. An HO-1 immunoblot is featured alongside the vimentin blot as a previously-established differentially expressed protein.

#### 4.4.1 Immunofluorescent imaging of vimentin in MEFs

Immunofluorescence was conducted to compare the vimentin staining patterns between wild-type and *As3mt*<sup>-/-</sup> MEFs (**Figure 10**). The MEFs presented unexpectedly with a non-filamentous staining pattern. Non-filamentous forms of vimentin have been reported in literature (Correia et al., 1999; Ivaska et al., 2007; Patteson et al., 2020). This non-filamentous pattern did not change when a z-stack was processed to create maximum intensity projection. It was observed that the relative intensities of the vimentin signal were consistent with the difference in total vimentin expression between the wild-type and *As3mt* knockout MEFs as gauged by the mass spectrometry and immunoblot (**Figure 6, Figure 9**).



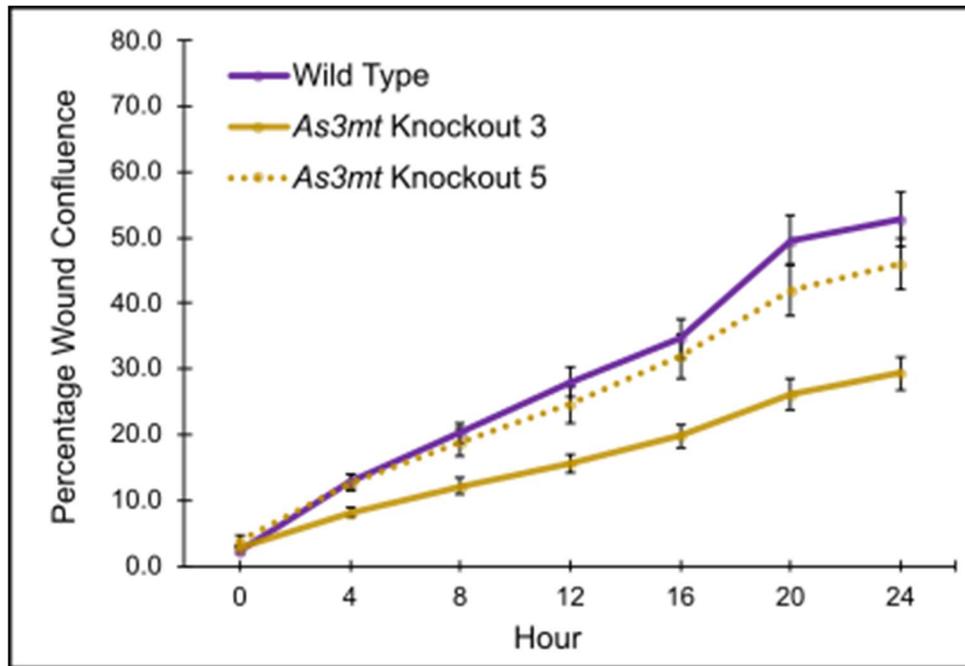
**Figure 10. Non-filamentous vimentin staining pattern in MEFs.** Images acquired were acquired using the ZEISS LSM 800 confocal microscope and using AiryScan deconvolution. The images depict anti-vimentin-stained, DAPI-stained wild-type and *As3mt* knockout MEFs (63X objective). The vimentin AF488 signal is green whereas the DAPI signal is cyan.

#### 4.4.2 Functional comparison between wild-type and *As3mt*<sup>-/-</sup> MEFs

*Vim*<sup>-/-</sup> MEFs have poor wound healing and collagen contractility compared to wild-type MEFs (Eckes et al., 1998; Ridge et al., 2022). Thus, a slower migration and collagen contractility phenotype was predicted in the *As3mt*<sup>-/-</sup> MEFs. This may be complimented by the downregulation of catenins in the *As3mt* knockout that would be predicted to affect the integrity of cell-matrix contacts that depend on adherens junctions. Anchoring junctions are necessary for the attachment of the cytoskeleton to the extracellular matrix, and so the lower expression of catenins gave rise to the prediction that the *As3mt* knockout MEFs would show impaired collagen contractility.

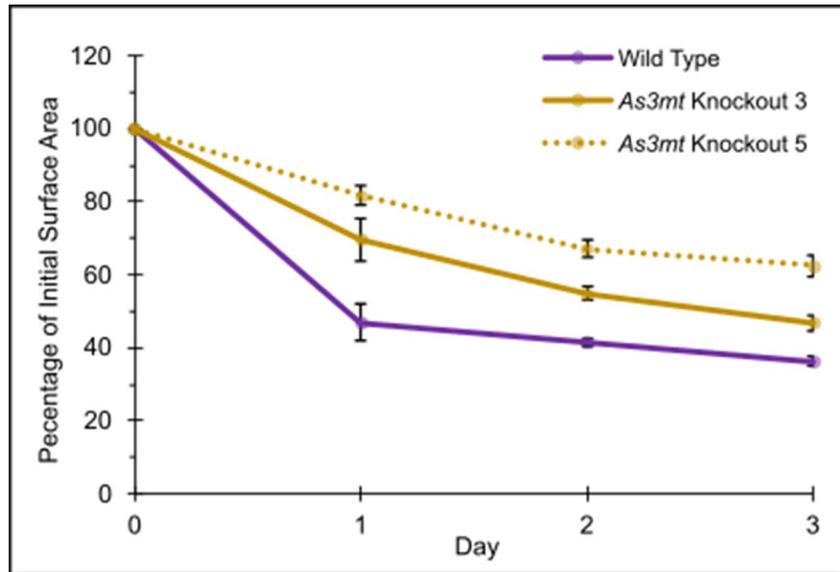
Migration of the wild-type and *As3mt* knockout MEFs was compared using a scratch migration assay (**Figure 11**), in which wound closure was monitored over 24 hours. Although the *As3mt* knockout 3 population showed a slower migration rate relative to the wild-type MEFs, there was not a migration difference between the wild-type and *As3mt* knockout populations.

Thus, these preliminary data suggested that the difference in migration was not correlated with AS3MT expression.



**Figure 11. Scratch wound migration of wild-type and *As3mt* knockout MEFs.** MEF migration was assessed by proxy as the percentage wound confluence over 24 hours after monolayers were scratched. Each population was observed across eight technical replicates. Images were acquired through the IncuCyte® Live-Cell Analysis. Error bars represent the SEM of eight technical replicates within a single biological experiment.

The collagen contractility assay tests the capacity of fibroblasts to contract and remodel the extracellular matrix, which is an important process in wound healing (Eckes et al., 2000; Q. Zhang et al., 2022). In an initial experiment, the wild-type MEF population had faster and greater collagen lattice contractility than the *As3mt* knockout MEF populations (**Figure 12**). These data are consistent with an AS3MT-associated difference in collagen contractility.



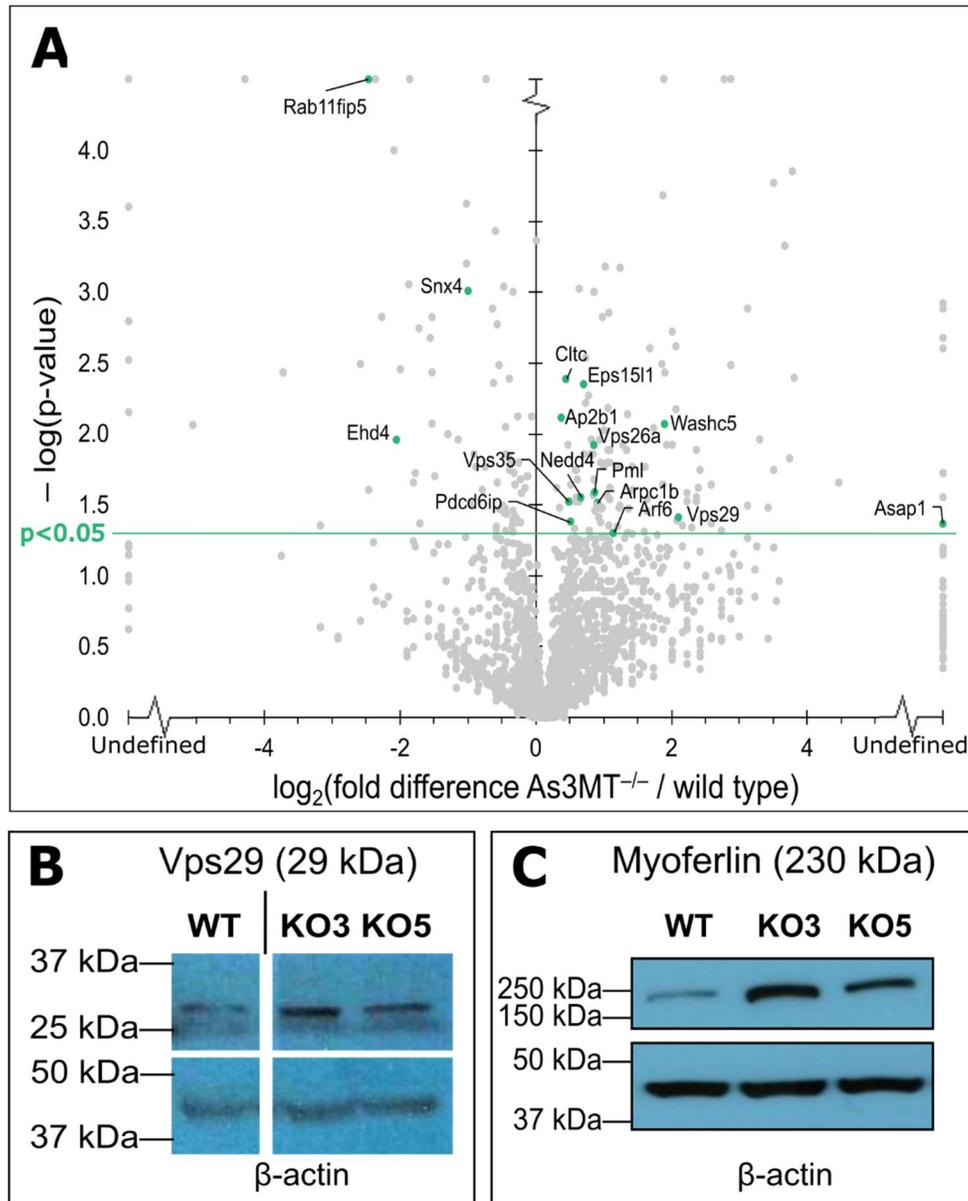
**Figure 12. Collagen contractility of wild-type and *As3mt* knockout MEFs.** Collagen contractility curve shows the areas of type I collagen lattices seeded with each MEF line over three days. Contractility was measured as a percentage of the surface area of the lattice at day 0 ( $\text{mm}^2$ ). Four technical replicate lattices were evaluated per type. All error bars represent the SEM of four technical replicates within a single biological experiment.

In summary, the lower expression of vimentin in the *As3mt* knockout MEFs was matched between the LC-MS data and the immunoblots, while it translated to the relative intensities of the non-filamentous staining pattern observed in the immunofluorescence experiment. Subsequent initial phenotypic assays suggested that the knockout MEFs did not show an overt migration defect but did show lesser collagen contractility compared to the wild-type control.

#### 4.5 Endocytic trafficking phenotype of wild-type and *As3mt*<sup>-/-</sup> MEFs

Based on the pathway enrichment analysis (**Table 5**), endocytic trafficking was identified and all three subunits of the retromer cargo-selective (CSC) complex, the VPS35–VPS29–VPS26 heterotrimer, were significantly upregulated in basal conditions among the *As3mt* knockout MEFs (**Figure 13A**). The differential expression of VPS29 was validated by immunoblotting (**Figure 13B**). In agreement with the LC-MS data, the two *As3mt*<sup>-/-</sup> populations showed higher VPS29 expression than the wild-type MEFs (**Figure 13B**). In addition, the transmembrane protein myoferlin was selected for immunoblot validation (**Figure 13C**) on

account of its roles in fusion and repair of membranes and endocytic trafficking (Bernatchez et al., 2009; Dong et al., 2019). Both *As3mt*<sup>-/-</sup> MEF populations showed higher expression of myoferlin, confirming the LC-MS results (**Figure 13C**).

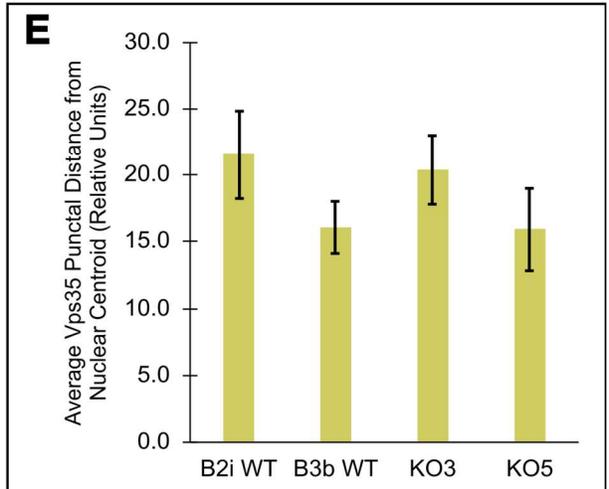
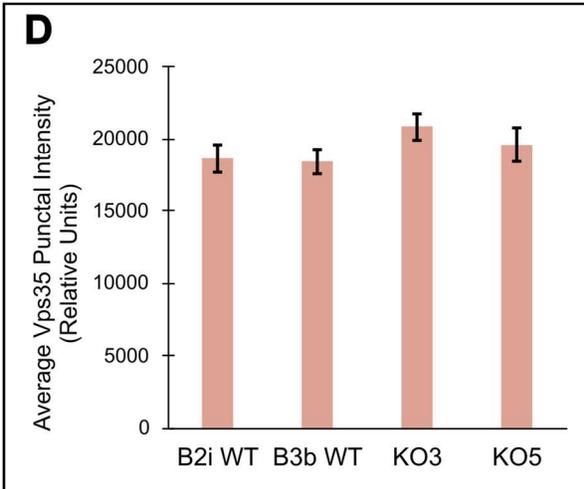
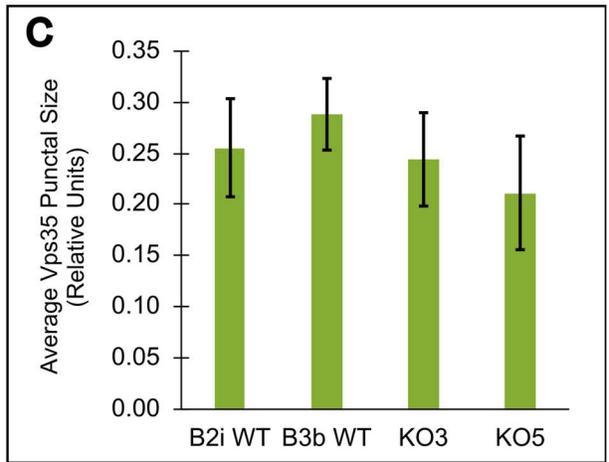
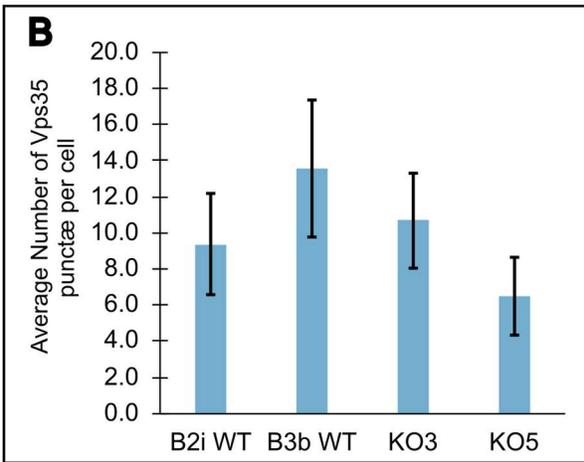
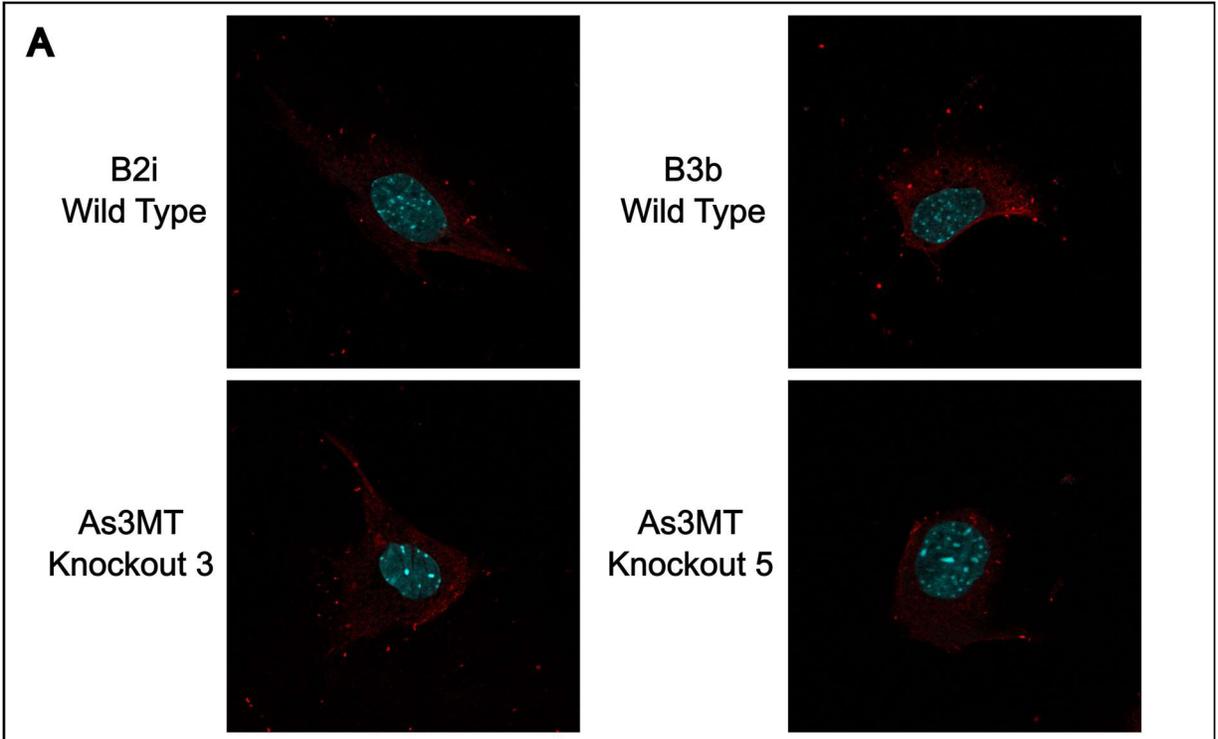


**Figure 13. Differential endocytic trafficking-related protein expression between wild-type and *As3mt* knockout MEFs.** (A) A volcano plot of LC-MS spectral count data displays fold differences of spectral count averages between wild-type versus *As3mt*<sup>-/-</sup> MEFs. The highlighted data represent the differentially expressed proteins selected by Enrichr corresponding with endocytic trafficking. (B) The VPS29 immunoblot suggests that there is higher basal VPS29 expression in the *As3mt*<sup>-/-</sup> MEFs compared with the wild type control. (C) The myoferlin immunoblot is representative of four individual experiments. It indicates that myoferlin is upregulated basally in *As3mt* knockout MEFs.

VPS35 immunoblots showed a doublet around 92 kDa in which the lower molecular weight band matched the LC-MS data (higher expression in *As3mt* knockout MEFs) while the higher molecular weight band did not show an intensity difference. Without an immunoblot with a *Vps35* knockout control, the immunoblots did not conclusively support the interpreted VPS35 upregulation in the *As3mt* knockout MEFs.

The distribution of VPS35 between wild-type and *As3mt*<sup>-/-</sup> MEFs was evaluated by confocal microscopy. VPS35 was prioritised because it is the central subunit of the retromer cargo-selective complex, whereas VPS29 and VPS26 participate in forming the inner and outer shells of the retromer coat (Norwood et al., 2011; Simonetti & Cullen, 2018). Here, an additional wild-type MEF population (“B3b”) was included as a control in addition to the initial wild-type population (“B2i”) that was used in the mass spectrometry, immunoblots, and phenotypic experiments.

In preliminary images, a punctal VPS35 staining pattern was observed (**Figure 14**). A larger sample of images was gathered to measure four variables: the number of VPS35 puncta per cell, average punctal size per cell, average punctal intensity per cell, and the average distance of puncta from the nuclear centroid. Analysis of the images was done on ImageJ with the Macro program (**Appendix B**) in order to quantify these parameters across cells in a consistent and unbiased manner. The results from the analysis of the images did not suggest a coherent, AS3MT-associated phenotype regarding the punctal staining pattern of the VPS35. There was neither a significant AS3MT-associated difference in VPS35 punctal intensity, size, number, nor distance from the nucleus. Thus, while total levels of VPS29 were increased in *As3mt*<sup>-/-</sup> MEFs, VPS35 levels were not significantly altered.



**Figure 14. Immunofluorescent VPS35 staining in wild-type and *As3mt*<sup>-/-</sup> MEFs.** (A) Images were acquired using the ZEISS LSM 800 confocal microscope. Anti-VPS35-stained, DAPI-stained MEFs were imaged using the 63X objective. The VPS35 AF594 signal is red whereas the DAPI signal is cyan. (Panels B–E) VPS35 punctal quantification was performed on single cells with the following sample sizes: B2i wild-type N = 23 cells, B3b wild-type N = 30 cells, KO3 N = 31 cells, KO5 N = 21 cells. (B) average number of puncta per cell. (C) average punctal size (average of intracellular averages). (D) average punctal signal intensity (average between intracellular averages). (E) average punctal distance from nuclear centroid (average between intracellular averages). All error bars represent 95 % confidence intervals.

Additionally, a qPCR experiment was performed to measure the expression of the CSC subunits at the mRNA level (**Supplementary Figure 2**). There were no differences in mRNA levels of any retromer subunit between the *As3mt* genotypes. Based on this information, it was concluded that any AS3MT-dependent differences in the expression of VPS35, VPS29, or VPS26A were not due to differences in mRNA levels.

The retromer complex recognises cargo receptors on endosomal membranes for retrieval back to the trans Golgi network (TGN) (Bonifacino & Rojas, 2006; Seaman, 2012). Retromer recycling is essential for the efficient targeting of cargo to the lysosome (Seaman, 2012). One of the cargo receptors retrieved by retromer recycling is the mannose-6-phosphate receptor (CI-MPR) (Seaman, 2012). At the TGN, the mannose 6-phosphate receptor (M6PR) binds nascent M6P-tagged lysosomal hydrolases for trafficking to endosomes (Seaman, 2004). Two examples of M6P-tagged cargo proteins are LAMP1 and LAMP2 (Saftig & Klumperman, 2009). Dysfunctional CI-MPR retrieval results in the loss of CI-MPR at the TGN membrane (Seaman, 2004). According to the LC-MS data, CI-MPR was present at lower levels in the two *As3mt* knockout MEF populations, although the expression difference was not significant. Lower expression of cargo receptors is consistent with a phenotype of less efficient retrieval from the endolysosomal membrane. At the same time, the LFQ data indicated that there was a significant 1.6-fold higher basal expression of LAMP1 in the *As3mt*<sup>-/-</sup> MEFs ( $p = 0.017$ ), while there was no

significant difference in LAMP2 levels. In summary, the LC-MS data suggested that all three retromer CSC subunits are expressed at significantly higher levels in the *As3mt* knockout MEFs, but the immunofluorescence image analysis of VPS35 puncta did not show higher signal intensity in the *As3mt* knockout MEFs.

#### **4.6 Hypothesis testing with expanded model and *As3mt* addback MEFs**

##### **4.6.1 Expanding MEF model with more wild-type controls**

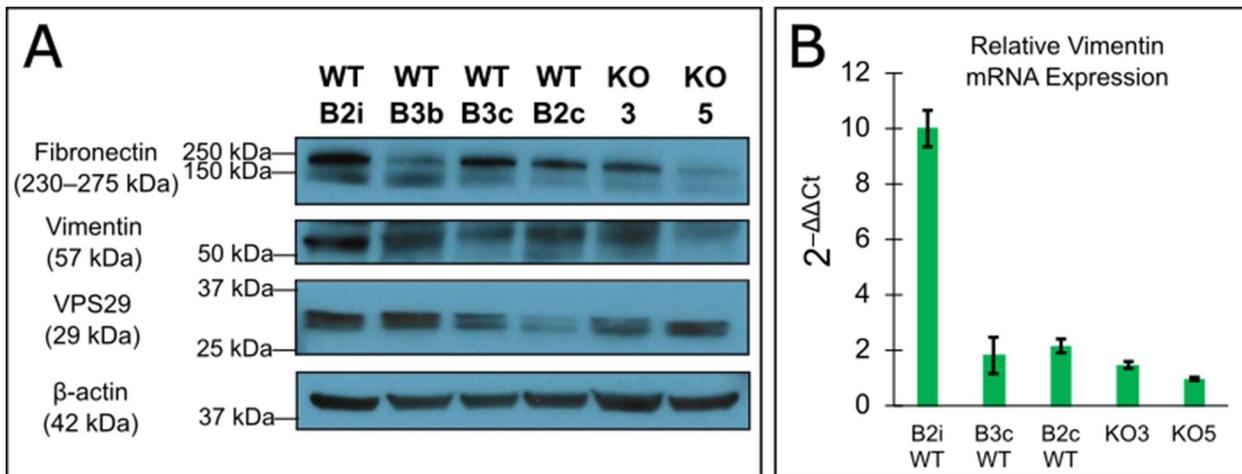
Most of the previous data was performed comparing the original wild-type population (B2i) with two *As3mt*<sup>-/-</sup> MEFs. However, differences were observed between the two *As3mt*<sup>-/-</sup> MEF populations and this could have been due to heterogeneity between cells of the same *As3mt* genotype or differences between sexes. Wild type B2i was female, KO3 was male, and KO5 was female (**Supplementary Figure 3**). Therefore we extended our analyses to an additional set of wild-type MEFs. Also, we wanted to determine whether add-back of AS3MT could reverse the proteomic changes in knockout MEFs and test whether methylation activity of AS3MT was required for differential expression of identified proteins.

Three new wild-type MEF populations (“B3b”, “B3c”, and “B2c”) were introduced as controls in addition to the initial wild-type population (“B2i”) that was used in the experiments presented thus far.

Fibronectin, an ECM glycoprotein, was introduced as another immunoblot target because it participates in cell growth, adhesion, and migration (Bradshaw, 2016; Patten & Wang, 2021), and it was therefore relevant with respect to the pathways and connected phenotypes investigated. In a preliminary immunoblot, the two *As3mt*<sup>-/-</sup> MEF populations showed lower expression of the fibronectin compared to the B2i wild-type MEF control (**Supplementary Figure 4**), which agreed with the LFQ data (KO/WT fold difference = 0.319,  $p = 3.71 \cdot 10^{-2}$ );

however, the newly-introduced B3b wild-type control showed much lower expression and contradicted the pattern of lower fibronectin expression in the *As3mt*<sup>-/-</sup> MEFs.

Immunoblotting for vimentin, fibronectin, and VPS29 was performed on the expanded set of wild-type MEFs and the KO MEFs (**Figure 15A**). For vimentin and fibronectin, the original finding of higher expression in the B2i wild-type MEFs than the KO MEFs remained, however we did not observe a significant increase in VPS29 in the *As3mt*<sup>-/-</sup> MEFs (**Figure 15A**). These immunoblots revealed extensive variability among wild-type MEF populations, making it impossible to conclude that there is any AS3MT-associated pattern in the expression of these targets.



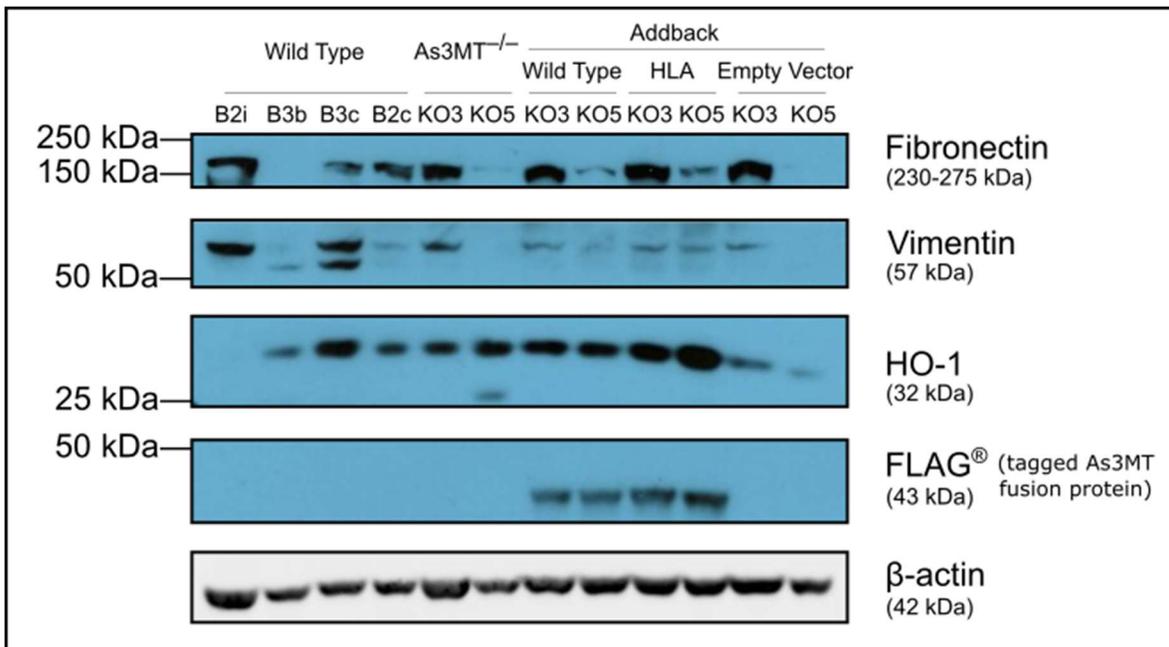
**Figure 15. Variability in expression of mass spectrometry hits among wild-type MEF populations.** (A) Immunoblots for vimentin, VPS29, and fibronectin. (B) Vimentin qPCR data reveals variability in *Vim* mRNA expression among wild-type MEF populations. Relative mRNA expression was quantified using the  $2^{-\Delta\Delta C_t}$  method and using Rn18S (18S rRNA) as the control transcript. qPCR was carried out with three technical triplicates. Error bars represent the SEM of technical triplicates.

According to the qPCR data (**Figure 15B**), there was a ten-fold higher expression of vimentin mRNA in the B2i wild-type MEFs compared with the *As3mt* knockout populations. This fold difference in vimentin mRNA expression between the B2i wild-type versus *As3mt* knockout MEFs aligns with the fold difference of 0.117 (KO/WT) observed for the vimentin protein according to the LC-MS label-free quantification data. However, the B2i wild-type MEF

population is an outlier relative to the two other wild-type populations, which showed a similar vimentin mRNA expression level relative to the *As3mt* knockout MEFs. This weakens the evidence for an AS3MT-associated difference in basal vimentin expression.

#### 4.6.2 Hypothesis testing with *As3mt* addback MEFs

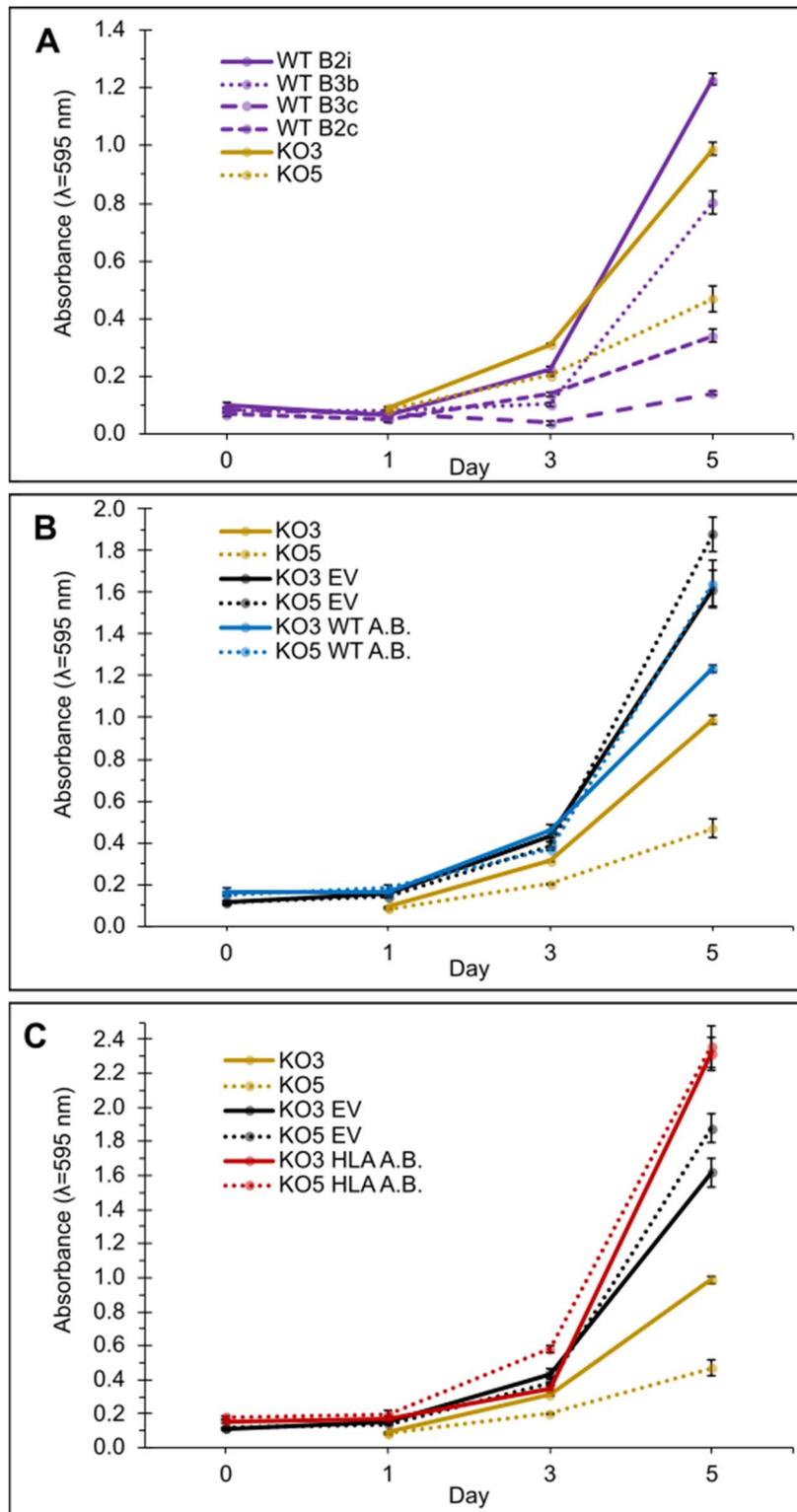
*As3mt* addback MEFs were studied to evaluate whether the LC-MS hits were basally expressed in an AS3MT-dependent manner. The knockout MEFs were reconstituted with either the wild-type murine *As3mt* construct or a methylation-dead mutant form of the gene that features codons encoding HLA in the place of DLG in exon 3. Immunoblots were performed with these MEFs for fibronectin, vimentin, and HO-1 (**Figure 16**). FLAG<sup>®</sup> expression was observed only in the MEFs that expressed the wild-type or the HLA mutant *As3mt* constructs. Reconstitution with *As3mt*, either wild-type or the methyltransferase-dead mutant, neither altered the expression of HO-1, vimentin, nor fibronectin within the *As3mt* knockout MEFs (**Figure 16**).



**Figure 16. Fibronectin, vimentin, and HO-1 expression do not show AS3MT dependence in MEFs.** Representative of two individual experiments. Immunoblots for HO-1, vimentin, fibronectin, and FLAG<sup>®</sup>. Wild-type and methylation-dead HLA *As3mt* addback populations are labelled “WT A. B.” and “HLA A. B.” and empty vector controls are labelled “EV”.

The evidence that *As3mt* knockout is associated with basal differences in HO-1, vimentin, and fibronectin expression was weakened by these data because they are inconsistent. The expression of these three targets varied greatly between wild-type MEFs. At the same time, the wild-type *As3mt* addback MEFs did not consistently undergo restoration of expression of any of these three targets to a level similar to the B2i wild-type MEF population.

Finally, growth curves were also performed to test whether there was also heterogeneity among wild-type MEF populations and whether reconstitution of knockout MEFs with *As3mt* increased growth rate. When compared to the original wild-type MEF line (B2i), all other wild-type MEFs had slower proliferation rates, and some were similar to those observed in the *As3mt* knockout MEFs (**Figure 17A**). Furthermore, while *As3mt* reconstitution increased the growth rates of the knockout MEFs from which they were established, the empty vector controls displayed comparable increases (**Figure 17B**). Interestingly, reconstitution with the methylation-dead *As3mt* led to the fastest growth rates in both *As3mt* knockout MEF lines (**Figure 17C**). Therefore, there is significant heterogeneity even amongst wild-type MEFs alone, while restoration of wild-type As3MT expression did not enhance growth rate.



**Figure 17. Growth rate does not show AS3MT dependence in MEFs.** Crystal violet growth curve displays the proliferation rates of (A) Proliferation rates of wild-type (WT) and *As3mt* knockout (KO), (B) wild-type *As3mt* addback MEFs (WT A. B.), and (C) methylation-dead mutant *As3mt* addback MEFs (HLA A. B.). Error bars represent the SEM of four technical replicates within one biological experiment.

## 5. DISCUSSION

AS3MT is the enzyme responsible for catalysing methylation of arsenic to expedite urinary arsenic elimination (Stýblo et al., 2000; Stýblo et al., 1995; Vahter, 1999). Apart from arsenic methylation, no other function of AS3MT has yet been established. The existence of an additional activity of AS3MT may explain why there are GWAS associations of the *AS3MT* locus with neuropsychiatric and cardiovascular diseases, why there is such unexpectedly high evolutionary conservation of the AS3MT protein, and why *As3mt* knockout mice present with higher adiposity and abnormal hepatic phosphatidylcholine metabolism (Huang et al., 2016; Douillet et al., 2017).

The hypothesis addressed in this thesis is that AS3MT has an additional function independent of arsenic methylation. This thesis aimed to pinpoint what pathway(s) this additional activity involves if a basal AS3MT activity indeed exists.

The hypothesis was approached through a proteomic lens. LC-MS was employed to screen for proteins that have differential expression associated with *As3mt* knockout in mouse embryonic fibroblasts. Pathway enrichment analysis was used to identify AS3MT-associated expression changes at the pathway level. A handful of protein hits were validated by immunoblot for expression differences between the one wild-type and *As3mt* two knockout MEF populations that were used in the LC-MS experiment. Proliferation, scratch wound migration, and collagen contractility assays were performed on the MEFs to evaluate whether there are AS3MT-associated phenotypic differences that relate with the suspected pathways affected by *As3mt* knockout. Lastly, *As3mt* addback MEFs were generated to test the AS3MT-dependence of the associated protein expression differences and growth phenotype. Separately, an IP-MS dataset was analysed to screen for putative interacting partners of AS3MT in AML12 hepatocytes.

According to the pathway enrichment analysis, basal AS3MT expression was associated with protein expression differences in adherens junction and endocytic trafficking pathways, but the finding was limited to the comparison with a single wild-type MEF population (“B2i WT”). When the hypotheses were tested with the *As3mt* addback MEFs, the differences in protein expression and proliferation rate were not reversed with exogenous AS3MT expression.

## ***5.1 Summary and interpretation of results***

### *5.1.1 IP-MS data*

Proteins related with cell adhesion and endocytic trafficking pathways were enriched in the pull-downs for AS3MT-specific interactions. Clathrin, dynein, myoferlin, and evectin-2 were observed as IP-MS hits related to trafficking. The trafficking-related hits attracted attention to retrograde endocytic trafficking and clathrin-dependent endocytosis pathways. Clathrin relates to retromer complex recycling because of the tight coupling between retromer and clathrin activities (Johannes & Popoff, 2008). It has been reported that clathrin and its adapter proteins are involved in cargo recognition for retrograde sorting on early endosomes, and that, upon clathrin depletion, there is a loss of the retromer coat from vesicle preparations and an inhibition of retrograde transport (Borner et al., 2006; Johannes & Popoff, 2008). Evectin-2, also known as Pleckstrin Homology Domain-Containing Family B Member 2, is located on the recycling endosome membrane and it participates in the retrograde transport from these endosomes to the trans-Golgi network (Okazaki et al., 2012).

Desmoyokin, plectin, afadin, supervillin, and tight junction proteins 1 (ZO-1) and 2 (ZO-2) were identified as IP-MS hits that related to adherens and tight junctions. In addition to its function as an adapter in tight junctions, tight junction protein 1 has been reported to have a role in adherens junction regulation as an adapter (Hartsock & Nelson, 2008). It has been

hypothesised that tight junction protein 1 serves as a scaffolding protein between cytoplasmic and transmembrane proteins involved in adherens junctions and tight junctions, and that it forms a link between these two types of adhesion complexes (Hartsock & Nelson, 2008). Afadin—another adherens junction adapter protein—is also present in the list. Afadin interacts with tight junction protein 1 while recruiting other proteins to adherens junctions (Birukova et al., 2012). Plectin is an adhesion-complex component that is relevant to cytoskeletal organisation and cell adhesion because it interlinks all three major cytoskeletal filament systems and it links the cytoskeleton to junctions at the plasma membrane (Přechová et al., 2023). Plectin also anchors mitochondria to vimentin filaments (Chang & Goldman, 2004; Winter et al., 2008). Since tight junction proteins 1 and 2, plectin, and afadin are all examples of anchoring junction adapters, the data also suggested that there was affinity between the FLAG<sup>®</sup>-AS3MT bait protein and an adherens junction or tight junction protein.

Having performed the experiment with AML12 hepatocytes, these preliminary IP-MS data independently reinforced the focus on cell adhesion and trafficking pathways, while also providing a perspective outside of the MEF model and its limitations. The correspondence between the LC-MS and IP-MS pathways was encouraging because it signified more expandability of the AS3MT-associated patterns across cell types. If the IP-MS hits are validated in future experiments and they belong to a pathway highlighted by the enrichment analysis in the MEFs, this would suggest that AS3MT interacts with one or more proteins belonging to a pathway that shows AS3MT-associated differences across multiple cell types.

### *5.1.2 Comparison to existing AS3MT knockout screening*

This proteomics investigation of *As3mt*<sup>-/-</sup> MEFs generated findings that share similarities with a prior shotgun screening of AS3MT-associated differences, especially with regard to ECM

pathways and cell adhesion. In an RNA-seq screening that compared *AS3MT*<sup>+/+</sup> and *AS3MT*<sup>-/-</sup> SH-SY5Y neurons in basal conditions, focal adhesions and cell adhesion were reported as AS3MT-associated pathways (Washer et al., 2022). There is overlap at the pathway level between the RNA-seq report and the results from the MEFs, although the top-ranked DEGs reported by the authors were not the same hits identified in our LC-MS data. The authors reported significant AS3MT-associated changes in gene expression related extracellular matrix formation, among other pathways, and hypothesised that the *AS3MT* knockout SH-SY5Y cells presented with an impairment in cell-cell contacts (Washer et al., 2022). In summary, Washer *et al.*'s findings from RNA-seq overlap with the results from the MEFs because these data both hint that disrupted cell adhesion is associated with *As3mt* knockout.

### 5.1.3 Adherens Junctions

The LC-MS data suggested that the *As3mt* knockout MEFs present with basally depleted levels of adherens junction proteins, namely catenins, cadherins, and adherens junction adaptors. For example,  $\beta$ -catenin—one of the validated hits—is an adherens junction adapter because it bridges cadherins with actin filaments (Brembeck et al., 2006; McCrea & Gu, 2010). The AS3MT-associated difference in  $\beta$ -catenin is important because  $\beta$ -catenin is at the crossroads of many pathways.  $\beta$ -catenin participates in cytoskeletal organisation and adhesion as a transcriptional coactivator in Wnt signalling, and in modulation of many signalling pathways influenced by the nuclear signalling pool of  $\beta$ -catenin (Brembeck et al., 2006; McCrea & Gu, 2010). Given the role in Wnt signalling, the lower basal  $\beta$ -catenin expression in the *As3mt* knockout MEFs is consistent with slower proliferation. Diseases of development and morphogenesis are predicted to result from  $\beta$ -catenin and p120-catenin deficiencies because they participate in cadherin-dependent adhesion, cell polarity, and motility (McCrea & Gu, 2010).

Additionally,  $\beta$ -catenin deficiency in endothelial cells has been reported to raise cell permeability and to cause blood-brain barrier breakdown in mice (Hussain et al., 2022). This relates to the differential  $\beta$ -catenin expression observed in *As3mt* knockout MEFs because of the implication that AS3MT deficiency could give rise to a vulnerability of the blood-brain barrier due to dysfunctional cell adhesion and cell membrane integrity, although validation of AS3MT-dependence of  $\beta$ -catenin expression in endothelial cells and subsequent *in vivo* experiments are required to investigate that hypothesis.

#### 5.1.4 Cytoskeletal Organisation and Vimentin

To explore the phenotype related to cytoskeletal organisation and ECM remodelling, the *As3mt*<sup>+/+</sup> and *As3mt*<sup>-/-</sup> MEFs were studied in scratch wound migration and collagen contractility experiments. These experiments were motivated by the LC-MS data (**Figure 6**) and the preliminary immunoblots (**Figure 7**), which suggested that the *As3mt* knockout was associated with lower basal expression of adherens junction proteins and vimentin.

Vimentin is an intermediate filament protein expressed in cells with ectodermal and mesenchymal origins (Ridge et al., 2022). Vimentin was an attractive target to explore in this thesis because it can be conceptualised as a nexus between many pathways. Vimentin provides mechanical stability on account of its load-bearing function (Ridge et al., 2022). It stabilises the cell by acting as a scaffold that interlinks actin filaments, microtubules, and focal adhesions while also anchoring organelles (Arrindell & Desnues, 2023; Gan et al., 2016; H. Wu et al., 2022). Vimentin filaments contribute tensile stiffness and mechanical stability because they can distort to absorb mechanical stresses and contribute to strain stiffening (Eckes et al., 1998). The vimentin network also integrates mechanical stimuli because it transmits mechanical signals through its anchoring contacts with adhesion complexes (Eckes et al., 1998).

Vimentin deficiency relates to pathology because vimentin knockout mice have reported to show impaired wound healing, leakiness of vascular endothelia, stunted differentiation of endothelial stem cells, increased arterial stiffness, decreased transmigration of leukocytes, and defective ovarian steroidogenesis (Ridge et al., 2022). The phenotype of *Vim*<sup>-/-</sup> MEFs has been established in literature (Eckes et al., 1998). Compared to wild-type MEFs, *Vim*<sup>-/-</sup> MEFs present with a flatter shape than wild-type MEFs, lower stiffness (stress-to-strain ratio) and lower resistance to mechanical shape distortion, slower migration toward chemoattractants and random instead of directional migration, impaired contractility of collagen lattices, defective organisation of collagen fibrils, and abnormal formation of the actin network and of focal contacts (Eckes et al., 1998). Based on the lower vimentin expression that was associated with the *As3mt* knockout MEFs, the *As3mt*<sup>-/-</sup> phenotype was expected to share some of these traits.

#### 5.1.4.1 Vimentin staining pattern in MEFs

A non-filamentous staining pattern in the MEFs was observed in the vimentin immunofluorescence experiment (**Figure 10**). The signal intensity corresponded with the relative expression levels that were indicated by LC-MS and immunoblotting, but the non-filamentous pattern was unexpected given that the protein has been reported to adopt a predominantly filamentous form (Ridge et al., 2022), and considering that MEFs are not an exception (Eckes et al., 1998).

Methodological concerns were ruled out regarding the immunofluorescence, and included:

- The primary anti-vimentin BD Pharmingen™ antibody has yielded filamentous signals in immunofluorescence experiments according to prior studies (Broers et al., 1986; Viebahn et al., 1988).

- The possibility was eliminated that the observed signal was simply background: 8% donkey serum was used for blocking, and the signal was proportional to the vimentin expression level between the wild-type and *As3mt* knockout 5 MEFs (the highest- and lowest-vimentin-expressing populations).
- 4 % PFA and 100 % methanol fixation methods were both attempted. The fixation method did not affect the appearance of the vimentin staining pattern.
- The quality of the microscopy was evaluated insofar as the ability to capture filamentous signals was verified with a  $\beta$ -actin control staining. *AiryScan* deconvolution was employed.
- A z-stack maximum intensity projection of anti-vimentin-stained MEFs was carried out to see whether the observed non-filamentous pattern was due to a limitation of two-dimensional image acquisition. The z-stack did not change the pattern.

The observation of the non-filamentous pattern raised the question about whether this form of vimentin has been reported. Non-filamentous vimentin has been documented *in vivo*, and it has been understood to serve non-mechanical functions both intracellularly and extracellularly (Patteson et al., 2020). Extracellular vimentin has been reported to promote migration and tissue repair (Bucki et al., 2023; Patteson et al., 2020). Additionally, studies on human monocytes and microvascular endothelial cells have reported that vimentin can be expressed at the cell surface and that it can also be excreted (Garg et al., 2006; Mor-Vaknin et al., 2003; M. B. Yu et al., 2018). It has been hypothesised that, in these cells, non-filamentous vimentin serves a function related to the innate immune response (Arrindell & Desnues, 2023). Thus, sometimes adopts a non-filamentous form, yet it is not known why the MEFs used in the experiments did not display the expected filamentous staining pattern.

#### 5.1.4.2 Scratch wound migration and collagen contractility assays

Preliminary scratch wound migration and collagen contractility experiments were carried out on the wild-type and *As3mt*<sup>-/-</sup> MEFs. These experiments were performed to assess whether predicted *As3mt*<sup>-/-</sup> MEFs would phenocopy *Vim*<sup>-/-</sup> MEFs based on the lower vimentin and adherens junction protein expression. In the scratch wound migration assay (**Figure 11**), the *As3mt* knockout 5 population showed a similar migration rate to the B2i wild-type MEFs, whereas the *As3mt* knockout 3 population showed slower migration. The two *As3mt*<sup>-/-</sup> MEF populations showed an unexpected difference in migration relative to each other. One possible source of the variation is that the B2i wild-type and KO5 MEFs populations are female whereas the KO3 population is male (**Supplementary Figure 3**). While sex-specific differences in SIRT3-mediated redox and detoxification of reactive oxidative species have been documented in MEFs (Belužić et al., 2024); however, a sex-specific migration difference in MEFs has not been reported. On account of the similar migration rates of the B2i wild-type and *As3mt* KO5 populations, the scratch wound migration assay did not reveal a consistent AS3MT-associated migration phenotype. On the other hand, the collagen contractility data (**Figure 12**) suggested that the B2i wild-type MEF population contracted collagen lattices at a significantly faster and greater rate. These data supported the hypothesis that impaired ECM remodelling is associated with *As3mt* knockout in MEFs. However, these experiments would need to be repeated in order to conclude whether migration or contractility are impaired in *As3mt*<sup>-/-</sup> MEFs.

#### 5.1.5 Endocytic Trafficking

Endocytic trafficking was another hit from the pathway enrichment analysis of the LC-MS data. The data suggested that increased expression of the three retromer CSC subunits was associated with *As3mt* knockout. Although the simultaneously lower expression of the cation-

independent mannose-6-phosphate receptor (CI-MPR) did not reach statistical significance, the higher VPS35, VPS29, and VPS26A protein expression raised the question about whether the *As3mt* knockout is associated with dysfunctional retromer recycling activity. A possible interpretation is that the retromer CSC subunits could be upregulated as an attempt to compensate for a decreased capacity to recycle cargo receptors from the endolysosomal membrane to the TGN. Altogether, these data inspired the hypothesis that there is a basal trafficking deficit in the *As3mt* knockout MEFs. Impaired CI-MPR retrieval results in the lower levels of CI-MPR in the TGN and can lead to improper sorting of lysosomal hydrolases (Seaman, 2004). Dysfunctional retrieval of the M6PR and trafficking of M6P-tagged cargo proteins results in an Inclusion-Cell Disease (ICD)-like phenotype in fibroblasts (Leroy & DeMars, 1967; Reitman et al., 1981). ICD is a lysosomal storage disorder that is caused by defective M6PR trafficking, which results in the accumulation of material in lysosomes (Reitman et al., 1981).

An intracellular trafficking pathway also emerged in Washer *et al.*'s analysis of RNA-seq data, albeit it was not the same pathway as the endocytic trafficking pathway identified in the pathway enrichment analysis of the proteomics data from the MEFs. The top hit reported in Washer *et al.*'s GO analysis was the signal recognition particle (SRP)-dependent cotranslational protein targeting to membranes (Washer et al., 2022). In eukaryotes, this pathway is essential for the localisation of nascent proteins to the endoplasmic reticulum, as it involves the recognition of the signal sequences of cargo proteins and the delivery of the cargo to target membranes (Akopian et al., 2013; Saraogi & Shan, 2011).

#### 5.1.5.1 Retromer complex phenotype of MEFs

The proteins of the retromer complex were explored because it was observed by LC-MS that all three subunits of retromer CSC (the VPS35–VPS29–VPS26 heterotrimer) were significantly upregulated in *As3mt*<sup>-/-</sup> MEFs under basal conditions. VPS35 immunofluorescence (**Figure 14**), VPS29 immunoblotting (**Figure 13B**), and RT-qPCR on all three subunits (**Supplementary Figure 2**) were conducted.

No differences in VPS35 staining pattern were associated with AS3MT expression. According to the VPS29 immunoblot (**Figure 13B**), two *As3mt*<sup>-/-</sup> MEF populations had higher VPS29 expression relative to the B2i wild-type MEFs. This expression difference matched the LC-MS data. Next, the expression of retromer subunits at the mRNA level was evaluated in a qPCR experiment that included more wild-type controls. The qPCR data indicated that differential mRNA expression of retromer subunits was not associated with As3MT expression. Ultimately, the negative results and inconsistency between wild-type MEFs favoured the null hypothesis concerning AS3MT expression and the retromer complex phenotype.

### 5.2 Limitations

#### 5.2.1 Narrow scope of the proteomics approach in MEFs

Even if the alternative hypothesis is true that there indeed is an additional activity of AS3MT, this screening approach would not detect the relevant target(s) or pathway(s) in the event that:

- the additional activity does not affect protein expression (e. g. if it concerns regulation at the DNA or RNA level, if it involves a posttranslational modification, if it represents a protein-protein interaction, or if it relates metabolism in an isolated way that not impact protein expression);

- the additional activity is cell type- or organ-specific and not relevant in MEFs, hence the MEFs do not present with a distinct *As3mt*<sup>-/-</sup> phenotype (for example, if the activity is only applicable to neuron or hepatocyte function);
- the additional activity relates only to the response to challenge (i. e. it is not necessary for homeostasis in basal conditions); or,
- any combination of these scenarios is true.

Although a proteomics scope is narrow and it has blind spots, it does provide a relatively wide and unbiased view of cell functions. For this reason, the shotgun proteomics was considered as an adequate starting point from which to investigate the hypothesis of additional AS3MT activity.

### *5.2.2 Limitations of the MEF model*

The inclusion of only one wild-type MEF population was an experimental design flaw in the LC-MS experiment, downstream pathway enrichment analysis, and early immunoblot hit validation. Ideally, more wild-type MEFs as well as *As3mt* addback MEF populations would also have been studied together at the outset of the project. This experimental design problem made the screening vulnerable to phenotypic variability among wild-type MEF populations, as the experiment was prone to steering toward any pathways where the single wild-type population may have been anomalous in ways independent of AS3MT activity. This flaw especially weakens the potential conclusions given the observed behaviour of the B2i wild-type population as an outlier. However, it should be noted that pathways such as adherens junctions were also found in the IP-MS dataset that was derived from the AML12 murine hepatocyte population.

The results from the immunoblots and phenotyping assays suggested that the original B2i wild-type MEF population, which was used as the control in the LC-MS experiment, was in fact

unique in multiple ways. This was the MEF population that was used as the control in the LC-MS experiment. The B2i wild-type MEFs showed much higher expression of vimentin and fibronectin compared to the other three wild-type controls, while it also expressed HO-1 at levels that were much lower than the other wild-type controls (**Figure 15**, **Figure 16**). This wild-type MEF population also had the highest proliferation rate (**Figure 17**) of all four wild-type MEF populations. This experimental flaw carried more weight because the LC-MS was the most upstream experiment and it provided that basis for downstream hypotheses.

Gene knockouts are advantageous compared with RNAi or morpholino knockdowns because they typically suffer from fewer off-target effects and they yield a complete loss of function (Zimmer et al., 2019). However, a potential weakness of the *As3mt* knockout MEF model concerns the rewiring of pathways that occurs in cells upon gene knockouts via genetic compensation. It has been reported that knockout mutants do not always show the same phenotype as knockdowns generated by RNAi or morpholino antisense oligomers (Zimmer et al., 2019). This limitation could be addressed by employing an *As3mt* siRNA knockdown in conjunction with the *As3mt* gene knockout so that results can be vetted for consistency between both methods.

Lastly, another potential limitation is the validity of the established *As3mt* addback MEFs as the model to verify the reversal of a phenotype. If AS3MT performs an additional function, it is possible that *As3mt*<sup>-/-</sup> MEFs adapted to the absence of AS3MT after having survived basal cell culture conditions without AS3MT for some time. If this is true, then the MEFs reconstituted with AS3MT would not necessarily phenocopy the wild-type MEFs, since the wild-type MEFs would not have experienced the pressure to adapt to homeostasis without AS3MT expression.

Primary non-immortalised MEFs may be useful as a model to minimise the concern of *As3mt*<sup>-/-</sup> adaptation.

### 5.2.3 LC-MS and data analysis

False discovery rate is a concern when hypothesis-testing is done by screening using shotgun omics (Burger, 2023). On account of its broad focus, the approach used in this thesis is an example of discovery proteomics rather than targeted proteomics. Omics data is often prone to having high noise (Ning & Lo, 2010). Discovery proteomics has also been viewed as a fishing expedition in some instances when it is used to aimlessly search for interesting differences (Ning & Lo, 2010). This danger arises due to the observer bias (or detection bias), as an observer will find likely differences when differences are searched for.

With these risks and limitations in mind, the purpose of the shotgun proteomic screening in this thesis was the generation of hypotheses and the narrowing of the focus, rather than the presentation of differential protein expression on its own as the sole evidence for conclusions. Thus, the uncorrected T-tests were used to identify hits as the subset of proteins to be used for pathway enrichment analysis. The exclusion of the vast majority of these proteins by a multiple test correction would not provide a viable input for the pathway enrichment analysis. Furthermore, within the experimental design there were downstream safeguards against false positives: LC-MS hits were validated by westerns, and after all, the data from the *As3mt* addback MEFs was intended to give the final word about AS3MT dependence.

Another limitation of LC-MS is the potential for batch effects that originate from technical variability. Batch effects are problematic because they introduce noise and hinder reproducibility (Čuklina et al., 2021). Sources of batch effects include ion intensity drift and missing peptide quantity values (Čuklina et al., 2021). The methods in this thesis aimed to

maximise the comparability between samples. All samples were prepared from MEFs cultured around the same time and at low passage numbers, and via the same protein harvesting protocol and lysis buffer preparation. Normalisation was performed in the spectral count analysis through the Scaffold 5 proteome software and through the established LFQ data analysis method (Aguilan et al., 2020).

#### 5.2.4 Immunoblot validation of LC-MS hits

According to the LC-MS data, vimentin,  $\beta$ -catenin, myoferlin, VPS29, and fibronectin showed significantly different basal expression between the B2i wild-type and *As3mt*<sup>-/-</sup> MEFs (**Figure 6**). Between these three MEF populations, the AS3MT-associated expression differences of these hits were validated by immunoblot (**Figure 7, Figure 9, Figure 13**). However, the introduction of more wild-type MEF controls in immunoblots revealed that these proteins are variably expressed among wild-type MEF populations (**Figure 15**). This raised the concern that the B2i wild-type MEF population is not phenotypically representative of wild-type MEFs. Therefore, the expression of these hits did not show a robust association with AS3MT expression when the experimental design was expanded to include more wild-type control populations. Furthermore, the immunoblots with *As3MT* addback MEFs indicated that HO-1, vimentin, and fibronectin do not show AS3MT-dependent basal expression patterns as they were predicted (**Figure 16**).

Overall, the results did not support the hypothesis that the LC-MS hits are differentially expressed in an AS3MT-dependent manner.

#### 5.2.5 Proliferation curves of *As3mt* addback MEFs

The growth phenotype was tested for AS3MT-dependence with the *As3mt* addback MEFs (**Figure 17**), and it was revealed that the proliferation rate did not follow an AS3MT-dependent relationship. The alternative hypotheses were not accepted with respect to these predicted *As3mt*

knockout MEF phenotypes. Furthermore, the data demonstrated that there was extensive variability between wild-type MEF populations. It was also observed that the B2i wild-type MEF population displayed the fastest proliferation rate of the four wild-type control populations.

#### *5.2.6 IP-MS and data analysis*

A limitation of the IP-MS method is that it aims to capture protein interactions in artificial conditions, for example:

- the protein interactions that only occur in a hepatocyte lysate cannot be distinguished from interactions that only take place in intact cells;
- there is overexpression of AS3MT instead of expression at physiological levels; and,
- the possible influence of the FLAG<sup>®</sup> tag on AS3MT interactions.

These limitations of the IP-MS experiment are compounded by the low experimental N of the preliminary data and the low spectral counts. Among the top five hits, the average differences between the spectral counts of the AS3MT bait IP and empty vector control IP were 10 to 17 spectral counts. Therefore, the analysis was crude due to the limited data. On account of the low experimental N and low spectral counts, a statistical analysis was not applied to the comparison. For these reasons, a definitive conclusion about interacting partners is not claimed based on these data. To establish an AS3MT interaction, the enrichment of a hit must also be validated by performing immunoblots on the pull-downs. These IP-MS results serve as preliminary data that can be compared with a future, more robust experiment (e. g. a phage display or biotinylation (BioID) experiment)).

### 5.3 Future Directions

After interpreting the data presented in this thesis as a whole, it was considered that it is not productive to continue parsing the LC-MS dataset and pathway enrichment analysis for associations between the one wild-type MEF control. Considering the outlying behaviour of the B2i wild-type MEFs and the problem of variability among wild-type MEF populations within the model, it would provide more value to perform LC-MS together with the other wild-type populations and the *As3mt* addback MEFs at the same time as the *As3mt* knockout MEFs. This screening would immediately narrow down the proteins, if there are any, that exhibit an AS3MT-dependent expression difference. The pursuit of any hits from this comparison would ensure that an investment is made only in the most promising candidates for AS3MT-dependent expression differences. Alternatively, shotgun proteomics could be carried out with an entirely different cell type, for example, *AS3MT*<sup>+/+</sup> and *AS3MT*<sup>-/-</sup> human hepatocytes or neurons. The results of either of these experiments would determine whether the focus is warranted on adherens junctions, cytoskeletal organisation, and endocytic trafficking.

### 5.4 Implications

The findings in this thesis highlight a unique focus on the relationship between basal AS3MT expression and cell adhesion, cytoskeletal organisation, and trafficking pathways in mouse embryonic fibroblasts. Before the results can be applied to future studies, further validation is needed by screening for proteomic differences in an expanded model. It is recommended that this future model adds multiple wild-type MEF populations together with *As3mt* addback MEF populations, or that it centres around a *AS3MT*<sup>+/+</sup> versus *AS3MT*<sup>-/-</sup> model in a different cell type.

GWAS and the unexplained conservation of AS3MT together hint that an additional, undefined AS3MT function may be related to pathology unrelated to arsenic toxicity. If AS3MT

plays a homeostatic role in basal conditions, it remains to be elucidated how genetic variation in the *AS3MT* gene alters disease risk related to the additional activity, and whether this influences the risk of the neuropsychiatric and cardiovascular conditions that emerged in the GWAS. This knowledge gap underscores the value of future investigation into *AS3MT*. Human *AS3MT* genotype analysis could be harnessed for diagnostic purposes as a personalised medicine approach—not only for predicting the risk level associated with arsenic methylation rate, but also the risks associated with a pathology related to the additional activity.

Furthermore, an additional activity of *AS3MT* could also prove to be pathologically relevant if it interacts with the effects of arsenic toxicity. Arsenic-induced changes that depend on *AS3MT* activity can be direct or indirect. The direct effects concern the clearance rate of inorganic arsenic as it depends on the conversion to more rapidly-excreted metabolites, whereas the indirect effects involve the potential effect on oxidative stress resilience due to glutathione depletion, the burden on one-carbon metabolism, and production of methylated arsenicals that have a unique toxicity profile (activation of arsenic). Another indirect effect of arsenic could be its interference with the additional activity of *AS3MT*. This would be the case if arsenic, while it occupies the active site of *AS3MT*, acts as a competitor against the unknown substrate or interacting partner involved in the additional *AS3MT* activity. An interaction between arsenic and this additional activity could partially explain unknown mechanisms of arsenic toxicity. Therefore, further investigation about a putative additional function of *AS3MT* could increase the understanding of the mechanisms underlying arsenic-promoted pathogenesis, and this would be useful in the design of therapies to prevent or counteract arsenic-induced pathogenesis.

## 6. CONCLUSION

The experiments in this thesis used an *As3mt* knockout mouse model to screen for proteomic differences between *As3mt*<sup>+/+</sup> and *As3mt*<sup>-/-</sup> MEFs. Pathway enrichment analysis was performed to gauge which pathways were associated with a disturbance in basal protein expression among the *As3mt*<sup>-/-</sup> MEFs. Cytoskeletal organisation, cell adhesions, and retrograde endocytic trafficking were identified as pathways that showed AS3MT-associated differences in basal protein expression. Further shotgun proteomic experiments with 1) more wild-type MEF control populations and 2) *As3mt* addback MEFs are required to verify that these pathways show association with and dependence on AS3MT activity respectively. The immunoblots of HO-1, vimentin, and fibronectin on *As3mt* addback MEFs indicated that these hits identified by LC-MS did not show AS3MT-dependent expression patterns. Preliminary data suggested that the *As3mt* knockout MEFs did not present with an overt growth or migration phenotype relative to the wild-type control populations but did suggest less collagen contractility in *As3mt*<sup>-/-</sup> MEFs. Based on these data alone, the null hypothesis about basal *As3mt*<sup>-/-</sup> phenotype can be neither accepted nor rejected, and therefore experiments with *AS3MT* knockout cells from other tissues are required to evaluate the hypothesis. IP-MS data from AML12 hepatocytes suggested that proteins involved in cell adhesion and endocytic trafficking were enriched in AS3MT pull-downs. Future validation of putative AS3MT-interactors is needed in order to make a conclusion about AS3MT interactions. Further investigation of an additional AS3MT activity could assess whether AS3MT deficiencies, polymorphisms, or mutations are responsible for a greater risk of neuropsychiatric, cardiovascular, or metabolic diseases. Additionally, if an additional AS3MT activity is susceptible to interference from arsenic exposure, these future studies could give rise to a more complete understanding about cell type- and organ-specific mechanisms of arsenic toxicity.

## 7. REFERENCES

- Aguilan, J. T., Kulej, K., & Sidoli, S. (2020). Guide for protein fold change and *p*-value calculation for non-experts in proteomics. *Molecular Omics*, *16*(6), 573–582. doi: 10.1039/D0MO00087F
- Agusa, T., Fujihara, J., Takeshita, H., & Iwata, H. (2011). Individual Variations in Inorganic Arsenic Metabolism Associated with AS3MT Genetic Polymorphisms. *International Journal of Molecular Sciences*, *12*(4), 2351–2382. doi: 10.3390/ijms12042351
- Akopian, D., Shen, K., Zhang, X., & Shan, S. (2013). Signal Recognition Particle: An Essential Protein-Targeting Machine. *Annual Review of Biochemistry*, *82*(1), 693–721. doi: 10.1146/annurev-biochem-072711-164732
- Arrindell, J. & Desnues, B. (2023). Vimentin: from a cytoskeletal protein to a critical modulator of immune response and a target for infection. *Frontiers in Immunology*, *14*. doi: 10.3389/fimmu.2023.1224352
- Belužić, R., Šimunić, E., Podgorski, I. I., Pinterić, M., Hadžija, M. P., Balog, T., & Sobočanec, S. (2024). Gene Expression Profiling Reveals Fundamental Sex-Specific Differences in SIRT3-Mediated Redox and Metabolic Signaling in Mouse Embryonic Fibroblasts. *International Journal of Molecular Sciences*, *25*(7), 3868. doi: 10.3390/ijms25073868
- Benjamini, Y. & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
- Bernatchez, P. N., Sharma, A., Kodaman, P., & Sessa, W. C. (2009). Myoferlin is critical for endocytosis in endothelial cells. *American Journal of Physiology-Cell Physiology*, *297*(3), C484–C492. doi: 10.1152/ajpcell.00498.2008
- Bertocchi, C., Vaman Rao, M., & Zaidel-Bar, R. (2012). Regulation of Adherens Junction Dynamics by Phosphorylation Switches. *Journal of Signal Transduction*, *2012*, 1–14. doi: 10.1155/2012/125295
- Bhumbla, D. K. & Keefer, R. F. (1994). *Advances in Environmental Science and Technology* (Vol. 26). New York, NY: John Wiley & Sons, Inc.
- Birukova, A. A., Fu, P., Wu, T., Dubrovskiy, O., Sarich, N., Poroyko, V., & Birukov, K. G. (2012). Afadin controls p120-catenin–ZO-1 interactions leading to endothelial barrier enhancement by oxidized phospholipids. *Journal of Cellular Physiology*, *227*(5), 1883–1890. doi: 10.1002/jcp.22916
- Bonifacino, J. S. & Rojas, R. (2006). Retrograde transport from endosomes to the trans-Golgi network. *Nature Reviews Molecular Cell Biology*, *7*(8), 568–579. doi: 10.1038/nrm1985
- Borner, G. H. H., Harbour, M., Hester, S., Lilley, K. S., & Robinson, M. S. (2006). Comparative proteomics of clathrin-coated vesicles. *The Journal of Cell Biology*, *175*(4), 571–578. doi: 10.1083/jcb.200607164
- Bradford, M. (1976). A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Analytical Biochemistry*, *72*(1–2), 248–254. doi: 10.1006/abio.1976.9999

- Bradshaw, A. D. (2016). The Extracellular Matrix. In *Encyclopedia of Cell Biology* (pp. 694–703). Elsevier. doi: 10.1016/B978-0-12-394447-4.20067-9
- Brembeck, F. H., Rosário, M., & Birchmeier, W. (2006). Balancing cell adhesion and Wnt signaling, the key role of  $\beta$ -catenin. *Current Opinion in Genetics & Development*, 16(1), 51–59. doi: 10.1016/j.gde.2005.12.007
- Broers, J. L. V., Carney, D. N., Rot, M. K., Schaart, G., Lane, E. B., Vooijs, G. P., & Ramaekers, F. C. S. (1986). Intermediate filament proteins in classic and variant types of small cell lung carcinoma cell lines: A biochemical and immunochemical analysis using a panel of monoclonal and polyclonal antibodies. *Journal of Cell Science*, 83(1), 37–60. doi: 10.1242/jcs.83.1.37
- Brown, T. A. (2002). Transcriptomes and Proteomes. In *Genomes* (2nd ed.). Oxford: Wiley-Liss.
- Buchet, J. P. & Lauwerys, R. (1985). Study of inorganic arsenic methylation by rat liver in vitro: relevance for the interpretation of observations in man. *Archives of Toxicology*, 57(2), 125–129. doi: 10.1007/BF00343122
- Bucki, R., Iwamoto, D. V., Shi, X., Kerr, K. E., Byfield, F. J., Suprewicz, Ł., Skłodowski, K., Sutaria, J., Misiak, P., Wilczewska, A. Z., Ramachandran, S., Wolfe, A., Thanh, M.-T. H., Whalen, E., Patteson, A. E., & Janmey, P. A. (2023). Extracellular vimentin is sufficient to promote cell attachment, spreading, and motility by a mechanism involving N-acetyl glucosamine-containing structures. *Journal of Biological Chemistry*, 299(8), 104963. doi: 10.1016/j.jbc.2023.104963
- Burger, T. (2023). Controlling for false discoveries subsequently to large scale one-way ANOVA testing in proteomics: Practical considerations. *PROTEOMICS*, 23(18). doi: 10.1002/pmic.202200406
- Canty, J. T., Tan, R., Kusakci, E., Fernandes, J., & Yildiz, A. (2021). Structure and Mechanics of Dynein Motors. *Annual Review of Biophysics*, 50(1), 549–574. doi: 10.1146/annurev-biophys-111020-101511
- Canuel, M., Korkidakis, A., Konnyu, K., & Morales, C. R. (2008). Sortilin mediates the lysosomal targeting of cathepsins D and H. *Biochemical and Biophysical Research Communications*, 373(2), 292–297. doi: 10.1016/j.bbrc.2008.06.021
- Challenger, F. (1947). Biological methylation. *Science Progress*, 35(139), 396–416.
- Chang, L. & Goldman, R. D. (2004). Intermediate filaments mediate cytoskeletal crosstalk. *Nature Reviews Molecular Cell Biology*, 5(8), 601–613. doi: 10.1038/nrm1438
- Chen, B., Arnold, L. L., Cohen, S. M., Thomas, D. J., & Le, X. C. (2011). Mouse Arsenic (+3 Oxidation State) Methyltransferase Genotype Affects Metabolism and Tissue Dosimetry of Arsenicals after Arsenite Administration in Drinking Water. *Toxicological Sciences*, 124(2), 320–326. doi: 10.1093/toxsci/kfr246
- Chen, J., Li, J., Jiang, X., & Rosen, B. P. (2017). Conserved cysteine residues determine substrate specificity in a novel As(III) S-adenosylmethionine methyltransferase from *Aspergillus fumigatus*. *Molecular Microbiology*, 104(2), 250–259. doi: 10.1111/mmi.13628

- Chen, S.-C., Sun, G.-X., Rosen, B. P., Zhang, S.-Y., Deng, Y., Zhu, B.-K., Rensing, C., & Zhu, Y.-G. (2017). Recurrent horizontal transfer of arsenite methyltransferase genes facilitated adaptation of life to arsenic. *Scientific Reports*, 7(1), 7741. doi: 10.1038/s41598-017-08313-2
- Chen, Y., Wu, F., Graziano, J. H., Parvez, F., Liu, M., Paul, R. R., Shaheen, I., Sarwar, G., Ahmed, A., Islam, T., Slavkovich, V., Rundek, T., Demmer, R. T., Desvarieux, M., & Ahsan, H. (2013). Arsenic Exposure From Drinking Water, Arsenic Methylation Capacity, and Carotid Intima-Media Thickness in Bangladesh. *American Journal of Epidemiology*, 178(3), 372–381. doi: 10.1093/aje/kwt001
- Correia, I., Chu, D., Chou, Y.-H., Goldman, R. D., & Matsudaira, P. (1999). Integrating the Actin and Vimentin Cytoskeletons. *The Journal of Cell Biology*, 146(4), 831–842. doi: 10.1083/jcb.146.4.831
- Čuklina, J., Lee, C. H., Williams, E. G., Sajic, T., Collins, B. C., Rodríguez Martínez, M., Sharma, V. S., Wendt, F., Goetze, S., Keele, G. R., Wollscheid, B., Aebersold, R., & Pedrioli, P. G. A. (2021). Diagnostics and correction of batch effects in large-scale proteomic studies: a tutorial. *Molecular Systems Biology*, 17(8). doi: 10.15252/msb.202110240
- Currier, J. M., Svoboda, M., de Moraes, D. P., Matoušek, T., Dědina, J., & Stýblo, M. (2011). Direct Analysis of Methylated Trivalent Arsenicals in Mouse Liver by Hydride Generation-Cryotrapping-Atomic Absorption Spectrometry. *Chemical Research in Toxicology*, 24(4), 478–480. doi: 10.1021/tx200060c
- Das, N., Giri, A., Chakraborty, S., & Bhattacharjee, P. (2016). Association of single nucleotide polymorphism with arsenic-induced skin lesions and genetic damage in exposed population of West Bengal, India. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, 809, 50–56. doi: 10.1016/j.mrgentox.2016.09.006
- de Rosemond, S., Xie, Q., & Liber, K. (2008). Arsenic concentration and speciation in five freshwater fish species from Back Bay near Yellowknife, NT, CANADA. *Environmental Monitoring and Assessment*, 147(1–3), 199–210. doi: 10.1007/s10661-007-0112-6
- DeMichele-Sweet, M. A. A., Weamer, E. A., Klei, L., Vrana, D. T., Hollingshead, D. J., Seltman, H. J., Sims, R., Foroud, T., Hernandez, I., Moreno-Grau, S., Tárraga, L., Boada, M., Ruiz, A., Williams, J., Mayeux, R., Lopez, O. L., Sibille, E. L., Kamboh, M. I., Devlin, B., & Sweet, R. A. (2018). Genetic risk for schizophrenia and psychosis in Alzheimer disease. *Molecular Psychiatry*, 23(4), 963–972. doi: 10.1038/mp.2017.81
- Dheeman, D. S., Packianathan, C., Pillai, J. K., & Rosen, B. P. (2014). Pathway of Human AS3MT Arsenic Methylation. *Chemical Research in Toxicology*, 27(11), 1979–1989. doi: 10.1021/tx500313k
- Dodmane, P. R., Arnold, L. L., Pennington, K. L., Thomas, D. J., & Cohen, S. M. (2013). Effect of dietary treatment with dimethylarsinous acid (DMAIII) on the urinary bladder epithelium of arsenic (+3 oxidation state) methyltransferase (As3mt) knockout and C57BL/6 wild type female mice. *Toxicology*, 305, 130–135. doi: 10.1016/j.tox.2013.01.015

- Dong, Y., Kang, H., Liu, H., Wang, J., Guo, Q., Song, C., Sun, Y., Zhang, Y., Zhang, H., Zhang, Z., Guan, H., Fang, Z., & Li, F. (2019). Myoferlin, a Membrane Protein with Emerging Oncogenic Roles. *BioMed Research International*, 2019, 1–9. doi: 10.1155/2019/7365913
- Douillet, C., Huang, M. C., Saunders, R. J., Dover, E. N., Zhang, C., & Stýblo, M. (2017). Knockout of arsenic (+3 oxidation state) methyltransferase is associated with adverse metabolic phenotype in mice: the role of sex and arsenic exposure. *Archives of Toxicology*, 91(7), 2617–2627. doi: 10.1007/s00204-016-1890-9
- Drobná, Z., Naranmandura, H., Kubachka, K. M., Edwards, B. C., Herbin-Davis, K., Stýblo, M., Le, X. C., Creed, J. T., Maeda, N., Hughes, M. F., & Thomas, D. J. (2009). Disruption of the Arsenic (+3 Oxidation State) Methyltransferase Gene in the Mouse Alters the Phenotype for Methylation of Arsenic and Affects Distribution and Retention of Orally Administered Arsenate. *Chemical Research in Toxicology*, 22(10), 1713–1720. doi: 10.1021/tx900179r
- Duarte, R. R. R., Troakes, C., Nolan, M., Srivastava, D. P., Murray, R. M., & Bray, N. J. (2016). Genome-wide significant schizophrenia risk variation on chromosome 10q24 is associated with altered *cis* -regulation of *BORCS7*, *AS3MT*, and *NT5C2* in the human brain. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 171(6), 806–814. doi: 10.1002/ajmg.b.32445
- Eckes, B., Dobic, D., Colucci-Guyon, E., Wang, N., Maniotis, A., Ingber, D., Merckling, A., Langa, F., Aumailley, M., Delouée, A., Koteliansky, V., Babinet, C., & Krieg, T. (1998). Impaired mechanical stability, migration and contractile capacity in vimentin deficient fibroblasts. *Journal of Cell Science*, 111(13), 1897–1907. doi: 10.1242/jcs.111.13.1897
- Eckes, B., Zigrino, P., Kessler, D., Holtkötter, O., Shephard, P., Mauch, C., & Krieg, T. (2000). Fibroblast-matrix interactions in wound healing and fibrosis. *Matrix Biology*, 19(4), 325–332. doi: 10.1016/S0945-053X(00)00077-9
- Fan, X., Chen, Y., Lu, J., Li, W., Li, X., Guo, H., Chen, Q., Yang, Y., & Xia, H. (2021). AS3MT Polymorphism: A Risk Factor for Epilepsy Susceptibility and Adverse Drug Reactions to Valproic Acid and Oxcarbazepine Treatment in Children From South China. *Frontiers in Neuroscience*, 15. doi: 10.3389/fnins.2021.705297
- Gan, Z., Ding, L., Burckhardt, C. J., Lowery, J., Zaritsky, A., Sitterley, K., Mota, A., Costigliola, N., Starker, C. G., Voytas, D. F., Tytell, J., Goldman, R. D., & Danuser, G. (2016). Vimentin Intermediate Filaments Template Microtubule Networks to Enhance Persistence in Cell Polarity and Directed Migration. *Cell Systems*, 3(3), 252-263.e8. doi: 10.1016/j.cels.2016.08.007
- Garg, A., Barnes, P. F., Porgador, A., Roy, S., Wu, S., Nanda, J. S., Griffith, D. E., Girard, W. M., Rawal, N., Shetty, S., & Vankayalapati, R. (2006). Vimentin Expressed on *Mycobacterium tuberculosis* -Infected Human Monocytes Is Involved in Binding to the NKp46 Receptor. *The Journal of Immunology*, 177(9), 6192–6198. doi: 10.4049/jimmunol.177.9.6192
- Gates, J. & Peifer, M. (2005). Can 1000 Reviews Be Wrong? Actin,  $\alpha$ -Catenin, and Adherens Junctions. *Cell*, 123(5), 769–772. doi: 10.1016/j.cell.2005.11.009

- Ghosh, J. & Sil, P. C. (2015). Mechanism for Arsenic-Induced Toxic Effects. In Handbook of Arsenic Toxicology (pp. 203–231). Elsevier. doi: 10.1016/B978-0-12-418688-0.00008-3
- Gong, G. & O’Byrant, S. E. (2012). Low-level arsenic exposure, AS3MT gene polymorphism and cardiovascular diseases in rural Texas counties. *Environmental Research*, 113, 52–57. doi: 10.1016/j.envres.2012.01.003
- Guilbert, C., Chou, H., Bolt, A. M., Wu, T. H., Luo, V. M., Orthwein, A., & Mann, K. K. (2020). A Functional Assay to Assess Toxicity During Murine B Cell Development In Vitro. *Current Protocols in Toxicology*, 83(1). doi: 10.1002/cptx.91
- Gusev, A., Mancuso, N., Won, H., Kousi, M., Finucane, H. K., Reshef, Y., Song, L., Safi, A.; Schizophrenia Working Group of the Psychiatric Genomics Consortium; McCarroll, S., Neale, B. M., Ophoff, R. A., O’Donovan, M. C., Crawford, G. E., Geschwind, D. H., Katsanis, N., Sullivan, P. F., Pasaniuc, B., & Price, A. L. (2018). Transcriptome-wide association study of schizophrenia and chromatin activity yields mechanistic disease insights. *Nature Genetics*, 50(4), 538–548. doi: 10.1038/s41588-018-0092-1
- Hannon, E., Dempster, E., Burrage, J., McQuillin, A., Clair, D. St., Morris, D., Di Forti, M., Gaughran, F., MacCabe, J., Breen, G., Collier, D., Murray, R., Schalkwyk, L., & Mill, J. (2019). An Integrated Genetic-Epigenetic Analysis of Schizophrenia: Evidence for Co-localization of Genetic Associations and Differential DNA Methylation from a Large Meta-Analysis of Whole Blood DNA. *European Neuropsychopharmacology*, 29, S782–S783. doi: 10.1016/j.euroneuro.2017.08.006
- Hartsock, A. & Nelson, W. J. (2008). Adherens and tight junctions: Structure, function and connections to the actin cytoskeleton. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1778(3), 660–669. doi: 10.1016/j.bbamem.2007.07.012
- Hayakawa, T., Kobayashi, Y., Cui, X., & Hirano, S. (2005). A new metabolic pathway of arsenite: arsenic-glutathione complexes are substrates for human arsenic methyltransferase Cyt19. *Archives of Toxicology*, 79(4), 183–191. doi: 10.1007/s00204-004-0620-x
- He, F., Ru, X., & Wen, T. (2020). NRF2, a Transcription Factor for Stress Response and Beyond. *International Journal of Molecular Sciences*, 21(13), 4777. doi: 10.3390/ijms21134777
- Huang, M. C., Douillet, C. C., & Stýblo, M. (2016). Knockout of arsenic (+3 oxidation state) methyltransferase results in sex-dependent changes in phosphatidylcholine metabolism in mice. *Archives of Toxicology*, 90(12), 3125–3128. doi: 10.1007/s00204-016-1844-2
- Huang, R., Grishagin, I., Wang, Y., Zhao, T., Greene, J., Obenauer, J. C., Ngan, D., Nguyen, D.-T., Guha, R., Jadhav, A., Southall, N., Simeonov, A., & Austin, C. P. (2019). The NCATS BioPlanet – An Integrated Platform for Exploring the Universe of Cellular Signaling Pathways for Toxicology, Systems Biology, and Chemical Genomics. *Frontiers in Pharmacology*, 10. doi: 10.3389/fphar.2019.00445
- Hussain, B., Fang, C., Huang, X., Feng, Z., Yao, Y., Wang, Y., & Chang, J. (2022). Endothelial  $\beta$ -Catenin Deficiency Causes Blood-Brain Barrier Breakdown via Enhancing the

- Paracellular and Transcellular Permeability. *Frontiers in Molecular Neuroscience*, 15. doi: 10.3389/fnmol.2022.895429
- Ivaska, J., Pallari, H.-M., Nevo, J., & Eriksson, J. E. (2007). Novel functions of vimentin in cell adhesion, migration, and signaling. *Experimental Cell Research*, 313(10), 2050–2062. doi: 10.1016/j.yexcr.2007.03.040
- Jeffery, C. J. (2003). Moonlighting proteins: old proteins learning new tricks. *Trends in Genetics*, 19(8), 415–417. doi: 10.1016/S0168-9525(03)00167-7
- Johannes, L. & Popoff, V. (2008). Tracing the Retrograde Route in Protein Trafficking. *Cell*, 135(7), 1175–1187. doi: 10.1016/j.cell.2008.12.009
- Kabiraj, A., Laha, A., Panja, A. S., & Bandopadhyay, R. (2023). In silico comparative structural and functional analysis of arsenite methyltransferase from bacteria, fungi, fishes, birds, and mammals. *Journal of Genetic Engineering and Biotechnology*, 21(1), 64. doi: 10.1186/s43141-023-00522-9
- Korte, N. E. & Fernando, Q. (1991). A review of arsenic (III) in groundwater. *Critical Reviews in Environmental Control*, 21(1), 1–39. doi: 10.1080/10643389109388408
- Lallemant-Breitenbach, V., Jeanne, M., Benhenda, S., Nasr, R., Lei, M., Peres, L., Zhou, J., Zhu, J., Raught, B., & de Thé, H. (2008). Arsenic degrades PML or PML–RAR $\alpha$  through a SUMO-triggered RNF4/ubiquitin-mediated pathway. *Nature Cell Biology*, 10(5), 547–555. doi: 10.1038/ncb1717
- Leroy, J. G. & DeMars, R. I. (1967). Mutant Enzymatic and Cytological Phenotypes in Cultured Human Fibroblasts. *Science*, 157(3790), 804–806. doi: 10.1126/science.157.3790.804
- Li, L., Chang, H., Peng, T., Li, M., & Xiao, X. (2016). Evidence of AS3MTd2d3-Associated Variants within 10q24.32-33 in the Genetic Risk of Major Affective Disorders. *Complex Psychiatry*, 2(4), 213–218. doi: 10.1159/000452998
- Lin, S., Shi, Q., Nix, F. B., Stýblo, M., Beck, M. A., Herbin-Davis, K. M., Hall, L. L., Simeonsson, J. B., & Thomas, D. J. (2002). A Novel S-Adenosyl-L-methionine:Artenic(III) Methyltransferase from Rat Liver Cytosol. *Journal of Biological Chemistry*, 277(13), 10795–10803. doi: 10.1074/jbc.M110246200
- Liu, M., Khasiyev, F., Sariya, S., Spagnolo-Allende, A., Sanchez, D. L., Andrews, H., Yang, Q., Beiser, A., Qiao, Y., Thomas, E. A., Romero, J. R., Rundek, T., Brickman, A. M., Manly, J. J., Elkind, M. S., Seshadri, S., Chen, C., Hilal, S., Wasserman, B. A., Toto, G., Fornage, M., & Gutierrez, J. (2023). Chromosome 10q24.32 Variants Associate With Brain Arterial Diameters in Diverse Populations: A Genome-Wide Association Study. *Journal of the American Heart Association*, 12(23). doi: 10.1161/JAHA.123.030935
- López-Menéndez, H. (2022). Mechanics of interactions of F-actin and vimentin networks. In *Mechanics of Fibrous Networks* (pp. 301–316). Elsevier. doi: 10.1016/B978-0-12-822207-2.00010-6
- Lotan, A., Fenckova, M., Bralten, J., Alttoa, A., Dixson, L., Williams, R. W., & van der Voet, M. (2014). Neuroinformatic analyses of common and distinct genetic components associated

- with major neuropsychiatric disorders. *Frontiers in Neuroscience*, 8. doi: 10.3389/fnins.2014.00331
- Mantha, M., Yeary, E., Trent, J., Creed, P. A., Kubachka, K., Hanley, T., Shockey, N., Heitkemper, D., Caruso, J., Xue, J., Rice, G., Wymer, L., & Creed, J. T. (2017). Estimating Inorganic Arsenic Exposure from U.S. Rice and Total Water Intakes. *Environmental Health Perspectives*, 125(5). doi: 10.1289/EHP418
- Marapakala, K., Qin, J., & Rosen, B. P. (2012). Identification of Catalytic Residues in the As(III) S-Adenosylmethionine Methyltransferase. *Biochemistry*, 51(5), 944–951. doi: 10.1021/bi201500c
- Martin, J. (2002). SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold. *Current Opinion in Structural Biology*, 12(6), 783–793. doi: 10.1016/S0959-440X(02)00391-3
- McCrea, P. D. & Gu, D. (2010). The catenin family at a glance. *Journal of Cell Science*, 123(5), 637–642. doi: 10.1242/jcs.039842
- Mersaoui, S. Y., Guilbert, C., Chou, H., Douillet, C., Bohle, D. S., Stýblo, M., Richard, S., & Mann, K. K. (2022). Arsenic 3 methyltransferase (AS3MT) automethylates on cysteine residues in vitro. *Archives of Toxicology*, 96(5), 1371–1386. doi: 10.1007/s00204-022-03248-8
- Mor-Vaknin, N., Punturieri, A., Sitwala, K., & Markovitz, D. M. (2003). Vimentin is secreted by activated macrophages. *Nature Cell Biology*, 5(1), 59–63. doi: 10.1038/ncb898
- Mulholland, D. J., Kobayashi, N., Ruscetti, M., Zhi, A., Tran, L. M., Huang, J., Gleave, M., & Wu, H. (2012). *Pten* Loss and RAS/MAPK Activation Cooperate to Promote EMT and Metastasis Initiated from Prostate Cancer Stem/Progenitor Cells. *Cancer Research*, 72(7), 1878–1889. doi: 10.1158/0008-5472.CAN-11-3132
- National Research Council (US) Subcommittee on Arsenic in Drinking Water. (1999). *Arsenic in Drinking Water*. Washington, D.C.: National Academies Press. doi: 10.17226/6444
- Negro Silva, L. F., Lemaire, M., Lemarié, C. A., Plourde, D., Bolt, A. M., Chiavatti, C., Bohle, D. S., Slavkovich, V., Graziano, J. H., Lehoux, S., & Mann, K. K. (2017). Effects of Inorganic Arsenic, Methylated Arsenicals, and Arsenobetaine on Atherosclerosis in the apoE<sup>-/-</sup> Mouse Model and the Role of As3mt-Mediated Methylation. *Environmental Health Perspectives*, 125(7). doi: 10.1289/EHP806
- Ning, M. & Lo, E. H. (2010). Opportunities and Challenges in Omics. *Translational Stroke Research*, 1(4), 233–237. doi: 10.1007/s12975-010-0048-y
- Nong, Q., Zhang, Y., Guallar, E., & Zhong, Q. (2016). Arsenic Exposure and Predicted 10-Year Atherosclerotic Cardiovascular Risk Using the Pooled Cohort Equations in U.S. Hypertensive Adults. *International Journal of Environmental Research and Public Health*, 13(11), 1093. doi: 10.3390/ijerph13111093
- Norwood, S. J., Shaw, D. J., Cowieson, N. P., Owen, D. J., Teasdale, R. D., & Collins, B. M. (2011). Assembly and Solution Structure of the Core Retromer Protein Complex. *Traffic*, 12(1), 56–71. doi: 10.1111/j.1600-0854.2010.01124.x

- Okazaki, S., Kato, R., Wakatsuki, S., Uchida, Y., Taguchi, T., & Arai, H. (2012). Structural Insights into the Phospholipid Binding Specificity of Human Evectin-2. *Nihon Kessho Gakkaishi*, *54*(2), 101–106. doi: 10.5940/jcrsj.54.101
- Pardiñas, A. F., Holmans, P., Pocklington, A. J., Escott-Price, V., Ripke, S., Carrera, N., Legge, S. E., Bishop, S., Cameron, D., Hamshere, M. L., Han, J., Hubbard, L., Lynham, A., Mantripragada, K., Rees, E., MacCabe, J. H., McCarroll, S. A., Baune, B. T., Breen, G., Byrne, E. M., Dannlowski, U., Eley, T. C., Hayward, C., Martin, N. G., McIntosh, A. M., Plomin, R., Porteous, D. J., Wray, N. R., Caballero, A., Geschwind, D. H., Huckins, L. M., Ruderfer, D. M., Santiago, E., Sklar, P., Stahl, E. A., Won, H., Agerbo, E., Als, T. D., Andreassen, O. A., Bækvad-Hansen, M., Mortensen, P. B., Pedersen, C. B., Børglum, A. D., Bybjerg-Grauholm, J., Durovic, S., Durmishi, N., Pedersen, M. G., Golimbet, V., Grove, J., Hougaard, D. M., Mattheisen, M., Molden, E., Mors, O., Nordentoft, M., Pejović-Milovančević, M., Sigurdsson, E., Silagadze, T., Hansen, C. S., Stefansson, K., Stefansson, H., Steinberg, S., Tosato, S., Werge, T.; GERAD1 Consortium; CRESTAR Consortium; Collier, D. A., Rujescu, D., Kirov, G., Owen, M. J., O'Donovan, M. C., & Walters, J. T. R. (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature Genetics*, *50*(3), 381–389. doi: 10.1038/s41588-018-0059-2
- Park, S., Park, J.-E., Yoo, H. J., Kim, J.-W., Cho, S.-C., Shin, M.-S., Cheong, J. H., Han, D. H., & Kim, B.-N. (2015). Family-based association study of the arsenite methyltransferase gene (AS3MT, rs11191454) in Korean children with attention-deficit hyperactivity disorder. *Psychiatric Genetics*, *25*(1), 26–30. doi: 10.1097/YPG.0000000000000064
- Patten, J. & Wang, K. (2021). Fibronectin in development and wound healing. *Advanced Drug Delivery Reviews*, *170*, 353–368. doi: 10.1016/j.addr.2020.09.005
- Patteson, A. E., Vahabikashi, A., Goldman, R. D., & Janmey, P. A. (2020). Mechanical and Non-Mechanical Functions of Filamentous and Non-Filamentous Vimentin. *BioEssays*, *42*(11). doi: 10.1002/bies.202000078
- Pearson, E., Hanz, S., Ben-Yaakov, K., Segal-Ruder, Y., Seger, R., & Fainzilber, M. (2005). Vimentin-Dependent Spatial Translocation of an Activated MAP Kinase in Injured Nerve. *Neuron*, *45*(5), 715–726. doi: 10.1016/j.neuron.2005.01.023
- Petrossian, T. C. & Clarke, S. G. (2009). Multiple Motif Scanning to Identify Methyltransferases from the Yeast Proteome. *Molecular & Cellular Proteomics*, *8*(7), 1516–1526. doi: 10.1074/mcp.M900025-MCP200
- Pierce, B. L., Kibriya, M. G., Tong, L., Jasmine, F., Argos, M., Roy, S., Paul-Brutus, R., Rahaman, R., Rakibuz-Zaman, M., Parvez, F., Ahmed, A., Quasem, I., Hore, S. K., Alam, S., Islam, T., Slavkovich, V., Gamble, M. V., Yunus, M., Rahman, M., Baron, J. A., Graziano, J. H., & Ahsan, H. (2012). Genome-Wide Association Study Identifies Chromosome 10q24.32 Variants Associated with Arsenic Metabolism and Toxicity Phenotypes in Bangladesh. *PLoS Genetics*, *8*(2), e1002522. doi: 10.1371/journal.pgen.1002522

- Pishva, E., Creese, B., Smith, A. R., Viechtbauer, W., Proitsi, P., van den Hove, D. L. A., Ballard, C., Mill, J., & Lunnon, K. (2020). Psychosis-associated DNA methylomic variation in Alzheimer's disease cortex. *Neurobiology of Aging*, *89*, 83–88. doi: 10.1016/j.neurobiolaging.2020.01.001
- Popoff, V., Mardones, G. A., Bai, S., Chambon, V., Tenza, D., Burgos, P. V., Shi, A., Benaroch, P., Urbé, S., Lamaze, C., Grant, B. D., Raposo, G., & Johannes, L. (2009). Analysis of Articulation Between Clathrin and Retromer in Retrograde Sorting on Early Endosomes. *Traffic*, *10*(12), 1868–1880. doi: 10.1111/j.1600-0854.2009.00993.x
- Přechová, M., Korelova, K., & Gregor, M. (2023). Plectin. *Current Biology*, *33*(4), R128–R130. doi: 10.1016/j.cub.2022.12.061
- Qin, J., Rosen, B. P., Zhang, Y., Wang, G., Franke, S., & Rensing, C. (2006). Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite *S*-adenosylmethionine methyltransferase. *Proceedings of the National Academy of Sciences*, *103*(7), 2075–2080. doi: 10.1073/pnas.0506836103
- Qiu, T., Wu, C., Yao, X., Han, Q., Wang, N., Yuan, W., Zhang, J., Shi, Y., Jiang, L., Liu, X., Yang, G., & Sun, X. (2023). AS3MT facilitates NLRP3 inflammasome activation by m6A modification during arsenic-induced hepatic insulin resistance. *Cell Biology and Toxicology*, *39*(5), 2165–2181. doi: 10.1007/s10565-022-09703-7
- Ratnaik, R. N. (2003). Acute and chronic arsenic toxicity. *Postgraduate Medical Journal*, *79*(933), 391–396. doi: 10.1136/pmj.79.933.391
- Reitman, M. L., Varki, A., & Kornfeld, S. (1981). Fibroblasts from patients with I-cell disease and pseudo-Hurler polydystrophy are deficient in uridine 5'-diphosphate-N-acetylglucosamine: glycoprotein N-acetylglucosaminylphosphotransferase activity. *Journal of Clinical Investigation*, *67*(5), 1574–1579. doi: 10.1172/JCI110189
- Ridge, K. M., Eriksson, J. E., Pekny, M., & Goldman, R. D. (2022). Roles of vimentin in health and disease. *Genes & Development*, *36*(7–8), 391–407. doi: 10.1101/gad.349358.122
- Ripke, S., O'Dushlaine, C., Chambert, K., Moran, J. L., Kähler, A. K., Akterin, S., Bergen S. E., Collins, A. L., Crowley, J. J., Fromer, M., Kim, Y., Lee, S. H., Magnusson, P. K., Sanchez, N., Stahl, E. A., Williams, S., Wray, N. R., Xia, K., Bettella, F., Borglum, A. D., Bulik-Sullivan, B. K., Cormican, P., Craddock, N., de Leeuw, C., Durmishi, N., Gill, M., Golimbet, V., Hamshere, M. L., Holmans, P., Hougaard, D. M., Kendler, K. S., Lin, K., Morris, D. W., Mors, O., Mortensen, P. B., Neale, B. M., O'Neill, F. A., Owen, M. J., Pejović-Milovančević, M., Posthuma, D., Powell, J., Richards, A. L., Riley, B. P., Ruderfer, D., Rujescu, D., Sigurdsson, E., Silagadze, T., Smit, A. B., Stefansson, H., Steinberg, S., Suvisaari, J., Tosato, S., Verhage, M., Walters, J. T.; Multicenter Genetic Studies of Schizophrenia Consortium; Levinson, D. F., Gejman, P. V., Kendler, K. S., Laurent, C., Mowry, B. J., O'Donovan, M. C., Owen, M. J., Pulver, A. E., Riley, B. P., Schwab, S. G., Wildenauer, D. B., Dudbridge, F., Holmans, P., Shi, J., Albus, M., Alexander, M., Campion, D., Cohen, D., Dikeos, D., Duan, J., Eichhammer, P., Godard, S., Hansen, M., Lerer, F. B., Liang, K. Y., Maier, W., Mallet, J., Nertney, D. A., Nestadt, G., Norton, N.,

- O'Neill, F. A., Papadimitriou, G. N., Ribble, R., Sanders, A. R., Silverman, J. M., Walsh, D., Williams, N. M., Wormley, B.; Psychosis Endophenotypes International Consortium; Arranz, M. J., Bakker, S., Bender, S., Bramon, E., Collier, D., Crespo-Facorro, B., Hall, J., Iyegbe, C., Jablensky, A., Kahn, R. S., Kalaydjieva, L., Lawrie, S., Lewis, C. M., Lin, K., Linszen, D. H., Mata, I., McIntosh, A., Murray, R. M., Ophoff, R. A., Powell, J., Rujescu, D., Van Os, J., Walshe, M., Weisbrod, M., Wiersma, D.; Wellcome Trust Case Control Consortium 2; Donnelly, P., Barroso, I., Blackwell, J. M., Bramon, E., Brown, M. A., Casas, J. P., Corvin, A. P., Deloukas, P., Duncanson, A., Jankowski, J., Markus, H. S., Mathew, C. G., Palmer, C. N., Plomin, R., Rautanen, A., Sawcer, S. J., Trembath, R. C., Viswanathan, A. C., Wood, N. W., Spencer, C. C., Band, G., Bellenguez, C., Freeman, C., Hellenthal, G., Giannoulatou, E., Pirinen, M., Pearson, R. D., Strange, A., Su, Z., Vukcevic, D., Donnelly, P., Langford, C., Hunt, S. E., Edkins, S., Gwilliam, R., Blackburn, H., Bumpstead, S. J., Dronov, S., Gillman, M., Gray, E., Hammond, N., Jayakumar, A., McCann, O. T., Liddle, J., Potter, S. C., Ravindrarajah, R., Ricketts, M., Tashakkori-Ghanbaria, A., Waller, M. J., Weston, P., Widaa, S., Whittaker, P., Barroso, I., Deloukas, P., Mathew, C. G., Blackwell, J. M., Brown, M. A., Corvin, A. P., McCarthy, M. I., Spencer, C. C., Bramon, E., Corvin, A. P., O'Donovan, M. C., Stefansson, K., Scolnick, E., Purcell, S., McCarroll, S. A., Sklar, P., Hultman, C. M., & Sullivan, P. F. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. (2013). *Nat Genet.*, 45(10), 1150–1159. doi: 10.1038/ng.2742
- Saftig, P. & Klumperman, J. (2009). Lysosome biogenesis and lysosomal membrane proteins: trafficking meets function. *Nature Reviews Molecular Cell Biology*, 10(9), 623–635. doi: 10.1038/nrm2745
- Saraogi, I. & Shan, S. (2011). Molecular Mechanism of Co-translational Protein Targeting by the Signal Recognition Particle. *Traffic*, 12(5), 535–542. doi: 10.1111/j.1600-0854.2011.01171.x
- Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. (2011). Genome-wide association study identifies five new schizophrenia loci. *Nat Genet.*, 43(10), 969–976. doi: 10.1038/ng.940
- Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. (2014). *Nature*, 511(7510), 421–427. doi: 10.1038/nature13595
- Schooling, C. M., Huang, J. V., Zhao, J. V., Kwok, M. K., Au Yeung, S. L., & Lin, S. L. (2018). Disconnect Between Genes Associated With Ischemic Heart Disease and Targets of Ischemic Heart Disease Treatments. *EBioMedicine*, 28, 311–315. doi: 10.1016/j.ebiom.2018.01.015
- Seaman, M. N. J. (2004). Cargo-selective endosomal sorting for retrieval to the Golgi requires retromer. *The Journal of Cell Biology*, 165(1), 111–122. doi: 10.1083/jcb.200312034
- Seaman, M. N. J. (2012). The retromer complex – endosomal protein recycling and beyond. *Journal of Cell Science*. doi: 10.1242/jcs.103440

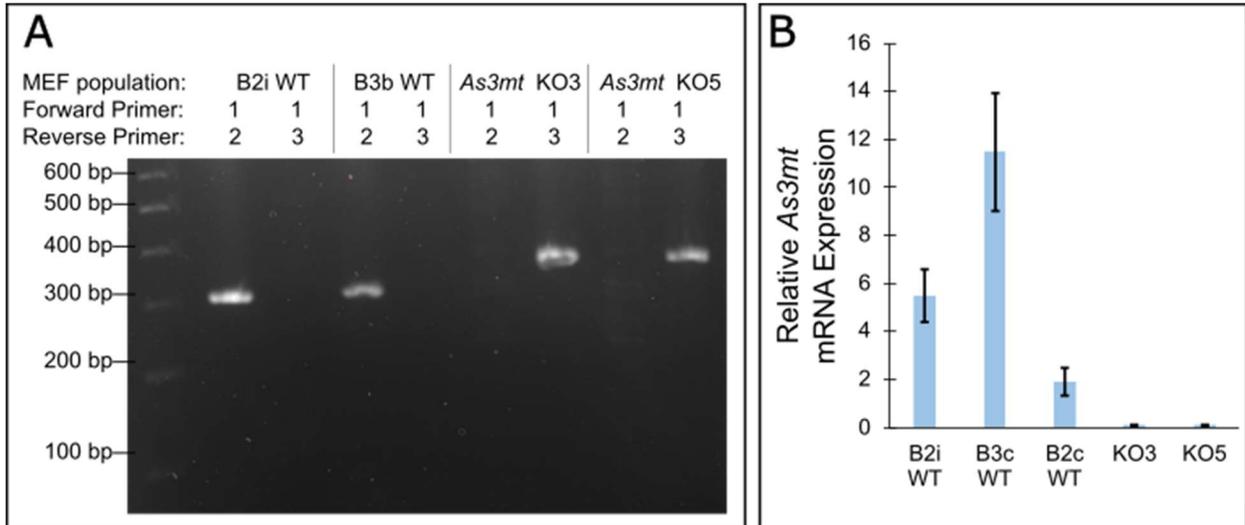
- Serres, M. P., Samwer, M., Truong Quang, B. A., Lavoie, G., Perera, U., Görlich, D., Charras, G., Petronczki, M., Roux, P. P., & Paluch, E. K. (2020). F-Actin Interactome Reveals Vimentin as a Key Regulator of Actin Organization and Cell Mechanics in Mitosis. *Developmental Cell*, 52(2), 210–222.e7. doi: 10.1016/j.devcel.2019.12.011
- Simeonova, P. P. & Luster, M. I. (2004). Arsenic and atherosclerosis. *Toxicology and Applied Pharmacology*, 198(3), 444–449. doi: 10.1016/j.taap.2003.10.018
- Simonetti, B. & Cullen, P. J. (2018). Endosomal Sorting: Architecture of the Retromer Coat. *Current Biology*, 28(23), PR1350–R1352. doi: 10.1016/j.cub.2018.10.040
- Smith, A. H., Arroyo, A. P., Mazumder, D. N., Kosnett, M. J., Hernández, A. L., Beeris, M., Smith, M. M., & Moore, L. E. (2000). Arsenic-induced skin lesions among Atacameño people in Northern Chile despite good nutrition and centuries of exposure. *Environmental Health Perspectives*, 108(7), 617–620. doi: 10.1289/ehp.00108617
- Snouwaert, J. N., Church, R. J., Jania, L., Nguyen, M., Wheeler, M. L., Saintsing, A., Mieczkowski, P., Manuel de Villena, F. P., Armao, D., Moy, S. S., Lorenzo, D. N., & Koller, B. H. (2018). A Mutation in the Borcs7 Subunit of the Lysosome Regulatory BORC Complex Results in Motor Deficits and Dystrophic Axonopathy in Mice. *Cell Reports*, 24(5), 1254–1265. doi: 10.1016/j.celrep.2018.06.118
- Song, X., Geng, Z., Zhu, J., Li, C., Hu, X., Bian, N., Zhang, X., & Wang, Z. (2009). Structure–function roles of four cysteine residues in the human arsenic (+3 oxidation state) methyltransferase (hAS3MT) by site-directed mutagenesis. *Chemico-Biological Interactions*, 179(2–3), 321–328. doi: 10.1016/j.cbi.2008.12.018
- Speer, R. M., Zhou, X., Volk, L. B., Liu, K. J., & Hudson, L. G. (2023). *Arsenic and cancer: Evidence and mechanisms* (Vol. 96, pp. 151–202). doi: 10.1016/bs.apha.2022.08.001
- Stýblo, M., Venkatratnam, A., Fry, R. C., & Thomas, D. J. (2021). Origins, fate, and actions of methylated trivalent metabolites of inorganic arsenic: progress and prospects. *Archives of Toxicology*, 95(5), 1547–1572. doi: 10.1007/s00204-021-03028-w
- Stýblo, M., Delnomdedieu, M., & Thomas, D. J. (1995). Biological Mechanisms and Toxicological Consequences of the Methylation of Arsenic. In R. A. Goyer & M. G. Cherian (Eds.), *Toxicology of Metals. Handbook of Experimental Pharmacology* (Vol. 115, pp. 407–433). Berlin, Heidelberg: Springer. doi: 10.1007/978-3-642-79162-8\_18
- Stýblo, M., Del Razo, L. M., Vega, L., Germolec, D. R., LeCluyse, E. L., Hamilton, G. A., Reed, W., Wang, C., Cullen, W. R., & Thomas, D. J. (2000). Comparative toxicity of trivalent and pentavalent inorganic and methylated arsenicals in rat and human cells. *Archives of Toxicology*, 74(6), 289–299. doi: 10.1007/s002040000134
- Sung, T.-C., Huang, J.-W., & Guo, H.-R. (2015). Association between Arsenic Exposure and Diabetes: A Meta-Analysis. *BioMed Research International*, 2015, 1–10. doi: 10.1155/2015/368087
- Thomas, D. J., Li, J., Waters, S. B., Xing, W., Adair, B. M., Drobná, Z., Devesa, V., & Stýblo, M. (2007). Arsenic (+3 oxidation state) methyltransferase and the methylation of arsenicals. *Experimental Biology and Medicine (Maywood, N.J.)*, 232(1), 3–13.

- Thomas, D. J., Stýblo, M., & Lin, S. (2001). The Cellular Metabolism and Systemic Toxicity of Arsenic. *Toxicology and Applied Pharmacology*, *176*(2), 127–144. doi: 10.1006/taap.2001.9258
- Tolins, M., Ruchirawat, M., & Landrigan, P. (2014). The Developmental Neurotoxicity of Arsenic: Cognitive and Behavioral Consequences of Early Life Exposure. *Annals of Global Health*, *80*(4), 303. doi: 10.1016/j.aogh.2014.09.005
- Uthus, E. O. (1992). Evidence for arsenic essentiality. *Environmental Geochemistry and Health*, *14*(2), 55–58. doi: 10.1007/BF01783629
- Uthus, E. & Poellot, R. (1991). Effect of dietary pyridoxine on arsenic deprivation in rats. *Magnesium and Trace Elements*, *10*(5–6), 339–347.
- Vahter, M. (1999). Methylation of Inorganic Arsenic in Different Mammalian Species and Population Groups. *Science Progress*, *82*(1), 69–88. doi: 10.1177/003685049908200104
- Vahter, M. & Concha, G. (2001). Role of Metabolism in Arsenic Toxicity. *Pharmacology and Toxicology*, *89*(1), 1–5. doi: 10.1034/j.1600-0773.2001.d01-128.x
- Valenta, T., Hausmann, G., & Basler, K. (2012). The many faces and functions of  $\beta$ -catenin. *The EMBO Journal*, *31*(12), 2714–2736. doi: 10.1038/emboj.2012.150
- Vergara-Gerónimo, C. A., León Del Río, A., Rodríguez-Dorantes, M., Ostrosky-Wegman, P., & Salazar, A. M. (2021). Arsenic-protein interactions as a mechanism of arsenic toxicity. *Toxicology and Applied Pharmacology*, *431*, 115738. doi: 10.1016/j.taap.2021.115738
- Viebahn, C., Lane, E. B., & Ramaekers, F. C. S. (1988). Keratin and vimentin expression in early organogenesis of the rabbit embryo. *Cell and Tissue Research*, *253*(3). doi: 10.1007/BF00219746
- Wang, L., Kou, M.-C., Weng, C.-Y., Hu, L.-W., Wang, Y.-J., & Wu, M.-J. (2012). Arsenic modulates heme oxygenase-1, interleukin-6, and vascular endothelial growth factor expression in endothelial cells: roles of ROS, NF- $\kappa$ B, and MAPK pathways. *Archives of Toxicology*, *86*(6), 879–896. doi: 10.1007/s00204-012-0845-z
- Washer, S. J., Flynn, R., Oguro-Ando, A., Hannon, E., Burrage, J., Jeffries, A., Mill, J., & Dempster, E. L. (2022). Functional characterization of the schizophrenia associated gene *AS3MT* identifies a role in neuronal development. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *189*(5), 151–162. doi: 10.1002/ajmg.b.32905
- Waters, S. B., Devesa, V., Del Razo, L. M., Stýblo, M., & Thomas, D. J. (2004). Endogenous Reductants Support the Catalytic Function of Recombinant Rat Cyt19, an Arsenic Methyltransferase. *Chemical Research in Toxicology*, *17*(3), 404–409. doi: 10.1021/tx0342161
- Wei, B., Yu, J., Kong, C., Li, H., Yang, L., Xia, Y., & Wu, K. (2018). Effects of arsenic methylation and metabolism on the changes of arsenic-related skin lesions. *Environmental Science and Pollution Research*, *25*(24), 24394–24402. doi: 10.1007/s11356-018-2512-2
- Winsvold, B. S., Bettella, F., Witoelar, A., Anttila, V., Gormley, P., Kurth, T., Terwindt, G. M., Freilinger, T. M., Frei, O., Shadrin, A., Wang, Y., Dale, A. M., van den Maagdenberg, A. M. J. M., Chasman, D. I., Nyholt, D. R., Palotie, A., Andreassen, O. A., & Zwart, J.-A.

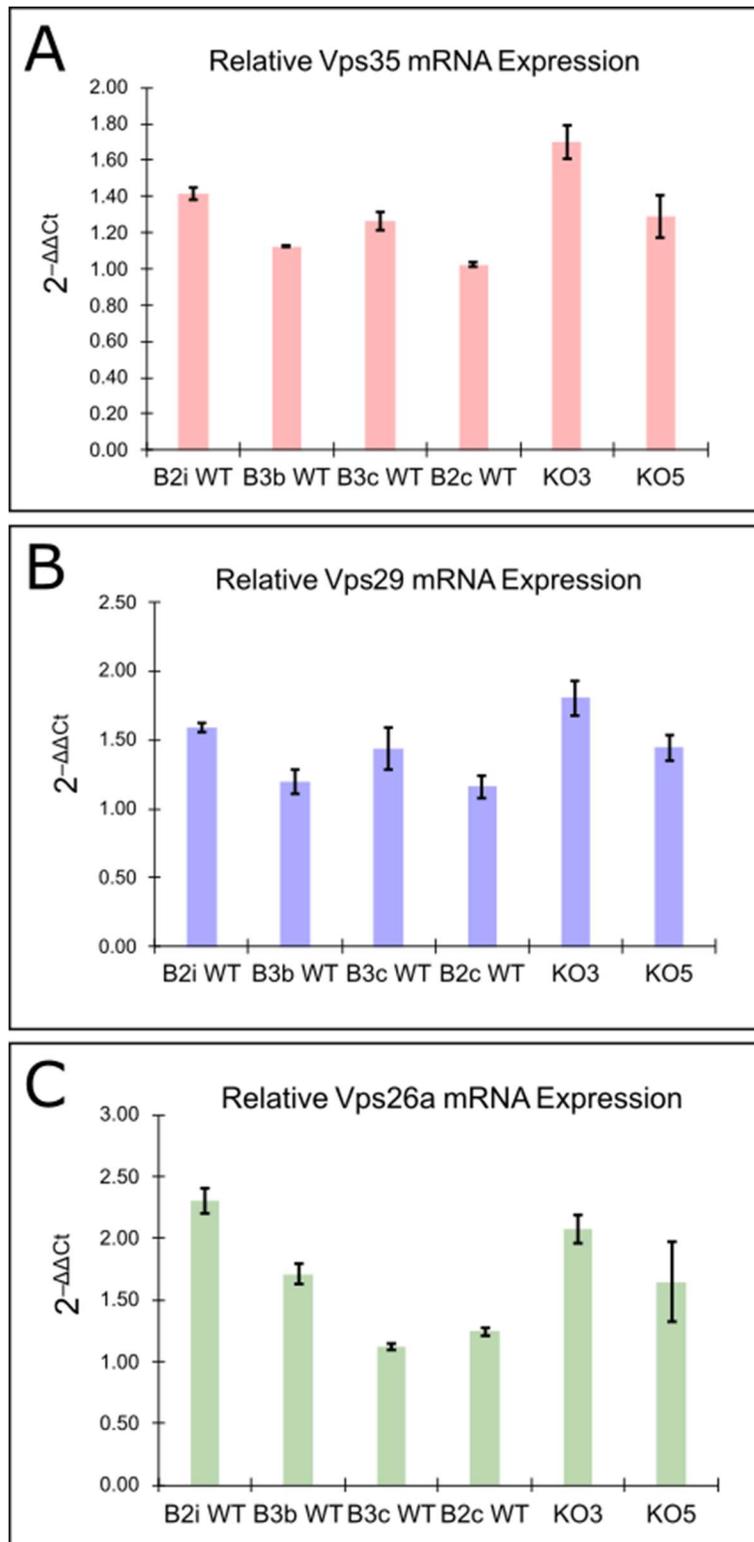
- (2017). Shared genetic risk between migraine and coronary artery disease: A genome-wide analysis of common variants. *PLOS ONE*, *12*(9), e0185663. doi: 10.1371/journal.pone.0185663
- Winter, L., Abrahamsberg, C., & Wiche, G. (2008). Plectin isoform 1b mediates mitochondrion–intermediate filament network linkage and controls organelle shape. *The Journal of Cell Biology*, *181*(6), 903–911. doi: 10.1083/jcb.200710151
- Wood, T. C., Salavagionne, O. E., Mukherjee, B., Wang, L., Klumpp, A. F., Thomae, B. A., Eckloff, B. W., Schaid, D. J., Wieben, E. D., & Weinshilboum, R. M. (2006). Human Arsenic Methyltransferase (AS3MT) Pharmacogenetics. *Journal of Biological Chemistry*, *281*(11), 7364–7373. doi: 10.1074/jbc.M512227200
- Wu, F., Jasmine, F., Kibriya, M. G., Liu, M., Cheng, X., Parvez, F., Islam, T., Ahmed, A., Rakibuz-Zaman, M., Jiang, J., Roy, S., Paul-Brutus, R., Slavkovich, V., Islam, T., Levy, D., VanderWeele, T. J., Pierce, B. L., Graziano, J. H., Ahsan, H., & Chen, Y. (2015). Interaction between Arsenic Exposure from Drinking Water and Genetic Polymorphisms on Cardiovascular Disease in Bangladesh: A Prospective Case-Cohort Study. *Environmental Health Perspectives*, *123*(5), 451–457. doi: 10.1289/ehp.1307883
- Wu, H., Shen, Y., Sivagurunathan, S., Weber, M. S., Adam, S. A., Shin, J. H., Fredberg, J. J., Medalia, O., Goldman, R., & Weitz, D. A. (2022). Vimentin intermediate filaments and filamentous actin form unexpected interpenetrating networks that redefine the cell cortex. *Proceedings of the National Academy of Sciences*, *119*(10). doi: 10.1073/pnas.2115217119
- Yamauchi, H. & Takata, A. (2021). Arsenic metabolism differs between child and adult patients during acute arsenic poisoning. *Toxicology and Applied Pharmacology*, *410*, 115352. doi: 10.1016/j.taap.2020.115352
- Yen, J.-C., Chang, F.-J., & Chang, S. (1995). A new criterion for automatic multilevel thresholding. *IEEE Transactions on Image Processing*, *4*(3), 370–378. doi: 10.1109/83.366472
- Yu, H., Yan, H., Li, J., Li, Z., Zhang, X., Ma, Y., Mei, L., Liu, C., Cai, L., Wang, Q., Zhang, F., Iwata, N., Ikeda, M., Wang, L., Lu, T., Chinese Schizophrenia Collaboration Group, Li, M., Xu, H., Wu, X., Liu, B., Yang, J., Li, K., Lv, L., Ma, X., Wang, C., Li, L., Yang, F., Jiang, T., Shi, Y., Li, T., Zhang, D., & Yue, W. (2017). Common variants on 2p16.1, 6p22.1 and 10q24.32 are associated with schizophrenia in Han Chinese population. *Molecular Psychiatry*, *22*(7), 954–960. doi: 10.1038/mp.2016.212
- Yu, M. B., Guerra, J., Firek, A., & Langridge, W. H. R. (2018). Extracellular vimentin modulates human dendritic cell activation. *Molecular Immunology*, *104*, 37–46. doi: 10.1016/j.molimm.2018.09.017
- Zaidel-Bar, R. (2013). Cadherin adhesome at a glance. *Journal of Cell Science*, *126*(2), 373–378. doi: 10.1242/jcs.111559
- Zeng, H., Uthus, E. O., & Combs Jr., G. F. (2005). Mechanistic aspects of the interaction between selenium and arsenic. *Journal of Inorganic Biochemistry*, *99*(6), 1269–1274. doi: 10.1016/j.jinorgbio.2005.03.006

- Zhang, Q., Wang, P., Fang, X., Lin, F., Fang, J., & Xiong, C. (2022). Collagen gel contraction assays: From modelling wound healing to quantifying cellular interactions with three-dimensional extracellular matrices. *European Journal of Cell Biology*, *101*(3), 151253. doi: 10.1016/j.ejcb.2022.151253
- Zhang, S., Wu, Q., Shan, Y., Zhao, Q., Zhao, B., Weng, Y., Sui, Z., Zhang, L., & Zhang, Y. (2016). Fast MS/MS acquisition without dynamic exclusion enables precise and accurate quantification of proteome by MS/MS fragment intensity. *Scientific Reports*, *6*(1), 26392. doi: 10.1038/srep26392
- Zhao, W., Zhang, Q., Chen, X., Li, Y., Li, X., Du, B., Deng, X., Ji, F., Wang, C., Xiang, Y.-T., Dong, Q., Chen, C., & Li, J. (2021). The VNTR of the *AS3MT* gene is associated with brain activations during a memory span task and their training-induced plasticity. *Psychological Medicine*, *51*(11), 1927–1932. doi: 10.1017/S0033291720000720
- Zhu, Y.-G., Yoshinaga, M., Zhao, F.-J., & Rosen, B. P. (2014). Earth Abides Arsenic Biotransformations. *Annual Review of Earth and Planetary Sciences*, *42*(1), 443–467. doi: 10.1146/annurev-earth-060313-054942
- Zimmer, A. M., Pan, Y. K., Chandrapalan, T., Kwong, R. W. M., & Perry, S. F. (2019). Loss-of-function approaches in comparative physiology: is there a future for knockdown experiments in the era of genome editing? *Journal of Experimental Biology*, *222*(7). doi: 10.1242/jeb.175737

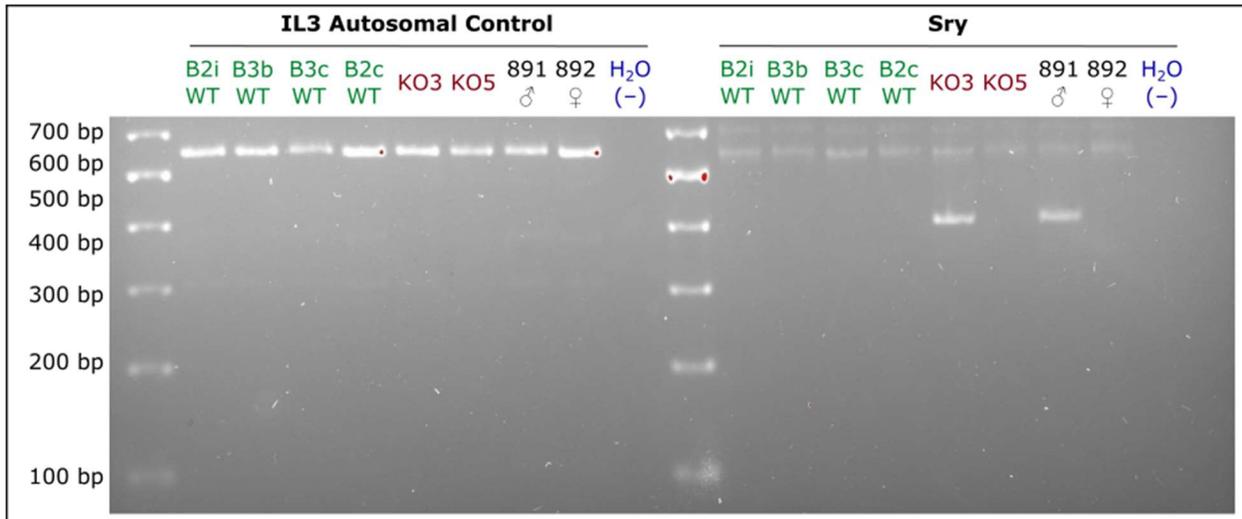
**APPENDIX A: Supplementary Figures and Tables**



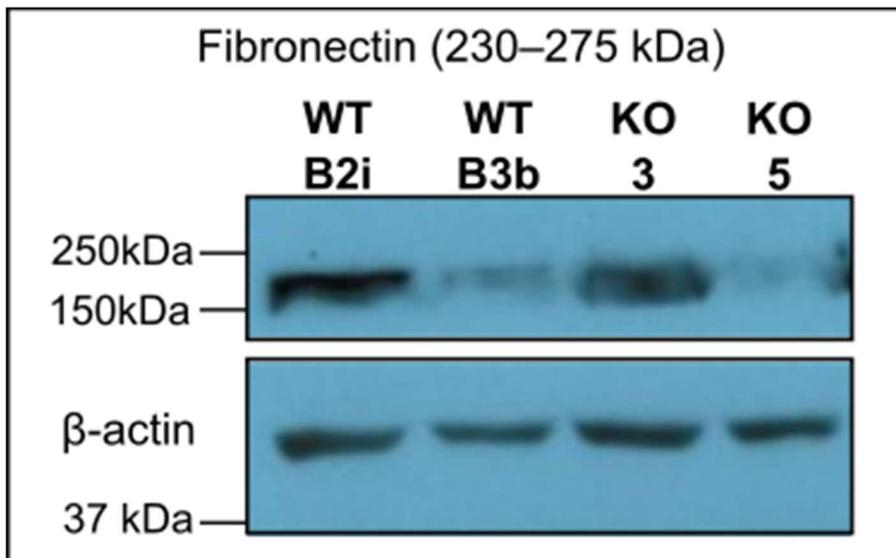
**Supplementary Figure 1. Confirmation of *As3mt* Genotype.** (A) Genotype analysis of MEFs: primers 1 and 2 amplify the wild-type *As3mt* gene and yield a 330 bp band, whereas PCR with the combination of primers 1 and 3 amplifies the modified (knockout) *As3mt* gene and results in a 400 bp band. (B) qPCR of *As3mt* mRNA: three technical replicates per population. Data analysed by  $2^{-\Delta\Delta C_t}$  method normalised based on 36B4 as the reference gene. Error bars represent the SEM of technical triplicates.



**Supplementary Figure 2. mRNA expression of retromer complex subunits in wild-type and *As3mt* knockout MEFs.** (A) *Vps35*, (B) *Vps29*, (C) *Vps26a*. Each graph is representative of two individual experiments. Relative mRNA expression quantified using the  $2^{-\Delta\Delta Ct}$  method and using Rn18S as the control gene. Error bars represent the SEM of technical triplicates.



**Supplementary Figure 3. Sex determination of MEFs by *Sry* genotype analysis.** The *Sry* primers amplify a 402 bp amplicon of the Y-specific *Sry* gene. The *IL3* autosomal control reactions yield a 544 bp amplicon. Samples 891 and 892 are male and female controls respectively.



**Supplementary Figure 4. Fibronectin expression in wild-type and *As3mt*<sup>-/-</sup> MEFs.** The fibronectin blot is representative of two individual experiments. Four MEF populations were tested: two wild-type MEF populations and two *As3mt* knockout populations.

**Supplementary Table 1. Primers used in PCR reactions**

PCR Purpose	Forward Primer Sequence (5' → 3')	Reverse Primer Sequence (5' → 3')
Amplification of WTm <i>As3mt</i> -FLAG <sup>®</sup> insert	AACTCGAGTGAGTCATGGCTGCTTCCC	ATGCCAGGGCCCATCACTTGTC
Site-directed <i>As3mt</i> mutagenesis	GGAAAAGTGCCGAATTTTGCA TCTGGCTAGTGGGAGTG	GCAATCCCTGCCACTCCC ACTAGCCAGATGCAAATTCG
Wild-type <i>As3mt</i> detection	GACGCTGATGAGACTCACAA (primer 1)	CCATAGTAGTTCTGACTCAGC (primer 2)
<i>As3mt</i> knockout detection		CTTCCTCGTGCTTTACGGTA (primer 3)
<i>Sry</i> genotype analysis	TGGGACTGGTGACAATTGTC	GAGTACAGGTGTGCAGCTCT
<i>IL3</i> autosomal control	GGGACTCCAAGCTTCAATCA	TGGAGGAGGAAGAAAAGCAA

**Supplementary Table 2. Primers used for RT-qPCR**

Target	Forward Primer Sequence (5' → 3')	Reverse Primer Sequence (5' → 3')
<i>As3mt</i> mRNA	GTTAGTTCGAGGTATTATGGC	GTGTTCAAGATAGGTTTTAGC
<i>Vim</i> mRNA	CGGCTGCGAGAGAAATTGC	CCACTTTCGGTTCAAGGTCAAG
<i>Vps35</i> mRNA	TCAAGGCCGAGCCGTGAGC	CCCTTTTACCTCCATGAAGCTCTTCCC
<i>Vps29</i> mRNA	GCCCTGGTGACAGGATGTTGGT	CACCGGTGCGGAATGTGCAG
<i>Vps26a</i> mRNA	GGCGGCGGTGACAATGAGTTTTC	AAAAGCTCTTTCCTGAGACAGACTCGC
<i>Rplp0</i> (36B4) mRNA	GGCACCGAGGCAACAGTT	TCATCCAGCAGGTGTTTGACA
18S rRNA	QuantiTech™ Mm_Rn18s_3_SG, (Qiagen, cat.: QT02448075)	

Primers were obtained from Integrated DNA Technologies (IDT) and purified by standard desalting. PCR primers were diluted serially in sterile, deionised water first to a concentration of 100 mM and then to a concentration of 10 μM. qPCR primers were diluted in sterile, deionised water to a concentration of 100 mM before the forward-reverse pairs were further diluted together into 4 μM primer mixes.

The vimentin RT-qPCR primer sequences were obtained from literature (Mulholland et al., 2012).

**Supplementary Table 3. List of adherens junction adapter proteins**

Adapter Protein	Detection in MEFs by LC-MS
$\alpha$ -actinin-1 (Actn1)	Yes
$\alpha$ -actinin-4 (Actn4)	Yes
$\alpha$ -adducin (Add1)	Yes
$\alpha$ E-catenin (Catenin $\alpha$ -1)	Yes
$\alpha$ N-catenin (Catenin $\alpha$ -2)*	Yes
$\beta$ 2 spectrin (Sptbn1)	Yes
$\beta$ -catenin (Catenin $\beta$ -1)	Yes
ADIP	No
Afadin (Afdn)	Yes
Ajuba (Jub)	No
AKAP5	No
Ankyrin G (Ank3)	Yes
ARVCF	No
Clip1	Yes
Dlg1	Yes
Eplin (Lima1)	Yes
Ezrin (Ezr)	Yes
ILK	Yes
IQGAP1	Yes
Krit1	No
LOMP	No
Magi1	No
Merlin (NF2)	No
Moesin (Msn)	Yes
Nezha	No
p120-catenin ( $\delta$ -Catenin or Ctnnd1)	Yes
Palladin (Palld)	Yes
Plakoglobin ( $\gamma$ -Catenin or Jup)	Yes
Plekha7	No
Protein 4.1 (Epb41)	Yes
Radixin (Rdx)	Yes
Supervillin (Svil)	Yes
Trip6	No
Vinculin (Vcl)	Yes
ZO-1 (Tjp1)	Yes

\* $\alpha$ N-catenin was added upon curation because it is another known adherens junction adapter

Blue: adapters that link the actomyosin cytoskeleton with adhesion receptors.

Red: adapters that link the microtubule network with adhesion receptors.

Purple: adapters that link both actin filaments and microtubules with adhesion receptors.

The list of cell junction adapters was gathered from a literature search (Zaidel-Bar, 2013).

## APPENDIX B: Macro program for VPS35 punctal analysis of immunofluorescent confocal microscopy images.

```
run("Bio-Formats Importer", "open=[] autoscale color_mode=Colorized rois_import=[ROI manager] split_channels view=Hyperstack stack_order=XYCZT");
filename = getInfo("image.filename");
filepath = getDir("file");
filepathname = filepath + filename;
print("[Now running " + filename + " through segmentation and quantification steps...]");
```

```
//Select Cell of Interest
```

```
selectWindow(filename+" - C=2");
//Colour balance is set as default (maximum DIC = 60,000)
run("Color Balance...");
setMinAndMax(0, 60000);
//Change selection tool, polygon is optimal for irregular shapes
setTool("polygon");
//Wait for manual confirmation of "cell shape" selection
waitForUser("Please Select Cell of Interest on DIC Image and Press OK");
//Adds "cell shape" defined by manual selection as an ROI to the manager
run("ROI Manager...");
roiManager("Add");
```

```
// Close DIC Image
```

```
if (isOpen(filename+" - C=2"))
    close(filename+" - C=2");
```

```
//Clear Background of Fluorescent Channel
```

```
selectWindow(filename+" - C=0");
//Select the "cell shape" ROI
roiManager("Select", 0);
//Clear everything except the cell shape
run("Clear Outside");
//Sets default background after clear to black, sometimes
//the background is white after clearing which interferes with downstream
Color.setBackground("Black")
run("Select None");
```

```
//Clear Background of DAPI Channel
```

```
selectWindow(filename+" - C=1");
```

```

//Same process as above just for a different channel
roiManager("Select", 0);
//Same process as above just for a different channel
run("Clear Outside");
//Same process as above just for a different channel
Color.setBackground("Black")
//Removes "cell shape" ROI from manager. Removing is necessary for proper
//segmentation in next steps
run("Select None");
roiManager("Delete");

//Segmentation and analysis of Nuclei
selectWindow(filename+" - C=1");
run("Color Balance...");
//Colour balance is set for optimal fluorescence (DAPI = 50,000)
setMinAndMax(0, 50000);
//Duplicates image of DAPI channel to have original image for analysis
// and working copy for macro to act on.
selectWindow(filename+" - C=1");
run("Duplicate...", " ");
selectWindow(filename+" - C=1-1");
// Uses default thresholding algorithm built into FIJI to pick up DAPI signal
run("Auto Threshold", "method=Default ignore_black white");
//Creates binary (black and white) output from thresholding
run("Convert to Mask");
// New ROI created that includes anything that passed through thresholding
run("Create Selection");
//Select original image and overlay ROI onto original image
selectWindow(filename+" - C=1");
run("Restore Selection");
//Wait for manual confirmation of nuclei selection
waitForUser("If the selection is correct press OK");
//Set measurements of interest for the area inside the nuclei ROI
run("ROI Manager...");
run("Set Measurements...", "area mean modal centroid display add redirect=None decimal=3");
roiManager("Add");
//Measure nuclei parameters and add to results table under label "Nuclei Results"
roiManager("Measure");
for (i = 0; i<nResults; i++)
    {
        oldLabel = getResultLabel(i);
        delimiter = indexOf(oldLabel, ":");
        newLabel = "Nuclei Results";
        setResult("Label", i, newLabel);
        irecall = i;
    }

```

```

    }

//Segmentation and analysis of Puncta

//Remove nuclei ROI from ROI manager
roiManager("Select", 0);
run("Select All");
roiManager("Deselect");
roiManager("Delete");
//Select channel with VPS35
selectWindow(filename+" - C=0");
run("Color Balance...");
//Optimize balance to see VPS35 puncta clearly (cosmetic only wont change recorded
flourescent values)
//Anyone using this code hereafter will probably need to change the 2 values below, value on left
= autoflourescence, value on right is brightness of image
setMinAndMax(0, 35000);
selectWindow(filename+" - C=0");
//Duplicates image of VPS35 channel to have original image for analysis
// and working copy for macro to act on.
run("Duplicate...", " ");
//For puncta auto-threshold using the “Yen” method (Yen et al., 1995)
run("Threshold...");
//Wait for manual confirmation of puncta selection.
//Provides stop in program to adjust the auto thresholding if clearly incorrect.
waitForUser("Please Select Puncta to Analyze");
//Creates binary (black and white) output from thresholding
run("Convert to Mask");
//Shrinks thresholded puncta by 1 size.
//Removes background from analysis to clean up binary image
run("Erode");
//Grows thresholded puncta by 1 size back to original size for analysis
run("Dilate");
// New ROI created that includes anything that passed through thresholding
run("Create Selection");
//Select original image and overlay ROI onto original image
selectWindow(filename+" - C=0");
run("Restore Selection");
//Wait for manual confirmation of puncta selection.
waitForUser("If the selection is correct press OK");
//Set measurements of interest for the area inside the VPS35 ROI.
run("ROI Manager...");
run("Set Measurements...", "area mean modal centroid display add redirect=None decimal=3");
roiManager("Add");
//up to this point the ROI is one item and will provide measurements for one item.
//splitting the ROI will generate a new ROI for each puncta allowing for individual assessment

```

```

roiManager("Split");
//remove the non-split ROI from analysis
roiManager("Select", 0);
roiManager("Delete");
//Measure the individual puncta for
//mean signal intensity, size, and distance from the center of the nucleus
//Add results to results table under label "Puncta Results"
roiManager("Measure");
for (i = irecall + 1; i<nResults; i++)
    {
        oldLabel = getResultLabel(i);
        delimiter = indexOf(oldLabel, ":");
        newLabel = "Puncta Results";
        setResult("Label", i, newLabel);
        irecall = i;
    }

//Save the results as a CSV file
saveAs("Results", filepathname+" Results.csv");

//End of Segmentation and quantification
//Close all open windows
print("Segmentation and quantification of " + filename + " complete :)");
run("Close All");
selectWindow("Results");
run("Close");
selectWindow("Log");
run("Close");
selectWindow("ROI Manager");
run("Close");
selectWindow("Color");
run("Close");
selectWindow("Threshold");
run("Close");

```

## ***Fiji* Macro program description**

The workflow of Macro program is proceeds as follows:

- Colour balance is set as default (maximum DIC = 60,000)
- A manual selection defines a closed area around each cell using the polygon tool
- The region of interest (ROI) is added to the selection manager
- The region outside of the ROI is cleared
- The DAPI (nuclear) channel is split from the Alexa Fluor™ 594 (VPS35) channel
- The following steps are carried out on the image of the VPS35 channel:
  - The balance range (0–50,000) is adjusted by zeroing at a higher value according to the background level
  - Auto-thresholding is done using the “Yen” method (Yen et al., 1995)
  - Converting to a mask
  - Selecting each thresholded punctum as an ROI
  - Overlaying these ROIs onto the image of the DAPI channel with the cut-out of the previously-defined cell area as the mask
- The program stops for manual approval of the selection
- The ROI manager is used to measure the mean signal intensity, size, and distance from the nuclear centroid of each ROI. These data were recorded in addition to the total number of puncta in the cell. Only puncta within the defined cell perimeter are analysed.
- These data are written into a spreadsheet as the output file (one spreadsheet per cell).