THE FUNCTIONAL CHARACTERIZATION OF TWO UNKNOWN GENES: *SLC2A14* AND *MFSD14A*

by

Haonan Zhouyao

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

In partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Food and Human Nutritional Sciences

University of Manitoba

Winnipeg Manitoba

Canada

ABSTRACT

The absorption of glucose, a fundamental energy source for all living cells requires membrane transporters. Currently, approximately 30% of the protein coding genes in humans are unknown. The glucose transporters appear to be attractive candidates for function investigation since many of them are associated with diseases such as cancer and are currently used as drug targets.

GLUT14 was reported as a duplicon of GLUT3 and is exclusively expressed in the testis. The substrate mediated by this transporter was unknown and variations in *SLC2A14* were associated to non-testicular diseases such as the Alzheimer's disease. Taken together, the function of GLUT14 needs to be examined in order to determine its substrate and extend prior knowledge on the disease mechanism.

MFSD14A was reported as a novel sugar transporter due to sequence similarities with mouse glucose transporters and a conserved sugar binding motif. Since no functional study was conducted on this gene, it is therefore hypothesized that MFSD14A might function as a novel sugar transporter, with glucose being one of the probable substrates.

A microarray study on the gills of *Carcinus Maenas* showed down-regulation of *Mfsd14a* by exposure to acidified sea water, revealing the possibility of MFSD14A mediating other substrates involved in acid-base balance. We therefore additionally hypothesized that MFSD14A could mediate the transport of ammonia, a critical compound in acid-base homeostasis.

These hypotheses and research gaps were addressed using bioinformatic studies utilizing deposited sequences from databases, functional analysis in the *Xenopus laevis* oocyte system using radiolabeled substrates, and a Mfsd14a-knockdown zebrafish (*Danio rerio*).

Our studies confirmed that GLUT14 mediates the transport of glucose and dehydroascorbic acid but no MFSD14A-mediated glucose uptake was found. However, the uptake and release of

radiolabelled methylamine (a proxy of ammonia) was observed. This result was further confirmed in the zebrafish model since Mfsd14a-knockdown larvae showed a reduction in both total and regional ammonia excretion. We therefore concluded that MFSD14A is a novel ammonia transporter. Since we have determined the substrates for two unknown transporters for the first time, our results will allow further explorations into their roles in glucose and ammonia metabolism, as well as the disease associations.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisors, AKA my academic parents, Drs. Peter Eck and Dirk Weihrauch. The effort they took to forge me into a molecular biologist / transport physiologist from "a cook with a diploma" is enormous. I still remember the feelings of seeing micropipettes for the first time in my life in Dr. Eck's lab and the first PCR reaction we've assembled together. I will also never forget the heat from the "hot seat" in Dr. Weihrauch's office, and drawing ammonia transport pathways under different scenarios will likely remain one of the most thrilling experiences in my research life. Without the patient guidance, encouragement, understanding and support from Drs. Eck and Weihrauch, I wouldn't have been able become who I am today. You are the scientists that I look up to. I will keep up to your expectations and perform research in the highest quality, ethics, and integrity.

I would like to sincerely thank my thesis advisory committee, Drs. Carla Taylor and Georg Hausner for their guidance and advice regarding my thesis. Your thoughtful inputs during committee meetings as well as informative comments on the thesis is greatly appreciated.

My special thanks go to Drs. Steve F. Perry and Alex M. Zimmer for providing me the invaluable opportunity to stay in their laboratory in the University of Ottawa to carry out part of my research project. Their knowledge in fish physiology added considerably to my research experience. I would also like to thank Drs. Sijo Joseph (Thandapilly), Lovemore N. Malunga and Nancy Ames from the Agriculture and Agri-Food Canada for giving me the opportunity to collaborate on multiple projects. The collaboration allowed me to gain insights from government research institution and greatly enhanced my work experience.

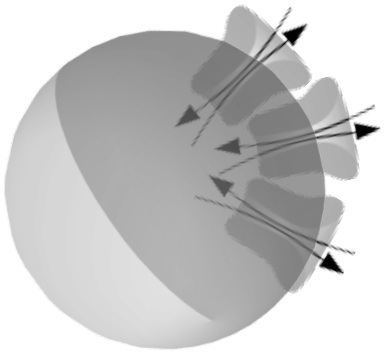
Large thanks are owned to all the members from the Eck and Crab lab: Ruotong Liu, Matthew Granger, Dr. Sandra Fehsenfeld, Alex Quijada-Roudriguez, Ashley Tripp, Garett Allen, Mikyla Nash and Maria Sachs for all your support during my Ph.D. program.

Furthermore, I want to thank the numerous students, professionals, technicians, Animal Care and Veterinary Service staff and the Environmental Health and Safety Staff who have provided training and support during my study. In addition, I thank the University of Manitoba for providing financial support to my Ph.D. program.

Last but not least, I thank my parents, Dr. Yi Yao and Mr. Yi Zhou for everything.

DEDICATION

This dissertation work is dedicated to Dr. David Oliver Richter, who has been a continuous source of support and encouragement to me. I am beyond thankful to have you in my life. Your ability to always bring the biggest smile to my face even during the most challenging times fill me up with joy. You enrich me through your fatherly companionship, and through the sharing of the enormous amount of knowledge you possess. Thank you for always having my back. You are my partner in crime (positive connotation, of course).



FOREWORD

This thesis is written in manuscript style and is composed of six chapters and four manuscripts. The first chapter includes an "Overall Introduction", which will explain the existing body of evidence surrounding the novel, uncharacterized transporters, Solute Carrier Family 2 member 14/Glucose Transporter 14 (SLC2A14/GLUT14) and the Major Facilitator Superfamily Domain Containing 14 A (MFSD14A, formally known as Hippocampus Abundant Gene Transcript 1, Hiat1). Research manuscripts, addressing specific objectives, are presented in the following chapters 2 to 5, with a one-page transitional bridge statement at the end of each manuscript. The final chapter discusses an "Overall Conclusion" of the thesis with concluding remarks of the work, its limitations, and proposed future directions of the research project.

List of Authors of Each Manuscript:

Manuscript 1 (Published in the American Journal of Clinical Nutrition 2017 Dec; 106(6): 1508–1513)

Mandana Amir Shaghaghi[1*], Haonan Zhouyao[1*], Hongbin Tu[2], Hani El-Gabalawy[4], Gary H. Crow[5], Mark Levine[2], Charles N. Bernstein[3,4], Peter Eck[1]

Manuscript 2 (Published in the Journal of Advances in Nutrition and Food Science, Volume 2022, Issue 01, DOI: 10.37722/ANAFS.2022101)

Haonan Zhouyao[1], Sandra Fehsenfeld[6], Dirk Weihrauch[7], Peter Eck[1]

Manuscript 3 (Published in the Journal of Advances in Nutrition and Food Science, Volume 2022, Issue 01, DOI: 10.37722/ANAFS.2022102)

Haonan Zhouyao[1], Dirk Weihrauch[7], Peter Eck[1]

Manuscript 4 (Submitted to the Journal of Experimental Biology. Submission number: JEXBIO/2022/244279.)

Haonan Zhouyao[1], Alex M. Zimmer[8,9], Sandra Fehsenfeld[6], Thomas Liebenstein[10], David O. Richter[10], Gerrit Begemann[10], Peter Eck[1], Steve F. Perry[9], Dirk Weihrauch[7]

Authors Affiliations:

*Shared First Authorship

[1]Department of Human Nutritional Sciences, University of Manitoba, Winnipeg, Canada.

[2]Molecular and Clinical Nutrition Section, National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institutes of Health (NIH), Bethesda, USA.

[3]IBD Clinical and Research Centre, University of Manitoba, Winnipeg, Canada.

[4]Department of Internal Medicine, University of Manitoba, Winnipeg, Canada.

[5]Department of Animal Sciences, University of Manitoba, Winnipeg, Canada.

[6]Université du Québec à Rimouski, Département de biologie, chimie et géographie, 300 Allée des Ursulines, Rimouski, QC G5L 3A1, Canada

[7]Department of Biological Sciences, University of Manitoba, Winnipeg, Canada.

[8]University of Alberta, Department of Biological Sciences, Edmonton, Canada

[9]University of Ottawa, Department of Biology, Ottawa, Canada

[10]Universität Bayreuth, Faculty of Biology, Chemistry and Earth Sciences, Bayreuth, Germany

Corresponding Author for Manuscripts 1-3: Peter K Eck, University of Manitoba, Faculty of Agriculture and Foods Sciences, Department of Food and Human Nutritional Sciences

W569 Duff Roblin Building, 190 Dysart Road, Winnipeg MB R3T 2N2, Canada

E-mail: peter.eck@umanitoba.ca.

Phone: +1-204- 291 2917

Corresponding Author for Manuscript 4: Dirk Weihrauch, University of Manitoba, Faculty of Sciences, Department of Biological Sciences

W467 Duff Roblin Building, 190 Dysart Road, Winnipeg MB R3T 2N2, Canada

E-mail: dirk.weihrauch@umanitoba.ca.

Phone: +1-204-474-6310

TABLE OF CONTENTS

Abstract II-III

Acknowledgements IV-V

Dedication VI

Foreword VII-IX

Table of contents X-XVII

List of all tables XVIII

List of all sequences XIX

List of all figures XX-XXII

List of abbreviations XXIII-XXVI

CHAPTER 1

OVERALL INTRODUCTION **1-32**

1 Introduction 1-2

1.1 Current knowledge on *SLC2A14* (GLUT14) 2-4

1.1.1 Genomic organization of *SLC2A14* (GLUT14) 6

1.1.2 Expression and cellular localization of *SLC2A14* (GLUT14) 6-7

1.1.3 Disease associations in the *SLC2A14* gene 6

1.1.3.1 Alzheimer's disease 6-7

1.1.3.2 Acute lymphoblastic leukemia 7-8

1.1.3.3 Rheumatoid arthritis 8

1.1.3.4 Primary open-angle glaucoma 9

1.2 Current knowledge on *MFSD14A*/MFSD14A 9-15

1.2.1 The genomic organization of *MmMfsd14a* 11

1.2.2 The tissue expression of *MmMfsd14a* 11

1.2.3 The cellular expression and subcellular location of *MmMfsd14a*/MmMfsd14a 12

1.2.4 The Neuronal expression of MmMfsd14a and its response to dietary interventions 13

1.2.5 Disease association of *MmMfsd14a* 14

1.2.5.1 Globozoospermia and infertility in male mice with a disrupted *MmMfsd14a* gene 14-15

1.2.6 *Mfsd14a* and its potential involvement in acid-base balance and ammonia excretion 15-16

1.3 Key methods for functional studies: the *Xenopus laevis* oocyte system and the zebrafish (*Danio rerio*) model system 16-19

1.3.1 The *Xenopus laevis* oocyte system 16-17

1.3.2 The zebrafish (*Danio rerio*) model system 17-19

1.4 Knowledge gaps 19

1.4.1 *SLC2A14* and GLUT14 19-20

1.4.2 *MFSD14A* and MFSD14A 20

1.5 Hypotheses 20

1.6 Objectives 21

1.7 References 22-32

CHAPTER 2

MANUSCRIPT 1

THE SLC2A14 GENE, ENCODING THE NOVEL GLUCOSE/DEHYDROASCORBATE TRANSPORTER GLUT14, IS ASSOCIATED WITH INFLAMMATORY BOWEL DISEASE **32-60**

2.1. Abstract 34-35

2.2 Introduction 36-37

2.3 Methods 38-40

2.3.1. Participants 38

2.3.2. Substrate transport 38-39

2.3.3 Single nucleotide polymorphism selection, genotyping, and association analysis 39

2.3.4. Statistical analysis 39-40

2.4 Results 41-46

2.4.1. GLUT14 mediates cellular dehydroascorbic acid and glucose uptake 41-42

2.4.2 *SLC2A14* single nucleotide polymorphisms independently associate with inflammatory bowel disease 43-46

2.5. Discussion 47-48

2.6. Acknowledgements 49

2.7. Competing interests 50

2.8 References 51-57

2.9 Supplementary Information 58-60

TRANSITIONAL STATEMENT 1 **61**

CHAPTER 3

MANUSCRIPT 2

THE MURINE MAJOR FACILITATOR SUPERFAMILY DOMAIN CONTAINING 14A (MFSD14A) GENE DOES NOT ENCODE A GLUCOSE TRANSPORTER **62-97**

3.1 Abstract 64

3.2 Introduction 65-66

3.3 Methods 67-71

3.3.1 Bioinformatics genomics analysis 67

3.3.2 Protein analysis 67

3.3.3 Expression patterns analysis 67

3.3.4 Plasmid preparation 68-69

3.3.5 Functional expression of Mfsd14a in *Xenopus laevis* oocytes 69-71

3.3.5.1 Chemicals and reagents 69

3.3.5.2 Plasmid preparation 69

3.3.5.3 Oocyte preparation 70

3.3.5.4 Microinjection of oocytes 70-71

3.3.6 Statistical analysis 71

3.4 Results 72-81

3.4.1 The mouse *Mfsd14a* gene locus 72-73

3.4.2 The mouse Mfsd14a protein 74-77

3.4.3 Tissues expression of the mouse *Mfsd14a* transcript 78-80

3.5 Discussion 81-82

3.6 Conclusion 83

3.7 References 84-86

3.8 Supplementary Information 87-97

CHAPTER 4

MANUSCRIPT 3

THE HUMAN MAJOR FACILITATOR SUPERFAMILY DOMAIN CONTAINING 14A (MFSD14A) GENE DOES NOT ENCODE A GLUCOSE TRANSPORTER 98-131

4.1 Abstract 100

4.2 Introduction 101

4.3 Methods 102-105

4.3.1 Bioinformatics genomics analysis 102

4.3.2 Protein analysis 102

4.3.3 Expression patterns analysis 102-103

4.3.4 Subcloning of *MFSD14A* 103-104

4.3.5 Oocyte experiments 104

4.3.6 cRNA injection 104

4.3.7 Glucose uptake study in oocytes expressing MFSD14A and GLUT3 105

4.4 Results 106-113

4.4.1 The human MFSD14A genomic locus 106-107

4.4.2 The human MFSD14A protein 108-112

4.4.3 Tissue expression of the human MFSD14A transcript 113

4.4.4 Human MFSD14A does not mediate glucose uptake into Xenopus laevis oocytes 113

4.5 Discussion 114-115

4.6 Conclusion 116

4.7 References 117-119

4.8 Supplementary Information 120-131

TRANSITIONAL STATEMENT 2 132

CHAPTER 5

MANUSCRUPT 4

CHARACTERIZATION OF TWO NOVEL AMMONIA TRANSPORTERS, HIAT1A AND HIAT1B, IN THE TELEOST MODEL SYSTEM *DANIO RERIO* 133-184

5.1 Abstract 135

5.2 Introduction 136-139

5.3 Methods 140-149

5.3.1 Sequence-based genetic structure analysis 140

5.3.2 Heterologous expression of DrHiat1a and DrHiat1b in *Xenopus oocytes* 141-142

5.3.3 [H^3] methylamine transport studies in oocytes expressing DrHiat1a and DrHiat1b 142-143

5.3.3.1 *[H^3] methylamine uptake experiments* 142-143

5.3.3.2 *[H^3] methylamine release experiments* 143

5.3.4 *In situ* hybridization 143-144

5.3.5 Ammonia excretion studies in zebrafish with *DrHiat1a* or *DrHiat1b* knockdown 145

5.3.5.1 Zebrafish 145

5.3.5.2 Morpholino injections 145-146

5.3.5.3 Whole-body ammonia flux 146-147

5.3.5.4 Scanning Ion-selective micro-Electrode technique (SIET) 147-149

5.3.6 Statistics 149

5.4 Results 150-161

5.4.1 Genetic analysis of Hiat1 150-152

5.4.2 DrHiat1-mediated methylamine (MA) transport in *Xenopus laevis* oocytes 153-155

5.4.3 *In situ* hybridization of Hiat1 isoforms in zebrafish embryos and larvae 156-158

5.4.4 Ammonia excretion in zebrafish larvae with *DrHiat1* knockdown 159-161

5.5 Discussion 162-165

5.5.1 Zebrafish Hiat1a and Hiat1b as novel ammonia transporters 163

5.5.2 Hiat1b-mediated ammonia excretion in zebrafish larvae 163-165

5.6 Conclusion 166

5.7 Acknowledgements 167

5.8 Competing interests 167

5.9 Funding 167

5.10 References 168-176

5.11 Supplementary Information 177-184

CHAPTER 6

DISCUSSION AND CONCLUSION **185-193**

6.1 Summary of the major findings 185-186

6.2 The discovery of *SLC2A14*/GLUT14 186-187

6.2.1 *SLC2A14*/GLUT14 and Alzheimer's disease 187-188

6.2.2 *SLC2A14*/GLUT14 and Rheumatoid arthritis 188-190

6.2.3 *SLC2A14*/GLUT14 and Inflammatory bowel disease 190-193

6.3 Discovering MFSD14A as a novel ammonium transporter 193-198

6.3.1 *MFSD14A*/MFSD14A and Globozoospermia 198

6.4 Strengths 199-200

6.5 Limitations 201

6.6 Conclusion and future direction 202-203

6.7 References 204-216

LIST OF ALL TABLES

Table 2.1 Genetic associations of SNPs in the SLC2A14 gene to UC and CD. 44

Supplementary Information Table 2.9.1 General characteristics of the study participants. 58

Supplementary Information Table 2.9.2 Summary of 8 tagging SNPs in GLUT 14 gene genotyped in this study. 59

Supplementary Information Table 3.8.1 Sequences of primers used in the cloning process of mouse *Mfsd14a*. Lower case letters indicate restriction enzyme recognition sites. 87

Supplementary Information Table 4.8.1 List of currently available published literature on MFSD14A. 120-122

Supplementary Information Table 4.8.2 Primers used in this study. 123

Supplementary Information Table 5.11.1 Protein sequences for phylogenetic tree as shown in **Figure 1.** 179-182

Supplementary Information Table 5.11.2 Primer sequences. 183-184

Table 6.2.1 A list of cells with confirmed *MFSD14A* mRNA expression and without the expression of *RhBG* and *RhCG*. 196

LIST OF ALL SEQUENCES

Supplementary Information Sequence 3.8.1 Amino acid sequence of the full mouse *Mfsd14a* protein (formally known as hippocampus abundant transcript 1, Hiat1). 91

Supplementary Information Sequence 4.8.1 The amino acid sequence of the human MFSD14A protein. 125

LIST OF ALL FIGURES

Figure 2.1 GLUT14 isoforms mediate DHA and deoxyglucose uptake. *Xenopus laevis* oocytes expressing. 42

Figure 2.2 Genetic linkage across the *SLC2A14* locus in individuals with Crohn disease. 46

Supplementary Information Figure 2.1 GLUT 14 isoforms do not mediate ascorbic acid (DHA) and fructose uptake. 60

Figure 3.1 The mouse mfsd14a genomic locus. 73

Figure 3.2 The predicted secondary (panel A) and tertiary (panel B) structures for the mouse Mfsd14a protein. 75-77

Figure 3.3 Mouse Mfsd14a expression data. 79

Figure 3.4 The uptake of radiolabeled 2-deoxyglucose into *Xenopus laevis* oocytes expressing a known glucose carrier, human GLUT3 (clone ID: HsCD00021270) and mouse mfsd14a. N=15-20 oocytes per group. 80

Supplementary Information Figure 3.8.1 Signature motifs in the mouse mfsd14a (hiat1) protein aligning with known monosaccharide transporters, as depicted in Matsuo et al. 88

Supplementary Information Figure 3.8.2 Alignment of the 5' transcripts (EST and refence sequence evidence form NCBI) in the mouse Mfsd14a locus. 89

Supplementary Information Figure 3.8.3 An overview of the alignments of the 3' transcripts (EST and refence sequence evidence form NCBI) in the mouse *Mfsd14a* locus. 90

Supplementary Information Figure 3.8.4 Predictions of membrane topology for the mouse Mfsd14a protein. 92-93

Supplementary Information Figure 3.8.5 NCBI protein report for Reference Sequence: NP_032272.2. 94

Supplementary Information Figure 3.8.6 Top ten tissues for *Mfsd14a* expression the adult mouse. 95

Supplementary Information Figure 3.8.7 *Mfsd14a* expression in adult mouse tissues. 96

Supplementary Information Figure 3.8.8 *Mfsd14a* expression in adult mouse brain from the Allen Mouse Brain Atlas 97

Figure 4.1 The human *MFSD14A* genomic locus and its encoded transcript 107

Figure 4.2 The predicted (panel A) secondary and (panel B) tertiary structures for the mouse Mfsd14a protein 109-111

Figure 4.3 Human MFSD14A expression data from the FANTOM5 project. Data Source: EBI expression Atlas. Expression levels in transcripts per million, TPM 112

Figure 4.4 The uptake of radiolabeled 2-deoxyglucose into *Xenopus laevis* oocytes expressing GLUT3 (known glucose carrier) and the human MFSD14A 113

Supplementary Information Figure 4.8.1 Alignment of the 5' transcripts (EST and reference mRNA sequence evidence form NCBI) in the human MFSD14A locus, identifying 155 additional 5' nucleotides for human exon 1 (beyond the annotated NCBI transcript NM_033055). 124

Supplementary Information Figure 4.8.2 Predictions of membrane topology for the mouse mfsd14a protein. Obtained from the Constrained Consensus TOPology prediction server (http://cctop.enzim.ttk.mta.hu), where 10 sources/algorithms are combined to create a model. 126

Supplementary Information Figure 4.8.3 NCBI protein report for Reference Sequence: NP_149044.2. 127

Supplementary Information Figure 4.3.3-A Expression of *MFSD14A* in adult human tissues. Data Source: EBI expression Atlas the GTEx project. Expression levels in transcripts per million, TPM. 128

Supplementary Information Figure 4.3.3-B Expression of *MFSD14A* in adult human tissues.

Data Source: EBI expression Atlas the NIH Epigenomics Roadmap project. Expression levels in transcripts per million, TPM. 129

Supplementary Information Figure 4.3.3-C Expression of *MFSD14A* in adult human tissues. Data Source: EBI expression Atlas the Kaessman project. Expression levels in transcripts per million, TPM. 130

Supplementary Information Figure 4.3.3-D Expression of *MFSD14A* in adult human tissues. Data Source: EBI expression Atlas the ENCODE project. Expression levels in transcripts per million, TPM. 131

Figure 5.1. Phylogenetic analysis of Hiat1-like, Hiat1a and Hiat1b across vertebrates. 151-152

Figure 5.2 DrHiat1a and DrHiat1b-mediated methylamine/methylammonium (MA) uptake in *Xenopus laevis* oocytes. 154

Figure 5.3 Relative methylamine/methylammonium (MA) release of pre-loaded *Xenopus laevis* oocytes expressing DrHiat1a or DrHiat1b. 155

Figure 5.4. Expression pattern of *DrHiat1a* and *DrHiat1b* in zebrafish embryos and larvae during early development. 158

Figure 5.5 Whole animal ammonia excretion in *DrHiat1a* and *DrHiat1b* knock-down (morpholino) larvae (4dpf). 159

Figure 5.6 Regional NH_4^+ flux in *DrHiat1b* knock-down larvae. 160

Supplementary Information Figure 5.11.1 Alignment of the three *Danio rerio* Hiat1 isoforms. 177

Supplementary Information Figure 5.11.2 *In situ* hybridization of *DrHiat1b* without proteinase K treatment. 178

LIST OF ABBREVIATIONS

ABCB1: The ATP-binding cassette transporter protein P-glycoprotein

CD: Crohn's disease

cDNA: Complementary DNA

Ci: Curie

Cm: Centimeter

CPM: Counts per minute

CRISPR/Cas9: Clustered regularly interspaced short palindromic repeats and clustered regularly interspaced short palindromic repeats associated protein 9.

cRNA: Capped messenger ribonucleic acid

DHA: Dehydroascorbic acid

Dm: *Drosophila melanogaster*

Dpf: Days post fertilization

DPY19L2: Developmental pluripotancy-associated family 19 Like 2

Dr: *Danio rerio*

EMBL-EBI: European Molecular Biology Laboratory's European Bioinformatics Institute

ENCODE: The Encyclopedia of DNA Elements

EST: Expressed sequence tag

GLUT: Glucose transporter family

GLUT1: Glucose transporter family member 1

GLUT2: Glucose transporter family member 2

GLUT3: Glucose transporter family member 3

GLUT8: Glucose transporter 8 family member

GLUT14: Glucose transporter family member 8

GXD: Gene Expression Database

Hiat1: Hippocampus abundant gene transcript 1

HPLC: High Performance Liquid Chromatography

Hs: *Homo sapiens*

IBD: Inflammatory bowel disease

JTT: Jones-Taylor-Thornton model

Kb: Kilobase

MA: Methylamine

mCi: Millicurie

MEGA X: Molecular Evolutionary Genetics Analysis for Linux Operation System

MFSD14A: Major Facilitator Superfamily Domain Containing 14A

MGI: The Materials Genome Initiative project

mL: Milliliter

Mm: *Mus musculus*

Mmol: Millimole

MO: Morpholino

mRNA: Messenger ribonucleic acid

NANOGP1: Nanog Homeobox Pseudogene 1

NCBI: The National Center for Biotechnology Information

NEB: New England Biolabs

ºC: Degree Celsius

OR: Odds ratio

OR2: Oocyte ringer 2

ORF: Open reading frame

P: Partial pressure gradient

PCR: Polymerase chain reaction

PICK1: Protein kinase C, alpha – binding protein

Pmol: Picomole

RPKM: Reads Per Kilobase Million

Rh: Rhesus factor

RhAG: Rhesus factor associated glycoprotein type A

RhBG: Rhesus factor associated glycoprotein type B

RhCG: Rhesus factor associated glycoprotein type C

RNA: Ribonucleic acid

RNA-seq: Ribonucleic acid sequencing

S.D.: Standard deviation

SAS: Statistical Analysis System

SDS: Sodium dodecyl sulfate

SEM: Standard error of the mean

SIET: Scanning Ion-Selective Electrode Technique

SLC: Solute carrier protein family

SLC2: Solute carrier protein family 2

SLC2A14: Solute carrier protein family 2 member 14

SLC22A4: Solute carrier protein family 22 member 4

SLC22A5: Solute carrier protein family 22 member 5

SLC22A23: Solute carrier protein family 22 member 23

SLC23A1: Solute carrier protein family 23 member 1

SLC35A3: Solute carrier protein family 35 member 3

SNP: Single nucleotide polymorphisms

SPATS16: Spermatogenesis-associated protein 16

SAAS6: Spindle assembly abnormal protein 6 homolog

Tet C: tetracycline efflux major facilitator superfamily transporter C

TM: Transmembrane domains

T_{amm}: Total ammonia concentration

TPM: Transcript per million

UC: Ulcerative colitis

UTR: Untranslated region

mg: Milligram

μM: Micromole

CHAPTER 1:

OVERALL INTRODUCTION

1 INTRODUCTION

Transporting ions and organic molecules across the lipid bilayers of cells accurately is crucial for the maintenance of health and homeostasis within the body. Most of the molecules are transported by the following three types of membrane transport proteins (transporters): the channels, the primary active transporters, and the secondary active transporters (Perland & Fredriksson, 2017). The secondary active transporters, commonly known as the solute carriers (SLCs) belong to the largest group of membrane-bound transporters in humans, consisting of at least 430 identified members (Fredriksson et al., 2008). The SLCs transport substrates by utilizing energy from coupled ions, such

as Na^+, or facilitative diffusion. The SLCs are essential throughout the body. For instance, they enable the absorption of nutrients, the excretion of wastes, and are involved in the maintenance of the acid-base homeostasis (Colas et al., 2016; Hediger et al., 2004; Nakhoul & Lee Hamm, 2013). Moreover, SLCs have significant medical values since they can be used as drug targets providing novel treatment options (Lin et al., 2015; Rives et al., 2017). Currently, around 30% of the 430 identified SLCs in humans are still with unknown characteristics and systematic research is required to reveal their functions and potential involvements in diseases (Colas et al., 2016).

In our research group, we are interested in sugar transporters, particularly glucose transporters, due to their involvements in numerous diseases and the possibility to target them directly by pharmaceutical agents, once identified (Dagogo-Jack et al., 2021). For instance, one of the glucose transporters expressed in the kidney is SGLT2. It plays a significant role in the renal glucose reabsorption. When the function of SGLT2 is inhibited by medications, filtered glucose in the kidneys will no longer be retrieved and will be excreted through the urine (Dagogo-Jack et al., 2021). Under normal physiological conditions, urinary glucose excretion is prevented by the actions of the renal glucose transporters. However, for people with diabetes, increasing urinary glucose excretion could work as a novel or alternative treatment option to improve their blood glucose control (Dagogo-Jack et al., 2021). To date, the discovery on sugar transporters continues, and among the unknown transporter genes, *SLC2A14* and *MFSD14A* appeared to be attractive candidates for further investigations. Both genes were predicted to encode membrane proteins that potentially mediate the transport of sugars based on motifs that showed sequence homologies to characterized glucose transporters (Matsuo et al., 1997; Wu & Freeze, 2002). Studying these two unknown transporters also holds value in medical related research as variations in the *SLC2A14* locus were found to be associated to various diseases such as Alzheimer's disease, cancers, and inflammatory bowel disease (IBD) (Amir Shaghaghi et al., 2016; Nag et al., 2013). On the other hand, the loss of function of *Mfsd14a* caused

infertility in homozygous male mice with a disease phenotype similar to globozoospermia in humans (Doran et al., 2016). In this thesis, I aim to characterize both *SLC2A14* and *MFSD14A* by using bioinformatics and *in vitro* studies.

1.1 Current knowledge on *SLC2A14* (GLUT14)

The *SLC2A14* gene encodes the GLUT14 protein. It was first described by Wu and Freeze in 2002 through the search within the human genome for uncharacterized proteins potentially belonging to the glucose transporter (GLUT) family (Wu & Freeze, 2002). The GLUT family members in humans are highly conserved, primarily transporting glucose (GLUT1-4 & GLUT6-12), fructose (GLUT5), and myoinositol (mammalian H^+-myoinositol symporter, HMIT) (Deng & Yan, 2016). Tissue and cell-type specific expression patterns are commonly found in the GLUT family. For example, GLUT1 can be found primarily in the erythrocytes and brain, and GLUT3 is expressed mainly in the brain, testis and specific cell lines such as the choriocarcinoma cell line (Jar) (Clarson et al., 1997; Deng & Yan, 2016). Due to the high sequence homology (94.5% DNA sequence identity) between GLUT14 and GLUT3, GLUT14 was proposed as a duplicon of GLUT3 exhibiting potentially similar function as a facilitated glucose transporter (Wu & Freeze, 2002).

1.1.1 The Genomic organization of *SLC2A14* (GLUT14)

The *SLC2A14* gene is located on chromosome 12, p13.3, 10 Mb away from the *SLC2A3* gene. The sequence identity between the matured mRNA of *SLC2A14* and *SLC2A3* is high (~94.5% identity between *SLC2A14* and *SLC2A3* cDNA). Wu and Freeze first described two isoforms of *SLC2A14*, the longer isoform (GLUT14-L), containing 11 exons and the shorter isoform (GLUT14-S), containing 10 exons. Shaghaghi and colleagues (2016) updated knowledge and described 10 additional exons in the *SLC2A14* gene, indicating that there are overall 20 exons contained in the *SLC2A14* gene.

Distinct tissue-specific exon-utilization was also observed in *SLC2A14*, and there are five distinct transcriptional start sites identified in exon 1, 4, 5, 10 and 12, respectively (Amir Shaghaghi et al., 2016). For neuronal cells originating from the hippocampus, the transcription of *SLC2A14* starts from exon 1 where the first four exons (exons 1-4) are exclusively utilized in neuronal cells. Exons 5-9 on the other hand were utilized in the testicular tissues, not in the neuronal tissue (Amir Shaghaghi et al., 2016). Exon 10 is a major alternative transcriptional start site for *SLC2A14* transcripts expressed in the testis. When the testis-specific transcript originates from the 10^{th} exon, the 11^{th} exon is skipped. This accounts for approximately 53% of all SLC2A14 transcript in the testis. This finding was the same for the neuronal-specific *SLC2A14* transcripts (Amir Shaghaghi et al., 2016). For the remaining testis-specific transcript originating from exon 5, exon 11 is not skipped. Lastly, exons 12-20 are utilized in all tissues where exon 12 represents the last identifiable transcriptional start site. 3.5% and 7% of the overall *SLC2A14* transcripts collected appeared to skip exon 13, and the first 198 nucleotides from exon 20, respectively (Amir Shaghaghi et al., 2016). The details of *SLC2A14* chromosomal location, overlapping and neighboring genes, overall genomic organization, and tissue-specific exon utilization are summarized in Figure 1.1.

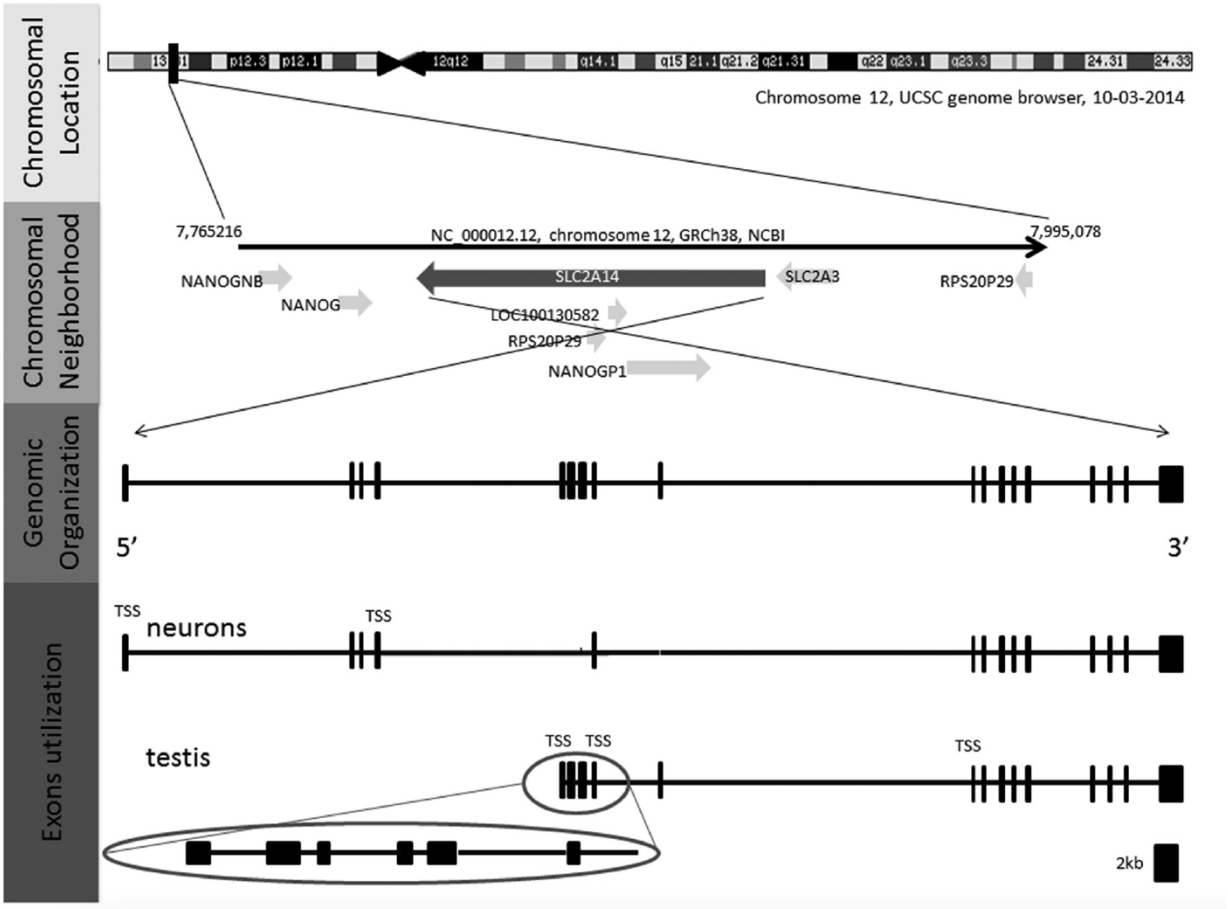


Figure 1.1 Chromosomal location, overlapping and neighboring genes, overall genomic organization, and tissue-specific exon utilization of the *SLC2A14* gene (Adapted from Shaghaghi et al., 2016 with author's premission).

1.1.2 The Expression and cellular localization of *SLC2A14* (GLUT14)

Initially, *SLC2A14* was described to be exclusively expressed in the testis (Wu and Freeze, 2002). By reviewing the publicly available RNA-sequencing datasets (Pipes et al., 2013) and results from experiments utilizing quantitative, real-time PCR, additional tissues expressing *SLC2A14* were identified by Shaghaghi and colleagues (2016). In addition to the testis, *SLC2A14* is also expressed in the heart, liver, kidney, small intestine, colon, lung, ovary, whole blood, brain, and skeletal muscle. Additional expression of *SLC2A14* found in the brain and neuronal cells may provide an explanation on the association of *SLC2A14* with neurological diseases such as Alzheimer's disease and Parkinson disease, however, the biological role of *SLC2A14* in these diseases remaines to be determined.

1.1.3 Disease associations with the *SLC2A14* gene

1.1.3.1 Alzheimer's disease

Alzheimer's disease is the most common type of dementia, affecting a total of 7.1% of the population aged above 65 in Canada (WHO, 2016). It is characterized as a progressive loss of memories which could possibly lead to the loss of the ability to interact with the surrounding environment (Adlimoghaddam et al., 2016; Amir Shaghaghi et al., 2016). Genome-wide association studies in humans found an association of the intronic rs10845990 polymorphism within the *SLC2A14* gene and Alzheimer's disease in a Caucasian population. Additionally, an independent, case-control association study in a Han Chinese population also revealed an association between rs10845990 polymorphism and late-onset Alzheimer's disease where carriers of the G genotype (GT+GG) showed a 1.4-fold increased risk compared to the TT genotype (Wang et al., 2012). Further, in common fruit flies, *Drosophila melanogaster,* the orthologue for human GLUT14 is named Glut1. The loss of function of the *Drosophila* ortholog of the human GLUT14, dmGlut1, was associated with the enhancement of Tau toxicity (Shulman et al., 2011). In patients with Alzheimer's disease, the Tau

proteins are misfolded and abnormally shaped (Brunello et al., 2020), disrupting the microtubule functions and thereby interrupting the transport of nutrients and other molecules across the nerve cells (Brunello et al., 2020).

1.1.3.2 Acute lymphoblastic leukemia

Acute lymphoblastic leukemia, also known as acute lymphocytic leukemia is a type of blood and bone marrow cancer and is the most common type of cancer among children (Hunger & Mullighan, 2015). Acute lymphoblastic leukemia is characterized by the rapid proliferation of immature lymphoblasts when the progenitors of the B cells or T cells fail to differentiate into mature cells (Hunger & Mullighan, 2015). The chromosomal translocations, as well as the modification of genetic or epigenetic functions can both contribute to the development of the acute lymphoblastic leukemia. One possible cause of the modification of the genetic or epigenetic functions is the aberrant methylation of gene promoter-related CpG islands (Taylor et al., 2007). If the gene targeted by the CGI is a tumor suppressor gene, the silencing can have deleterious effects (Taylor et al., 2007). In order to identify which gene promotor is presented in the acute lymphoblastic leukemia, Taylor et al (2007) developed methylation microarray profiles using a CpG islands microarray representing more than 3000 unique sequences in the human chromosomes and confirmed that the aberrant methylation in *SLC2A14* is associated to the development of the acute lymphoblastic leukemia (Taylor et al., 2007). Further, in the human B cell precursor leukemia cell line, NALM-6, and Jurkat cells (immortalized human T lymphocyte cell line) treated with 5-aza-2-deoxycytidine (a nucleic acid synthesis inhibitor), trichostatin (a potent and reversible inhibitor of Histone Deacetylase) or a combination of both, the expression of *SLC2A14* was up-regulated except in the NALM-6 cell line treated with 5-aza-2-deoxycytidine or trichostatin, suggesting the possibility that *SLC2A14* may present an epigenetic

biomarker that could be useful for the diagnosis and the treatment for the acute lymphoblastic leukemia, and the detection of leukemic relapse (Taylor et al., 2007).

1.1.3.3 Rheumatoid arthritis

Rheumatoid arthritis is a chronic autoimmune disease that affects approximately 1% of the global population, predominantly in women and the elderly population (Veal et al., 2014). It is characterized as inflammation of the synovium and can cause severe joint damage. Genetic factors attribute to almost 60% of the overall risk in the development of the rheumatoid arthritis. Veal and colleagues (2014) found a substantial protective effect against Rheumatoid arthritis when a 129-kb deletion occurs on chromosome 12. As the result of this deletion, *SLC2A3* is deleted entirely, as well as a portion of *SLC2A14* (Veal et al., 2014). Since *SLC2A3* encodes a high-affinity glucose transporter important in the immune response and chondrocyte metabolism, deleting its function could potentially subside the inflammatory response of the synovium and thereby reduce the severity of the rheumatoid arthritis. *NANOGP1* and *SLC2A14* on the other hand are not obvious candidates for a direct role in rheumatoid arthritis according to Veal and colleagues since *NANOGP1* is expressed but untranslated, and the only tissue known to express *SLC2A14* at that time was the testis, far away from the joints (Veal et al., 2014). However, Shaghaghi and colleagues (2016) found additional tissues besides the testis expressing *SLC2A14*, making its association with rheumatoid arthritis possible again due to the high sequence homology (94.5% DNA sequence identity) between *SLC2A14* and *SLC2A3* (Amir Shaghaghi et al., 2016; Wu & Freeze, 2002). However, since the function of *SLC2A14* remained unknown, the exact mechanism of its contribution to RA remained unclear.

1.1.3.4 Primary open-angle glaucoma

As the most common cause of irreversible vision loss, glaucoma is the result of degeneration of the retinal ganglion cells and the loss of nerve axons passing through the optic nerve head (Quigley & Broman, 2006). Slowly, the loss of the peripheral visual field can progress to the loss of the central visual field, and eventually blindness. In the Caucasian population, open-angle glaucoma is the most common form of glaucoma, responsible for approximately 50% of all the cases of vision loss and affecting up to 12% of the population over the age of 60 (Friedman et al., 2006; Nag et al., 2013). The cause of the open-angle glaucoma can be environmental and/or genetic, as well as elevated intraocular pressure (Carbonaro et al., 2008; Charlesworth et al., 2010; Nag et al., 2013; van Koolwijk et al., 2007). Nag and colleagues (2013) have identified that the copy number variations at chromosome 12p.13.3 had a moderate effect on the intraocular pressure in people with European ancestry, and both genes *SLC2A14* and *SLC2A3* are overlapped in this chromosomal region. Due to the complexity of genes in this region influencing the susceptibility of glaucoma, the causal gene at this locus has not yet been identified, and the disease mechanism is currently unknown. It is known, however, that diabetes and chronic elevated blood glucose levels are both significant risk factors for the open-angle glaucoma (Song et al., 2016). Since *SLC2A3* encodes a well-characterized glucose transporter, GLUT3, disrupting its function could lead to imbalance of the regional glucose level, leading to the development of open-angle glaucoma. Again, due to the high similarity between *SLC2A14* and *SLC2A3* (94.5% DNA sequence homology), the possibility of *SLC2A14* being associated with glaucoma in a similar way as *SLC2A3* cannot be omitted.

1.2 Current knowledge on *MFSD14A*/MFSD14A

In humans, the major facilitator superfamily domain containing 14A, MFSD14A, formally known as HIAT1 (hippocampus abundant gene transcript 1) is an unknown membrane protein

previously categorized under the atypical solute carrier (SLC) superfamily. In this chapter, MFSD14A originating from different organisms will be discussed. To avoid confusion, the abbreviation of the scientific name of each species will be added in front of the name of the gene. For example, mouse *Mfsd14a* gene will be written as *MmMfsd14a*.

After the initial cloning and sequencing of the *MmMfsd14a* gene by Matsuo et al (2003), MmMfsd14a protein was proposed to function as a novel sugar transporter since it is sharing several characteristic structural motifs which are commonly found in the glucose transporter family (GLUT). One of the characteristics is the sugar transporter-specific motif, D-R/K-X-G-R-R/K which can be found between the second and third putative transmembrane domains, and a sequence segment similar to the P-E-S-P-R motif, commonly found in the GLUT family, is present at the end of the sixth putative membrane spanning region (Matsuo et al., 1997).

The function of HsMFSD14A is currently unknown. But HsMFSD14A seems to play an important role in the maintenance of the overall homeostasis since the mRNA expression level of *CmMfsd14a* is affected by chronic exposure to acidified sea water (found in the gills of *Carcinus maenas*), suggesting the possibility of CmMfsd14a and perhaps it's human orthologue, HsMFSD14A, is mediating the transport of a substrate involved in acid-base balance. Moreover, *MmMfsd14a* mRNA expression levels were altered in both a mouse primary cortex cell line maintained in media deprived of common amino acids as well as whole animals exposed to various dietary stressors. This observation suggests the possible involvement of MmMfsd14a in the transport of nutrients or waste products resulting from nutrient metabolism. Interestingly, when eliminating *MmMfsd14a* completely, male homozygous mice became infertile with a disease phenotype similar to humans with globozoospermia, suggesting the possibility of MmMfsd14a mediating the transport of a substrate critical for the development of the spermatozoa. In this chapter, the current knowledge on MFSD14A

(HIAT1) will be reviewed, including the genomic organization, tissue expression, mRNA expression and disease association.

1.2.1 The genomic organization of *MmMfsd14a*

The cDNA of *MmMfsd14a* was first cloned and sequenced by Matsuo and colleagues from the hippocampus of the neonatal mouse (Matsuo et al., 1997). Accordingly, the mRNA sequence is 2.7 Kb long, translating into a protein consisting of 490 amino acids with 12 putative transmembrane domains. However, no functional tests were done to validate this statement.

1.2.2 The tissue expression of *MmMfsd14a*

To date, there is no publication indicating the protein and mRNA expression profile of HsMFSD14A. In mice, the mRNA expression of *MmMfsd14a* was found to be relatively even across the brain and peripheral tissues (Doran et al., 2016; Lekholm et al., 2017; Matsuo et al., 1997). *MmMfsd14a* mRNA is detected in various tissues including the heart, skeletal muscle, blood, brain, eye, intestine, kidney, liver, lung, olfactory bulb, ovary, spinal cord, spleen, uterus, and testis (Doran et al., 2016; Matsuo et al., 1997). Among the tissues listed above, the mouse brain was further dissected to reveal the mRNA expression of *MmMfsd14a* in-depth. In the mouse brain, *MmMfsd14a* was found in the brainstem, cerebellum, cortex, hippocampus, hypothalamus, thalamus and striatum (Lekholm et al., 2017).

The protein expression level of MmMfsd14a was also investigated in the mouse brain. The localization of MmMfsd14a protein is ubiquitous, and the expression pattern correlates well with the expression pattern of other neuronal transporters found when studying the expression patterns of the 307 *Slc* genes in the Allen Brain Atlas and comparing their patterns with neuronal, interneuronal, and astrocylic markers (Dahlin et al., 2009; Lekholm et al., 2017).

1.2.3 The cellular expression and subcellular location of *MmMfsd14a*/MmMfsd14a

The cellular and subcellular expression of *MmMfsd14a* in mice was investigated by eliminating *MmMfsd14a* in the mouse (Doran et al., 2016) and utilizing immunohistochemistry (Lekholm et al., 2017). Doran and colleagues (2016) generated the *MmMfsd14a*-mutant mice by deleting a section in exon 4 in *MmMfsd14a* consisting of 70bp and inserting an internal ribosome entry sequence containing the *Lac-Z* transgene. This procedure disrupts the function of *MmMfsd14a* and enables visualization of cells naturally expressing *MmMfsd14a* since the expression of *MmMfsd14a* is being replaced by the β-galactosidase encoded by the *Lac-Z* transgene, which can then be visualized when the β-galactosidase activity is stained. In the testicular tissues isolated from the *MmMfsd14a*-mutant mice, *MmMfsd14a* gene expression was detected in the cytoplasm of the Sertoli cells, which are essential for the formation of the testis and for spermatogenesis (Doran et al., 2016).

Further, cultures from the mouse embryonic cortex were prepared from embryos at days 14-16 and were used by Lekholm and colleagues to study the subcellular localization of the MmMfsd14a protein. By double staining with the Pan neuronal marker, MmMfsd14a was localized to neurons, however, not all Pan-positive cells were positive with mfsd14a (Lekholm et al., 2017). Overall, the localization of MmMfsd14a appeared to these authors to be intracellular since there was no mfsd14a localization on the plasma membrane (Lekholm et al., 2017). When double-stained with the anti-Giantin antibody, an intracellular organelle marker for Golgi, the staining overlaped (Lekholm et al., 2017). To a limited extend, the staining of MmMfsd14a was also found to overlap with anti-KDEL staining, an antibody targeting the signal peptide for retention and retrieval of proteins to the ER (Lekholm et al., 2017). Taken together, these results suggested the localization of the MmMfsd14a protein is in the Golgi apparatus, and the ER.

1.2.4 The neuronal expression of *MmMfsd14a* and its response to dietary interventions

The brain and the central nervous system have extremely high energy demands, requiring efficient transport systems to ensure proper nutrient supply, pH balance, maintenance of neurotransmitter levels, and waste removal (Dahlin et al., 2009; Simpson et al., 2007). As a putative sugar transporter, the mRNA expression of *MmMfsd14a* was investigated in a mouse cell line, in tissue samples, and also in tissues of adult mice which had been exposed to various dietary stressors (Lekholm et al., 2017). When deprived from common amino acids, mouse primary cortex cells showed a transient up-regulation of *MmMfsd14a* mRNA, returning back to control values after 7 hours of common amino acids deprivation (Lekholm et al., 2017). The mRNA expression levels of *MmMfsd14a* were also measured in tissue samples harvested from adult mice which had been challenged with 24-hour total starvation. A significant down-regulation of *MmMfsd14a* mRNA expression was found in both hypothalamus and brainstem samples, but not in the cortex (Lekholm et al., 2017). Lastly, when feeding adult mice with a high fat diet consisting of 21% fat by weight, *MmMfsd14a* mRNA expression was significantly up-regulated in the striatum, which is an area involved in "motivation and reward" responses (Lekholm et al., 2017; Yager et al., 2015). These results strongly suggest the involvement of *MmMfsd14a* in cellular energy metabolism. In addition to the enormous energy demand from the brain, its function is also complex. Other interactions, in addition to the involvements of *MmMfsd14a* and energy metabolism cannot be dismissed.

1.2.5 Disease association of *MmMfsd14a*

1.2.5.1 Globozoospermia and infertility in male mice with a disrupted *MmMfsd14a* gene

Infertility affects nearly 1 in 7 couples globally while male reproductive issues are responsible for around 30% of the overall infertility cases (Modarres et al., 2019). A rare and severe condition estimated to affect only 1 in 65,000 men, named globoozospermia, is accounting for approximately 1 in 100 cases of male infertility (Modarres et al., 2019). The phenotype globozoospermia is characterized by round-headed spermatozoa, an absence of the acrosome, and a cap-like structure covering the anterior portion of the head of the spermatozoa (Modarres et al., 2019). Without the acrosome, the spermatozoa can no longer penetrate the oocytes since one of the key functions of the acrosome is to release digestive enzymes (Modarres et al., 2019). So far several genes have been associated with this disease including *DPY19L2* (Pierre et al., 2012), *PICK1* (Xiao et al., 2009) and *SPATA16* (Dam et al., 2007). When Doran and colleagues disrupted the *MmMfsd14a* gene in mice, the male homozygous mice became infertile, with a reduced total number of spermatozoa. The developed spermatozoa were round-headed and lacking the acrosome (Doran et al., 2016). Despite being entirely sterile, the male *MmMfsd14a*-mutant mice appeared to be overall healthy since their development, hormone levels and mating behavior were all normal compared to control mice (Doran et al., 2016).

After further investigations, Doran and colleagues found severe disruption in the process of spermiogenesis in male *MmMfsd14a*-mutant mice (Doran et al., 2016). Normally, during spermiogenesis, the haploid round spermatids undergoe a series of events to mature and become mobile (Dreef et al., 2007). This process is divided into 14 steps (Marshall et al., 2005), and there was no significant difference observed initially between the control and *MmMfsd14a*-mutant mice (Doran et al., 2016). By step 6, however, deformation in spermatids in *MmMfsd14a*-mutant mice began to be apparent since there was no acrosome formation in contrast to the control mice (Doran et al., 2016).

By step 9, the spermatids are normally undergoing substantial remodeling where the head of the spermatids elongate, and the excess cytoplasm is reduced (Doran et al., 2016). In *MmMfsd14a*-mutant mice, however, there was still no acrosome formation, and the remodeling of the nucleus was also disrupted with abnormal nucleus vacuolation (Doran et al., 2016). By step 13, spermatids from the control mice appear in a typical elongated shape with a clear acrosome and no residual cytoplasm; on the other hand, *MmMfsd14a*-mutant mice produced round headed spermatids with residual cytoplasm (Doran et al., 2016). Collectively, Doran and coworkers suggested that *MmMfsd14a* plays a critical role during spermiogenesis, since *MmMfsd14a*-mutant mice produced a reduced number of spermatids with round heads, similar the findings in humans with globozoospermia (Doran et al., 2016). Based on the protein sequence analysis (Matsuo et al., 1997), Doran and coworkers speculated that the underlying mechanism between *MmMfsd14a* function and globozoospermia is related to a disrupted glycosylation in the spermatids since *MmMfsd14a* was considered a putative sugar transporter (Doran et al., 2016; Yassine et al., 2015). From a medical perspective it is most important to note that there are currently no treatment options to prevent or cure globozoospermia.

1.2.6 *Mfsd14a* and its potential involvement in acid-base balance and ammonia excretion

A microarray study investigating the impact of ocean acidification in the green shore crab *Carcinus maenas* revealed that the crab orthologue of *HsMFSD14A, CmHIAT1* (Genbank accession # DW250260.1), was downregulated in the branchial tissue of the animals after one week of exposure to elevated pCO_2 levels in the seawater (Fehsenfeld et al., 2011).

A follow-up study further showed that, when expressed in *Xenopus* oocytes, CmHIAT1 demonstrated ammonia transport capability. It was speculated that, due to its ubiquitous tissue expression pattern, CmHIAT1 could play an essential role in the general cellular detoxification/excretion mechanism of ammonia. This hypothesis needed to be tested to reevaluate

our understanding of ammonia handling in key model systems and tissues, including the mammalian kidneys and fish gills (Fehsenfeld pers. communication).

1.3 Key methods for functional studies: The *Xenopus laevis* oocyte system and the zebrafish (*Danio rerio*) model system

In the sections below, the utilization of the *Xenopus laevis* oocyte system and the zebrafish model system in the functional characterization of membrane transporters will be discussed.

1.3.1 The *Xenopus laevis* oocyte system

Xenopus laevis, commonly known as the African clawed frog, as a model organism has been used in a wide range of scientific studies ranging from pregnancy tests and the functional studies of the kidney and the muscles (Dawid & Sargent, 1988). In addition to that, the oocytes harvested from the *Xenopus laevis* oocytes were found to be an excellent system for the investigation of heterologous expressed proteins (Gurdon et al., 1971). When the capped mRNA of the gene of interest is delivered into the oocytes *via* microinjection, oocytes can utilize the mRNA rapidly to generate proteins. After the first protein expression experiment using the oocytes to express hemoglobin (Gurdon et al., 1971), many membrane transporters, particularly the glucose transporters were characterized in the oocyte system (Long et al., 2018; Pike et al., 2019; Rogers et al., 2003; Rumsey et al., 1997).

As a well-established model to study protein expression, the oocyte system offers unparalleled advantages in the functional studies of membrane transporters. First of all, it is an excellent model for studying human genetics since the genomic structure of oocytes is very simple compared with human and other experimental animals such as mice or rats (Tandon et al., 2017). Second, the oocytes provide a blank experimental background for studying of glucose transport since there are no endogenous glucose transporter are expressed in the oocyte membrane (Dawid & Sargent, 1988). Moreover, the

handling of the oocytes is easy. The size of the oocytes ranges between 1.1-1.3 mm. Due to their colossal sizes, they can be manipulated and injected under a basic dissection microscope From each frog, up to 3,000 mature oocytes can be harvested, ensuring adequate number of oocytes for the generation of accurate and reliable data (Fortriede et al., 2019). Moreover, the oocytes are extremely resilient against contamination and an artificial growth environment due to their essential nutrient storage, which not only allows them to survive for several days in artificial growth medium, but also for expressing proteins rapidly (Wagner et al., 2000).

Like all other methods, the awareness of some disadvantages when employing the oocyte system is essential. First, as a poikilothermic animal, the preferred environmental temperature for *Xenopus laevis* is between 16 to 22 °C (Fortriede et al., 2019). However, most of the oocyte experiments are done at room temperature, around 25 °C instead. The elevated environmental temperature could potentially impose risks of protein expression (Wagner et al., 2000). Second, the protein expression in oocytes is transient, limiting the design of downstream experiments since they must be done within a short period of time following the microinjection (Beck & Slack, 2001). Additionally, the oocytes are different from epithelial cells since they do not distinguish between the apical and basolateral membrane sites. Therefore, the usefulness of the oocyte system is limited to the functional characterization of heterologously expressed proteins, such as the identification of substrate and transport kinetics (Pike et al., 2019). The localization of proteins needs to be determined using complimentary, alternative methods, such as immunohistochemistry using tissues or polarized epithelial cell cultures.

1.3.2 The zebrafish (*Danio rerio*) system

Zebrafish (*Danio rerio*) are tropical fresh-water fish belonging to the minnow family, originally found in the rivers and ponds in India (Graham et al., 2018; Schilling & Webb, 2007). Nowadays, they

are often seen in pet stores and are considered a suitable model to investigate development, genetics, immunity, physiology, behavior and nutrition (Bradford et al., 2017; Briggs, 2002; Trede et al., 2004; Ulloa et al., 2011).

Even though there are other models that are evolutionarily more similar to humans, such as mice and rats, zebrafish share at least 70% genetic similarities to humans (Howe et al., 2013). At the same time, using zebrafish to study genetics and physiology offer several advantages over the other model organisms. One of the most significant advantages is that adult zebrafish are small and prefer to be housed in large groups, requiring less space (Keck et al., 2015). This makes them cheaper to maintain compared to rodents (Keck et al., 2015; MacRae & Peterson, 2003). Zebrafish also breed readily, producing 50 to 300 embryos every two weeks (Nasiadka & Clark, 2012). In contrast, the gestation time for rodents is much longer, and the overall offspring generated from rodents are much less (Pritchett-Corning et al., 2009). Since scientific experiments are often repeated multiple times to ensure the accuracy of the results, having a large quantity of offspring available repetitively is extremely helpful. Moreover, the embryos of zebrafish are laid and fertilized externally, allowing them to be collected and manipulated easily in various ways (Nasiadka & Clark, 2012). For example, the one-cell-stage fertilized eggs can be easily injected with DNA, RNA, or in our case, morpholino to either permanently or transiently modify their genetic makeup to generate transgenic, knock-out or knock-down zebrafish lines (Gut et al., 2017; Stainier et al., 2017). Morpholino or morpholino oligonucleotides are small, synthetic oligomers commonly used to modify gene expression by binding to the pre-mRNA and alter its splicing, resulting in non-functioning proteins as the end result (Gut et al., 2017). Since the zebrafish embryos are clear, their developments can be observed under a microscope (Rosen et al., 2009; Stainier et al., 2017). At the same time, the transparency also enables the visualization of tissues if they are fluorescently labeled in genetically modified zebrafish lines (Stainier et al., 2017).

However, the zebrafish model has its limit when it comes to human disease investigations. The most obvious disadvantage of using zebrafish as a model for human disease is that it is not a mammalian model (Vaz et al., 2019). Similarly, to the *Xenopus laevis* oocyte system, zebrafish are also poikilothermic animals, meaning that their optimal environmental temperature is different from the human body temperature (Vaz et al., 2019). These leads to potential differences in metabolism of certain substrates between humans and zebrafish, reducing its relevance to humans during disease modeling (Vaz et al., 2019). Since many genetic diseases of human do not exist in zebrafish, alternative models might be needed (Adamson et al., 2018). Additionally, many key metabolic pathways in zebrafish are critically different compared to humans. For example, zebrafish is ammoniotelic, excreting ammonia directly through their gills as nitrogenic waste (Litman et al., 2009; Veauvy et al., 2009). In contrast, humans convert ammonia into urea first prior to excretion (Weihrauch et al., 2009). When investigating the physiology of such metabolic pathways, the difference between human and zebrafish must be kept in mind and alternative methods needs to be used in order to generate comprehensive results.

1.4 Knowledge gaps

1.4.1 *SLC2A14* and GLUT14

Even though GLUT14 has been proposed as a duplicon of a well-characterized glucose transporter, GLUT3, due to their high sequence homology (~94% DNA sequence identity) (Wu & Freeze, 2002), conflicting evidence was found between the disease association and the expression pattern described by Wu and colleagues (2002). According to Wu and Freeze (2002), *SLC2A14* is expressed exclusively in the testis. On the other hand, however, this gene has been associated with diseases such as Alzheimer's disease (Wang et al., 2012). The substrate mediated by GLUT14 is also

unknown. Therefore, we have set out to describe characterize *SLC2A14* by heterologously expressing it in the oocytes of *Xenopus laevis* to determine its substrate/substrates.

1.4.2 *MFSD14A* and MFSD14A

MFSD14A has been proposed as a novel sugar transporter based on the sequence analysis of mouse *MmMfsd14a*, since a glucose binding motif commonly found in the GLUT family is present. However, the *MFSD14A* genetic structure in humans as well as its tissue expression remain undescribed. There is also currently no experimental verification of the substrate/substrates mediated by MFSD14A. After analyzing the gene, we found high protein sequence homology between HsMFSD14A and MmMfsd14a, as well as with CmMfsd14a. We therefore tested the transport of glucose after expressing all three proteins in frog oocytes. In addition, due to the fact that Fehsenfeld and colleague found the expression of *CmMfsd14a* in the gills of green crab was down-regulated when the crabs were exposed to elevated environmental *p*CO2, it is possible that CmMfsd14a is mediating the transport of a substrate playing critical roles in acid-base regulation, besides glucose (Fehsenfeld et al., 2011). We have therefore modified our research questions and proceeded with investigating HsMFSD14A and its function as an ammonia transporter, by expressing it in the *Xenopus laevis* oocytes as well as using genetically modified zebrafish (*Danio rerio*).

1.5 Hypotheses

Hypothesis 1: the *SLC2A14* gene encodes for a glucose transporter.

Hypothesis 2: the *HsMFSD14A* gene encodes for a glucose transporter.

Hypothesis 3: the *DrMfsd14a* gene encodes for an ammonium transporter.

Hypothesis 4: *DrMfsd14a* gene knockdown affects the excretion of ammonia in zebrafish.

1.6 Objectives

Objective 1: Determine if the two GLUT14 protein isoforms mediate glucose transport.

Objective 2: Determine *HsMFSD14A* and *MmMfsd14a* genetic structure, isoforms, and tissue expression via a bioinformatics approach in order to set up experiments to determine substrate/substrates mediated by HsMFSD14A and MmMfsd14a.

Objective 3: Determine if HsMFSD14 and MmMfsd14a mediate glucose transport.

Objective 4: Determine if HsMFSD14A and MmMfsd14a mediate ammonium transport.

Objective 5: Determine if the genetic knockdown of DrMfsd14a and its isoforms affects ammonia excretion using zebrafish (*Danio rerio*) as a common model organism used to study the biological process and human diseases.

1.7 References

Adamson, K. I., Sheridan, E., & Grierson, A. J. (2018). Use of zebrafish models to investigate rare human disease. *Journal of Medical Genetics*, *55*(10), 641-649. https://doi.org/10.1136/jmedgenet-2018-105358

Adlimoghaddam, A., Sabbir, M. G., & Albensi, B. C. (2016). Ammonia as a Potential Neurotoxic Factor in Alzheimer's Disease. *Front Mol Neurosci*, *9*, 57. https://doi.org/10.3389/fnmol.2016.00057

Amir Shaghaghi, M., Murphy, B., & Eck, P. (2016). The SLC2A14 gene: genomic locus, tissue expression, splice variants, and subcellular localization of the protein. *Biochem Cell Biol*, *94*(4), 331-335. https://doi.org/10.1139/bcb-2015-0089

Beck, C. W., & Slack, J. M. (2001). An amphibian with ambition: a new role for Xenopus in the 21st century. *Genome Biol*, *2*(10), Reviews1029. https://doi.org/10.1186/gb-2001-2-10-reviews1029

Bradford, Y. M., Toro, S., Ramachandran, S., Ruzicka, L., Howe, D. G., Eagle, A., . . . Westerfield, M. (2017). Zebrafish Models of Human Disease: Gaining Insight into Human Disease at ZFIN. *Ilar j*, *58*(1), 4-16. https://doi.org/10.1093/ilar/ilw040

Briggs, J. P. (2002). The zebrafish: a new model organism for integrative physiology. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *282*(1), R3-R9. https://doi.org/10.1152/ajpregu.00589.2001

Brunello, C. A., Merezhko, M., Uronen, R. L., & Huttunen, H. J. (2020). Mechanisms of secretion and spreading of pathological tau protein. *Cell Mol Life Sci*, *77*(9), 1721-1744. https://doi.org/10.1007/s00018-019-03349-1

Dawid, I. B., & Sargent, T. D. (1988). Xenopus laevis in developmental and molecular biology. *Science*, *240*(4858), 1443-1448. https://doi.org/10.1126/science.3287620

Carbonaro, F., Andrew, T., Mackey, D. A., Spector, T. D., & Hammond, C. J. (2008). Heritability of intraocular pressure: a classical twin study. *Br J Ophthalmol*, *92*(8), 1125-1128. https://doi.org/10.1136/bjo.2007.133272

Charlesworth, J., Kramer, P. L., Dyer, T., Diego, V., Samples, J. R., Craig, J. E., . . . Wirtz, M. K. (2010). The path to open-angle glaucoma gene discovery: endophenotypic status of intraocular pressure, cup-to-disc ratio, and central corneal thickness. *Invest Ophthalmol Vis Sci*, *51*(7), 3509-3514. https://doi.org/10.1167/iovs.09-4786

Clarson, L. H., Glazier, J. D., Sides, M. K., & Sibley, C. P. (1997). Expression of the facilitated glucose transporters (GLUT1 and GLUT3) by a choriocarcinoma cell line (JAr) and cytotrophoblast cells in culture. *Placenta*, *18*(4), 333-339. https://doi.org/10.1016/s0143-4004(97)80068-9

Colas, C., Ung, P. M., & Schlessinger, A. (2016). SLC Transporters: Structure, Function, and Drug Discovery. *Medchemcomm*, *7*(6), 1069-1081. https://doi.org/10.1039/C6MD00005C

Dagogo-Jack, S., Pratley, R. E., Cherney, D. Z. I., McGuire, D. K., Cosentino, F., Shih, W. J., . . . Gantz, I. (2021). Glycemic efficacy and safety of the SGLT2 inhibitor ertugliflozin in patients with type 2 diabetes and stage 3 chronic kidney disease: an analysis from the VERTIS CV randomized trial. *BMJ Open Diabetes Res Care*, *9*(1). https://doi.org/10.1136/bmjdrc-2021-002484

Dahlin, A., Royall, J., Hohmann, J. G., & Wang, J. (2009). Expression profiling of the solute carrier gene family in the mouse brain. *J Pharmacol Exp Ther*, *329*(2), 558-570. https://doi.org/10.1124/jpet.108.149831

Dam, A. H., Koscinski, I., Kremer, J. A., Moutou, C., Jaeger, A. S., Oudakker, A. R., . . . Viville, S. (2007). Homozygous mutation in SPATA16 is associated with male infertility in human globozoospermia. *Am J Hum Genet*, *81*(4), 813-820. https://doi.org/10.1086/521314

Deng, D., & Yan, N. (2016). GLUT, SGLT, and SWEET: Structural and mechanistic investigations of the glucose transporters. *Protein Sci*, *25*(3), 546-558. https://doi.org/10.1002/pro.2858

Doran, J., Walters, C., Kyle, V., Wooding, P., Hammett-Burke, R., & Colledge, W. H. (2016). Mfsd14a (Hiat1) gene disruption causes globozoospermia and infertility in male mice. *Reproduction*, *152*(1), 91-99. https://doi.org/10.1530/REP-15-0557

Dreef, H. C., Van Esch, E., & De Rijk, E. P. (2007). Spermatogenesis in the cynomolgus monkey (Macaca fascicularis): a practical guide for routine morphological staging. *Toxicol Pathol*, *35*(3), 395-404. https://doi.org/10.1080/01926230701230346

Fehsenfeld, S., Kiko, R., Appelhans, Y., Towle, D. W., Zimmer, M., & Melzner, F. (2011). Effects of elevated seawater pCO(2) on gene expression patterns in the gills of the green crab, Carcinus maenas. *BMC Genomics*, *12*, 488. https://doi.org/10.1186/1471-2164-12-488

Fortriede, J. D., Pells, T. J., Chu, S., Chaturvedi, P., Wang, D., Fisher, M. E., . . . Vize, P. D. (2019). Xenbase: deep integration of GEO & SRA RNA-seq and ChIP-seq data in a model organism database. *Nucleic Acids Research*, *48*(D1), D776-D782. https://doi.org/10.1093/nar/gkz933

Fredriksson, R., Nordstrom, K. J., Stephansson, O., Hagglund, M. G., & Schioth, H. B. (2008). The solute carrier (SLC) complement of the human genome: phylogenetic classification reveals four major families. *FEBS Lett*, *582*(27), 3811-3816. https://doi.org/10.1016/j.febslet.2008.10.016

Friedman, D. S., Jampel, H. D., Munoz, B., & West, S. K. (2006). The prevalence of open-angle glaucoma among blacks and whites 73 years and older: the Salisbury Eye Evaluation

Glaucoma Study. *Arch Ophthalmol*, *124*(11), 1625-1630. https://doi.org/10.1001/archopht.124.11.1625

Graham, C., von Keyserlingk, M. A. G., & Franks, B. (2018). Zebrafish welfare: Natural history, social motivation and behaviour. *Applied Animal Behaviour Science*, *200*, 13-22. https://doi.org/https://doi.org/10.1016/j.applanim.2017.11.005

Gurdon, J. B., Lane, C. D., Woodland, H. R., & Marbaix, G. (1971). Use of frog eggs and oocytes for the study of messenger RNA and its translation in living cells. *Nature*, *233*(5316), 177-182. https://doi.org/10.1038/233177a0

Gut, P., Reischauer, S., Stainier, D. Y. R., & Arnaout, R. (2017). LITTLE FISH, BIG DATA: ZEBRAFISH AS A MODEL FOR CARDIOVASCULAR AND METABOLIC DISEASE. *Physiol Rev*, *97*(3), 889-938. https://doi.org/10.1152/physrev.00038.2016

Hediger, M. A., Romero, M. F., Peng, J. B., Rolfs, A., Takanaga, H., & Bruford, E. A. (2004). The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteinsIntroduction. *Pflugers Arch*, *447*(5), 465-468. https://doi.org/10.1007/s00424-003-1192-y

Howe, K., Clark, M. D., Torroja, C. F., Torrance, J., Berthelot, C., Muffato, M., . . . Stemple, D. L. (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature*, *496*(7446), 498-503. https://doi.org/10.1038/nature12111

Hunger, S. P., & Mullighan, C. G. (2015). Acute Lymphoblastic Leukemia in Children. *N Engl J Med*, *373*(16), 1541-1552. https://doi.org/10.1056/NEJMra1400972

Keck, V. A., Edgerton, D. S., Hajizadeh, S., Swift, L. L., Dupont, W. D., Lawrence, C., & Boyd, K. L. (2015). Effects of Habitat Complexity on Pair-Housed Zebrafish. *J Am Assoc Lab Anim Sci*, *54*(4), 378-383.

Lekholm, E., Perland, E., Eriksson, M. M., Hellsten, S. V., Lindberg, F. A., Rostami, J., & Fredriksson, R. (2017). Putative Membrane-Bound Transporters MFSD14A and MFSD14B Are Neuronal and Affected by Nutrient Availability. *Front Mol Neurosci*, *10*, 11. https://doi.org/10.3389/fnmol.2017.00011

Lin, L., Yee, S. W., Kim, R. B., & Giacomini, K. M. (2015). SLC transporters as therapeutic targets: emerging opportunities. *Nat Rev Drug Discov*, *14*(8), 543-560. https://doi.org/10.1038/nrd4626

Litman, T., Sogaard, R., & Zeuthen, T. (2009). Ammonia and urea permeability of mammalian aquaporins. *Handb Exp Pharmacol*(190), 327-358. https://doi.org/10.1007/978-3-540-79885-9_17

Long, W., O'Neill, D., & Cheeseman, C. I. (2018). GLUT Characterization Using Frog Xenopus laevis Oocytes. *Methods Mol Biol*, *1713*, 45-55. https://doi.org/10.1007/978-1-4939-7507-5_4

MacRae, C. A., & Peterson, R. T. (2003). Zebrafish-Based Small Molecule Discovery. *Chemistry & Biology*, *10*(10), 901-908. https://doi.org/https://doi.org/10.1016/j.chembiol.2003.10.003

Marshall, G. R., Ramaswamy, S., & Plant, T. M. (2005). Gonadotropin-independent proliferation of the pale type A spermatogonia in the adult rhesus monkey (Macaca mulatta). *Biol Reprod*, *73*(2), 222-229. https://doi.org/10.1095/biolreprod.104.038968

Matsuo, N., Kawamoto, S., Matsubara, K., & Okubo, K. (1997). Cloning of a cDNA encoding a novel sugar transporter expressed in the neonatal mouse hippocampus. *Biochem Biophys Res Commun*, *238*(1), 126-129. https://doi.org/10.1006/bbrc.1997.7252

Modarres, P., Tavalaee, M., Ghaedi, K., & Nasr-Esfahani, M. H. (2019). An Overview of The Globozoospermia as A Multigenic Identified Syndrome. *Int J Fertil Steril*, *12*(4), 273-277. https://doi.org/10.22074/ijfs.2019.5561

Nasiadka, A., & Clark, M. D. (2012). Zebrafish breeding in the laboratory environment. *Ilar j*, *53*(2), 161-168. https://doi.org/10.1093/ilar.53.2.161

Nag, A., Venturini, C., Hysi, P. G., Arno, M., Aldecoa-Otalora Astarloa, E., Macgregor, S., . . . Hammond, C. J. (2013). Copy number variation at chromosome 5q21.2 is associated with intraocular pressure. *Invest Ophthalmol Vis Sci*, *54*(5), 3607-3612. https://doi.org/10.1167/iovs.13-11952

Nakhoul, N. L., & Lee Hamm, L. (2013). Characteristics of mammalian Rh glycoproteins (SLC42 transporters) and their role in acid-base transport. *Mol Aspects Med*, *34*(2-3), 629-637. https://doi.org/10.1016/j.mam.2012.05.013

Perland, E., & Fredriksson, R. (2017). Classification Systems of Secondary Active Transporters. *Trends Pharmacol Sci*, *38*(3), 305-315. https://doi.org/10.1016/j.tips.2016.11.008

Pierre, V., Martinez, G., Coutton, C., Delaroche, J., Yassine, S., Novella, C., . . . Arnoult, C. (2012). Absence of Dpy19l2, a new inner nuclear membrane protein, causes globozoospermia in mice by preventing the anchoring of the acrosome to the nucleus. *Development*, *139*(16), 2955-2965. https://doi.org/10.1242/dev.077982

Pike, S., Matthes, M. S., McSteen, P., & Gassmann, W. (2019). Using Xenopus laevis Oocytes to Functionally Characterize Plant Transporters. *Curr Protoc Plant Biol*, *4*(1), e20087. https://doi.org/10.1002/cppb.20087

Pritchett-Corning, K. R., Chang, F. T., & Festing, M. F. (2009). Breeding and housing laboratory rats and mice in the same room does not affect the growth or reproduction of either species. *J Am Assoc Lab Anim Sci*, *48*(5), 492-498.

Quigley, H. A., & Broman, A. T. (2006). The number of people with glaucoma worldwide in 2010 and 2020. *Br J Ophthalmol*, *90*(3), 262-267. https://doi.org/10.1136/bjo.2005.081224

Rives, M. L., Javitch, J. A., & Wickenden, A. D. (2017). Potentiating SLC transporter activity: Emerging drug discovery opportunities. *Biochem Pharmacol*, *135*, 1-11. https://doi.org/10.1016/j.bcp.2017.02.010

Rogers, S., Chandler, J. D., Clarke, A. L., Petrou, S., & Best, J. D. (2003). Glucose transporter GLUT12-functional characterization in Xenopus laevis oocytes. *Biochem Biophys Res Commun*, *308*(3), 422-426. https://doi.org/10.1016/s0006-291x(03)01417-7

Rosen, J. N., Sweeney, M. F., & Mably, J. D. (2009). Microinjection of zebrafish embryos to analyze gene function. *J Vis Exp*(25). https://doi.org/10.3791/1115

Rumsey, S. C., Kwon, O., Xu, G. W., Burant, C. F., Simpson, I., & Levine, M. (1997). Glucose transporter isoforms GLUT1 and GLUT3 transport dehydroascorbic acid. *J Biol Chem*, *272*(30), 18982-18989. https://doi.org/10.1074/jbc.272.30.18982

Schilling, T. F., & Webb, J. (2007). Considering the zebrafish in a comparative context. *J Exp Zool B Mol Dev Evol*, *308*(5), 515-522. https://doi.org/10.1002/jez.b.21191

Shulman, J. M., Chipendo, P., Chibnik, L. B., Aubin, C., Tran, D., Keenan, B. T., . . . De Jager, P. L. (2011). Functional screening of Alzheimer pathology genome-wide association signals in Drosophila. *Am J Hum Genet*, *88*(2), 232-238. https://doi.org/10.1016/j.ajhg.2011.01.006

Simpson, I. A., Carruthers, A., & Vannucci, S. J. (2007). Supply and demand in cerebral energy metabolism: the role of nutrient transporters. *J Cereb Blood Flow Metab*, *27*(11), 1766-1791. https://doi.org/10.1038/sj.jcbfm.9600521

Song, B. J., Aiello, L. P., & Pasquale, L. R. (2016). Presence and Risk Factors for Glaucoma in Patients with Diabetes. *Curr Diab Rep*, *16*(12), 124. https://doi.org/10.1007/s11892-016-0815-6

Stainier, D. Y. R., Raz, E., Lawson, N. D., Ekker, S. C., Burdine, R. D., Eisen, J. S., . . . Moens, C. B. (2017). Guidelines for morpholino use in zebrafish. *PLoS Genet*, *13*(10), e1007000. https://doi.org/10.1371/journal.pgen.1007000

Tandon, P., Conlon, F., Furlow, J. D., & Horb, M. E. (2017). Expanding the genetic toolkit in Xenopus: Approaches and opportunities for human disease modeling. *Dev Biol*, *426*(2), 325-335. https://doi.org/10.1016/j.ydbio.2016.04.009

Taylor, K. H., Pena-Hernandez, K. E., Davis, J. W., Arthur, G. L., Duff, D. J., Shi, H., . . . Caldwell, C. W. (2007). Large-scale CpG methylation analysis identifies novel candidate genes and reveals methylation hotspots in acute lymphoblastic leukemia. *Cancer Res*, *67*(6), 2617-2625. https://doi.org/10.1158/0008-5472.CAN-06-3993

Tandon, P., Conlon, F., Furlow, J. D., & Horb, M. E. (2017). Expanding the genetic toolkit in Xenopus: Approaches and opportunities for human disease modeling. *Dev Biol*, *426*(2), 325-335. https://doi.org/10.1016/j.ydbio.2016.04.009

Trede, N. S., Langenau, D. M., Traver, D., Look, A. T., & Zon, L. I. (2004). The Use of Zebrafish to Understand Immunity. *Immunity*, *20*(4), 367-379. https://doi.org/https://doi.org/10.1016/S1074-7613(04)00084-6

Ulloa, P. E., Iturra, P., Neira, R., & Araneda, C. (2011). Zebrafish as a model organism for nutrition and growth: towards comparative studies of nutritional genomics applied to aquacultured fishes. *Reviews in Fish Biology and Fisheries*, *21*(4), 649-666. https://doi.org/10.1007/s11160-011-9203-0

van Koolwijk, L. M., Despriet, D. D., van Duijn, C. M., Pardo Cortes, L. M., Vingerling, J. R., Aulchenko, Y. S., . . . Lemij, H. G. (2007). Genetic contributions to glaucoma: heritability of intraocular pressure, retinal nerve fiber layer thickness, and optic disc morphology. *Invest Ophthalmol Vis Sci*, *48*(8), 3669-3676. https://doi.org/10.1167/iovs.06-1519

Vaz, R., Hofmeister, W., & Lindstrand, A. (2019). Zebrafish Models of Neurodevelopmental Disorders: Limitations and Benefits of Current Tools and Techniques. *Int J Mol Sci*, *20*(6). https://doi.org/10.3390/ijms20061296

Veal, C. D., Reekie, K. E., Lorentzen, J. C., Gregersen, P. K., Padyukov, L., & Brookes, A. J. (2014). A 129-kb deletion on chromosome 12 confers substantial protection against rheumatoid arthritis, implicating the gene SLC2A3. *Hum Mutat*, *35*(2), 248-256. https://doi.org/10.1002/humu.22471

Veauvy, C. M., Walsh, P. J., & McDonald, M. D. (2009). Effect of elevated ammonia on tissue nitrogen metabolites in the ureotelic gulf toadfish (Opsanus beta) and the ammoniotelic midshipman (Porichthys notatus). *Physiol Biochem Zool*, *82*(4), 345-352. https://doi.org/10.1086/588829

Wang, W., Yu, J. T., Zhang, W., Cui, W. Z., Wu, Z. C., Zhang, Q., & Tan, L. (2012). Genetic association of SLC2A14 polymorphism with Alzheimer's disease in a Han Chinese population. *J Mol Neurosci*, *47*(3), 481-484. https://doi.org/10.1007/s12031-012-9748-y

Wagner, C. A., Friedrich, B., Setiawan, I., Lang, F., & Bröer, S. (2000). The use of Xenopus laevis oocytes for the functional characterization of heterologously expressed membrane proteins. *Cell Physiol Biochem*, *10*(1-2), 1-12. https://doi.org/10.1159/000016341

WHO. (2016). Dementia - Fact Sheet.

Weihrauch, D., Wilkie, M. P., & Walsh, P. J. (2009). Ammonia and urea transporters in gills of fish and aquatic crustaceans. *J Exp Biol*, *212*(Pt 11), 1716-1730. https://doi.org/10.1242/jeb.024851

Wu, X., & Freeze, H. H. (2002). GLUT14, a duplicon of GLUT3, is specifically expressed in testis as alternative splice forms. *Genomics*, *80*(6), 553-557. https://doi.org/10.1006/geno.2002.7010

Xiao, N., Kam, C., Shen, C., Jin, W., Wang, J., Lee, K. M., . . . Xia, J. (2009). PICK1 deficiency causes male infertility in mice by disrupting acrosome formation. *J Clin Invest*, *119*(4), 802-812. https://doi.org/10.1172/JCI36230

Yager, L. M., Garcia, A. F., Wunsch, A. M., & Ferguson, S. M. (2015). The ins and outs of the striatum: role in drug addiction. *Neuroscience*, *301*, 529-541. https://doi.org/10.1016/j.neuroscience.2015.06.033

Yassine, S., Escoffier, J., Martinez, G., Coutton, C., Karaouzene, T., Zouari, R., . . . Arnoult, C. (2015). Dpy19l2-deficient globozoospermic sperm display altered genome packaging and DNA damage that compromises the initiation of embryo development. *Mol Hum Reprod*, *21*(2), 169-185. https://doi.org/10.1093/molehr/gau099

CHAPTER 2

MANUSCRIPT 1

This manuscript was published in the American Journal of Clinical Nutrition

(Am J Clin Nutr. 2017 Dec; 106(6): 1508–1513)

THE *SLC2A14* GENE, ENCODING THE NOVEL GLUCOSE/DEHYDROASCORBATE TRANSPORTER GLUT14, IS ASSOCIATED WITH INFLAMMATORY BOWEL DISEASE

Mandana Amir Shaghaghi[1*], Haonan Zhouyao[1*], Hongbin Tu, Hani El-Gabalawy[4], Gary H. Crow[5], Mark Levine[2], Charles N. Bernstein[3,4], Peter Eck[$1]

Authors Affiliations:

*Shared First Authorship

$Corresponding Author

[1]Department of Food and Human Nutritional Sciences, University of Manitoba, Winnipeg, Canada.

[2]Molecular and Clinical Nutrition Section, National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institutes of Health (NIH), Bethesda, USA.

[3]IBD Clinical and Research Centre, University of Manitoba, Winnipeg, Canada.

[4]Department of Internal Medicine, University of Manitoba, Winnipeg, Canada.

[5]Department of Animal Sciences, University of Manitoba, Winnipeg, Canada.

Author's contribution

My contribution:

I conducted the glucose uptake studies and added the results to the manuscript.

Other author's contribution:

PE: was the principal investigator, responsible for the conception and design of the project, the acquisition of financial support, and the writing and final editing of the manuscript; MAS: conducted the genetic association studies and created the first draft of the manuscript; HZ: conducted the glucose uptake studies and added the results to the manuscript; CNB: was responsible for conducting the Manitoba Inflammatory Bowel Disease Cohort Study and collecting the samples, and contributed to revisions of the final manuscript; HE-G: was responsible for sample collection from healthy controls; HT: conducted the dehydroascorbic acid uptake experiments; GHC: consulted on the statistical analyses; ML: contributed to the experimental design data interpretation and manuscript writing; and all authors: reviewed, read, and approved the final manuscript. CNB has served on the advisory boards of or consulted to Abbott Canada, Abbvie Canada, Bristol-Myers Squibb, Forest Canada, Takeda Canada, Janssen Canada, Hospira, and Vertex Pharmaceuticals, and has received research grants from Abbott Canada and Abbvie Canada, and an educational grant from Aptalis. He is supported in part by the Bingham Chair in Gastroenterology. None of the authors reported a conflict of interest related to the study.

2.1 ABSTRACT

Background: Variations in intestinal antioxidant membrane transporters are implicated in the initiation and progression of inflammatory bowel disease (IBD). Facilitated glucose transporter member 14 (GLUT14), encoded by the Solute Carrier Family 2 Member 14 (*SLC2A14*) gene, is a putative transporter for dehydroascorbic acid and glucose. Although very limited information on the gene exists, a shorter and longer GLUT14 isoform had been identified. We hypothesized that GLUT14 mediates glucose and dehydroascorbic acid uptake. If this function could be validated, genetic variations might associate with IBD.

Objectives: This study aimed to determine the substrate(s) for the GLUT14 protein and interrogated genetic associations of *SLC2A14* with IBD.

Design: The uptake of radiolabeled substrates into *Xenopus laevis* oocytes expressing the two GLUT14 isoforms was assessed. A gene targeted genetic association study was conducted through genotyping of single nucleotide polymorphisms (SNP) representing linkage blocks of the *SLC2A14* gene in the Manitoba IBD Cohort.

Results: Both GLUT14 isoforms mediated uptake of dehydroascorbic acid and glucose into *Xenopus laevis* oocytes. Three alleles in the SLC2A14 gene associated independently with IBD. The odds of having ulcerative colitis (UC) or Crohn's disease (CD) were elevated in carriers of the *SLC2A14* SNP *rs2889504-T* allele (OR=3.60, 95% CI 1.95, 6.64; OR=4.68, 95% CI 2.78, 8.50, respectively). Similarly, the SNP *rs10846086-G* allele was associated with an increased risk of both UC and CD (OR=2.91, 95% CI 1.49, 5.68; OR=3.00, 95% CI 1.55, 5.78, respectively). Moreover, the SNP *rs12815313-T* allele associated with increased susceptibility to CD and UC (OR: 2.12, 95% CI 1.33, 3.36; 1.61, 95% CI1.01, 2.57; respectively).

Conclusion: These findings strengthen the hypothesis that genetically determined local dysregulation of dietary vitamin C or antioxidants transport contributes to IBD development. These transporter

proteins are targetable by dietary interventions, opening the avenue to a precision intervention for IBD patients of specific genotypes.

Keywords: Genetic Association; *SLC2A14*; Vitamin C; Antioxidants; Inflammatory bowel disease.

2.2 INTRODUCTION

Inflammatory bowel disease (IBD), including ulcerative colitis (UC) and Crohn's disease (CD), are chronic, relapsing, immune-mediated intestinal inflammatory diseases of complex etiology (Green et al., 2006; Hong et al., 2016; Ng et al., 2013; Nøgård et al., 2016), predominantly found in the Caucasian population (Green et al., 2006; 2016; Ng et al., 2013, and incurring significant socioeconomic costs, e.g. in excess of annually $1.7 billion in Canada (Canadian Digestive Health Foundation, 2017). The exact contributions of environmental factors (Torres et al., 2016), immune dysregulations (Blumberg., 2016), and genetics (Cleynen et al., 2016) are unknown, which impairs therapeutic successes. Therefore, current research efforts aim to derive genetic markers to guide diagnostic and treatment decisions.

Variations in the genes *SLC23A1* (Amir Shagahi et al., 2014), *SLC22A23* (León., 2014; Barett et al., 2008) *SLC22A4*/*SLC22A5* (Nakahara et al., 2008; Ferraris et al., 2006; Newman et al., 2005; Peltekova et al., 2004), and *ABCB1* (Anderson et al., 2015; Urcelay et al., 2006; Brant et al., 2003) had been associated with IBD. They encode for intestinal membrane transporters of antioxidants, and it was hypothesis that genetically determined dysregulation of dietary antioxidants, prominently vitamin C, might contribute to IBD (Amir Shagahi et al., 2014; Corpe et al., 2013; Padayatty et al., 2003).

Vitamin C exists in two stable forms, reduced ascorbic acid and oxidized dehydroascorbic acid, each utilizing distinct pathways of transmembrane transport. The uptake of dehydroascorbic acid is mediated by selected facilitative hexose transporters (GLUT1-3, 8) (Corpe et al., 2013), however, several genes in this family remain unknown and need to be investigated in regard to function and disease association.

The primate specific *SLC2A14* gene encodes GLUT14 (Amir Shaghaghi et al., 2016), consists of 20 exons, shows tissue specific exon utilization in the 5' portion, where the first four exons are neuron specific, and exons five to nine are exclusively utilized in extraneural tissues, most prominently

testis. Expression is limited to the testis, small intestine, colon, lung, ovaries, brain, skeletal muscle, heart, kidney, liver, blood, and placenta. Two major splice variants exist, and the resulting two protein isoform have distinct N-termini, where the longer form (L-GLUT14) has twenty-nine amino acids, and the shorter form (S-GLUT14) has six (Wu & Freeze., 2003). Both isoforms locate to the cell membrane; however, no substrates are identified (Amir Shaghaghi et al., 2016; Wu & Freeze., 2003). Due to a high similarity to GLUT3 it could be hypothesized that both mediate the transmembrane transport of glucose and dehydroascorbic acid (Amir Shaghaghi et al., 2016; Wu & Freeze., 2003).

Genetic variations of *SLC2A14* are associated with diseases of the central nervous system (Shulman et al., 2011; Wang et al., 2012), lymphatic cancer (Taylor et al., 2007), rheumatoid arthritis (Veal et al., 2014), and intraocular pressure in primary open-angle glaucoma(Nag et al., 2013). These diseases are consistent with tissue expression, and by virtue of expression in the intestinal tract, it can be hypothesized that variations in *SLC2A14* could be associated with IBD.

Taken together, if dehydroascorbic acid could be confirmed as a GLUT14 substrate, *SLC2A14* would be a candidate gene for an association with IBD. Here we report on both hypotheses.

2.3 METHODS

2.3.1 Participants

The Manitoba IBD Cohort Study, initiated in 2002 included 388 persons drawn from a population based registry. Participants were required to be between 18-80 years of age and diagnosed within the previous 7 yr. The diagnosis and extent of IBD was determined based on surgical, endoscopic, radiologic, and histologic data and participating individuals were phenotyped using the Montreal classification (Silverberg et al., 2005). Controls, drawn from the general population, included healthy individuals with no personal or first-degree relatives with chronic immune diseases. Of the 388 persons in the Cohort 311 IBD patients (162 CD and 149 UC patients) were Caucasian and 142 age- and sex-matched healthy Caucasian controls were included in the study. More details on the study design and creation of this study population are previously provided (Ediger et al., 2007). The Manitoba IBD Cohort Study was approved by the Research Ethics Board of the University of Manitoba.

2.3.2 Substrate transport

As previously described elsewhere, the two major GLUT14 isoforms were subcloned (Amir Shaghaghi et al., 2016). [^{14}C]Fructose was purchased from Amersham Biosciences. Radiolabeled [^{14}C]Dehydroascorbic acid was prepared from crystalline [^{14}C]Ascorbic acid (6.6 mCi/mmol, PerkinElmer Life Sciences) (Rumsey et al., 1997). Total conversion of [^{14}C]Ascorbic acid to DHA was confirmed using HPLC with electro-coulometric detection (Washko., 1989). Radiolabeled Deoxy-D-glucose, 2-[1,2-3H(N)] (25-50Ci/mmol, PerkinElmer Life Sciences) was adjusted to required concentrations in transport buffer. Previously described *Xenopus laevis* oocyte isolation and injection techniques were used to express the GLUT14 isoforms (Soreq & Seidman, 1992).

Transport of radiolabeled substrates was determined using groups of 10–20 oocytes in OR-2 buffer at 21°C. Individual oocytes were dissolved in 500 µl of 10% SDS, and internalized radioactivity was determined using scintillation spectrometry. Transport was analyzed and plotted using Microsoft Excel, Student's *t* test was used to determined statistical differences. Data are expressed as the arithmetic mean ± S.D. of 10–20 oocytes analyzed at each data point.

2.3.3 Single nucleotide polymorphism selection, genotyping, and association analyses

A haplotype-based tag-SNP (National Institute of Environmental Health Sciences., 2013) approach was implemented using the eight tagging SNPs: rs10846086, rs12815313, rs2889504, rs10845990, rs11612319, rs7132415, rs2376904, and rs73007730. These seven intronic and one noncoding exonic SNPs (5'UTR) captured all the common variation with a minor allelic frequency >5% in *SLC2A14* (**Supplementary information, Table 1**).

Genomic DNA was isolated from peripheral white blood cells by absorption onto QIamp silics-gel following QAIGEN protease digestion (QIAGEN, Mississauga, Canada). Subsequent to column elution, the purity and concentration of DNA was determined by UV spectroscopy with a NanoDrop 2000 Spectrophotometer (Thermo Fisher, Mississauga, Canada).

Genotyping was performed for all subjects using TaqMan Real-Time polymerase chain reaction assays (Applied Biosystems, Grand Island, New York, USA) (supplementary information, TABLE 1). The genotype concordance rate was 100% in duplicate samples.

2.3.4 Statistical analysis

All data were initially processed using Microsoft Excel 2010. The functional data for radiolabels' uptake in the *Xenopus oocyte* system were processed in Microsoft Excel 2010, and the T.TEST function was used interrogate probability associated with a Student's t-Test.

Genetic data were analyzed using SAS version 9.2 (SAS Institute Inc, Cary, NC); $P<0.05$ was considered statistically significant. The association of genotypes with CD or UC risks was examined by logistic regression to estimate ORs and 95% CIs, using a three levels genotypic model (two homozygotes and one heterozygote). Overdominance was tested when the heterozygote was not intermediate in effect between the two homozygotes. The genotype-phenotype association for individuals with CD and UC was determined using multinomial logistic regression. Linkage disequilibrium and haplotype blocks were determined using Haploview 4.2 software (Broad Institute) applying default parameters (Barret et al., 2005).

2.4 RESULTS

2.4.1 GLUT14 mediates cellular dehydroascorbic acid and glucose uptake

The shorter and longer GLUT14 isoforms (S-GLUT14 and L-GLUT14) locate to the plasma membrane in mammalian cells (Amir Shaghaghi et al., 2016) . Upon expression in *Xenopus laevis* oocytes these two major GLUT14 isoforms mediate the uptake of radiolabeled deoxyglucose and dehydroascorbic acid (**Figure 2.1**). Ascorbic acid as well as fructose did not get transported by GLUT14 (**Supplementary information, Figure 2.1**).

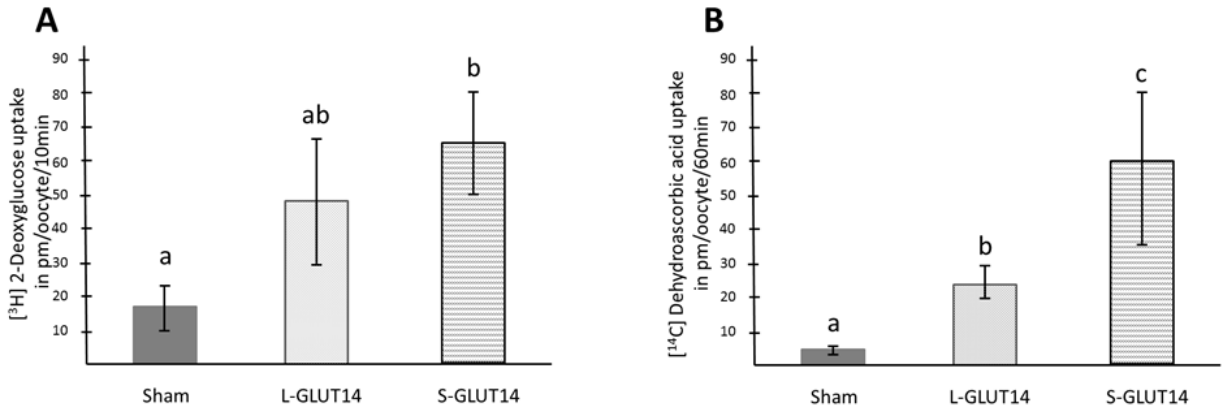


Figure 2.1

GLUT14 isoforms mediate DHA and deoxyglucose uptake. *Xenopus laevis* oocytes expressing S-GLUT14 and L-GLUT14 isoforms and exhibit uptake of radiolabeled deoxyglucose

(A) and dehydroascorbic acid (B). Incubations were performed with 300 μM [^{14}C]DHA or 300 μM deoxy-d-glucose,2-[1,2-3H(N)] on 20 oocytes in each group in 3 independent experiments. Data are expressed as means ± SDs. Lowercase letters indicate statistically significant differences determined through 1-factor ANOVA, $P < 0.05$. GLUT14, facilitated glucose transporter 14; L-GLUT14, long GLUT14 isoform; S-GLUT, short GLUT14 isoform.

2.4.*2 SLC2A14* single nucleotide polymorphism independently associate with inflammatory bowel disease

Baseline characteristics of participants are presented in the **supplementary information, Table 2**.

Genetic variations in SNPs *rs10846086*, *rs2889504* and *rs12815313* associated with UC and CD, (**Table 2.1)**. No linkage was observed for the eight tag-SNPs in the *SLC2A14* gene (**Figure 2.2**), and the pattern of inheritance for the three SNPs associated to any disease phenotype differed significantly (**Table 2.1**).

First, the susceptibility for CD and UC was elevated in individuals carrying the SNP *rs12815313*-T allele (OR: 2.12 and 1.62 respectively, TABLE 1), with equal impact sizes for T-homozygotes and CT-heterozygotes. Elevated susceptibility for CD was established for *rs12815313* T-homozygotes and CT-heterozygotes (OR: 2.04 for CT, and OR: 2.41 for TT, TABLE 1).

Second, the *rs10846086*-G allele elevated risks for UC and CD (ORs: 2.91 and 3.00, respectively, **Table 2.1**). An additive allele dosage effect was demonstrated, where highest susceptibility for UC and CD was observed for *rs10846086*-GG homozygotes (ORs: 4.30 and 4.20, respectively, TABLE 1), while the impact size was about halved for *rs10846086*-AG heterozygotes (ORs: 2.14 and 2.33, respectively, **Table 2.1**).

Third, the presence of the SNP *rs2889504*-T allele increased susceptibility to UC and CD (ORs: 3.60 and 4.68, respectively, TABLE 2.1). *Rs2889504*-GT heterozygotes exhibited the highest risk for UC and CD (ORs: 5.36 and 6.49, respectively, **Table 2.1**) compared to T-allele homozygotes (ORs: 2.37, 3.41, respectively, TABLE 2.1). However, the *rs2889504*-GT heterozygosity was not over-dominant, indicating similar effect sizes of the T allele in both genotypes.

No significant correlations were found to specific CD and UC sub-phenotypes as defined in the Montreal Classification (data not shown).

	Controls, *n* = 142	UC, *n* = 149	UC vs. control, OR (95% CI)	CD, *n* = 162	CD vs. control, OR (95% CI)
rs12815313					
CC	73 (51.4)	59 (39.6)	Ref	54 (33.4)	Ref
CT	55 (38.7)	71 (47.6)	1.59 (0.98, 2.61)	83 (51.2)	2.04 (1.25, 3.33)*
TT	14 (9.9)	19 (12.7)	1.68 (0.78, 3.63)	25 (15.4)	2.41 (1.15, 5.07)*
T carrier	69 (48.59)	90 (60.4)	1.61 (1.01, 2.57)*	108 (66.7)	2.12 (1.33, 3.36)*
rs10845990					
TT	34 (23.9)	29 (19.5)	Ref	32 (19.7)	Ref
GT	68 (47.9)	64 (42.9)	1.10 (0.60, 2.01)	76 (46.9)	1.19 (0.66, 2.13)
GG	40 (28.2)	56 (37.6)	1.64 (0.86, 3.11)	54 (33.4)	1.43 (0.76, 2.70)
G carrier	108 (76.1)	120 (80.5)	1.30 (0.74, 2.28)	130 (80.25)	1.28 (0.74, 2.21)
rs11612319					
GG	66 (46.5)	65 (43.6)	Ref	79 (48.8)	Ref
GA	66 (46.5)	69 (46.3)	1.06 (0.66, 1.72)	65 (40.1)	0.82 (0.51, 1.32)
AA	10 (7.0)	15 (10.1)	1.52 (0.64, 3.64)	18 (11.1)	1.50 (0.65, 3.48)
A carrier	79 (53.5)	84 (56.4)	1.12 (0.71, 1.78)	83 (51.2)	0.91 (0.58, 1.43)
rs7132415					
GG	37 (26.1)	36 (24.2)	Ref	31 (19.1)	Ref
GT	70 (49.3)	76 (51.0)	1.12 (0.64, 1.96)	72 (44.5)	1.23 (0.69, 2.19)
TT	35 (24.6)	37 (24.8)	1.09 (0.57, 2.08)	59 (36.4)	2.01 (1.07, 3.79)
T carrier	105 (73.9)	113 (75.8)	1.11 (0.65, 1.88)	131 (80.9)	1.49 (0.87, 2.56)
rs10846086					
AA	128 (90.1)	113 (75.8)	Ref	122 (75.4)	Ref
AG	9 (6.3)	17 (11.4)	2.14 (0.92, 4.99)	20 (12.3)	2.33 (1.02, 5.32)*
GG	5 (3.5)	19 (12.8)	4.30 (1.56, 11.9)*	20 (12.3)	4.20 (1.53, 11.5)*
G carrier	14 (9.9)	36 (24.2)	2.91 (1.49, 5.68)*	40 (24.7)	3.00 (1.55, 5.78)*
rs2376904					
GG	93 (65.5)	83 (55.7)	Ref	113 (69.8)	Ref
GA	43 (30.3)	58 (38.9)	1.51 (0.92, 2.47)	37 (22.8)	0.70 (0.42, 1.19)
AA	6 (4.2)	8 (5.4)	1.49 (0.50, 4.48)	12 (7.4)	1.65 (0.59, 4.55)
A carrier	49 (34.5)	66 (44.3)	1.51 (0.94, 2.42)	49 (30.2)	0.82 (0.51, 1.33)
rs7300773					
TT	50 (35.2)	47 (31.5)	Ref	57 (35.2)	Ref
CT	74 (52.1)	74 (49.7)	1.06 (0.64, 1.78)	76 (46.9)	0.90 (0.54, 1.48)
CC	18 (12.7)	28 (18.8)	1.65 (0.81, 3.37)	29 (17.9)	1.41 (0.70, 2.85)
CT + CC	92 (64.8)	102 (68.5)	1.18 (0.72, 1.92)	105 (64.8)	1.00 (0.62, 1.60)
rs2889504					
GG	125 (88.0)	100 (67.1)	Ref	99 (61.1)	Ref
GT	7 (4.9)	30 (20.1)	5.36 (2.26, 12.71)*	36 (22.2)	6.49 (2.77, 15.2)*
TT	10 (7.1)	19 (12.8)	2.37 (1.06, 5.34)*	27 (16.7)	3.41 (1.57, 7.38)*
T carrier	17 (12.0)	49 (32.9)	3.60 (1.95, 6.64)*	63 (38.9)	4.68 (2.78, 8.50)*

TABLE 2.1 Genetic associations of SNPs in the SLC2A14 gene to UC and CD[1]

[1]Values are n (%) unless otherwise indicated. Per-allele effects were derived from binary logistic regression performed with SAS version 9.2 (SAS Institute Inc.) through the use of a 3-level genotypic model (2 homozygotes and 1 heterozygote). Overdominance was tested when the heterozygote was not intermediate in effect between the 2 homozygotes. *Significant differences in the ORs to the reference SNPs. CD, Crohn disease; Ref, reference; *SLC2*, solute carrier family 2; SNP, single nucleotide polymorphism; UC, ulcerative colitis.

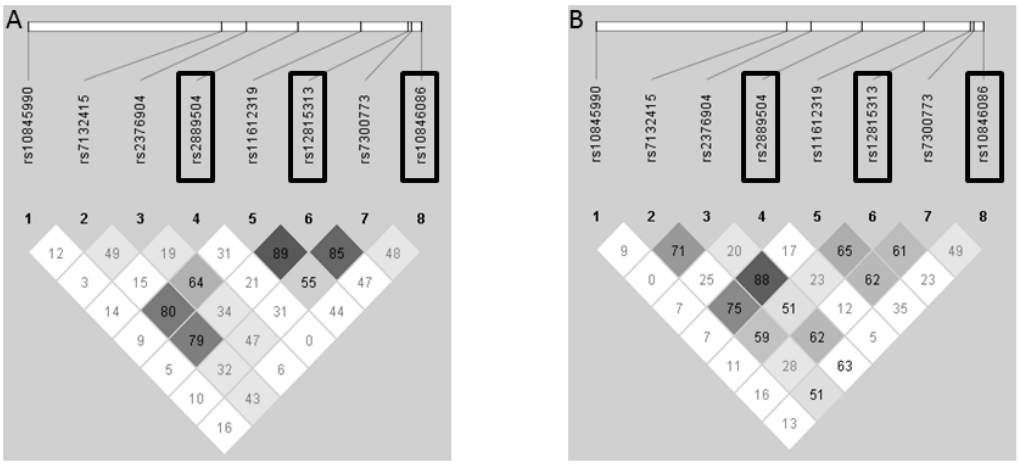


Figure 2.2 Genetic linkage across the *SLC2A14* locus in individuals with Crohn disease

(A) and ulcerative colitis (B). The degree of linkage (diamonds) is given as a percentage, and the darker colour indicates a higher degree of linkage. SNPs with relevant associations to any form of inflammatory bowel disease are indicated by rectangles. The single nucleotide polymorphisms associated with inflammatory bowel disease are not in genetic linkage. *SLC2*, solute carrier family 2; SNP, single nucleotide polymorphism.

2.5. DISCUSSION

The presented data establish GLUT14 as a membrane transporter for glucose and dehydroascorbic acid. Moreover, three *SLC2A14* SNPs associate independently with IBD in a well phenotyped Caucasian cohort of moderate size, in which major genetic associations with IBD identified by GWAS have been replicated (Ryan et al., 2013). All SNPs interrogated in this targeted genetic association study are noncodding and have been chosen based on their ability to tag haplotype blocks in the *SLC2A14* gene. This method of selection did not consider the possible impact of the SNPs on the genes or proteins function, such as transcription, translation, or protein activity, but their ability to serve as markers for disease association. In this regard it was somewhat remarkable that three variations in distinct linkage blocks strongly associated with IBD (ORs between 2.1 and 4.3), strengthening the validity of each association. However, this does not indicate causality of these SNPs, which is unlikely for the intronic SNPs *rs10846086*, and *rs2889504*, which do not fall into regions of high genetic conservation or known genetic enhancers of transcription (Gao et al., 2016). Neither does SNP *rs12815313* in the 5'UTR (NM_001286233.1:c.-171G>A) effect any known transcription factors or enhancers binding sites (Gao et al., 2016; Matys et al., 2006). Therefore, future studies to identify causal SNPs are warranted.

These findings strengthen the evidence that a local mucosal vitamin C imbalance could contribute to IBD, as previously *SLC23A1*, the gene encoding an intestinal trans-epithelial ascorbic acid transporter, was also associated with CD (Amir Shaghaghi et al., 2014). Low mucosal tissue levels and lower plasma concentrations of vitamin C have been reported in individuals with IBD, even when dietary intake was adequate (Wendland et al., 2001; Hoffenberg et al., 1997; Fernandez-Banares et al., 1989; Buffinton & Doe., 1995), suggesting increased antioxidants consumption during the inflammation process. Consequently, since genetic variations in the two intestinal vitamin C transporters genes *SLC23A1* (Amir Shaghaghi et al., 2014) and *SLC2A14* are now associated with

IBD, we speculate that a localized vitamin C deficiency due to decreased transmembrane transport could be a causative or contributing factor in the etiology of intestinal inflammation. While this study was not intended to define the underlying mechanisms as to how mutations in these transporter genes lead to IBD, two scenarios could contribute to the weakening of the intestinal barrier function.

First, variations in *SLC2A14* could decrease the capacity to provide dehydroascorbic acid to the enterocytes during the inflammatory oxidative burst, where dehydroascorbic acid is produced in the extracellular environment, immediately transported into the cell and reduced to ascorbic acid (Buffinton & Doe., 1995; Ramsey et al., 2000; Ramsey et al., 1998; Washko et al., 1998). This mechanism, called ascorbate recycling or the bystander effect greatly elevates the cells antioxidant capacity (Nualart et al., 2003; Ki et al., 2011), and if compromised, could decrease the intestinal barrier function, leading to increased bacterial invasion and inflammation.

Second, a decrease in vitamin C transport could impact the functioning of immune cells themselves through redox changes described above or changes in gene expression, enhancing the severity of an existing inflammation (Chatterjee et al., 2008; Chapple et al., 2013).

In either case, the local vitamin C imbalance caused by impaired transport of ascorbic acid through SLC23A1 or dehydroascorbic acid through GLUT14 could be remedied by dietary supplementation of the complementary forms. In the future, individuals for such interventions could be identified by their genotypes. While this may be relevant for only a minority of individual affected by IBD, genotype specific dietary prevention and intervention strategies would likely be easy to implement and less costly than current treatments.

We emphasize the fact that the underlying mechanism for the presented genetic associations ought to be validated through biological and clinical intervention studies in order to devise precision genotype specific IBD intervention strategies.

2.6 ACKNOWLEDGMENTS

Mandana Amir Shaghaghi and Haonan Zhouyao share first authorship.

The authors gratefully acknowledge the participants of the Manitoba IBD Cohort.

The authors' responsibilities were as follows: PE was the principal investigator, responsible for the conception and design of the project, seeking financial support, manuscript writing and final editing. MAS conducted the genetic association studies and drafted the first manuscript. HZ conducted the glucose uptake studies and added these results to the manuscript; CB was responsible for conducting the Manitoba IBD Cohort Study (funded by the Canadian Institutes of Health Research), sample collection, and contributed to revisions of the final manuscript; HEG was responsible for sample collection for healthy controls; HT conducted the dehydroascorbic acid uptake experiments; GHC consulted on the statistical analyses; ML contributed to experimental design data interpretation and manuscript writing.

All the authors reviewed the final version of the manuscript.

The Manitoba IBD Cohort Study is registered at the clinical trials registry, protocol NCT03262649.

2.7 COMPETING INTERESTS

Dr. Bernstein has served on the advisory boards or has consulted to Abbott Canada, Abbvie Canada, Bristol Myers Squibb, Forest Canada, Takeda Canada, Janssen Canada, Hospira, and Vertex Pharmaceuticals and has received research grants from Abbott Canada and Abbvie Canada and an education grant from Aptalis. He is supported in part by the Bingham Chair in Gastroenterology. All other authors had no conflicts of interest to declare.

2.8 REFERENCES

Amir Shaghaghi M, Bernstein CN, Serrano Leon A, El-Gabalawy H, Eck P. Polymorphisms in the sodium-dependent ascorbate transporter gene SLC23A1 are associated with susceptibility to Crohn disease. Am J Clin Nutr 2014;99(2):378-83. doi: 10.3945/ajcn.113.068015.

Amir Shaghaghi MA, Murph B, Eck P. The SLC2A14 gene: genomic locus, tissue expression, splice variants, and subcellular localization of the protein. Biochemistry and Cell Biology 2016. doi: 10.1139/bcb-2015-0165.

Andersen V, Svenningsen K, Knudsen LA, Hansen AK, Holmskov U, Stensballe A, Vogel U. Novel understanding of ABC transporters ABCB1/MDR/ P-glycoprotein, ABCC2/MRP2, and ABCG2/BCRP in colorectal pathophysiology. World J Gastroentero 2015;21(41):11862-76. doi: 10.3748/wjg.v21.i41.11862.

Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21(2):263-5. doi: 10.1093/bioinformatics/bth457.

Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, Rioux JD, Brant SR, Silverberg MS, Taylor KD, Barmada MM, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nat Genet 2008;40(8):955-62. doi: 10.1038/ng.175.

Brant SR, Panhuysen CIM, Nicolae D, Reddy DM, Bonen DK, Karaliukas R, Zhang L, Swanson E, Datta LW, Moran T, et al. MDR1 Ala893 Polymorphism Is Associated with Inflammatory Bowel Disease. American Journal of Human Genetics 2003;73(6):1282-92. doi: 10.1086/379927.

Blumberg RS. Environment and Genes: What Is the Interaction? Digestive Diseases 2016;34(1-2):20-6.

Buffinton GD, Doe WF. Altered Ascorbic-Acid Status in the Mucosa from Inflammatory Bowel-Disease Patients. Free Radical Res 1995;22(2):131-43. doi: Doi 10.3109/10715769509147535.

Chatterjee M, Saluja R, Kumar V, Jyoti A, Jain GK, Barthwal MK, Dikshit M. Ascorbate sustains neutrophil NOS expression, catalysis, and oxidative burst. Free Radical Bio Med 2008;45(8):1084-93. doi: 10.1016/j.freeradbiomed.2008.06.028.

Chapple ILC, Matthews JB, Wright HJ, Scott AE, Griffiths HR, Grant MM. Ascorbate and alpha-tocopherol differentially modulate reactive oxygen species generation by neutrophils in response to Fc gamma R and TLR agonists. Innate Immun-London 2013;19(2):152-9. doi: 10.1177/1753425912455207.

Cleynen I, Boucher G, Jostins L, Schumm LP, Zeissig S, Ahmad T, Andersen V, Andrews JM, Annese V, Brand S, et al. Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study. The Lancet;387(10014):156-67. doi: http://dx.doi.org/10.1016/S0140-6736(15)00465-1.

Corpe CP, Eck P, Wang J, Al-Hasani H, Levine M. Intestinal Dehydroascorbic acid (DHA) transport mediated by the facilitative sugar transporters, GLUT2 and GLUT8. Journal of Biological Chemistry 2013. doi: 10.1074/jbc.M112.436790.

Ediger JP, Walker JR, Graff L, Lix L, Clara I, Rawsthorne P, Rogala L, Miller N, McPhail C, Deering K, et al. Predictors of medication adherence in inflammatory bowel disease. American Journal of Gastroenterology 2007;102(7):1417-26. doi: 10.1111/j.1572-0241.2007.01212.x.

Fernandezbanares F, Abadlacruz A, Xiol X, Gine JJ, Dolz C, Cabre E, Esteve M, Gonzalezhuix F, Gassull MA. Vitamin Status in Patients with Inflammatory Bowel-Disease. American Journal of Gastroenterology 1989;84(7):744-8.

Ferraris A, Torres B, Knafelz D, Barabino A, Lionetti P, de Angelis GL, Iacono G, Papadatou B, D'Amato G, Di Cionmo V, et al. Relationship between CARD15, SLC22A4/5, and DLG5 polymorphisms and early-onset inflammatory bowel diseases: An Italian multicentric study.

Inflammatory Bowel Diseases 2006;12(5):355-61. doi: Doi 10.1097/01.Mib.0000217338.23065.58.

Green C, Elliott L, Beaudoin C, Bernstein CN. A population-based ecologic study of inflammatory bowel disease: Searching for etiologic clues. American Journal of Epidemiology 2006;164(7):615-23. doi: 10.1093/aje/kwj260.

Gao T, He B, Liu S, Zhu H, Tan K, Qian J. EnhancerAtlas: A resource for enhancer annotation and analysis in 105 human cell/tissue types. Bioinformatics 2016;32(23):3543-51. doi: 10.1093/bioinformatics/btw495.

Hoffenberg EJ, Deutsch J, Smith S, Sokol RJ. Circulating antioxidant concentrations in children with inflammatory bowel disease. American Journal of Clinical Nutrition 1997;65(5):1482-8.

Hong SN, Park C, Park SJ, Lee CK, Ye BD, Kim YS, Lee S, Chae J, Kim JI, Kim YH. Deep resequencing of 131 Crohn's disease associated genes in pooled DNA confirmed three reported variants and identified eight novel variants. Gut 2016;65(5):788-96. doi: 10.1136/gutjnl-2014-308617.

Internet: http://www.cdhf.ca/en/statistics (accessed 08-15-2017)

Ki MR, Lee HR, Park JK, Hong IH, Han SY, You SY, Lee EM, Kim AY, Lee SS, Jeong KS. Ascorbate promotes carbon tetrachloride-induced hepatic injury in senescence marker protein 30-deficient mice by enhancing inflammation. J Nutr Biochem 2011;22(6):535-42. doi: 10.1016/j.jnutbio.2010.04.008.

Leon AS, Shaghaghi MA, Yurkova N, Bernstein CN, El-Gabalawy H, Eck P. Single-nucleotide polymorphisms in SLC22A23 are associated with ulcerative colitis in a Canadian white cohort. American Journal of Clinical Nutrition 2014;100(1):289-94. doi: 10.3945/ajcn.113.080549.

Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, Reuter I, Chekmenev D, Krull M, Hornischer K, et al. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic acids research 2006;34(Database issue).

Nag A, Venturini C, Hysi PG, Arno M, Aldecoa-Otalora Astarloa E, Macgregor S, Hewitt AW, Young TL, Mitchell P, Viswanathan AC, et al. Copy number variation at chromosome 5q21.2 is associated with intraocular pressure. Invest Ophth Vis Sci 2013;54(5):3607-12.

Nakahara S, Arimura Y, Saito K, Goto A, Motoya S, Shinomura Y, Miyamoto A, Imai K. Association of SLC22A4/5 polymorphisms with steroid responsiveness of inflammatory Bowel disease in Japan. Dis Colon Rectum 2008;51(5):598-603. doi: 10.1007/s10350-008-9208-5.

Newman B, Gu X, Wintle R, Cescon D, Yazdanpanah M, Liu X, Peltekova V, Van Oene M, Amos CI, Siminovitch KA. A risk haplotype in the Solute Carrier Family 22A4/22A5 gene cluster influences phenotypic expression of Crohn's disease. Gastroenterology 2005;128(2):260-9.

Ng SC, Bernstein CN, Vatn MH, Lakatos PL, Loftus Jr EV, Tysk C, O'Morain C, Moum B, Colombel JF. Geographical variability and environmental risk factors in inflammatory bowel disease. Gut 2013;62(4):630-49. doi: 10.1136/gutjnl-2012-303661.

Nørgård BM, Larsen PV, Fedder J, De Silva PS, Larsen MD, Friedman S. Live birth and adverse birth outcomes in women with ulcerative colitis and Crohn's disease receiving assisted reproduction: A 20-year nationwide cohort study. Gut 2016;65(5):767-76. doi: 10.1136/gutjnl-2015-311246.

Nualart FJ, Rivas CI, Montecinos VP, Godoy AS, Guaiquil VH, Golde DW, Vera JC. Recycling of vitamin C by a bystander effect. Journal of Biological Chemistry 2003;278(12):10128-33. doi: 10.1074/jbc.M210686200.

Peltekova VD, Wintle RF, Rubin LA, Amos CI, Huang Q, Gu X, Newman B, Van Oene M, Cescon D, Greenberg G, et al. Functional variants of OCTN cation transporter genes are associated with Crohn disease. Nat Genet 2004;36(5):471-5. doi: 10.1038/ng1339.

Rumsey SC, Kwon O, Xu GW, Burant CF, Simpson I, Levine M. Glucose transporter isoforms GLUT1 and GLUT3 transport dehydroascorbic acid. Journal of Biological Chemistry 1997;272(30):18982-9.

Rumsey SC, Levine M. Absorption, transport, and disposition of ascorbic acid in humans. J Nutr Biochem 1998;9(3):116-30.

Rumsey SC, Zarnowski M, Simpson IA, Levine M. DHA transport by GLUT4 in Xenopus oocytes and isolated rat adipocytes. FASEB Journal 1998;12(4).

Rumsey SC, Daruwala R, Al-Hasani H, Zarnowski MJ, Simpson IA, Levine M. Dehydroascorbic acid transport by GLUT4 xenopus oocytes and isolated rat adipocytes. Journal of Biological Chemistry 2000;275(36):28246-53.

Ryan JD, Silverberg MS, Xu W, Graff LA, Targownik LE, Walker JR, Carr R, Clara I, Miller N, Rogala L, et al. Predicting complicated Crohn's disease and surgery: phenotypes, genetics, serology and psychological characteristics of a population-based cohort. Aliment Pharm Ther 2013;38(3):274-83. doi: 10.1111/apt.12368.

Sciences NIoEH. Internet: http://snpinfo.niehs.nih.gov/snpinfo/snptag.htm (accessed January 5 2013).

Shulman JM, Chipendo P, Chibnik LB, Aubin C, Tran D, Keenan BT, Kramer PL, Schneider JA, Bennett DA, Feany MB, et al. Functional Screening of Alzheimer Pathology Genome-wide Association Signals in Drosophila. The American Journal of Human Genetics 2011;88(2):232-8. doi: 10.1016/j.ajhg.2011.01.006.

Silverberg MS, Satsangi J, Ahmad T, Arnott IDR, Bernstein CN, Brant SR, Caprilli R, Colombel JF, Gasche C, Geboes K, et al. Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: Report of a Working Party of the 2005 Montreal

World Congress of Gastroenterology. Canadian Journal of Gastroenterology and Hepatology 2005;19:5a-36a.

Soreq H, Seidman S. Xenopus Oocyte Microinjection - from Gene to Protein. Methods in Enzymology 1992;207:225-65.

Taylor KH, Pena-Hernandez KE, Davis JW, Arthur GL, Duff DJ, Shi H, Rahmatpanah FB, Sjahputera O, Caldwell CW. Large-Scale CpG Methylation Analysis Identifies Novel Candidate Genes and Reveals Methylation Hotspots in Acute Lymphoblastic Leukemia. Cancer Research 2007;67(6):2617-25. doi: 10.1158/0008-5472.can-06-3993.

Torres J, Colombel J-F. Genetics and phenotypes in inflammatory bowel disease. The Lancet;387(10014):98-100. doi: http://dx.doi.org/10.1016/S0140-6736(15)00464-X.

Padayatty SJ, Katz A, Wang Y, Eck P, Kwon O, Lee JH, Chen S, Corpe C, Dutta A, Dutta SK, et al. Vitamin C as an antioxidant: Evaluation of its role in disease prevention. Journal of the American College of Nutrition 2003;22(1):18-35.

Urcelay E, Mendoza JL, Martín MC, Mas A, Martínez A, Taxonera C, Fernandez-Arquero M, Díaz-Rubio M, De La Concha EG. MDR1 Gene: Susceptibility in Spanish Crohn's disease and ulcerative colitis patients. Inflammatory Bowel Diseases 2006;12(1):33-7. doi: 10.1097/01.MIB.0000194184.92671.78.

Veal CD, Reekie KE, Lorentzen JC, Gregersen PK, Padyukov L, Brookes AJ. A 129-kb Deletion on Chromosome 12 Confers Substantial Protection Against Rheumatoid Arthritis, Implicating the Gene SLC2A3. Hum Mutat 2014;35(2):248-56. doi: 10.1002/humu.22471.

Wang W, Yu JT, Zhang W, Cui WZ, Wu ZC, Zhang Q, Tan L. Genetic Association of SLC2A14 Polymorphism with Alzheimer's Disease in a Han Chinese Population. Journal of Molecular Neuroscience 2012;47(3):481-4. doi: DOI 10.1007/s12031-012-9748-y.

Washko PW, Hartzell WO, Levine M. Ascorbic acid analysis using high-performance liquid chromatography with coulometric electrochemical detection. Anal Biochem 1989;181(2):276-82.

Washko PW, Wang Y, Levine M. Ascorbic acid recycling in human neutrophils. Journal of Biological Chemistry 1993;268(21):15531-5.

Wendland BE, Aghdassi E, Tam C, Carrrier J, Steinhart AH, Wolman SL, Baron D, Allard JP. Lipid peroxidation and plasma antioxidant micronutrients in Crohn disease. American Journal of Clinical Nutrition 2001;74(2):259-64.

Wu XH, Freeze HH. GLUT14, a duplicon of GLUT3, is specifically expressed in testis as alternative splice forms. Genomics 2002;80(6):553-7. doi: DOI 10.1006/geno.2002.7010.

2.9 SUPPLEMENTARY INFORMATION

SNP	SNP Type	Position (Chromosome 12)	Alleles Major/minor	Catalog ID[1]
rs11612319	Intron	8016925 bp	G/A	ID: C___2875382_10
rs7132415	Intron	7997547 bp	G/T	ID: C__29025049_10
rs10846086	Intron	8025266 bp	A/G	ID: C__12095733_10
rs2376904	Intron	8000995 bp	G/A	ID: C__26191692_10
rs7300773	Intron	8023943 bp	T/C	ID: C__29292592_20
rs2889504	Intron	8008132 bp	G/T	ID: C___2875393_10
rs12815313	Exon	8023530 bp	C/T	ID: C__31787372_10
rs10845990	Intron	7970721 bp	T/G	ID: C__30894136_10

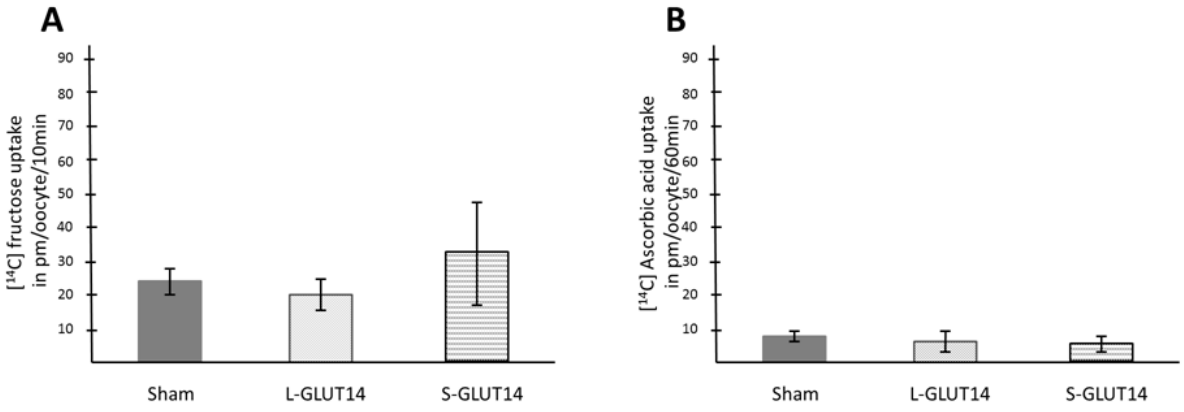
Supplementary Information table 2.9.1: Summary of 8 tagging SNPs in GLUT 14 gene genotyped in this study

Abbreviations: SNP: single nucleotide polymorphisms; bp: base pair

[1] TaqMan® SNP Genotyping Assay (Life Technologies)

Parameters	Ulcerative Colitis (n= 149)	Crohn´s Disease (n=162)	Controls (n=142)
Sex			
Female	87 (58.4%)	97 (59.9%)	80 (56.3%)
Male	62 (41.6%)	65 (40.1%)	62 (43.7%)
Age at Diagnosis			
A1(<17 yr)	12 (8.1%)	17 (10.5%)	-
A2 (17-40 yr)	78 (52.3%)	101 (62.3%)	-
A3 (>40 yr)	59 (39.6%)	44 (27.2%)	-
Disease Location			
Terminal ileum	-	69 (42.6%)	-
Colon	-	37 (22.8%)	-
Ileocolon	-	51 (31.5%)	-
Proximal GI tract	-	5 (3.1%)	-
Proctitis	11 (7.4%)	-	-
Left-sided colitis	68 (45.6%)	-	-
Pancolitis	70 (47.0%)	-	-
Disease Type			
Inflammatory	-	69 (42.6%)	-
Fibrostenotic	-	54 (33.3%)	-
Penetrating/Fistulizing	-	39 (24.1%)	-

Supplementary Information Table 2.9.2: General characteristics of the study participants



Supplementary Information Figure 2.9.1 GLUT 14 isoforms do not mediate ascorbic acid (DHA) and fructose uptake. *Xenopus laevis* oocytes expressing the short and long GLUT 14 isoforms (S-GLUT14 and L-GLUT14) do not exhibit uptake of radiolabeled fructose (A) and ascorbic acid (B). Incubations were performed with 300μM [^{14}C]Ascorbic acid or 300μM [^{14}C]Fructose on 20 oocytes each group, in three independent experiments. Data are expressed as averages ± SD. Small letters would indicate statistical significance determined through the student's t-test, $p<0.05$; however, no statistic differences were established.

TRANSITIONAL STATEMENT 1

In contrast to GLUT14 where the genomic locus, tissue expression splice variants and subcellular localization were previously determined, there was only one publication reporting the subcloning of mouse Mfsd14a and its sequence analysis. Since no further information was available on the human MFSD14A gene or protein, we have conducted bioinformatic analysis and analysed the gene locus, protein structures, and tissue expression profiles for both human MFSD14A and mouse Mfsd14a. Additionally, we tested our hypothesis where MFSD14A is a novel glucose transporter with the established and optimized *Xenopus laevis* heterologous expression system. We conducted functional studies on MFSD14A by expressing the human MFSD14A and mouse Mfsd14a in oocytes and investigating the uptake of radiolabeled glucose. The results are presented in the following two chapters.

CHAPTER 3:

MANUSCRIPT 2

This manuscript was published in the Journal of Advances in Nutrition and Food Sciences

(Advances in Nutrition and Food Science, Volume 2022, Issue 01, DOI:

10.37722/ANAFS.2022101)

THE MURINE MAJOR FACILITATOR SUPERFAMILY DOMAIN CONTAINING 14A (MFSD14A) GENE DOES NOT ENCODE A GLUCOSE TRANSPORTER

Haonan Zhouyao[1], Sandra Fehsenfeld[2], Dirk Weihrauch[3], and Peter Eck*[1]

Authors Affiliations:

[1]Department of Food and Human Nutritional Sciences, University of Manitoba, Winnipeg, Canada.

[2]Université du Québec à Rimouski, Département de biologie, chimie et géographie, 300 Allée des Ursulines, Rimouski, QC G5L 3A1, Canada.

[3]University of Manitoba, Department of Biological Sciences, 190 Dysart Road, Winnipeg, MB R3T 2N2, Canada

*Corresponding Author

My contribution:

I conducted the literature search, the glucose uptake studies and wrote the manuscript.

Other author's contribution:

PE was the principal investigator, responsible for the conception and design of the project, the acquisition of financial support, and the writing and final editing of the manuscript.

SF cloned the mouse Mfsd14a and provided the plasmid for the heterologous expression studies.

DW helped editing the manuscript.

3.1. ABSTRACT

Despite the fact that the Mfsd14a knock-out mouse shows a phenotype resembling human globozoospermia, the literature record on the mouse mfsd14a gene and its encoded products is extremely limited. This report therefore aims to establish the baseline on the mouse genomic locus.

Utilizing visually supervised bioinformatics of the Mfsd14a, the genomics locus and its encoded products are annotated. Our analysis revealed a single protein isoform with all the hallmarks of a functional membrane bound solute carrier. The gene's sole transcript is almost ubiquitously expressed with highest expression in reproductive tissues followed by nervous tissues. Despite previous suggestions, the murine mfsd14a protein does not mediate transmembrane glucose transport when expressed in *Xenopus laevis* oocytes.

3.2. INTRODUCTION

The public literature record on the mouse mfsd14a gene, its encoded protein products and function(s) is virtually nonexistent. Only six credible publications can be identified in PubMed or in Scopus referring to murine mfsd14a, none of them identifying the encoded product and its function (Doran et al., 2016; Lekholm et al., 2017; Luco et al., 2012; Matsuo et al., 1997; S. Sreedharan et al., 2011; Wang et al., 2017).

Two publications provide very limited information about the gene's transcript in murine species. Specifically, Matsu et al. (1997) cloned and sequenced the gene transcript (cDNA) from neonatal mouse hippocampus and named it "hiat1 (hippocampus abundant transcript 1)". The authors predicted that the encoded product is a transmembrane protein. Based on sequence similarities of the predicted protein and the presence of "signature motifs" (**Supplementary Information Figure 1**), the authors further extrapolated that the mouse mfsd14a might encode for a novel sugar transporter (Matsuo et al., 1997), however, no functional data were provided.

Matsu et al. (1997) showed gene expression of the murine mfsd14a in almost all mouse tissues they investigated, which was mirrored by Sreedharan et al. (2011) for tissues examined in rat (Smitha Sreedharan et al., 2011). However, both studies did not provide details on standardized expression levels and only examined a limited spectrum of tissues.

Altered spermatogenesis and male infertility is observed in mfsd14a$^{-/-}$ mice (Doran et al., 2016), where homozygous mutant mice were viable and healthy, but males were sterile due to a 100-fold reduction in the number of spermatozoa in the vas deferens. The few spermatozoa that were formed showed rounded head defects like those found in humans with globozoospermia. The expression of mfsd14a in Sertoli cells ("nurse cells") led Doran and coworkers to suggest that mfsd14a may transport a nutrient or metabolite from the bloodstream into the cells that is required for spermatogenesis.

Lekholm and coworkers described mfsd14a gene expression in the mouse central nervous system, and the location of the protein in the Golgi system and endoplasmatic reticulum (Lekholm et al., 2017). Based on the differential regulation data, the authors also predicted an organic substrate with a possible involvement in energy homeostasis, such as glucose. This was deduced because an upregulation was reported after 3 hours of amino acid starvation in mouse embryonic primary cultures. By contrast, in mice subjected to 24 hours of starvation, a downregulation of mfsd14a was shown in the hypothalamus and the brainstem. Moreover, a high fat diet caused upregulation in the striatum (Lekholm et al., 2017).

The emerging disease associations precipitate the need for a detailed annotation of the mouse mfsd14a locus and a characterization of its predicted function as a putative glucose transporter. Based on visually supervised analysis of available data we annotate mouse mfsd14a and its encoded products. Its functions as a solute carrier for glucose is investigated in the heterologous *Xenopus laevis* oocytes expression system.

3.3 METHODS

3.3.1 Bioinformatics genomics analysis

The National Center for Biotechnology Information (NCBI) database was used to obtain the genomic sequence (Chromosome 3 – NC_000069.7), reference, nonreference full length RNA, and the expressed sequence tags (ESTs) sequences for mouse mfsd14a. The DNA sequence analysis software Sequencher 4.8 (GeneCodes Corporation; Ann Arbor, Michigan, USA) was used to create alignments between the reference RNA sequence and the genomic sequence to determine the structures of exons and introns. The alignment parameters used were large gaps, 20 minimum overlap, and minimum match of 80%. These alignments were visually curated to determine exons and introns and to create the scaled illustrations.

3.3.2 Protein analysis

Protein Homology/analogY Recognition Engine V 2.0 (Phyre2, (Kelley et al., 2015)) and the AlphaFold Model using 3D ligand site (Jumper et al., 2021; Wass et al., 2010) were used to analyse the function and structure of the mfsd14a protein. During the analysis, parameters from the standard/normal modes were applied.

3.3.3 Expression patterns analysis

Experimental RNA-Seq data were interrogated at the following resources: Gene Expression Database (GXD), Mouse Genome Informatics Web Site; World Wide Web (URL: http://www.informatics.jax.org/expression.shtml) (January 23, 2022); European Molecular Biology Laboratory's European Bioinformatics Institute (EMBL-EBI) Baseline Expression Atlas (URL: https://www.ebi.ac.uk/gxa/home) (Moreno et al., 2022) and The Encyclopedia of DNA Elements (ENCODE) (URL: https://www.encodeproject.org, (Consortium, 2012)).

3.3.4 Plasmid preparation

Mouse Mfsd14a: Initially, the complete open reading frame (ORF; 1701 bp) for the mouse mfsd14a (GenBank accession no. NM_008246.2) was amplified from mouse brain cDNA with Phusion high fidelity DNA polymerase (New England Biolabs, Ipswich, Massachusetts, USA) and subcloned into the p426 expression vector for amplification in yeast applying sticky-end ligation with T4 ligase (New England Biolabs, Ipswich, Massachusetts, USA) using the restriction enzymes SpeI and SmaI (please see **Table 3.8.1** in **Supplementary Information** for sequence details). Subsequently, the ORF was sent for sequencing at the DNA Sequencing Facility of the Robarts Research Institute (London, Ontario, Canada) using T7/SP6 primers. For functional expression in *Xenopus laevis* oocytes, the verified mfsd14a ORF was then subcloned from the p426 vector into the pGEM®-HE plasmid, a modified pGEM®-3Z vector for the cloning site to be flanked by *Xenopus* beta globin 5'- and 3'-UTR sequence, using T4 ligase as described above. Both primers for this subcloning contained a restriction site for SmaI (forward primer including T7/SmaI-restriction site/ATG and reverse primer containing SmaI-restriction site). The insertion and the correct orientation of the ORF was verified by PCR (T7 plus mfsd14a internal reverse primer) and subsequent sequencing with T7/SP6 primers (supplied by the Centre for Applied Genoics (TCAG) sequencing facility, Toronto, Canada).

Human GLUT3: Plasmid containing the full ORF of human GLUT3 (positive control for glucose transport study) was purchased from Harvard PlasmID (clone ID: HsCD00021270) and transformed into DH5α competent cells (New England Biolabs, Ipswich, Massachusetts, USA) following the manufactures' guidance for propagation, followed by column purification (QIAprep Spin Miniprep Kit, Qiagen, Hilden, Germany). The ORFs were then amplified with Q5 High Fidelity DNA Polymerase (New England Biolabs) and tagged with T7 RNA polymerase promoter sequence. The GLUT3 amplicon containing GLUT3 ORF and T7 promoter sequence was used directly for cRNA synthesis as described previously. cRNA products were column purified (RNeasy MinElute

Cleanup Kit, Qiagen) prior to the spectrophotometric quantification (Nanodrop, ND-1000, Thermofisher, Waltham, MA, USA) and the integrity assessments (MOPS agarose gel containing formaldehyde).

3.3.5 Functional expression of Mfsd14a in *Xenopus laevis* oocytes

3.3.5.1 Chemicals and reagents

[^{3}H] 2-deoxyglucose (25.5 Ci/mmol in water) was purchased from PerkinElmer (Waltham, MA, USA). All chemicals were purchased from Sigma-Aldrich (St. Louis, MI, USA) unless otherwise noted. The standard oocyte ringer (OR2) contained (in mmol L^{-1}) 82.5 NaCl, 2.5 KCl, 1 $CaCl_2$, 1 $MgCl_2$, 1 Na_2HPO_4, 5 HEPES, pH 7.4 and sterilized by vacuum bottle-top filters (EMD MilliporeTM Steritop^{TM} sterile vacuum bottle-top filters, ThermoFisher, Waltham, Massachusetts, USA). After sterilization, OR2 was then supplemented with sodium pyruvate (2.5 mmol L^{-1}), penicillin-streptomycin (1 mg mL^{-1}) (Gibco, Long Island, NY, USA), and gentamicin (50 μg mL^{-1}) for long term storage of sorted oocytes.

3.3.5.2 Plasmid preparation

The pGEM®-HE plasmid containing the ORF for mfsd14a as described above, was transferred into DH5-alpha cells for amplification, column purified (Plasmid miniprep kit, Qiagen, Hilden, Germany) and linearized with SphI (New England Biolabs (NEB), Ipswich, Massachusetts, United States). The HiScribe™ T7 ARCA mRNA Kit (with tailing) (NEB) was used for the in-vitro transcription of the capped mRNA (cRNA) following the manufacturer's suggestion, column purified (RNeasy MinElute Cleanup Kit, Qiagen, Hilden, Germany) and eluted in nuclease free water.

3.3.5.3 Oocyte preparation

Stage VI-V oocytes were collected from mature female *Xenopus laevis* (VWR International, Randor, PA, USA) as previously described (Soreq, 1992). Briefly, the frogs were anesthetized with MS-222 (2 g L^{-1}) followed with decapitation prior to the collection of the ovary. The ovary was placed in Ca^{2+}-free OR2 solution containing collagenase type VI (1 mg mL^{-1}) (Gibco, Waltham, Massachusetts, USA) while gently agitated for 90 minutes at room temperature. The activity of collagenase was terminated by rinsing the oocytes three times with standard OR2. Oocytes were then sorted manually, rinsed three additional times with standard OR2 and allowed to recover in sterile OR2 overnight at 16 °C. All procedures used were approved by the University of Manitoba Animal Research Ethics Board and are in accordance with the Guidelines of the Canadian Council on Animal Care.

3.3.5.4 Microinjection of oocytes

After recovering overnight, 18.4 ng (36.8 nL with 0.5 ng nL^{-1}) of mfsd14a cRNA, human glucose transporter 3 (GLUT3) cRNA (as positive control), or nuclease-free water (as negative control, sham) was delivered into the oocytes using the Nanoject II auto-nanoliter injector (Drummond Scientific, Broomall, Pennsylvania, USA). The glucose uptake experiments were done at room temperature two days after the injections, where oocytes were randomly divided into groups containing 20 oocytes each either injected with water (sham) or cRNA (mfsd14a or GLUT3). Oocytes were incubated in 200 µL standard OR2 with 125 pmol [^{3}H] 2-deoxyglucose for 30 minutes. Excess ice-cold standard OR2 was used terminate the transport, and was used to rinse oocytes four times to remove the external radioactivity. Oocytes were then lysed individually in 200 µL of 10% sodium dodecyl sulfate (SDS) after which 5 mL Ultima Gold scintillation cocktail (PerkinElmer) was added.

Internal radioactivity was quantified by liquid scintillation spectrometry (Tri-Carb 2900 TR; PerkinElmer) as counts per minute (CPM)/oocyte.

The washing efficiency was ensured by assessing 50 µL of the radioactivity of the buffer collected from the fourth washing post radioactive exposure compared to fresh, radioactivity-free standard OR2 at the same volume. If excessive amount of radioactivity was found, the oocytes were discarded.

3.3.6 Statistical analysis

Data were tested for a normal distribution and homogeneity of variances with the Bartlett's test and an one way ANOVA test with Tukey's post hoc analysis was used to compare multiple means in the glucose uptake studies. All results with $P < 0.05$ were considered significant. Error bars indicate standard error of the mean (SEM) and different letters indicates p values less than 0.05. Graphs were generated using the software GraphPad/Prism 9 for Mac OS, GraphPad Software, San Diego, California USA, www.graphpad.com).

3.4 RESULTS

3.4.1 The mouse *Mfsd14a* gene locus

The mouse *mfsd14a* gene is listed in NCBI as gene ID 15247 and resides on the mouse chromosome 3qG1, reverse strand. It is immediately flanked by the Spindle assembly abnormal protein 6 homolog (sass6) and the mir3473h and solute carrier family 35 member a3 (slc35a3) genes (**Figure 1**).

The visually supervised alignment of transcripts shows that the mouse mfsd14a gene spans 31,496 bases from its transcription initiation to the termination site. The transcript starts 11 nucleotides upstream of the site currently identified for the NCBI Reference Transcript NM_008246 (**supplementary information, Figure 3.8.2**). In the terminal 3' exon, two potential transcriptional termination sites are visualized (**Supplementary information, Figure 3.8.3**).

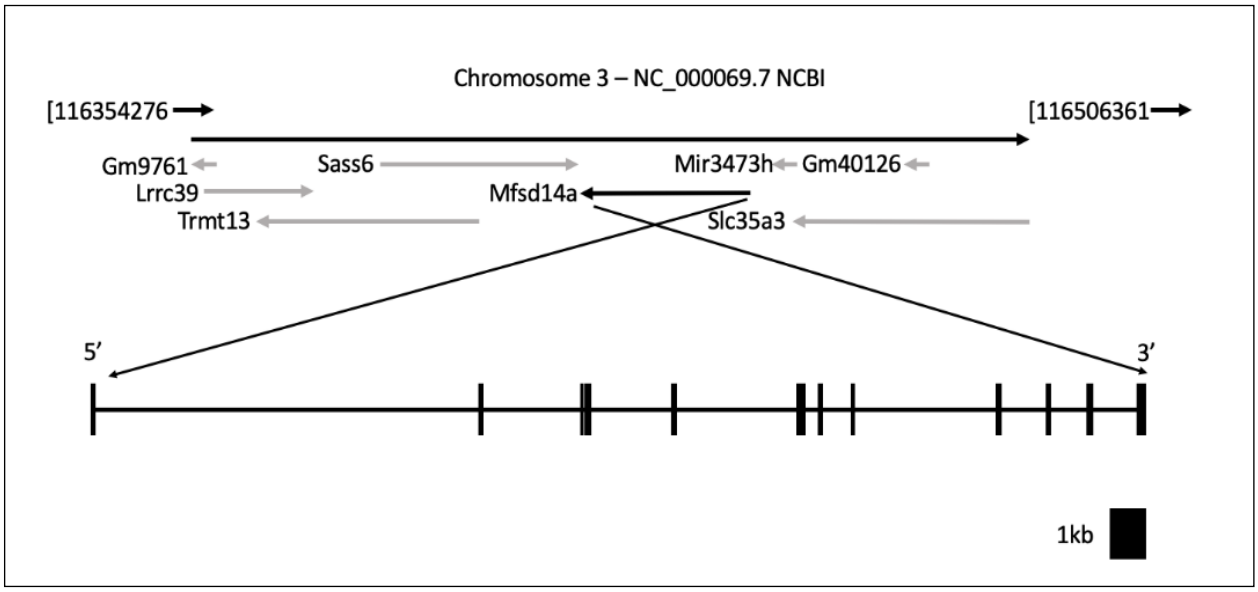


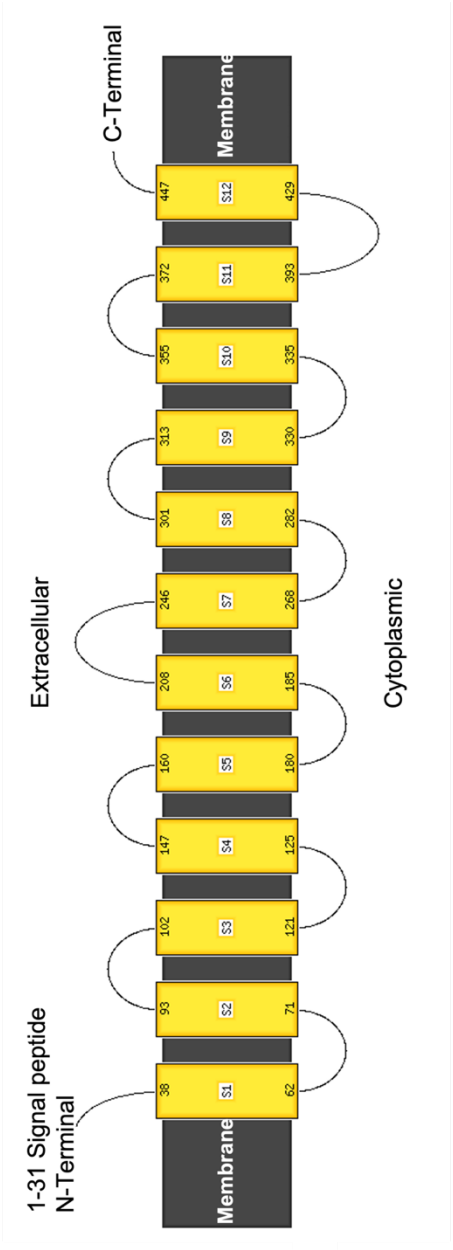
Figure 3.1 The mouse *Mfsd14a* genomic locus. The location of the transcript on Chromosome 3 is indicated by the black arrow in the top panel. The exon-intron structure of the sole transcript is depicted on the bottom panel.

3.4.2 The mouse Mfsd14a protein

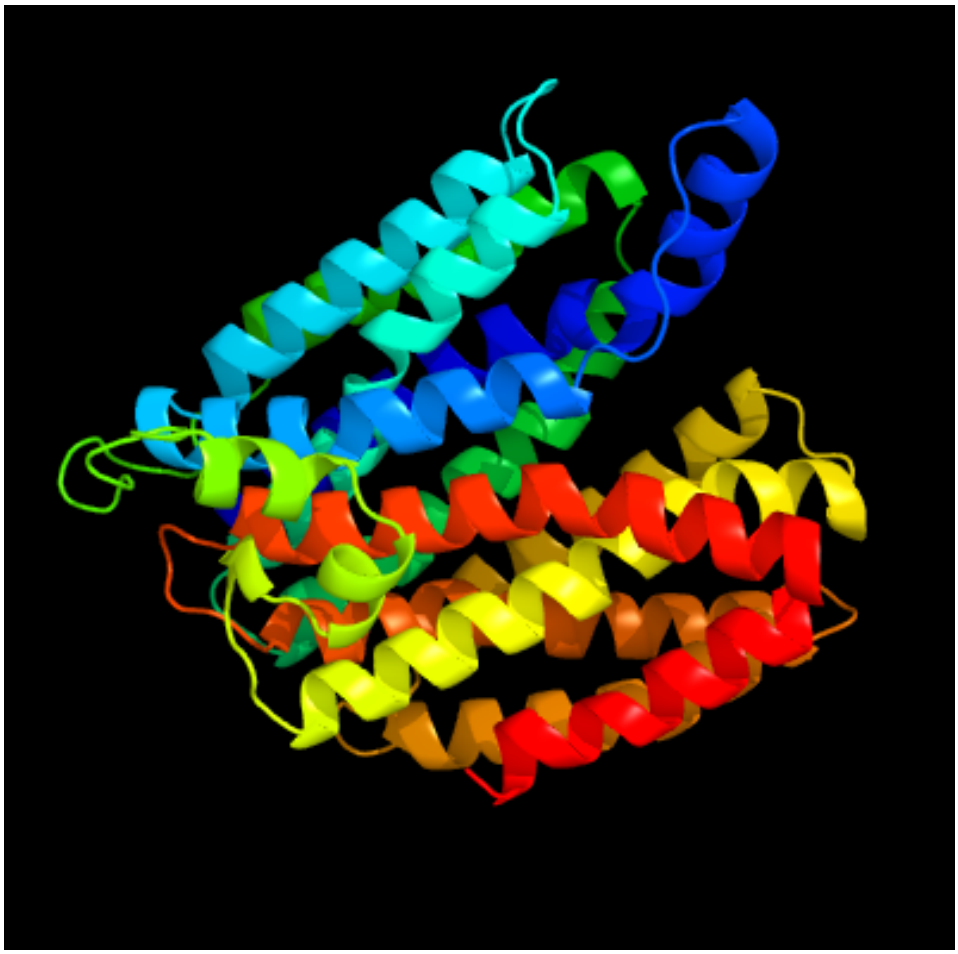
The mouse Mfsd14a protein consists of 490 amino acids, as represented by NCBI reference protein NP_032272.2 (**Supplementary information, Sequence 3.8.1**). The protein is predicted to be a solute membrane carrier with twelve transmembrane domains (**Figure 3.2A & B** and **Supplementary information, Figure 3.8.4**).

The Mfsd14a protein clusters to the conserved protein domain family MFS, subfamily MFSD14, which contains only one additional member, the orthologous mfsd14b (**Supplementary Information Figure 3.8.5**). A unique putative chemical substrate binding pocket is predicted. As putative organic substrates, glucose, xylose, maltose, alanine, chloramphenicol, and (3alpha,5beta,12alpha)- 3,12-dihydroxy Cholan-24-oic acid are predicted to bind in a pore-like intramolecular binding pocket (**Figure 3.2 C**). This corroborates the speculations that glucose might be a main substrate.

A)


C-Terminal
Membrane
447
S12
429
372
S11
393
355
S10
335
313
S9
330
301
S8
282
246
S7
268
208
S6
185
160
S5
180
147
S4
125
102
S3
121
93
S2
71
38
S1
62
Membrane
1-31 Signal peptide
N-Terminal
Extracellular
Cytoplasmic

B)



C)

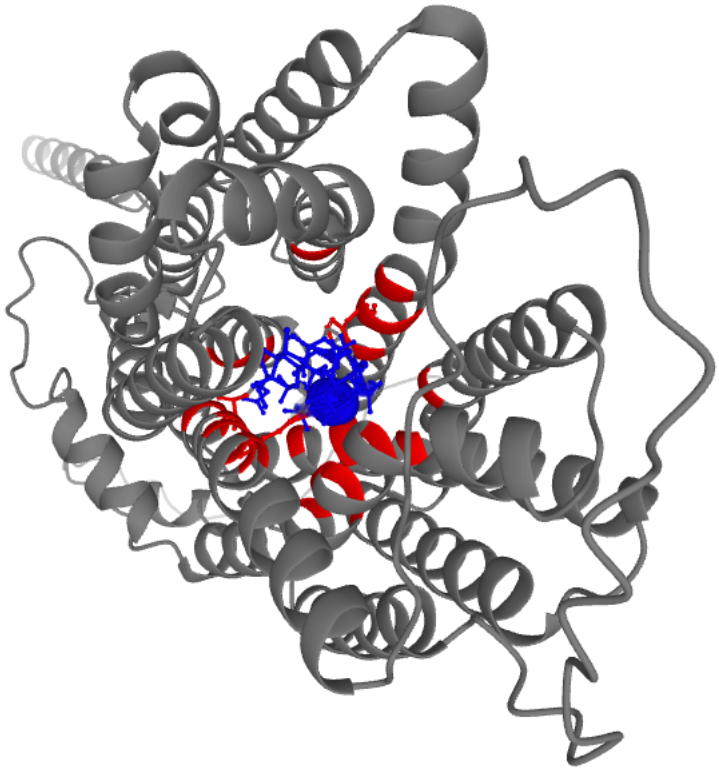


Figure 3.2 The predicted secondary (panel A) and tertiary (panel B & C) structures for the mouse Mfsd14a protein (Image colored by rainbow N→C terminus; Model dimensions (Å):X:58.386 Y:50.944 Z:54.127). Models created using the Protein Homology/Analogy Recognition Engine V 2.0 (Phyre2) web-server (Kelley et al., 2015). Panel C depicts the putative ligand binding pocked aligned in a "pore-like" structure created as an AlphaFold Model using the 3D ligand site (Wass et al., 2010). The ligands are coloured in blue and the binding residues are coloured in red; the remaining protein structure is coloured in grey.

3.4.3 Tissue expression of the mouse *Mfsd14a* transcript

Mfsd14a tissue expression is almost ubiquitous, with highest expression in the testis, oocyte, corpus callosum, spermatid, thymus, mesenteric adipose tissue, brainstem, ovary (**Figure 3** and **Supplementary information, Figures 3.8.6** and **3.8.7**). Depending on the project analysed, almost all interrogated tissues showed some degree of expression (**Figure 3** and **Supplementary information, Figures 3.8.6** and **3.8.7**). Lowest expression was observed in muscle tissues.

High expression in the mouse central nervous system was reported. A more detailed interrogation shows a very distinct Mfsd14a's spatial distribution, located to single cell structures in the Main Olfactory Bulb inner plexiform layer, the pyramidal cell layer of the Hippocampal Formation, and the cerebellum ganglionic layer or Purkinje cell layer (**Supplementary information, Figure 3.8.8**).

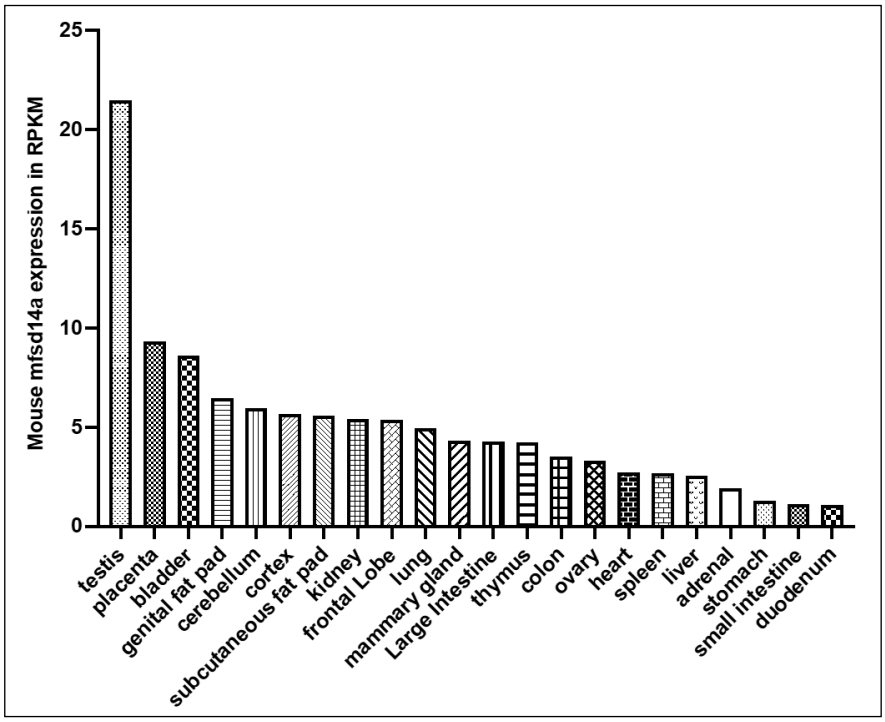


Figure 3.3 Mouse *Mfsd14a* expression data (displayed as Reads Per Kilobase Million, RPKM) from the ENCODE project. Data were obtained via NCBI gene page, Gene ID: 15247.

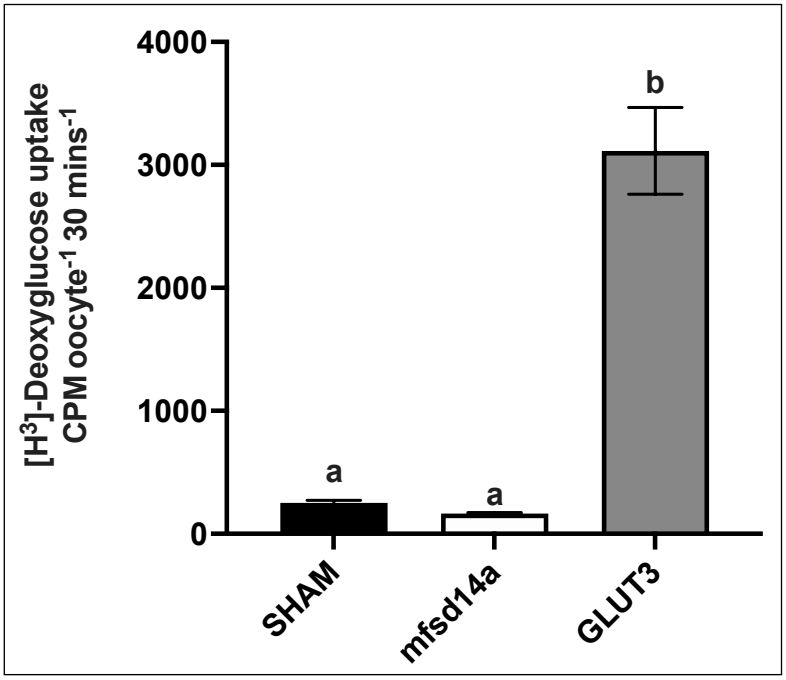


Figure 3.4 The uptake of radiolabeled 2-deoxyglucose into *Xenopus laevis* oocytes expressing a known glucose carrier, mouse Mfsd14a, or human GLUT3 (clone ID: HsCD00021270) N=15-20 oocytes per group. Error bars indicate standard error of the mean (SEM) and different letters indicates p values less than 0.05.

3.5 DISCUSSION

The *mfsd14a* gene is currently underreported, specifically with respect to the mfsd14a$^{-/-}$ mouse phenotype resembling human globozoospermia. This report establishes the baseline annotation for the mouse gene and encoded products. Significantly, we identified one sole transcript to encode for one protein isoform, which is predicted to be a novel glucose membrane transporter both in the literature and with current bioinformatics tools.

However, the mouse mfsd14a did not mediate uptake of glucose, the canonical substrate of many members of the *Slc2a* (Glut) and *Slc5a* (Sglt) family of transporters (Deng & Yan, 2016; Holman, 2020). The lack of glucose transport does not necessarily rule out transport of alternative saccharoses, but should emphasise the need to look beyond sugars as substrates. Supporting the potential for alternative substrates is the fact that a decrease in mRNA abundance of the mfsd14a paralogue in the gills of the European green crabs, *Carcinus maenas*, had been observed in response to elevated environmental *p*CO2 (Fehsenfeld et al., 2011). This observation might provide an avenue to explore further substrate candidates involved in acid-base homeostasis and regulation.

About 95% of human multi-exonic genes are alternatively spliced and there are in excess of 100,000 alternative splicing events in major human tissues (Pan et al., 2008). Changes in splicing are known to affect the function and regulation of genes (Mazin et al., 2018), but splicing aberrations negatively correlate with fitness costs (Saudemont et al., 2017). Therefore, with no identified splice variants, the mouse mfsd14a gene can be considered to be under significant selective pressure resulting in high conservation in regard to splicing, but also in regard to nucleotide and amino acid conservation (Luo et al., 2015). The high conservation further corroborates that mfsd14a might have an essential nutrient/metabolite as a substrate, or to be a key protein in detoxifications or acid-base balance.

Mfsd14a's widespread tissue expression also indicates a role in pathways critical for cell viability. However, the only deleterious phenotype reported for the Mfsd14a$^{-/-}$ mouse is male

infertility. This might be owed to the fact that the mfsd14a$^{-/-}$ mouse had not been fully phenotyped and had not been subjected to metabolic challenges reported to result in differential regulation in the brain (Lekholm et al. 2017). Further exploration of the mfsd14a$^{-/-}$ mouse model might reveal phenotypes related to brain functioning, since expression levels in brain tissues is comparable to testis. However, considering Mfsd14a's distinct spatial expression (**Supplementary information, Figure 3.8.8**), existing reports on differential regulation in mouse brain under nutrient deprivation and in high fat diets, could be viewed with caution, since tissue isolation might have contributed to the observed differential regulation. Those experimental problems could be resolved utilizing single cell RNA sequencing technology in future experiments.

3.6 CONCLUSION

The mouse *mfsd14a* gene is annotated to express a single transcript encoding one protein and is almost ubiquitously expressed. It does not mediate glucose uptake into *Xenopus laevis* oocytes. The genomic and functional baseline to explore disease phenotypes is established.

3.7 REFERENCES

Consortium E P (2012). An integrated encyclopedia of DNA elements in the human genome. Nature, 489:57-74.

Deng D & Yan N (2016). GLUT, SGLT, and SWEET: Structural and mechanistic investigations of the glucose transporters. Protein Sci, 25:546-558.

Doran J, Walters C, Kyle V, Wooding P, Hammett-Burke R, et al. (2016). Mfsd14a (Hiat1) gene disruption causes globozoospermia and infertility in male mice. Reproduction, 152:91-99.

Fehsenfeld S, Kiko R, Appelhans Y, Towle D W, Zimmer M, et al. (2011). Effects of elevated seawater pCO (2) on gene expression patterns in the gills of the green crab, Carcinus maenas. BMC Genomics, 12, 488.

Holman G D (2020). Structure, function and regulation of mammalian glucose transporters of the SLC2 family. Pflugers Arch, 472:1155-1175.

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596:583-589.

Kall L, Krogh A, Sonnhammer E L (2004). A combined transmembrane topology and signal peptide prediction method. J Mol Biol, 338:1027-1036.

Kelley L A, Mezulis S, Yates C M, Wass M N, Sternberg M J (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc, 10:845-858.

Krogh A, Larsson B, von Heijne G, Sonnhammer E L (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol, 305:567-580.

Lekholm E, Perland E, Eriksson M M, Hellsten S V, Lindberg F A, et al. (2017). Putative membrane-bound transporters MFSD14A and MFSD14B are neuronal and affected by nutrient availability. Frontiers in Molecular Neuroscience, 10.

Luco S, Delmas O, Vidalain P O, Tangy F, Weil R, et al. (2012). RelAp43, a Member of the NF-ϰB Family Involved in Innate Immune Response against Lyssavirus Infection. PLoS Pathogens, 8(12), Article e1003060.

Luo H, Gao F & Lin Y (2015). Evolutionary conservation analysis between the essential and nonessential genes in bacterial genomes. Sci Rep, 5, 13210.

Matsuo N, Kawamoto S, Matsubara K, Okubo K (1997). Cloning of a cDNA encoding a novel sugar transporter expressed in the neonatal mouse hippocampus. Biochemical and Biophysical Research Communications, 238:126-129.

Mazin P V, Jiang X, Fu N, Han D, Guo M, et al. (2018). Conservation, evolution, and regulation of splicing during prefrontal cortex development in humans, chimpanzees, and macaques. RNA, 24:585-596.

Moreno P, Fexova S, George N, Manning J R, Miao Z, et al. (2022). Expression Atlas update: gene and protein expression in multiple species. Nucleic Acids Res, 50:D129-D140.

Pan Q, Shai O, Lee L J, Frey B J, Blencowe B J (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet, 40:1413-1415.

Peters C, Tsirigos K D, Shu N, Elofsson A (2016). Improved topology prediction using the terminal hydrophobic helices rule. Bioinformatics, 32:1158-1162.

Reynolds S M, Kall L, Riffle M E, Bilmes J A, Noble W S (2008). Transmembrane topology and signal peptide prediction using dynamic bayesian networks. PLoS Comput Biol, 4:e1000213.

Saudemont B, Popa A, Parmley J L, Rocher V, Blugeon C, et al. (2017). The fitness cost of mis-splicing is the main determinant of alternative splicing patterns. Genome Biol, 18:208.

Soreq H, Seidman S (1992). Xenopus oocyte microinjection: from gene to protein. Methods in Enzymology, 207, 225-265.

Sreedharan S, Stephansson O, Schioth H B, Fredriksson R (2011). Long evolutionary conservation and considerable tissue specificity of several atypical solute carrier transporters. Gene, 478:11-18.

Viklund H, Elofsson A (2004). Best alpha-helical transmembrane protein topology predictions are achieved using hidden Markov models and evolutionary information. Protein Sci, 13:1908-1917.

Viklund H, Elofsson A (2008). OCTOPUS: improving topology prediction by two-track ANN-based preference scores and an extended topological grammar. Bioinformatics, 24:1662-1668.

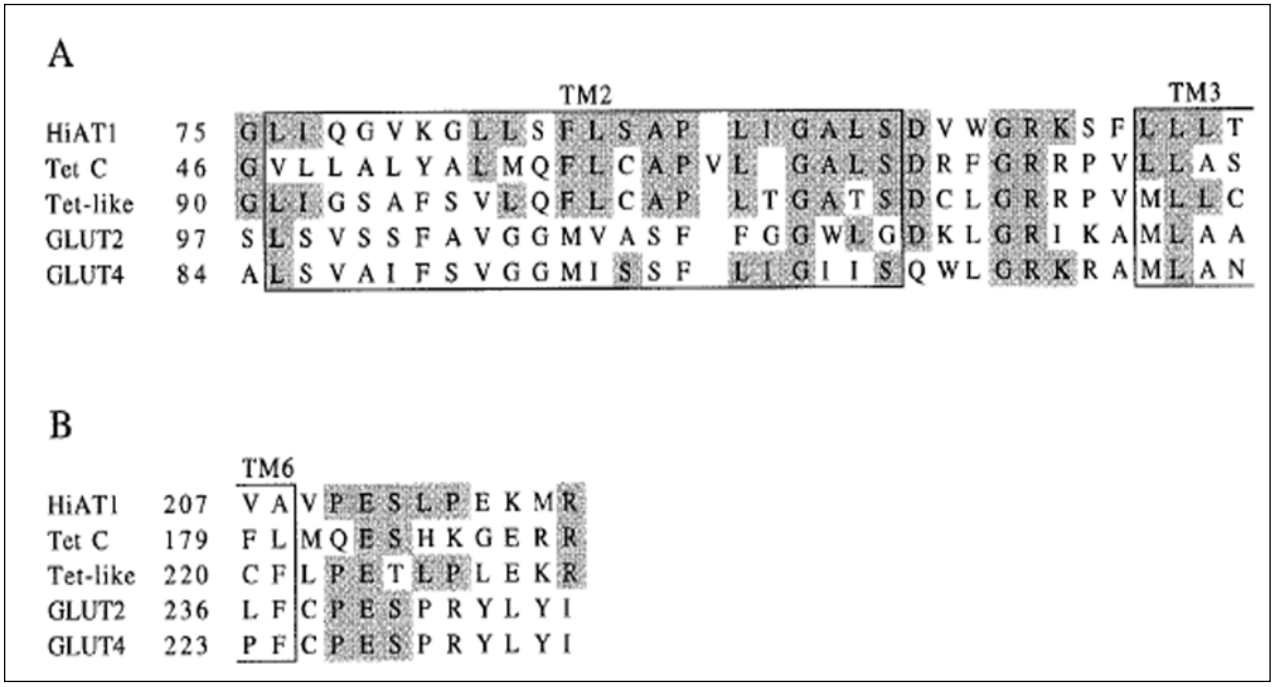
Wang K, Sun Y, Tao W, Fei X, Chang C (2017). Androgen receptor (AR) promotes clear cell renal cell carcinoma (ccRCC) migration and invasion via altering the circHIAT1/miR-195-5p/29a-3p/29c-3p/CDC42 signals. Cancer Lett. 394:1-12.

Wass M N, Kelley L A, Sternberg M J (2010). 3DLigandSite: predicting ligand-binding sites using similar structures. Nucleic Acids Res, 38:W469-473.

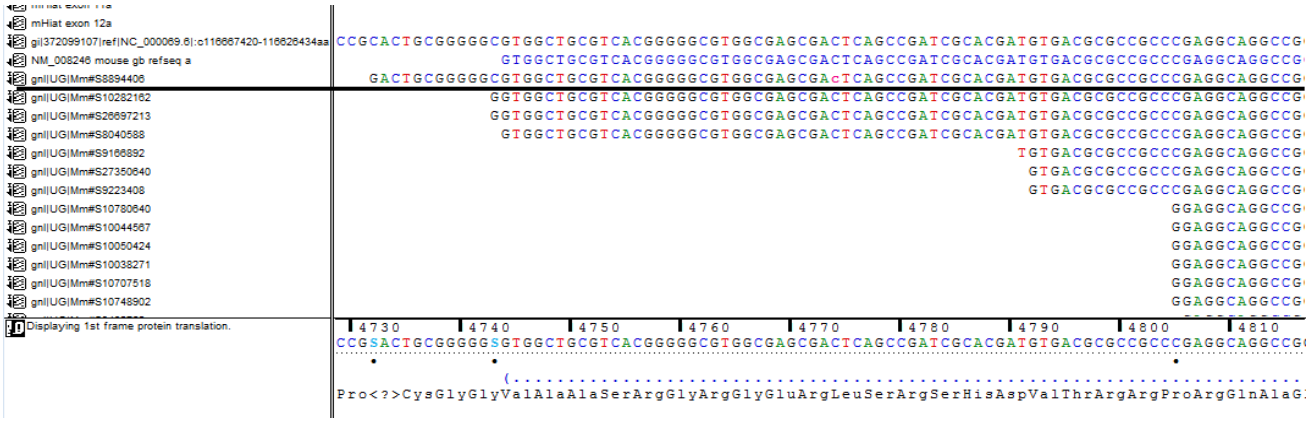
3.8 SUPPLEMENTARY INFORMATION

Name	Sequence 5'→ 3'
mmfsd14a SpeI forward	ATAactagtATGACCCAGGGGAAGAAA
mmfsd14a SmaI reverse	ATAcccgggAGTGGCAAGAGAGTGGTGCT
mmfsd14a SmaI T7 forward	TAATACGACTCACTATAGGGcccgggATGAAGAACAGAGTTGCAGG
mmfsd14a SmaI reverse 2	AATcccgggAGTGGCAAGAGAGTGGTGCT.
mmfsd14a T7 verification forward	AGGGTGTTGTTCAAGGGAT
mmfsd14a internal reverse	AGTGGCAAGAGAGTGGTGCT
GLUT3 ORF forward	ATGGGGACACAGAAGGTCACCCCAG
GLUT3 ORF reverse	TTAGACATTGGTGGTGGTCTCCTTA
GLUT3 T7 forward	AAAATAATACGACTCACTATAGACCATGGGGACACAGAAGGTCAC

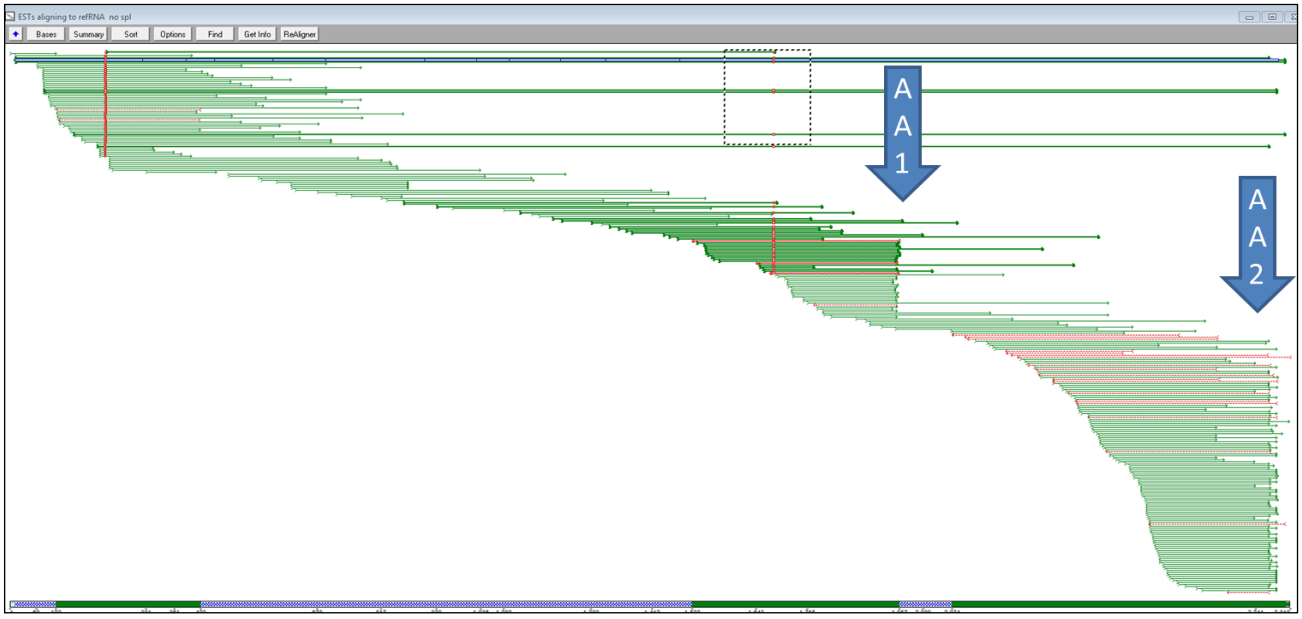
Supplementary Information Table 3.8.1 Sequences of primers used in the cloning process of mouse *Mfsd14a*. Lower case letters indicate restriction enzyme recognition sites.



Supplementary Information Figure 3.8.1 Signature motifs in the mouse *Mfsd14a* (*Hiat1*) protein aligning with known monosaccharide transporters, as depicted in Matsuo et al. (1997, Figure 3). The D-R/K-X-G-R-R/K motif shown in panel A and the P-E-S-P-R motif shown in panel B of mouse *Mfsd14a* (formally known as *Hiat1*, or hippocampus abundant gene transcript 1), tetracycline efflux major facilitator superfamily transporter C (Tet C), the human tetracycline transporter-like protein (Tet-like), mouse glucose transporter type 2 (GLUT2) and mouse glucose transporter type 4. Homologous regions are shaded, and the putative transmembrane domains (TM) are boxed. Numbers on the left indicates amino acid numbers.



Supplementary Information Figure 3.8.2 Alignment of the 5' transcripts (EST and refence sequence evidence form NCBI) in the mouse *Mfsd14a* locus. EST Mm#S8894406 (marked with a black line) shows eleven additional 5' nucleotides currently not included in the reference sequence NM_008246.

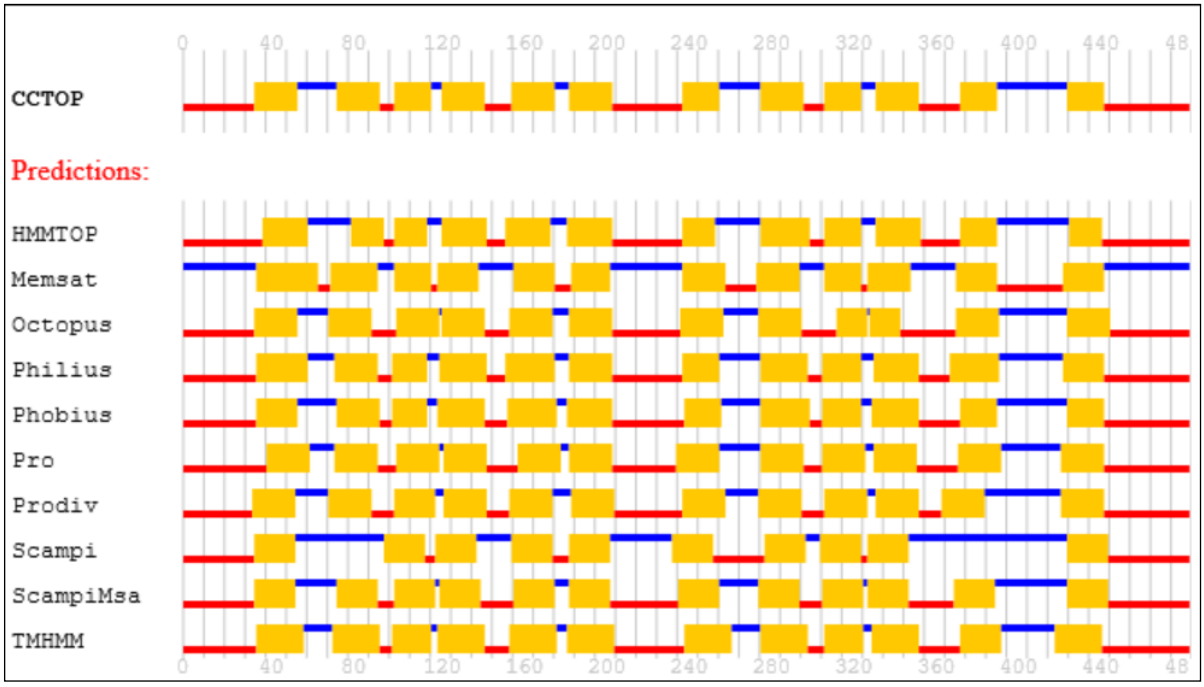


Supplementary Information Figure 3.8.3 An overview of the alignments of the 3' transcripts (EST and refence sequence evidence form NCBI) in the mouse *Mfsd14a* locus. Green arrows represent EST sequences sequenced from the 5' to 3' direction and the red arrows represent EST sequences sequenced from the 3' to 5' direction. After all the available EST sequences were downloaded, they were aligned and visually curated to ensure the quality and accuracy of the alignment. Two transcription termination sites (also called poly a-tails, marked as arrow AA1 and AA2) are visualised. The most 5' site does not impact on the protein coding region.

>NP_032272.2, Mfsd14a, formerly known as hippocampus abundant transcript 1 protein [*Mus musculus*]

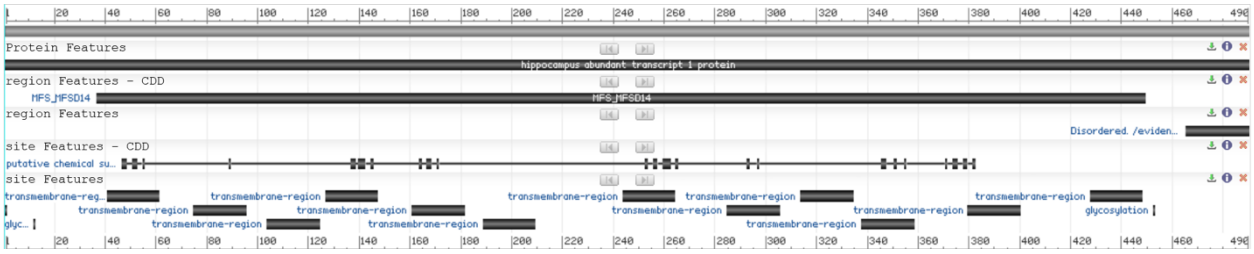
MTQGKKKKRAANRSIMLAKKIIIKDGGTPQGIGSPSVYHAVIVIFLEFFAWGLLTAPTLVV
LHETFPKHTFLMNGLIQGVKGLLSFLSAPLIGALSDVWGRKSFLLLTVFFTCAPIPLMKISP
WWYFAVISVSGVFAVTFSVVFAYVADITQEHERSMAYGLVSATFAASLVTSPAIGAYLGRV
YGDSLVVVLATAIALLDICFILVAVPESLPEKMRPASWGAPISWEQADPFASLKKVGQDSI
VLLICITVFLSYLPEAGQYSSFFLYLRQIMKFSPESVAAFIAVLGILSIIAQTIVLSLLMRSIGNK
NTILLGLGFQILQLAWYGFGSEPWMMWAAGAVAAMSSITFPAVSALVSRTADADQQGVV
QGMITGIRGLCNGLGPALYGFIFYIFHVELKELPITGTDLGTNTSPQHHFEQNSIIPGPPFL
FGACSVLLALLVALFIPEHTNLSLRSSSWRKHCGSHSHPHSTQAPGEAKEPLLQDTNV

Supplementary Information Sequence 3.8.1 Amino acid sequence of the full mouse Mfsd14a protein (formally known as hippocampus abundant transcript 1, Hiat1)

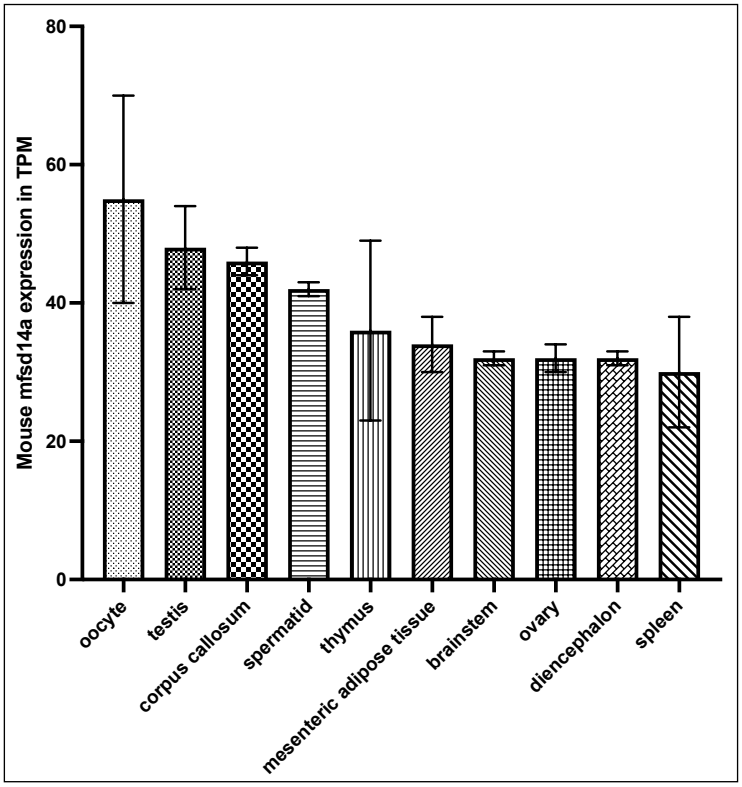


Supplementary Information Figure 3.8.4 Predictions of membrane topology for the mouse Mfsd14a protein. Obtained from the Constrained Consensus TOPology prediction server (http://cctop.enzim.ttk.mta.hu), where 10 sources/algorithms are combined to create a model. HMMTOP: Hidden Markov Model for Topology & Prediction; Memsat: MemSatSVM, Membrane helix prediction with support vector machines; Octopus: obtainer of correct topologies for uncharacterized sequences, a novel topology predictor (Viklund & Elofsson, 2008); Philius: a combined transmembrane topology and signal peptide predictor that extends Phobius by exploiting the power of dynamic Bayesian networks (DBN) (Reynolds et al., 2008); Phobius: A combined transmembrane topology and signal peptide predictor (Kall et al., 2004); Pro: Prodiv: profile-based hidden Markov model (Viklund & Elofsson, 2004); Scampi & ScampiMSA: first-principle-based topology predictors (Peters et al., 2016); TMHMM: Transmembrane Helices; Hidden Markov Model

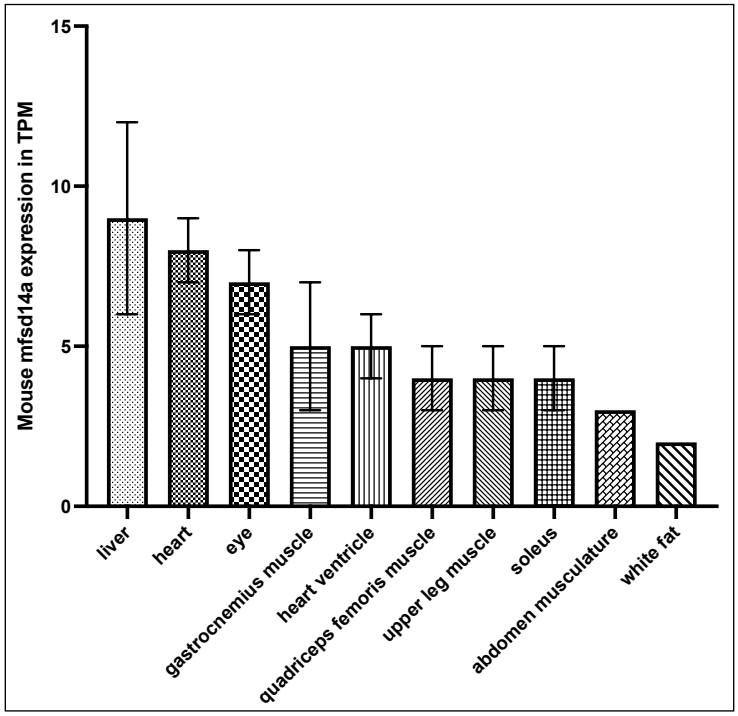
(protein topology), a tool used to predict the presence of transmembrane helices in proteins (Krogh et al., 2001).



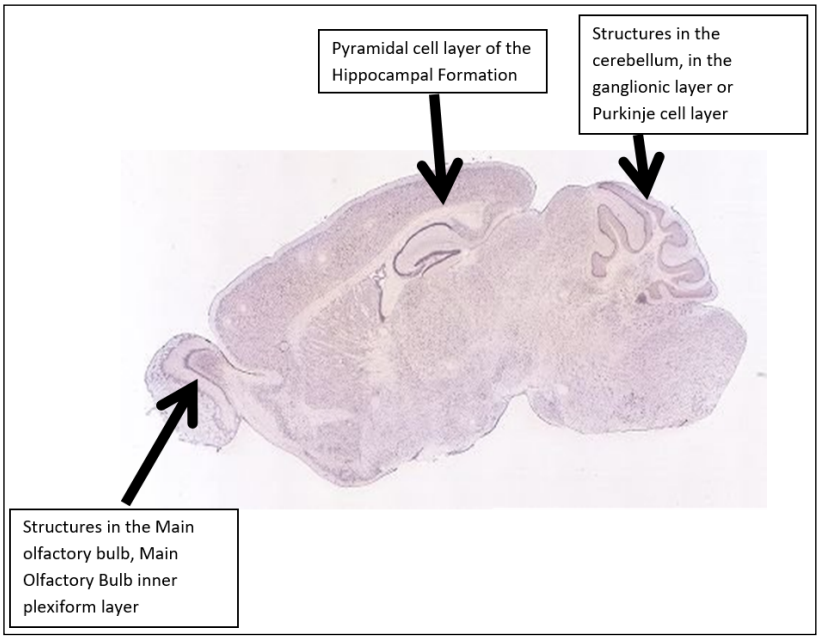
Supplementary Information Figure 3.8.5 NCBI protein report for Reference Sequence: NP_032272.2. Obtained on January 20th 2022, from https://www.ncbi.nlm.nih.gov/protein/NP_032272.2?report=graph



Supplementary Information Figure 3.8.6 Top ten tissues for *Mfsd14a* expression (displayed in Transcripts Per Million, TPM) in the adult mouse. Data Source: MGI mouse tissue expression database.



Supplementary Information Figure 3.8.7 *Mfsd14a* expression (displayed in Transcripts Per Million, TPM) in adult mouse tissues. The ten tissues of lowest expression. Data Source: MGI mouse tissue expression database.



Supplementary Information Figure 3.8.8 *Mfsd14a* expression in adult mouse brain from the Allen Mouse Brain Atlas. Data retrieved from: the Allen Mouse Brain Atlas, http://mouse.brain-map.org/experiment/show/69734989)

CHAPTER 4:

MANUSCRIPT 3

This manuscript was published in the Journal of Advances in Nutrition and Food Sciences

(Advances in Nutrition and Food Sciences, Volume 2022, Issue 01, DOI:

10.37722/ANAFS.2022102)

THE HUMAN MAJOR FACILITATOR SUPERFAMILY DOMAIN CONTAINING 14A (MFSD14A) GENE DOES NOT ENCODE A GLUCOSE TRANSPORTER

Haonan Zhouyao[1], Dirk Weihrauch[2], Peter Eck*[1],

Authors Affiliations:

[1]Department of Food and Human Nutritional Sciences, University of Manitoba, Winnipeg, Canada.

[2]University of Manitoba, Department of Biological Sciences, 190 Dysart Road, Winnipeg, MB R3T 2N2, Canada

*Corresponding Author

My contribution:

I conducted the literature search, the glucose uptake studies and wrote the manuscript.

Other author's contribution:

PE was the principal investigator, responsible for the conception and design of the project, the acquisition of financial support, and the writing and final editing of the manuscript.

DW helped editing the manuscript.

4.1 ABSTRACT

The literature record on the human *MFSD14A* gene is extremely limited. However, disease associations with infertility and cancer severity are emerging, which warrants a thorough interrogation of the genomic locus to establish a baseline annotation. It had been suggested that MFSD14A is a putative glucose solute carrier, however, this was never tested.

This report establishes the baseline on the human *MFSD14A* genomic locus, encoding one ubiquitously expressed transcript which translates into a protein of 490 amino acids. The protein is predicted to contain a binding pocket for glucose and similar monosaccharides, however, when expressed in *Xenopus laevis* oocytes does not mediate uptake of radiolabeled 2-deoxy-glucose.

4.2 INTRODUCTION

To date, eight publications refer to the human *MFSD14A* gene **(Supplementary information, Table 4.8.1)**. It is described to contain an "off target" target site in CRISPR/Cas9 intervention (Osborn et al., 2016), and a circular RNA derived from the human *MFSD14A* gene locus is implicated in the development of gastric (Quan et al., 2020) and renal cell carcinomas (Wang et al., 2017). In T-cell lymphoblastic lymphoma a novel fusion transcript containing SLC35A3-MFSD14A is identified (Lopez-Nieva et al., 2019). Significantly, no baseline data on the human *MFSD14A* genomic locus and its encoded products are in the publication record. This report aims to establish this.

Based on the analysis of the mouse protein, it is extrapolated that mouse mfsd14a might encode for a novel sugar transporter (Matsuo et al., 1997), however, no functional data are published. This report interrogates the glucose transport mediated by human MFSD14A. This might illuminate on the clinically relevance of the phenotype of male infertility resembling human globozoospermia in mfsd14a$^{-/-}$ mice (Doran et al., 2016).

Overall, in the light of emerging clinical relevance, this report aims to establish baseline data on the human MFSD14A genomic locus and interrogates its ability to mediate transmembrane transport of glucose.

4.3 METHODS

4.3.1 Bioinformatics genomics analysis

The genomic sequence of human *MFSD14A* (Chromosome 1 - NC_000001.11), full length RNA sequences including both reference and non-reference RNA sequences, and the expressed sequence tags (ESTs) sequences were retrieved from the National Center for Biotechnology Information (NCBI) database. Alignments between reference RNA sequence and the genomic DNA sequence was made in the DNA Sequence Analysis Software – Sequencher 4.8 (GeneCodes Corporation; Ann Arbor, Michigan, USA) to determine the structures of the exons and the introns (with large gap alignment parameter, 20 minimum overlap and 80% minimum matches). A scaled illustration was created after the exons and introns are manually curated to demonstrate the genomic organization and the structures of the exon and introns.)

4.3.2 Protein analysis

The function and structure of the MFSD14A protein was analysed with the Protein Homology/analogY Recognition Engine V 2.0 (Phyre2, (Kelley et al., 2015)) and the AlphaFold Model using 3D ligand site (Jumper et al., 2021; Wass et al., 2010) with parameters from the standard/normal modes.

4.3.3 Expression patterns analysis

The following resources are used to obtain the experimental RNA-seq data on panels of adult human tissues: the European Molecular Biology Laboratory's European Bioinformatics Institute (EMBL-EBI) Baseline Expression Atlas (URL: https://www.ebi.ac.uk/gxa/home) (Moreno et al., 2022) and The Encyclopedia of DNA Elements (ENCODE) (URL: https://www.encodeproject.org, (Consortium, 2012). RNA-seq data from the following projects were interrogated: the Genotype-

Tissue Expression (GTEx) project containing 53 samples (Consortium, 2015); The National Institutes of Health (NIH) Roadmap Epigenomics Mapping Consortium containing 19 samples (Roadmap Epigenomics et al., 2015), The Functional Annotation of the Mammalian Genome 5 (FANTOM5) project containing 56 samples (URL: https://fantom.gsc.riken.jp/data/), the Mammalian Kaessmann project containing 8 samples (Cardoso-Moreira et al., 2019) and the Encyclopedia of DNA Elements (ENCODE) project containing 13 samples (Lin et al., 2014). The RNA expression level of MFSD14A is presented in transcript per million (TPM).

4.3.4 Subcloning of *MFSD14A*

Plasmids containing the full open reading frames (ORF) of MFSD14A and GLUT3 (positive control for glucose transport study) was purchased from Harvard PlasmID (MFSD14A clone ID: HsCD00399291; GLUT3 clone ID: HsCD00021270) and transformed into DH5α competent cells (New England Biolabs, Ipswich, Massachusetts, USA) following the manufactures guidance for propagation, followed by column purification (QIAprep Spin Miniprep Kit, Qiagen, Hilden, Germany). The ORFs were then amplified with Q5 High Fidelity DNA Polymerase (New England Biolabs) and either tagged with restriction enzyme recognition sites (BamHI and XbaI) suitable for *Xenopus laevis* expression vector, PGEM-HE © (for MFSD14A) or tagged with T7 RNA polymerase promoter sequence (for GLUT3) **(primer sequences in Supplementary Information Table 4.8.2)**. The plasmid containing human MFSD14A was then linearized by SphI prior to the in vitro transcription of capped mRNA (cRNA) (HiScribe™ T7 ARCA mRNA kit, New England Biolabs) whereas the GLUT3 amplicon containing GLUT3 ORF and T7 promoter sequence was used directly for cRNA synthesis as described previously. cRNA products were column purified (RNeasy MinElute Cleanup Kit, Qiagen) prior to the spectrophotometric quantification (Nanodrop, ND-1000, Thermofisher, Waltham, MA, USA) and the integrity assessments (MOPS agarose gel containing

formaldehyde).

4.3.5 Oocyte experiments

Stage VI-V oocytes were collected from mature female *Xenopus laevis* (VWR International, Randor, PA, USA) as previously described ((Soreq & Seidman, 1992). Briefly, the frogs were euthanized via decapitation prior to the collection of the ovaries. The ovaries were placed in calcium-free OR2 solution (contained (in mmol l^{-1}) 82.5 NaCl, 2.5 KCl, 1 $MgCl_2$, 1 Na_2HPO_4, 5 HEPES, pH 7.4) with collagenase type VI (1 mg ml^{-1}) (Gibco, Waltham, MA, USA), gently agitated for 90 minutes at room temperature. 1 mmol l^{-1} $CaCl_2$ was added to terminate the collagenase activity. Oocytes were then manually sorted, rinsed three additional times with standard OR2 solution (contained (in mmol l^{-1}) 82.5 NaCl, 2.5 KCl, 1 $MgCl_2$, 1 Na_2HPO_4, 5 HEPES, 1 $CaCl_2$, pH 7.4). For long term storage of isolated oocytes, the standard OR2 solution was supplemented with 2.5 mmol l^{-1} sodium pyruvate, 1 mg ml^{-1} penicillin-streptomycin (Gibco, Long Island, NY, USA) and 50 μg ml^{-1} gentamicin and held at 16 °C without the exposure of light. All procedures used were approved by the University of Manitoba Office of Research Ethics & Compliance and Animal Care Committee. Animal care procedures were based on guidelines described in the Canadian Council for Animal Care.

4.3.6 cRNA injection

Isolated oocytes after the overnight recovery were injected with 18.4 ng of cRNA (36.8 nL with 0.5 ng nl^{-1}) or nuclease-free water as negative control using the Nanoject II auto-nanoliter injector (Drummond Scientific, Broomall, Pennsylvania, USA). Experiments were conducted 48 hours post-injection.

4.3.7 Glucose uptake study in oocytes expressing MFSD14A and GLUT3

Oocytes expressing MFSD14A, GLUT3 (positive control) and SHAM (oocytes injected with water, negative control) were equilibrated at room temperature in standard OR2 solution two days post-injection, and randomly divided into groups containing 20 oocytes per group. Experiments were performed by incubating groups of oocytes in 200 μl standard OR2 solution which consisting of 125 pmol [^{3}H] 2-deoxyglucose (25.5 Ci/mmol, PerkinElmer, Waltham, MA, USA) for 30 minutes. Excess ice-cold standard OR2 solution was added to terminate the transport, followed by four washes with the same solution. Oocytes were then individually lysed in scintillation vials (6 mL Pony Vial, PerkinElmer) with 200 μl of 10% sodium dodecyl sulfate (SDS) before the addition of 5 ml Ultima Gold scintillation cocktail (PerkinElmer). Internal radioactivity was quantified by liquid scintillation spectrometry as CPM oocyte^{-1} 30min^{-1}.

4.4 RESULTS

4.4.1 The human *MFSD14A* genomic locus

The human *MFSD14A* gene is listed in NCBI as gene ID 64645, residing on the sense strand of human chromosome 1, more specific chr1:100,503,789-100,548,929 in 1p31 (GRCh37/hg19) **(Figure 1).** It is immediately flanked by the SAS-6 centriolar assembly protein (SASS6) and the solute carrier family 35 member A3 (SLC35A3) genes, which might explain the detected fusion protein (Lopez-Nieva et al., 2019).

The visually curated alignment of transcripts shows that the human *MFSD14A* gene spans 45,285 bases from its transcription initiation to the termination site. The transcript starts 155 nucleotides upstream of the site currently identified NCBI Reference Transcript NM_033055 (**Supplementary Information Figure 4.8.1)**, while the transcriptional termination sites is correctly annotated.

The gene contains 12 exons transcribing into one transcript (NM_033055) **(Figure 1)**, encoding Reference Protein NP_149044 **(Supplementary Information Sequence 4.8.1).**

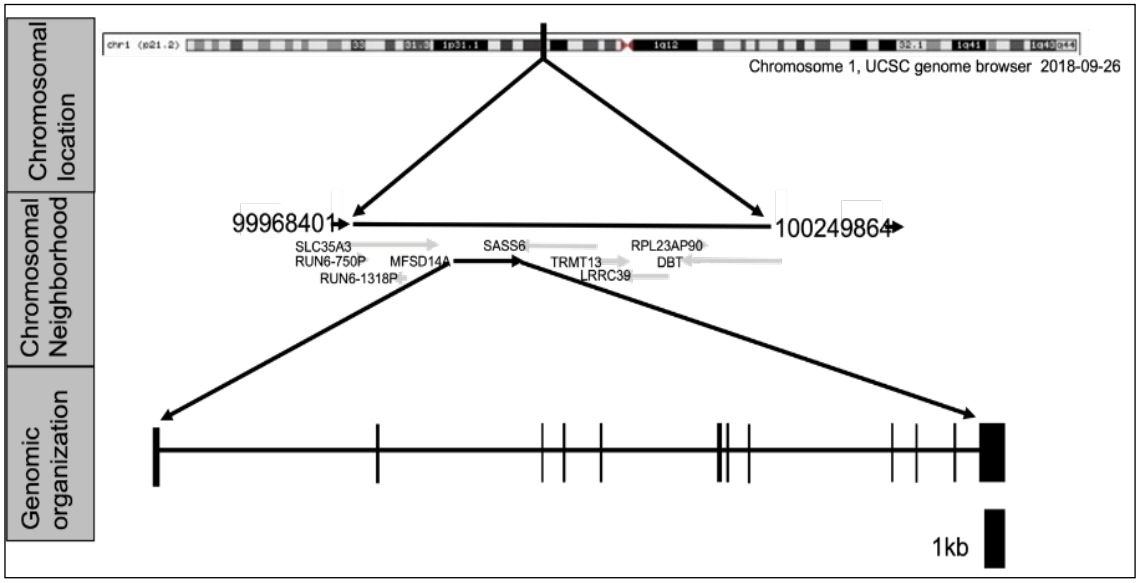


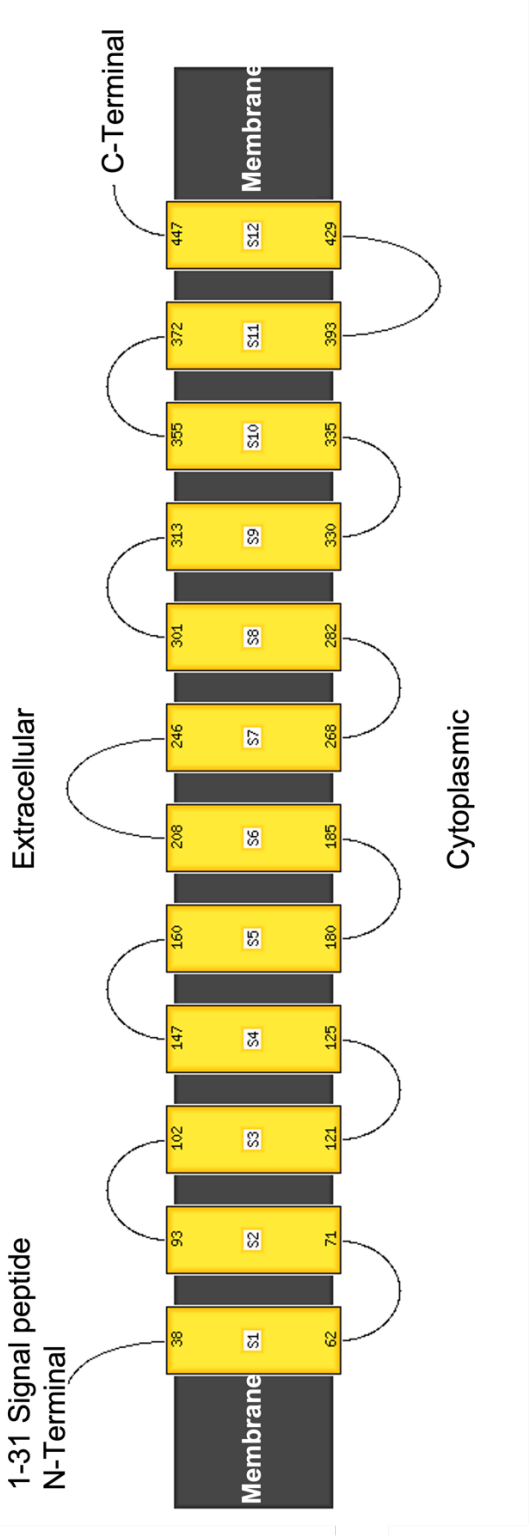
Figure 4.1 The human *MFSD14A* genomic locus and its encoded transcript.

4.4.2 The human MFSD14A protein

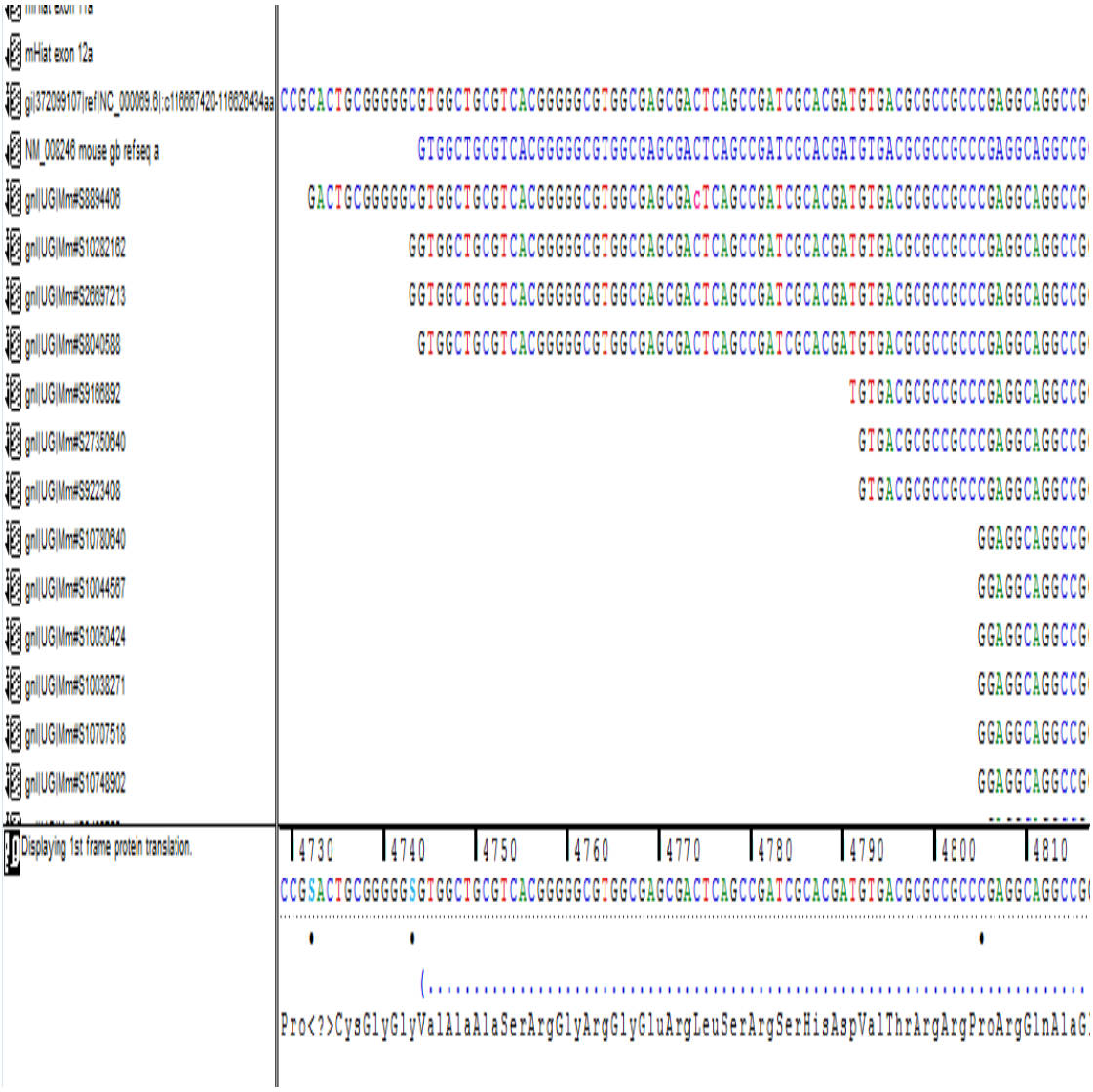
The human MFSD14A protein contains 490 amino acids, as represented by NCBI reference protein NP_149044 (**Supplementary Information Sequence 4.8.1**). Consistent with prior literature reports, the protein is predicted to be a solute membrane carrier spanning twelve transmembrane domains (**Figure 2A and Supplementary Information Figure 4.8.2**).

The human MFSD14A protein clusters to the protein family MFS, subfamily MFSD14 (**Supplementary Information Figure 4.8.2C**). This subfamily contains only one additional member, MFSD14B. Both family members contain a unique putative chemical substrate binding pocket, setting it somewhat aside from the glucose transporter families. However, the putative organic substrates, glucose, xylose, maltose, alanine, chloramphenicol, and 3-alpha, 5-beta, 12-alpha)-3, 12-dihydroxycholan-24-oic acid are predicted to bind in a pore like intramolecular binding pocket (**Figure 2C),** supporting others' speculations that glucose might be a main substrate.

A)


C-Terminal
Membrane
Extracellular
Cytoplasmic
1-31 Signal peptide
N-Terminal
Membrane
S1
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11
S12
38
62
93
71
102
121
147
125
160
180
208
185
246
268
301
282
313
330
355
335
372
393
447
429

B)


mHist exon 12a
gi|372099107|ref|NC_000069.6|:c118667420-118628434aa
NM_008246 mouse gb refseq a
gnl|UG|Mm#S8894408
gnl|UG|Mm#S10282162
gnl|UG|Mm#S28697213
gnl|UG|Mm#S8040888
gnl|UG|Mm#S9166892
gnl|UG|Mm#S27350840
gnl|UG|Mm#S9223408
gnl|UG|Mm#S10780840
gnl|UG|Mm#S10044567
gnl|UG|Mm#S10050424
gnl|UG|Mm#S10038271
gnl|UG|Mm#S10707518
gnl|UG|Mm#S10748902
Displaying 1st frame protein translation.
4730
4740
4750
4760
4770
4780
4790
4800
4810
Pro<?>CysGlyGlyValAlaAlaSerArgGlyArgGlyGluArgLeuSerArgSerHisAspValThrArgArgProArgGlnAlaG

C)

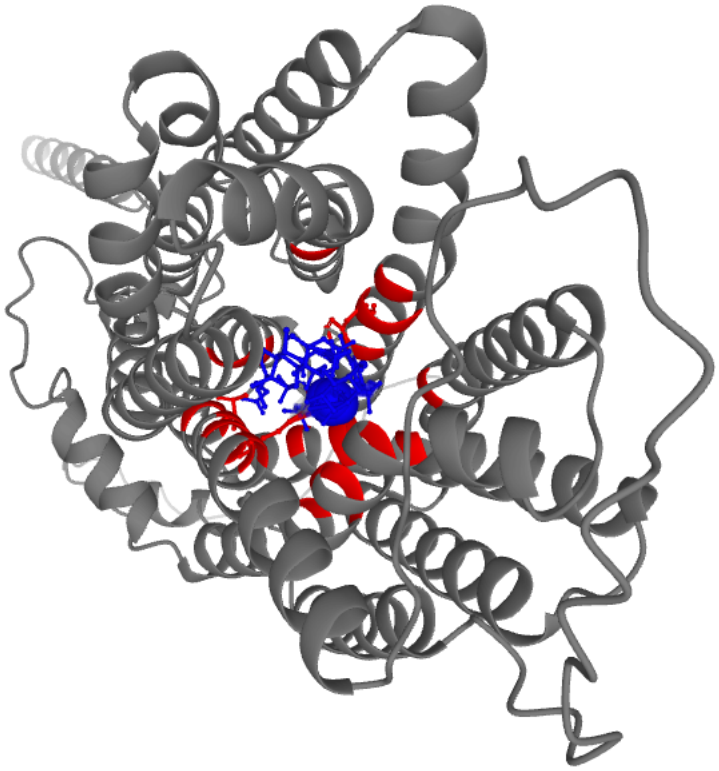


Figure 4.2: The predicted (panel A) secondary and (panel B & C) tertiary structures for the mouse mfsd14a protein (Image colored by rainbow N→C terminus; Model dimensions (Å): X:58.386 Y:50.944 Z:54.127). Models created using the phyre[2] web-server. Panel C depicts the putative ligand binding pocked aligned in a "pore like" structure created as an AlphaFold Model using the 3D ligand site (https://www.wass-michaelislab.org/3dlig/web_results/61eae944c331a/61eae944c331a.html).

4.4.3 Tissues expression of the human *MFSD14A* transcript

Human *MFSD14A* expression is ubiquitous; no tissue examined is without expression evidence **(Figure 3 and Supplementary Information Figures 4.8.3 A-D)**. Highest expression was observed in breast, spinal cord, seminal vesicles, testis **(Figure 4.3)**, and renal areas **(Supplementary Information Figures 4.8.3 A-D)**. In general, significant expression was exhibited by all tissues assessed, with lowest levels in muscle, liver, placenta, and thymus **(Figure 4.3 and Supplementary Information Figures 4.8.3 A-D)**.

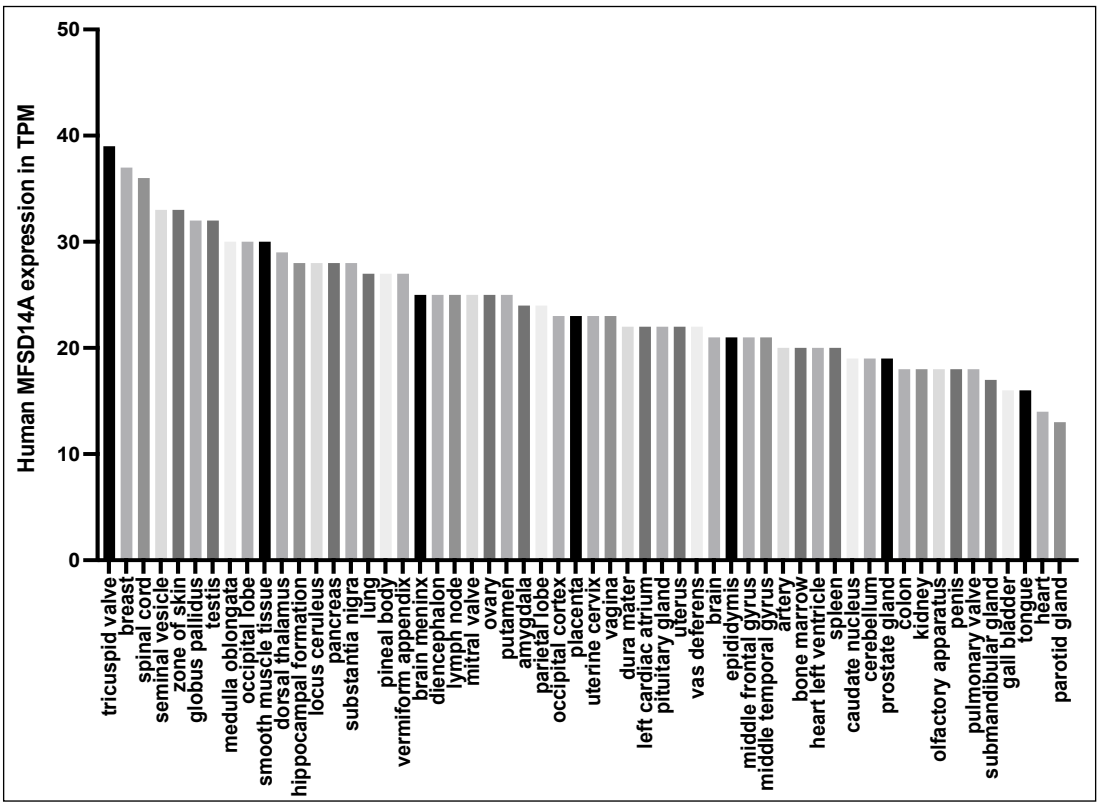


Figure 4.3 Human MFSD14A expression data from the FANTOM5 project. Data Source: EBI expression Atlas. Expression levels in transcripts per million, TPM.

4.4.4 Human MFSD14A does not mediate glucose uptake into *Xenopus laevis* oocytes.

When human MFSD14A is expressed in *Xenopus laevis* oocytes, no glucose uptake was detected (**Figure 4**). This provides evidence that glucose or similar saccharides are not a substrate.

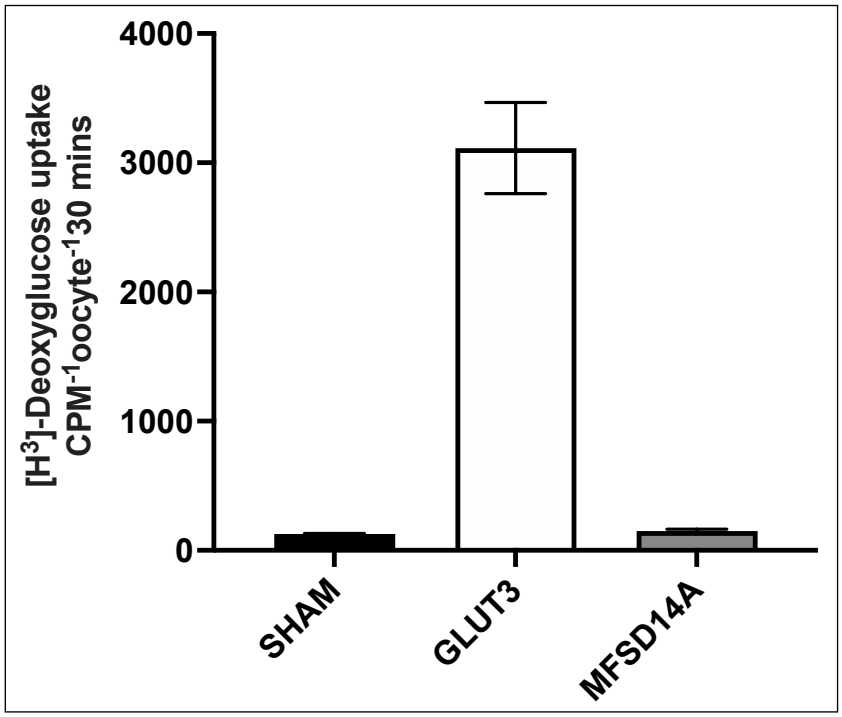


Figure 4.4 The uptake of radiolabeled 2-deoxyglucose into *Xenopus laevis* oocytes expressing GLUT3 (known glucose carrier) and the human MFSD14A. Compared to the SHAM oocytes (water-injected, negative control), there is no significant glucose uptake in oocytes expressing human MFSD14A. N=15-20 oocytes.

4.5 DISCUSSION

This report establishes the basic annotation of the human MFSD14A locus, containing one transcript, and encoding one protein of 490 amino acids. This visually curated annotation corroborates the records in the NCBI, EMBL and UCSC genomic databases, except that the transcriptional start site is located 155 nucleobases upstream of the currently annotated. Tissue expression is ubiquitous with some of the highest levels detected in testis, which might indicate a relevance to the globozoospermia-like phenotype of the *mfsd14a*$^{-/-}$ mice (Doran et al., 2016). The baseline genomics established here should serve as a nucleus to interrogate the genetics of this very rare human condition (Modarres et al., 2019).

Evolutionary conservation of MFSD14A is high (Doran et al., 2016; Lekholm et al., 2017), which is confirmed by the presence of a sole transcript and protein, since alternative splicing seems to negatively correlate with fitness of the locus (Saudemont et al., 2017). This also indicates a basic ubiquitous function, and it could be expected that more than fertility might be affected in gene disruption. More mechanistic insight will be possible upon the identification of the MFSD14A substrate.

The human MFSD14A does not mediate transmembrane glucose uptake. This dismisses the strong suggestion voiced in prior report (Matsuo et al., 1997) and indicated by here presented bioinformatics prediction, that glucose could be a canonical substrate of the MFSD14A protein. Although this does not rule out other monosaccharides as MFSD14A substrates, it limits the substrate range to fructose/fructose-like molecules, since the multi-specific substrate recognition reported for the SLC2A (GLUT) and SLC5A (SGLT) transporters family members include glucose, xylose, arabinose, and mannose (Deng & Yan, 2016). This severely limits the likelihood of any of these substrates been recognized by MFSD14A.

In the quest to functionally characterize the MFSD14A gene/protein, data from other species might give novel leads: the mouse neuronal *mfsd14a* is differentially regulated in response to amino acid and food deprivation, as well as high fat diet (Lekholm et al., 2017). Although this compels a strong connection to energy/carbohydrate regulation, it could also point towards a function in amino acid or lipid pathways. Moreover, *mfsd14a* was differentially regulated in the gills of the crab *Carcinus maenas* acclimated to brackish water (Fehsenfeld et al., 2011). Specifically, after challenges with high pCO_2 and ammonia, *mfsd14a* was downregulated (Fehsenfeld et al., 2011). This indicates a potential involvement in detoxification/acid-base pathways, with ammonium as a prime candidate.

Ubiquitous expression in relevant levels could link *MFSD14A* to cancers beyond gastric cancer and renal cell carcinomas (Li et al., 2017; Quan et al., 2020; Wang et al., 2017), since the expression of the disease modulating circular RNA circHIAT1 is correlated with the general expression of the gene. Therefore, this report should establish the baseline expression to inform future research avenues.

4.6 CONCLUSION

The ubiquitously expressed human *MFSD14A* gene transcribes into a single transcript encoding one protein. It does not mediate glucose uptake into *Xenopus laevis* oocytes. The genomic and functional baseline to inform on human diseases is established.

4.7 REFERENCES

Cardoso-Moreira, M., Halbert, J., Valloton, D., Velten, B., Chen, C., Shao, Y., Kaessmann, H. (2019). Gene expression across mammalian organ development. Nature, 571(7766), 505-509. https://doi.org/10.1038/s41586-019-1338-5

Consortium, E. P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature, 489(7414), 57-74. https://doi.org/10.1038/nature11247

Consortium, G. T. (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science, 348(6235), 648-660. https://doi.org/10.1126/science.1262110

Deng, D., & Yan, N. (2016). GLUT, SGLT, and SWEET: Structural and mechanistic investigations of the glucose transporters. Protein Sci, 25(3), 546-558. https://doi.org/10.1002/pro.2858

Doran, J., Walters, C., Kyle, V., Wooding, P., Hammett-Burke, R., & Colledge, W. H. (2016). Mfsd14a (Hiat1) gene disruption causes globozoospermia and infertility in male mice [Article]. Reproduction, 152(1), 91-99. https://doi.org/10.1530/REP-15-0557

Fehsenfeld, S., Kiko, R., Appelhans, Y., Towle, D. W., Zimmer, M., & Melzner, F. (2011). Effects of elevated seawater pCO(2) on gene expression patterns in the gills of the green crab, Carcinus maenas. BMC Genomics, 12, 488. https://doi.org/10.1186/1471-2164-12-488

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., . . . Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596(7873), 583-589. https://doi.org/10.1038/s41586-021-03819-2

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc, 10(6), 845-858. https://doi.org/10.1038/nprot.2015.053

Lekholm, E., Perland, E., Eriksson, M. M., Hellsten, S. V., Lindberg, F. A., Rostami, J., & Fredriksson, R. (2017). Putative Membrane-Bound Transporters MFSD14A and MFSD14B Are Neuronal and Affected by Nutrient Availability. Front Mol Neurosci, 10, 11. https://doi.org/10.3389/fnmol.2017.00011

Li, P., Chen, H., Chen, S., Mo, X., Li, T., Xiao, B., Yu, R., Guo, J. (2017). Circular RNA 0000096 affects cell growth and migration in gastric cancer. Br J Cancer, 116(5), 626-633. https://doi.org/10.1038/bjc.2016.451

Lin, S., Lin, Y., Nery, J. R., Urich, M. A., Breschi, A., Davis, C. A., . . . Snyder, M. P. (2014). Comparison of the transcriptional landscapes between human and mouse tissues. Proc Natl Acad Sci U S A, 111(48), 17224-17229. https://doi.org/10.1073/pnas.1413624111

Lopez-Nieva, P., Fernandez-Navarro, P., Grana-Castro, O., Andres-Leon, E., Santos, J., Villa-Morales, M., Cobos-Fernández, C., González-Sánchez, L., Malumbres, M., Salazar-Roa, M., Fernandez-Piqueras, J. (2019). Detection of novel fusion-transcripts by RNA-Seq in T-cell lymphoblastic lymphoma. Sci Rep, 9(1), 5179. https://doi.org/10.1038/s41598-019-41675-3

Matsuo, N., Kawamoto, S., Matsubara, K., & Okubo, K. (1997). Cloning of a cDNA encoding a novel sugar transporter expressed in the neonatal mouse hippocampus [Article]. Biochemical and Biophysical Research Communications, 238(1), 126-129. https://doi.org/10.1006/bbrc.1997.7252

Modarres, P., Tavalaee, M., Ghaedi, K., & Nasr-Esfahani, M. H. (2019). An Overview of The Globozoospermia as A Multigenic Identified Syndrome. Int J Fertil Steril, 12(4), 273-277. https://doi.org/10.22074/ijfs.2019.5561

Moreno, P., Fexova, S., George, N., Manning, J. R., Miao, Z., Mohammed, S., . . . Papatheodorou, I. (2022). Expression Atlas update: gene and protein expression in multiple species. Nucleic Acids Res, 50(D1), D129-D140. https://doi.org/10.1093/nar/gkab1030

Osborn, M. J., Webber, B. R., Knipping, F., Lonetree, C. L., Tennis, N., DeFeo, A. P., . . . Blazar, B. R. (2016). Evaluation of TCR Gene Editing Achieved by TALENs, CRISPR/Cas9, and megaTAL Nucleases. Mol Ther, 24(3), 570-581. https://doi.org/10.1038/mt.2015.197

Quan, J., Dong, D., Lun, Y., Sun, B., Sun, H., Wang, Q., & Yuan, G. (2020). Circular RNA circHIAT1 inhibits proliferation and epithelial-mesenchymal transition of gastric cancer cell lines through downregulation of miR-21. J Biochem Mol Toxicol, 34(4), e22458. https://doi.org/10.1002/jbt.22458

Roadmap Epigenomics, C., Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., . . . Kellis, M. (2015). Integrative analysis of 111 reference human epigenomes. Nature, 518(7539), 317-330. https://doi.org/10.1038/nature14248

Saudemont, B., Popa, A., Parmley, J. L., Rocher, V., Blugeon, C., Necsulea, A., . . . Duret, L. (2017). The fitness cost of mis-splicing is the main determinant of alternative splicing patterns. Genome Biol, 18(1), 208. https://doi.org/10.1186/s13059-017-1344-6

Wang, K., Sun, Y., Tao, W., Fei, X., & Chang, C. (2017). Androgen receptor (AR) promotes clear cell renal cell carcinoma (ccRCC) migration and invasion via altering the circHIAT1/miR-195-5p/29a-3p/29c-3p/CDC42 signals. Cancer Lett, 394, 1-12. https://doi.org/10.1016/j.canlet.2016.12.036

Wass, M. N., Kelley, L. A., & Sternberg, M. J. (2010). 3DLigandSite: predicting ligand-binding sites using similar structures. Nucleic Acids Res, 38(Web Server issue), W469-473. https://doi.org/10.1093/nar/gkq406

4.8 SUPPLEMENT INFORMATION

Authors	Title	Year	Journal
Quan, J., Dong, D., Lun, Y., Sun, B., Sun, H., Wang, Q., & Yuan, G.	Circular RNA circHIAT1 inhibits proliferation and epithelial-mesenchymal transition of gastric cancer cell lines through downregulation of miR-21.	2020	The Journal of Biochemical and Molecular Toxicology
Lekholm, E., Perland, E., Eriksson, M. M., Hellsten, S. V., Lindberg, F. A., Rostami, J., & Fredriksson, R.	Putative Membrane-Bound Transporters MFSD14A and MFSD14B Are Neuronal and Affected by Nutrient Availability.	2017	Frontiers in Molecular Neuroscience
Li, P., Chen, H., Chen, S., Mo, X., Li, T., Xiao, B., Yu, R., Guo, J.	Circular RNA 0000096 affects cell growth and migration in gastric cancer	2017	British Journal of Cancer
Wang, K., Sun, Y., Tao, W., Fei, X., & Chang, C.	Androgen receptor (AR) promotes clear cell renal	2017	Cancer Letters

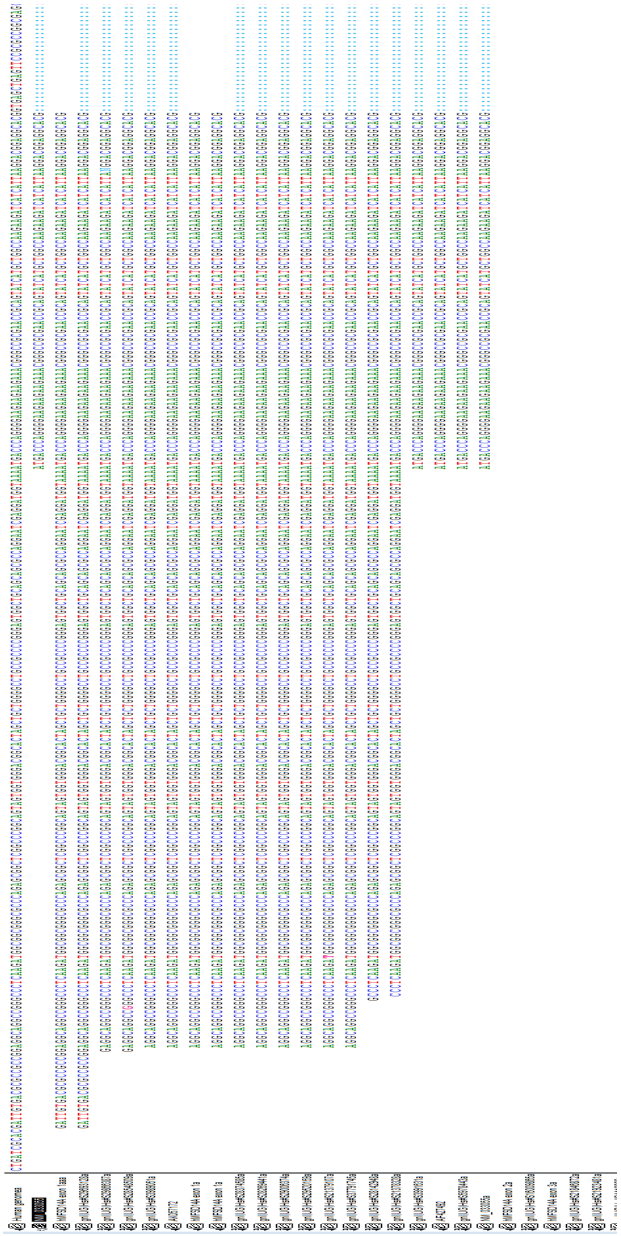
	cell carcinoma (ccRCC) migration and invasion via altering the circHIAT1/miR-195-5p/29a-3p/29c-3p/CDC42 signals.		
Doran, J., Walters, C., Kyle, V., Wooding, P., Hammett-Burke, R., & Colledge, W. H.	Mfsd14a (Hiat1) gene disruption causes globozoospermia and infertility in male mice	2016	Reproduction
Osborn, M. J., Webber, B. R., Knipping, F., Lonetree, C. L., Tennis, N., DeFeo, A. P., . . . Blazar, B. R.	Evaluation of TCR Gene Editing Achieved by TALENs, CRISPR/Cas9, and megaTAL Nucleases.	2016	Molecular Therapy
Fehsenfeld, S., Kiko, R., Appelhans, Y., Towle, D. W., Zimmer, M., & Melzner, F.	Effects of elevated seawater pCO(2) on gene expression patterns in the gills of the green crab, Carcinus maenas.	2011	BMC Genomics

Matsuo, N., Kawamoto, S., Matsubara, K., & Okubo, K.	Cloning of a cDNA encoding a novel sugar transporter expressed in the neonatal mouse hippocampus	1997	Biochemical and Biophysical Research Communication

Supplementary Information Table 4.8.1 List of currently available published literature on MFSD14A

Name	Sequence 5' -> 3'
GLUT3 ORF forward	ATGGGGACACAGAAGGTCACCCCAG
GLUT3 ORF reverse	TTAGACATTGGTGGTGGTCTCCTTA
GLUT3 T7 forward	AAAATAATACGACTCACTATAGACCATGGGGACACAGAAGGTCAC
MFSD14A ORF forward	ATGACCCAGGGGAAGAAGAAGAAAC
MFSD14A ORF reverse	TCACACATTTGTGTCCTGGAGTAAA
MFSD14A BamHI forward	AAAAGGATCCACCATGACCCAGGGGAAGAAGAAGAAAC
MFSD14A XbaI reverse	AAAATCTAGATCACACATTTGTGTCCTGGAGTAAA

Supplementary Information Table 4.8.2: Primers used in this study

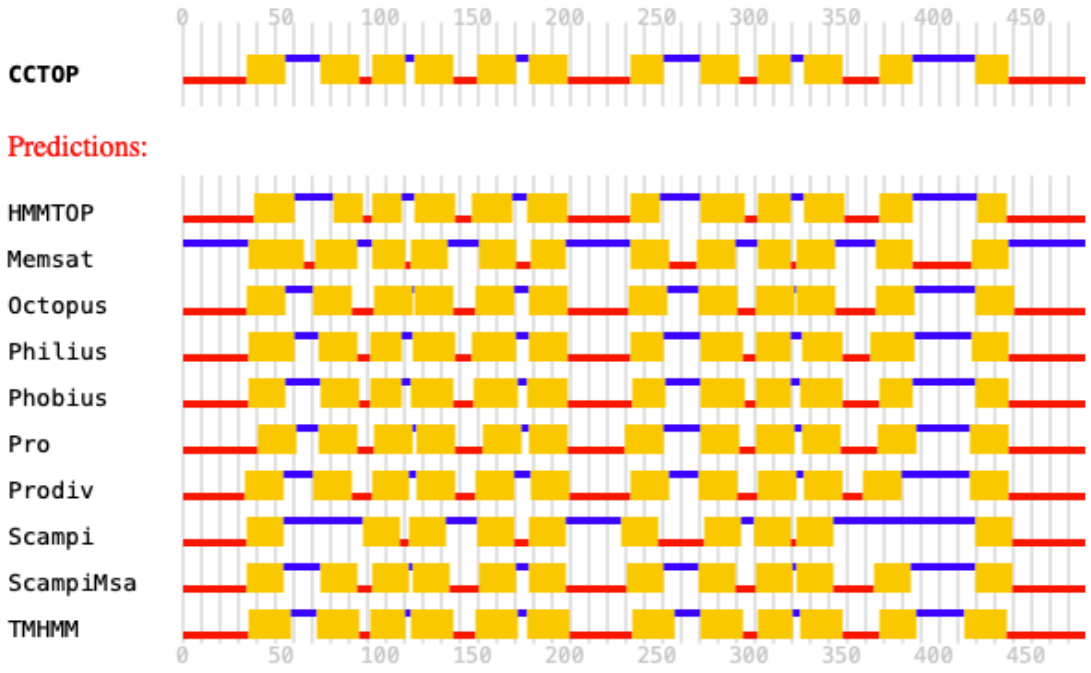


Supplementary Information Figure 4.8.1 Alignment of the 5' transcripts (EST and reference mRNA sequence evidence form NCBI) in the human MFSD14A locus, identifying 155 additional 5' nucleotides for human exon 1 (beyond the annotated NCBI transcript NM_033055).

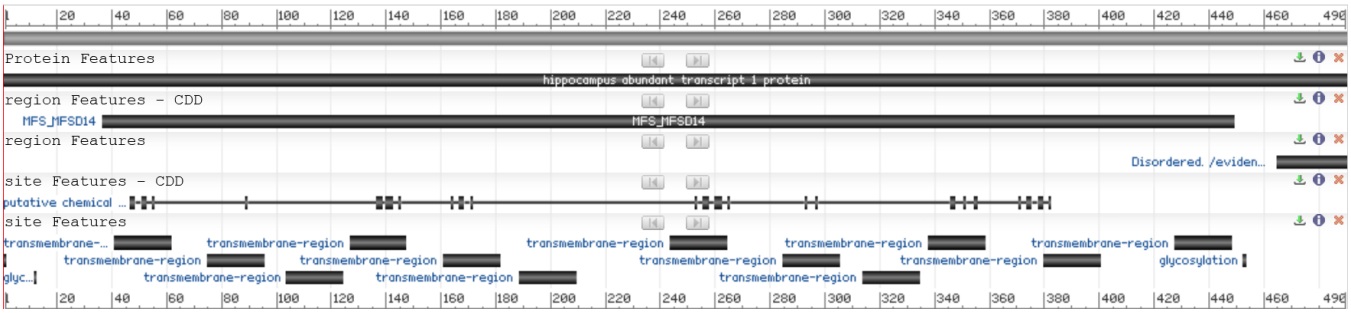
>NP_149044.2 MFSD14A [Homo sapiens]

MTQGKKKKRAANRSIMLAKKIIIKDGGTPQGIGSPSVYHAVIVIFLEFFAWGLLTAPTLVV
LHETFPKHTFLMNGLIQGVKGLLSFLSAPLIGALSDVWGRKSFLLLTVFFTCAPIPLMKISP
WWYFAVISVSGVFAVTFSVVFAYVADITQEHERSMAYGLVSATFAASLVTSPAIGAYLGRV
YGDSLVVVLATAIALLDICFILVAVPESLPEKMRPASWGAPISWEQADPFASLKKVGQDSI
VLLICITVFLSYLPEAGQYSSFFLYLRQIMKFSPESVAAFIAVLGILSIIAQTIVLSLLMRSIGNK
NTILLGLGFQILQLAWYGFGSEPWMMWAAGAVAAMSSITFPAVSALVSRTADADQQGVV
QGMITGIRGLCNGLGPALYGFIFYIFHVELKELPITGTDLGTNTSPQHHFEQNSIIPGPPFL
FGACSVLLALLVALFIPEHTNLSLRSSSWRKHCGSHSHPHNTQAPGEAKEPLLQDTNV

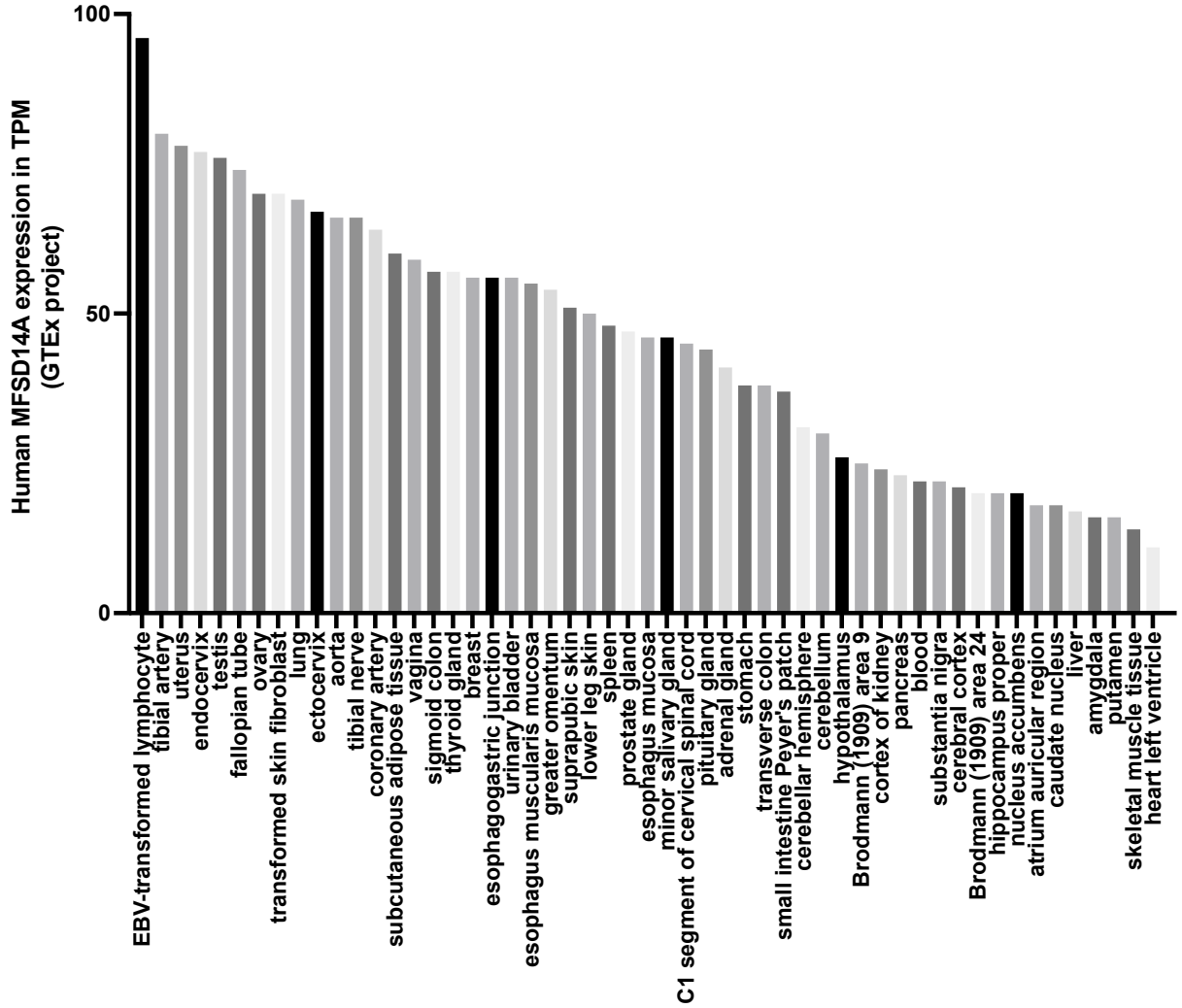
Supplementary Information Sequence 4.8.1 the amino acid sequence of the human MFSD14A protein.



Supplementary Information Figure 4.8.2 Predictions of membrane topology for the human MFSD14A protein. Obtained from the Constrained Consensus TOPology prediction server (http://cctop.enzim.ttk.mta.hu), where 10 sources/algorithms are combined to create a model.

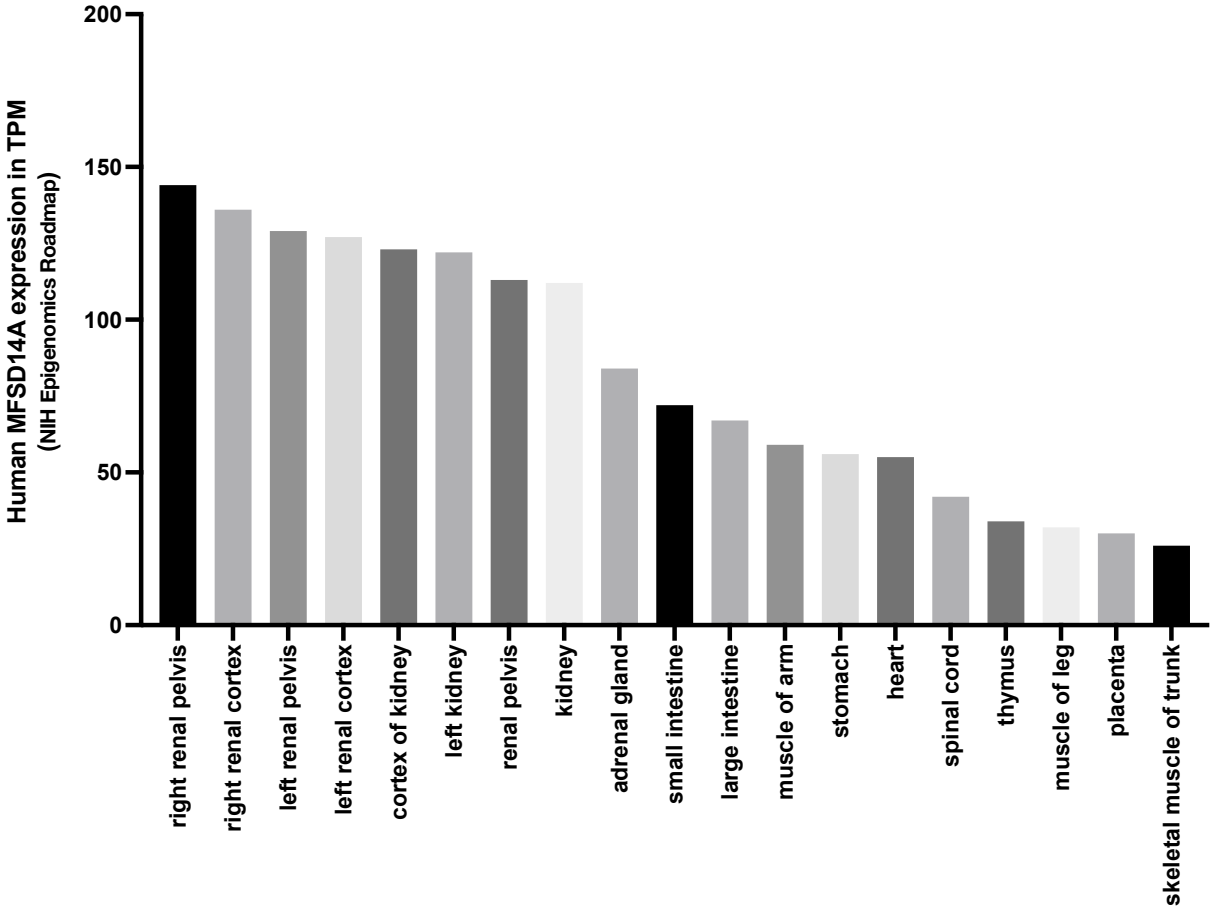


Supplementary Information Figure 4.8.3 NCBI protein report for Reference Sequence: NP_149044.2. Obtained on January 22nd 2022, from https://www.ncbi.nlm.nih.gov/protein/NP_149044.2?report=graph

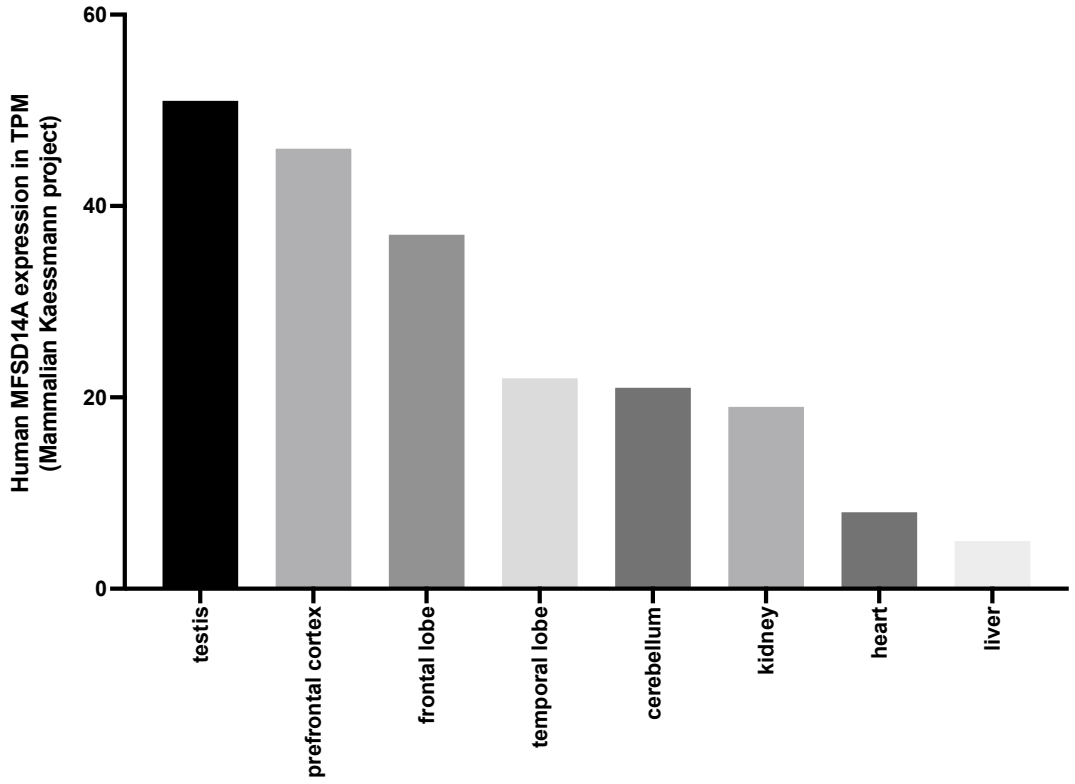


Supplementary Information Figure 4.8.3 A Expression of *MFSD14A* in adult human tissues.

Data Source: EBI expression Atlas the GTEx project. Expression levels in transcripts per million, TPM.

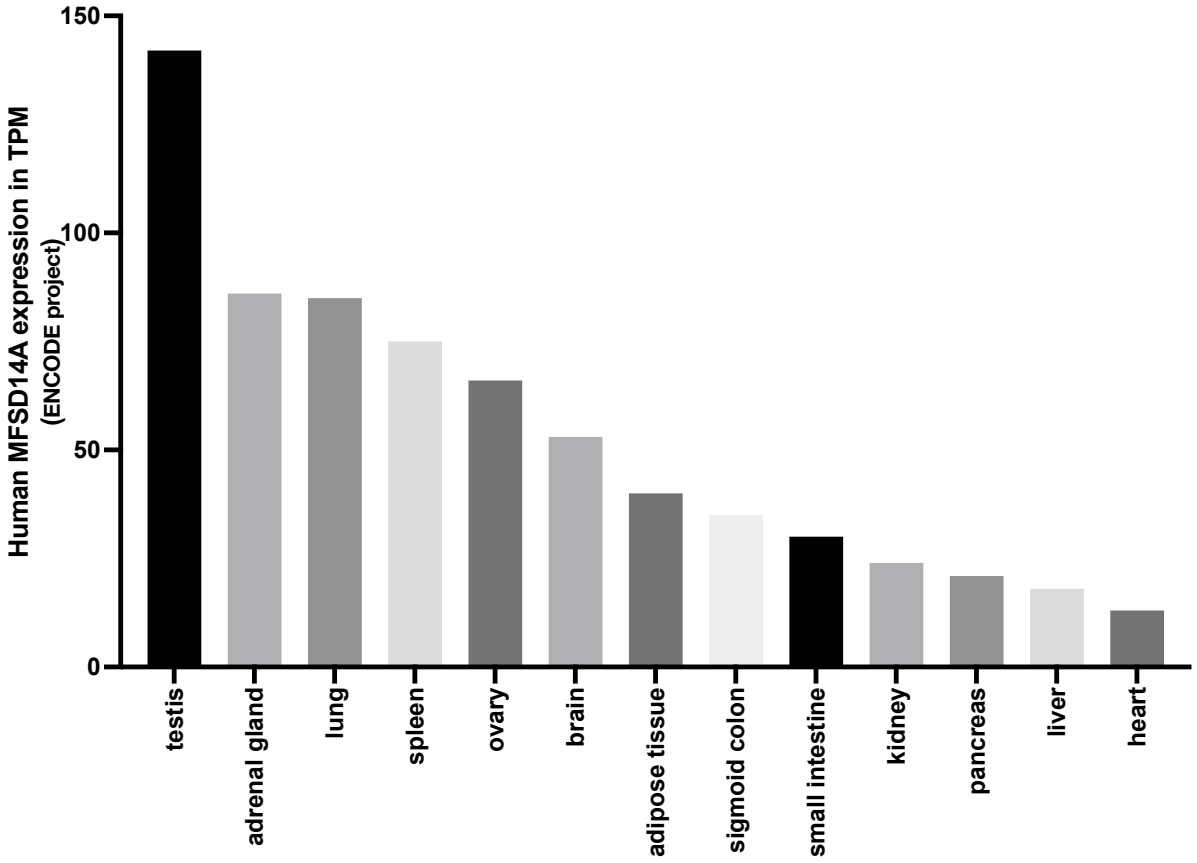


Supplementary Information Figure 4.8.3 B Expression of *MFSD14A* in adult human tissues. Data Source: EBI expression Atlas the NIH Epigenomics Roadmap project. Expression levels in transcripts per million, TPM.



Supplementary Information Figure 4.8.3 C Expression of *MFSD14A* in adult human tissues.

Data Source: EBI expression Atlas the Kaessman project. Expression levels in transcripts per million, TPM.



Supplementary Information Figure 4.8.3 D Expression of *MFSD14A* in adult human tissues.

Data Source: EBI expression Atlas the ENCODE project. Expression levels in transcripts per million, TPM.

TRANSITIONAL STATEMENT 2

In the previous two experimental chapters, we examined the genomic structure and tissue expression, and conducted transport experiments on mouse Mfsd14a and human MFSD14A. The expression profile of both genes appeared to be ubiquitous in all tested tissues with high amino acid conservation between human and mouse (99.4% amino acid identity). After expressing both mouse Mfsd14a and human MFSD14A in the *Xenopus laevis* oocytes and testing glucose transport using radiolabeled glucose, no glucose uptake was observed. Hence, we concluded that MFSD14A is not a glucose transporter and proceeded to review additional literature regarding this gene to look for alternative substrate/substrates. Based on our search, the function of MFSD14A appeared to be involved in the regulation of acid-base balance. We therefore proceeded to investigate MFSD14A as a novel ammonia transporter. The detailed experimental process and the results are presented in the next chapter.

CHAPTER 5:

MANUSCRIPT 4

CHARACTERIZATION OF TWO NOVEL AMMONIA TRANSPORTERS, HIAT1A AND HIAT1B, IN THE TELEOST MODEL SYSTEM *DANIO RERIO*

(Submitted to the Journal of Experimental Biology. Submission # JEXBIO/2022/244279)

Haonan Zhouyao[1], Alex M. Zimmer[2,3], Sandra Fehsenfeld[4], Thomas Liebenstein[5], David O. Richter[5], Gerrit Begemann[5], Peter Eck[1], Steve F. Perry[3], Dirk Weihrauch*[6]

[1]University of Manitoba, Department of Food and Human Nutritional Sciences, 66 Chancellors Circle, Winnipeg, MB - R3T 2N2, Canada

[2]University of Alberta, Department of Biological Science, 11355 Saskatchewan Drive, Edmonton, AB - T6G 2E9, Canada

[3]University of Ottawa, Department of Biology, 30 Marie Curie, Ottawa ON - K1N 6N5, Canada

[4]Université du Quebec à Rimouski, Département de biologie, chimie et géographie, 300 Allée des Ursulines, Rimouski, QC - G5L 3A1, Canada

[5]Universität Bayreuth, Faculty for Biology, Chemistry, and Earth Science, Building NW I, 95440 Bayreuth, Germany

[6]University of Manitoba, Department of Biological Sciences, 66 Chancellors Circle, Winnipeg, MB R3T 2N2, Canada

*Corresponding Author

My contribution:

I searched the literature, conducted all experiments except the in-situ hybridization, the construction of the phylogenetic tree and the phylogenic analysis. I coordinated the communications between authors, and I wrote the first draft of the manuscript.

Other author's contribution:

AZ supervised HZ during her visit in SFP's lab and provided support and training to HZ for zebrafish ammonia transport experiments and wrote the manuscript with HZ.

SF constructed the phylogenetic tree, conducted phylogenic analysis and wrote the manuscript with HZ.

SFP hosted HZ as a visiting student and provided resources and infrastructures for zebrafish ammonia transport experiments. SFP helped editing the manuscript.

TL, DOR and GB provided resources and infrastructures and conducted zebrafish embryo in-situ hybridization experiments. TL and GB helped editing the manuscript.

PE supervised HZ and provided resources and infrastructures for HZ to conduct heterologous expression studies of zebrafish Hiat transporters in the *Xenopus laevis* oocytes

DW was the principal investigator, responsible for the conception and design of the project, the acquisition of financial support, the writing and final editing of the manuscript.

5.1 ABSTRACT

Ammonia excretion in fish excretory epithelia is a complex interplay of multiple membrane transport proteins and mechanisms. With the recent identification of Hippocampus-abundant transcript 1 (Hiat1) as a novel ammonia transporter in the green crab, *Carcinus maenas,* existing models might need to be re-evaluated. Using the model system of zebrafish (*Danio rerio*) larvae we here identified three different Hiat1-paralogues found also in most vertebrates. Phylogenetic analysis revealed Hiat1b likely to be the most ancient isoform to be duplicated in elasmobranchs (Hiat1a) and again in early bony fishes (Hiat1-like). When functionally expressed in *Xenopus laevis* oocytes, DrHiat1a and DrHiat1b were promoting methylamine transport in a competitive manner with NH_4^+. *In situ* experiments showed that both proteins were ubiquitously expressed very early on in zebrafish embryos, whereas DrHiat1b showed an overall higher abundance and particularly in the intestine of the larvae. When knocked down, DrHiat1b morpholino-treated larvae exhibited significantly lower whole animal ammonia excretion rates compared to control larvae. Drastically decreased site-specific NH_4^+ flux of up to 85% could be verified for both the pharyngeal region (site of developing gills) and the yolk sac (region shown to have the highest NH_4^+ flux). In conclusion, our study is the first to identify DrHiat1a and DrHiat1b as important contributors to ammonia detoxification in early zebrafish. Being evolutionary highly conserved, these proteins are likely involved in multiple other general ammonia-handling mechanisms, making them interesting candidates for future studies on nitrogen regulation in fishes and across the animal kingdom.

5.2 INTRODUCTION

Ammonia is a toxic waste product generated by the catabolism of proteins and nucleic acids, but also *via* ureolytic and uricolytic pathways (Larsen et al., 2014) (in the current study, NH_4^+ refers to the ammonium ion, NH_3 to gaseous ammonia, and ammonia to the sum of both). Ammonia is a weak base with a pK_a of 9.0 - 9.5 and occurs accordingly in physiological solutions dominantly (> 90%) in its ionic form, NH_4^+, while only a small portion is present as the membrane permeable NH_3 (Cameron and Heisler, 1983). In fishes, elevated plasma ammonia levels caused by exercise or feeding (Bucking, 2017; Fehsenfeld and Wood, 2018; Sinha et al., 2012a; Sinha et al., 2012b) need to be tightly regulated to avoid negative physiological implications. Furthermore, even when exposed to non-lethal high environmental ammonia (HEA), the resulting internal plasma ammonia load can cause a wide range of deleterious effects in fish, including disruption of general ion homeostasis (Diricx et al., 2013; Wilkie, 1997), reduced swimming performance (McKenzie et al., 2009; McKenzie et al., 2003; Shingles et al., 2001), changes in energy metabolism (Arillo et al., 1981; Sinha et al., 2012a), and a reduced growth rate (Dosdat et al., 2003). To avoid its negative side effects, ammonia needs to be rapidly excreted from the body fluids. With a few exceptions, teleost fishes are ammoniotelic and excrete up to 90% of their metabolically produced ammonia directly through the gills (Sayer and Davenport, 1987; Smith et al., 2012; Smith, 1929; Zimmer et al., 2014).

In 2009, Wright and Wood published a working model for the branchial ammonia excretion mechanisms in a freshwater fish with the contribution of Rhesus-glycoproteins. It is assumed that plasma ammonia enters the epithelial cell in its gaseous form NH_3 *via* a basolateral localized Rhesus-glycoprotein, Rhbg, along its partial pressure gradient. Further, Mallery (1983) and Nawata et al. (2010) showed that similarly to invertebrates (Adlimoghaddam et al., 2015; Hans et al., 2018; Masui et al., 2002) and the mammalian kidney (Wall and Koger, 1994), teleost Na^+/K^+-ATPase (NKA) is capable of accepting NH_4^+ instead of K^+ as a substrate. Consequently, in addition to NH_3 uptake *via* Rhbg,

the basolateral NKA may mediate an active uptake of extracellular NH_4^+ into the brachial epithelial cell. The capacity of NKA to accept NH_4^+ as a substrate, however, seems to be dependent on external salinity and varies substantially among fishes (Wood and Nawata, 2011).

According to the model, it is further assumed that cytoplasmatic ammonia is excreted apically in a sodium-dependent manner *via* an apical Na^+/NH_4^+ exchange complex (Ito et al., 2013; Wright and Wood, 2009). This complex is thought to act as a metabolon that promotes the acidification of the unstirred boundary layer by means of a V-type H^+-ATPase and Na^+/H^+-exchanger 3, with both transporters creating a partial pressure gradient for gaseous NH_3. NH_3 follows this gradient utilizing another Rhesus-glycoprotein, Rhcgb (formerly Rhcg1, Nakada et al., 2007), a process commonly referred as ammonia trapping. A membrane-bound carbonic anhydrase (CA2a and/or CA15a, Ito et al., 2013) supports this metabolon by catalyzing the hydration of CO_2, providing the required H^+ for the process (Wright and Wood, 2009). A similar ammonia excretion mechanism has also been proposed for freshwater planarians (Weihrauch et al., 2012).

An according metabolon, however, could not be verified for freshwater leeches, where the cutaneous ammonia excretion process was shown to be independent of functional apical Na^+ channels and NHEs (Quijada-Rodriguez et al., 2017). Furthermore, in the nematode *Caenorhabditis elegans*, a knock-out of the apical localized Rh-protein RHR-2 caused a reduction in ammonia excretion rates, the remaining flux was, however, independent of the environmental pH, in contrast to the wild-type (Adlimoghaddam et al., 2016). More importantly, knocking out the apical Rhcgb in zebrafish *via* CRISPR/Cas9 did not reduce ammonia excretion and/or Na^+- uptake. Rather, an increase in ammonia excretion was observed (Zimmer and Perry, 2020). The diversity of ammonia-excretory mechanisms across fishes and other aquatic animals indicates the importance of these processes to effectively regulate ammonia waste, and at the same time warrants to consider additional routes and mechanisms to be involved.

In invertebrates, for example, NH_4^+-transporting ammonium transporters (AMTs, also members of the AMT/MEP/Rh-protein ammonia transporter family) driven by the cytoplasm-to-environment electrochemical gradient of the cation, have been suggested to be involved in the epithelial ammonia excretion process (Boo et al., 2018; Chasiotis et al., 2016; Durant and Donini, 2018; Pitts et al., 2014; Thiel et al., 2017). In vertebrates, however, AMTs have not been shown to be expressed.

Alternatively, a potential candidate for participating in the branchial ammonia excretory process may be the hippocampus-abundant transcript 1 (Hiat1). This potential ammonia transporter was initially associated with acid-base regulation in decapod crabs, as it was down regulated in green crabs, *Carcinus maenas*, when the animals were acclimated to elevated P_{CO2} levels (Fehsenfeld et al., 2011). Just recently, CmHIAT1 has been identified to indeed act as an ammonia (likely NH_4^+) transporter in this species (Fehsenfeld et al., 2022 under review) being differentially expressed in response to high environmental ammonia. In preliminary experiments, Hiat1 was shown to be expressed also in the gills of Arctic charr, *Salvelinus alpinus*. A heterologous expression analysis employing *Xenopus laevis* oocytes revealed methylamine transport capabilities of this membrane protein (Weihrauch and Zhouyao, unpublished observations).

According to these recent findings, we here hypothesized that Hiat1 is involved in ammonia excretion in zebrafish. Based on its high level of evolutionary conservation and endogenous expression in *Xenopus laevis* oocytes, we further hypothesized Hiat1 to play an important role in ammonia detoxification early on in development. To address these hypotheses, we investigated Hiat1 function in zebrafish (*Danio rerio*) larvae, applying multiple physiological and molecular techniques. After identifying and characterizing Hiat1 paralogues in a phylogenetic analysis, we expressed DrHiat1a and DrHiat1b proteins in frog oocytes to verify their capability of transporting methylamine (ammonia analog) competitively with NH_4Cl. We then determined expression patterns of both paralogs

throughout early zebrafish development (4-cell stage to 4 days post fertilization) by *in situ* hybridization analysis. Lastly, we generated morpholino knock-downs for DrHiat1a and DrHiat1b and used the Scanning Ion-selective micro-Electrode Technique (SIET) to identify regions of interest of Hiat1-related ammonia transport in the *D. rerio* larvae in early development.

5.3 METHODS

5.3.1 Sequence-based genetic structure analysis

DrHiat1-like (GenBank accession no. XP_002663499.2; 500 aa), DrHiat1b (GenBank accession no. AAH97075.1, 485 aa), and the translated DrHiat1a protein based on the open reading frame (ORF) of zebrafish *DrHiat1a* (GenBank accession no. BC056817, 493 aa) were aligned with Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/). Protter (Omasits et al., 2014) was used to predict transmembrane domains (TM) and the potential phosphorylation sites for all three Hiat1 proteins, and Scansite 4.0 (Obenauer et al., 2003) was used to identify the binding motifs. All transmembrane models and proteoform predictions were performed using the system's default settings.

DrHiat1b was then used to identify HIAT1 isoforms in other species *via* GenBank-BLASTp search (**Supplementary Information Table 5.11.1**) (Altschul et al., 1997). Proteins were aligned with the MUSCLE algorithm as provided by the Molecular Evolutionary Genetics Analysis across computing platforms (MEGA X (Kumar et al., 2018)) using default settings. The evolutionary history was inferred by using the Maximum Likelihood method and JTT matrix-based model (Jones et al., 1992) as identified by the included testing function provided by MEGA X. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.3614)). All positions with less than 95% site coverage were eliminated, i.e., fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position (partial deletion option).

5.3.2 Heterologous expression of DrHiat1a and DrHiat1b in *Xenopus* oocytes

cDNA clones containing the full ORFs of *DrHiat1a* (clone ID: 2640969) and *DrHiat1b* (clone ID: 2600163) were purchased from Dharmacon (Lafayette, CO, USA). DNA miniprep kits (Qiagen, Hilden, Germany) were used to extract the plasmids following the manufacturer's guidance. Regular PCR employing Q5 High-Fidelity DNA Polymerase was used to obtain the full ORFs of *DrHiat1a* and *DrHiat1b* (see **Supplementary Information Table 5.11.2**, ORF). The annealing temperature used for the amplification of *DrHiat1a* and *DrHiat1b* ORFs were 60°C and 62°C, respectively with a total of 35 cycles. Subsequently, regular PCR was used to add restriction enzyme sites to the 5' and 3' ends of the ORFs of *DrHiat1a* and *DrHiat1b* (see **Supplementary Information table 5.11.2,** oocytes). The annealing temperature used for the amplification of *DrHiat1a* and *DrHiat1b* ORFs were both 65°C with a total of 35 cycles. The size and the integrity of the DNA were assessed by gel electrophoresis and the concentration of DNA was determined spectrophotometrically after purification (GeneJET PCR purification kit, ThermoFisher, Waltham, MA, USA). The ORFs were then cloned into the pGEM®-HE plasmid, a modified pGEM®-3Z vector containing the *Xenopus* beta globin 5'- and 3'- UTR sequences using T4 ligase following the manufacturer's guidance (New England Biolabs (NEB), Ipswich, MA, USA). The *in vitro* transcription of capped mRNA (cRNA) was performed with HiScribe™ T7 ARCA mRNA kit (NEB) followed with column purification (RNeasy MinElute Cleanup Kit, Qiagen, Hilden, Germany). The cRNA was quantified spectrophotometrically (Nanodrop, ND-1000, Thermofisher, Waltham, MA, USA) and integrity was assessed on a MOPS agarose gel containing formaldehyde.

Stage VI-V oocytes were collected from mature female *Xenopus laevis* (VWR, Radnor, PA, USA) as previously described (Soreq and Seidman, 1992) (Protocol Reference# F20-021). Briefly, the frogs were euthanized *via* decapitation prior to the collection of the ovary. The ovary was placed in Ca^{2+}-free OR2 (contained (in mmol l^{-1}) 82.5 NaCl, 2.5 KCl, 1 $MgCl_2$, 1 Na_2HPO_4, 5 HEPES, pH 7.2)

solution with collagenase type VI (1 mg ml^{-1}) (Gibco, Waltham, MA, USA) and gently agitated for 90 minutes at room temperature. 1 mmol l^{-1} $CaCl_2$ was added to terminate the collagenase activity. Oocytes were then manually sorted and rinsed three additional times with standard OR2 solution. For overnight storage of isolated oocytes, OR2 was supplemented with 2.5 mmol l^{-1} sodium pyruvate, 1 mg mL^{-1} penicillin-streptomycin (Gibco, Long Island, NY, USA) and 50 µg ml^{-1} gentamicin at 16 °C. All procedures used were approved by the University of Manitoba Animal Care Committee. Animal care procedures were based on the guidelines described in the Canadian Council for Animal Care.

Isolated oocytes after the overnight recovery were injected with 18.4 ng of cRNA (36.8 nl with 0.5 ng nl^{-1}) or nuclease-free water as control using the Nanoject II auto-nanoliter injector (Drummond Scientific, Broomall, PA, USA). Experiments were conducted 48 hours post-injection.

5.3.3 [H^3] methylamine transport studies in oocytes expressing DrHiat1a and DrHiat1b

5.3.3.1[H^3] methylamine uptake experiments

Experiments were conducted at room temperature in 200 µl of standard OR2, pH 7.2, containing 3 µmol l^{-1} (9.3 µCi) [H^3]-methylamine (Moravek Inc., USA) with 60 minutes of incubation time. For each experiment 20 oocytes injected with either with cRNA or water (sham) were randomly selected. At the end of each experiment, ice-cold standard OR2 was used to terminate [H^3] methylamine uptake by washing oocytes three times to remove the external radioactivity. The washed oocytes were immediately solubilized individually in 200 µl of 10% SDS (Sigma) to assess the internal radioactivity employing 5 ml of Ultima-Gold scintillation cocktail (PerkinElmer) and liquid scintillation counting (Tri-Carb 2900 TR; PerkinElmer). The effect of washing was determined by assessing the radioactivity in 200 µl standard OR2 used for the third washing post radioactive exposure compared to fresh, radioactivity-free standard OR2 at the same volume. For NH_4Cl competitive

uptake experiments, the same procedures described above were conducted, however in the presence of 1 mmol l^{-1} NH_4Cl added to the incubation solution.

5.3.3.2[H^3] methylamine release experiments

To determine whether Hiat proteins mediate bi-directional methylamine transport, the efflux of methylamine from SHAM oocytes, DrHiat1a-, and DrHiat1b-expressing oocytes were measured. Experiments were performed at room temperature in 200 µl of standard OR2 containing 3 µmol l^{-1} (9.3 µCi) [H^3] methylamine adjusted to pH 7.2 with 60 minutes of incubation time. Two groups of cRNA injected oocytes (n = 20) and two groups of water injected oocytes (n = 20) were assessed for each experimental point. Each group was considered as one replicate and overall, two replicates were done. After termination of [H^3] methylamine uptake as described above, one group of cRNA injected oocytes and one group of water injected oocytes were immediately solubilized individually to assess the uptake of [H^3] methylamine as described above, to determine the starting point for the following release experiment. The remaining groups of cRNA injected oocytes and water injected oocytes were placed in 200 µl of room temperature, radioactivity-free standard OR2 for 60 minutes for release and afterwards washed three times with ice-cold standard OR2, solubilized individually in 200 µl of 10% SDS and assess for remaining internal radioactivity. The efficiency of the washing steps was determined by assessing 200 µl of the radioactivity of the third washing post radioactive exposure compared to fresh, radioactivity-free standard OR2 at the same volume.

5.3.4 *In situ* hybridization

Dr Hiat1a and *DrHiat1b* ORFs were amplified and cloned as described above with Q5 High Fidelity DNA Polymerase (NEB) and primers located at both ends of the ORFs with restriction enzyme sites (**Supplementary Information Table 5.11.2,** *in situ*).

The deletion of the *Xenopus* beta globin 5' UTR was done by removing a 63 bp fragment containing the 5' *Xenopus* UTR with KpnI and XmaI followed by blunting with Quick Blunting Kit (NEB) followed by re-ligation. The *Xenopus* beta globin 3' UTR was removed in the course of the linearization of the vector for the preparation of antisense-RNA probe. The antisense-RNA probe was synthesized as previously described (Thisse and Thisse, 2008) and experiments were conducted at the University of Bayreuth, Germany, using the Casper strain (White et al., 2008). Adult fish were maintained under standard conditions (14:10 h light:dark-cycle; 0.1 g l^{-1} salinity; 300 µS conductivity; pH 7.5) at 27.5 - 28.5 °C. Embryos were obtained from pairing events of 3 male and 2 female fish and were reared in E3-medium (contained (in mmol l^{-1}) 5 NaCl, 0.17 KCl, 0.33 $CaCl_2$, 0.33 $MgSO_4$, 10-5% methylene blue; pH 7.2) at 28.5 °C.

Zebrafish embryos and larvae were euthanized on ice, fixed in 4 % PFA/PBS, washed in PBS/0,1 % Tween 20 (PBTw), transferred to 100 % methanol and stored at least overnight at -20 °C. After rehydration using a methanol/PBTw series and post-fixation in 4 % PFA/PBS, older developmental stages were treated with 10 µg ml^{-1} proteinase K (Fermentas, Waltham, MA, USA) for 10 min (2 dpf), 15 min (3 dpf) and 20 min (4 dpf), respectively, post-fixed in 4% PFA/PBS and washed several times in PBTw. After hybridization with the corresponding antisense probes overnight at 68 °C, the samples were washed in a formamide solution/PBTw series, blocked in 0.5% blocking reagent, incubated in a 1:2000 dilution of anti-DIG-AP antibody in 0.5% blocking reagent for 4 h at room temperature and washed several times in PBTw. For detection, samples were equilibrated in BCL buffer and stained with NBT and BCIP/X-phos until staining was clearly visible.

5.3.5 Ammonia excretion studies in zebrafish with DrHiat1a or DrHiat1b knockdown

5.3.5.1 Zebrafish

Danio rerio originally purchased from the pet trade (Big Al's Aquarium, Ottawa, ON, Canada) were maintained in flow-through aquaria receiving dechloraminated city of Ottawa tap water ("system water"; in mmol l^{-1}: 0.8 Na^+, 0.4 Cl^-, 0.25 Ca^{2+}; pH = 7.6) and held at 28.5 ºC with a photoperiod of 14:10 h light:dark. Embryos were obtained from breeding events between 1 male and 2 females held in static tanks. Embryos were reared in 60 ml Petri dishes containing system water with 0.05% methylene blue at a density of approximately 50 embryos per dish held at 28 ºC. Handling and experimentation of zebrafish was conducted in compliance with the University of Ottawa Animal Care and Veterinary Service (ACVS) guidelines under protocol BL-1700 and followed the recommendations from the Canadian Council for Animal Care (CCAC).

5.3.5.2 Morpholino injections

Splice-blocking morpholino oligonucleotides (Gene Tools, LLC, Philomath, OR, USA) targeting the intron-exon junction between intron and exon 4 of *DrHiat1a* (NCBI accession: NM_199584.1; 5'-CATGGTGAGAGCCTCTGAAATCAAG-3'), targeting the intron-exon junction between intron and exon 4 of *DrHiat1b* (NCBI accession: NM_213527.2; 5'-AAAGTGAGGAGCGTAACGAACCATG-3'), and a standard control morpholino with no biological target in zebrafish (5'-CCTCTTACCTCAGTTACAATTTATA-3') were suspended to a final concentration of 2-4 ng morpholino nl^{-1} in Danieau buffer (in mmol l^{-1}: 58 NaCl, 0.7 KCl, 0.4 $MgSO_4$, 0.6 $Ca(NO_3)_2$, and 5.0 HEPES buffer; pH 7.6) containing 0.05% phenol red for injection visualization. 1-cell stage embryos were injected with 1 nl of this injection solution [4 ng/embryo for DrHiat1a morpholino (MO), 4 ng/embryo for DrHiat1b MO, and 4 ng/embryo for control (sham)

MO] using an IM 300 microinjection system (Narishige, Amityville, NY, USA). However, we observed a large proportion (>50%) of morphological abnormalities (curved tail/spine) with the 4 ng DrHiat1b dose and therefore used a dose of 2 ng for DrHiat1b and observed virtually no morphological abnormalities (<5-10%). Injections were verified at 1 dpf by examining the distribution of carboxyfluorescein (tagged to each morpholino) in individual embryos under a fluorescence dissecting microscope (Nikon SMZ1500, Nikon Instruments, Melville, NY, USA). Embryos that were successfully injected, containing a homogenous distribution of fluorescein signal, were reared as described above.

Knockdown was confirmed by reverse transcription PCR (RT-PCR). At 4 days post-fertilization (dpf), approximately 20 pooled larvae (n=1) from each treatment group (DrHiat1a MO, DrHiat1b MO, and sham MO) were euthanized by an overdose of neutralized tricaine methanesulfonate (MS-222; Syndel Canada, Nanaimo, BC, Canada), flash frozen in liquid nitrogen, and stored at -80 ºC. RNA was extracted from pooled samples using Trizol (ThermoFisher, Waltham, MA, USA) following the manufacturer protocol. Extracted RNA was treated with DNAse (ThermoFisher) and cDNA was then synthesized from 700 ng total RNA using the iScript cDNA synthesis kit (Bio-Rad, Hercules, CA, USA). RT-PCR was performed using DreamTaq DNA polymerase (ThermoFisher) following manufacturer's guidance and primers specific to DrHiat1a and DrHiat1b (see **Supplementary Information Table 5.11.2**) The reaction volume for each PCR reaction was 20 µl with 55°C annealing temperature and the reaction was terminated after 40 cycles.

5.3.5.3 Whole-body ammonia flux

At 4 dpf, DrHiat1a, DrHiat1b, and sham MO-treated larvae were placed into 2 ml centrifuge tubes containing 1.5 mL system water (6 larvae/tube; n=1) maintained at 28-30 ºC in a water bath.

Larvae were allowed to adjust for 30 min where thereafter, initial 0.5 ml water samples were drawn from each tube, marking the beginning of the flux period. Following a 6 h flux period, final 0.5 ml water samples were drawn from each tube. Samples were frozen and stored at -20 ºC for no longer than 1 week. Total ammonia concentration (T_{Amm}) was measured using the indophenol method (Verdouw et al., 1978) and ammonia excretion rate (nmol larva^{-1} h^{-1}) was calculated using the following equation:

$$\text{Ammonia excretion} = [(T_{Amm}f - T_{Amm}i)*V]/t/n \qquad \text{(Eqn 1)}$$

where $T_{Amm}f$ and $T_{Amm}i$ (nmol ml^{-1}) are the final and initial total ammonia concentrations, respectively, V is volume (ml), t is flux duration (h), and n is the number of larvae in the tube.

5.3.5.4 Scanning Ion-selective micro-Electrode technique (SIET)

SIET was performed to assess NH_4^+ fluxes by 4 dpf larvae treated with sham and DrHiat1b morpholino. Measurements were made in a K^+-free medium which was made by adding stock solutions of Na_2SO_4, $CaCl_2$, $MgSO_4 \cdot 7H_2O$, Na_2HPO_4, and NaH_2PO_4 to doubly distilled water (in mmol l^{-1}: 0.84 Na^+, 0.5 Cl^-, 0.25 Ca^{2+}, 0.15 Mg^{2+}; pH = 7.6). The omission of K^+ in the medium is necessary because this ion interferes with NH_4^+ signals from the ion-selective probe. In addition, the medium contained MS-222 anaesthetic (0.2 g l^{-1}) and 0.05 mmol l^{-1} $(NH_4)_2SO_4$, the latter being necessary because the Nernstian slope of the NH_4^+ tends to fall at concentrations below 0.1 mmol l^{-1}. NH_4^+-selective probes were constructed as described previously (Donini and O'Donnell, 2005). Glass capillary tubes (World Precision Instruments, Sarasota, FL, USA) were pulled to microelectrodes with a tip diameter of approximately 5 μm using a P-2000 micropipette puller (Sutter Instrument, Novato, CA, USA) and silanized with N,N-Dimethyltrimethylsilylamine (Sigma-Aldrich, St. Louis, MO, USA) on a hot plate covered with a glass Petri dish. Silanized microelectrodes were then back-loaded with

100 mmol l^{-1} NH_4Cl and front-loaded with an approximately 250 μm ionophore column of NH_4^+ Ionophore I Cocktail A (Sigma-Aldrich). Probes were calibrated in the same K^+-free medium containing MS-222 described above to which NH_4Cl was added to final concentrations of 0.1, 1, or 10 mmol l^{-1}. pH was maintained at 7.6 across all calibration solutions using NaOH or H_2SO_4. The slope (in mV) for a tenfold change in NH_4^+ concentration was 59.2 ± 1.1 (n = 3).

Larvae were allowed to adjust to the measurement solution for 15-20 min prior to flux measurements where thereafter they were restrained in a measurement chamber described previously (Hughes et al., 2019) that prevented larvae from drifting during measurements. NH_4^+ fluxes were measured at the apex of the yolk sac epithelium, similar to previous studies (Shih et al., 2012) and at the pharyngeal region, the latter measurement used as an approximation of gill NH_4^+ flux. Voltage from the NH_4^+-selective probe was measured approximately 2-5 μm away from the epithelial surface (origin) and at an excursion distance of 100 μm away from the epithelium (excursion). Each scan consisted of 5 replicate origin and extrusion measurements and scans were replicated 3 times at the yolk sac and at the jaw for each fish (n = 1). For each fish, a background scan was conducted at the beginning and end of the measurement at a location approximately 1 cm away from the larva. Voltage gradients were converted to concentration gradients using the following equation (Donini and O'Donnell, 2005):

$$\Delta C = C_B \times 10^{(\Delta V/S)} - C_B \qquad \text{(Eqn 2)}$$

where ΔC is the concentration gradient (μmol/cm3) between the origin and excursion points, C_B is the average background ion concentration measured by the probe at all points, ΔV is the voltage difference (mV) between the origin and excursion points, and S is the slope obtained from the probe calibration (mV). Note that ΔV of the background solution was subtracted from that measured at the epithelium. Flux was then calculated using Fick's law of diffusion:

$$J = D \times \Delta C / \Delta x \qquad \text{(Eqn 3)}$$

where J is NH_4^+ flux (pmol/cm^2/s), D is the diffusion coefficient for NH_4^+ (2.09 x 10^{-5} cm^2/s), and Δx is the excursion distance (cm).

5.3.6 Statistics

All data were tested for normal distribution by Shapiro-Wilk's test. In case data were not normally distributed, data were log transformed. Homogeneity of variances was assessed by Levene's test. If data did not meet the criteria for parametric testing including normal distribution and homogeneity of variances, non-parametric testing wase applied. Statistical analyses were performed with PAST3 (Hammer et al., 2001). Graphs were generated with GraphPad Prism 9.0 for Windows (GraphPad Software, San Diego, California USA, www.graphpad.com) and Inkscape (https://inkscape.org).

5.4 RESULTS

5.4.1 Genetic analysis of Hiat1

The three isoforms of DrHiat1 (DrHiat1a with 493aa, DrHiat1b with 485aa, DrHiat-like with 500aa) exhibited a conservation of 323 identical amino acids (ca. 66%), 80 residues as strong groups (ca. 16%), and 19 residues as weak groups (ca. 4%). The structure analysis of DrHiat1a, DrHiat1b, and DrHiat1-like by Protter (Omasits et al., 2014) identified 12 predicted transmembrane domains (TM). Furthermore, a sugar binding motif (aa91-108 for DrHiat1a, aa89-106 for DrHiat1b, aa103-120 for Hiat1-like) seems to be associated with TM3 in all three cases. While DrHiat1a and DrHiat1b contained two equivalent phosphoserine/threonine binding groups (aa5-21 / aa5-19, N-terminal; aa299-312 / aa297-310, between TM8 and TM9), DrHiat1b contained an additional one between TM1 and TM2 (aa55-70), but DrHiat1-like did not contain any at all. DrHiat1b was the only isoform to contain a proline-dependent serine/threonine kinase group (Cdk5, aa456-469) (**Supplementary Figure 5.11.1**).

The phylogenetic analysis clearly confirmed three distinct clades according to the three different paralogues. Lampreys, as the most ancient species included, possess only one isoform, Hiat1b. Hiat1a first appeared with the elasmobranchs, while Hiat1-like could be found in teleosts onwards but seemed to be lost in terrestrial reptilians as well as terrestrial and aquatic mammals (**Figure 5.1**).

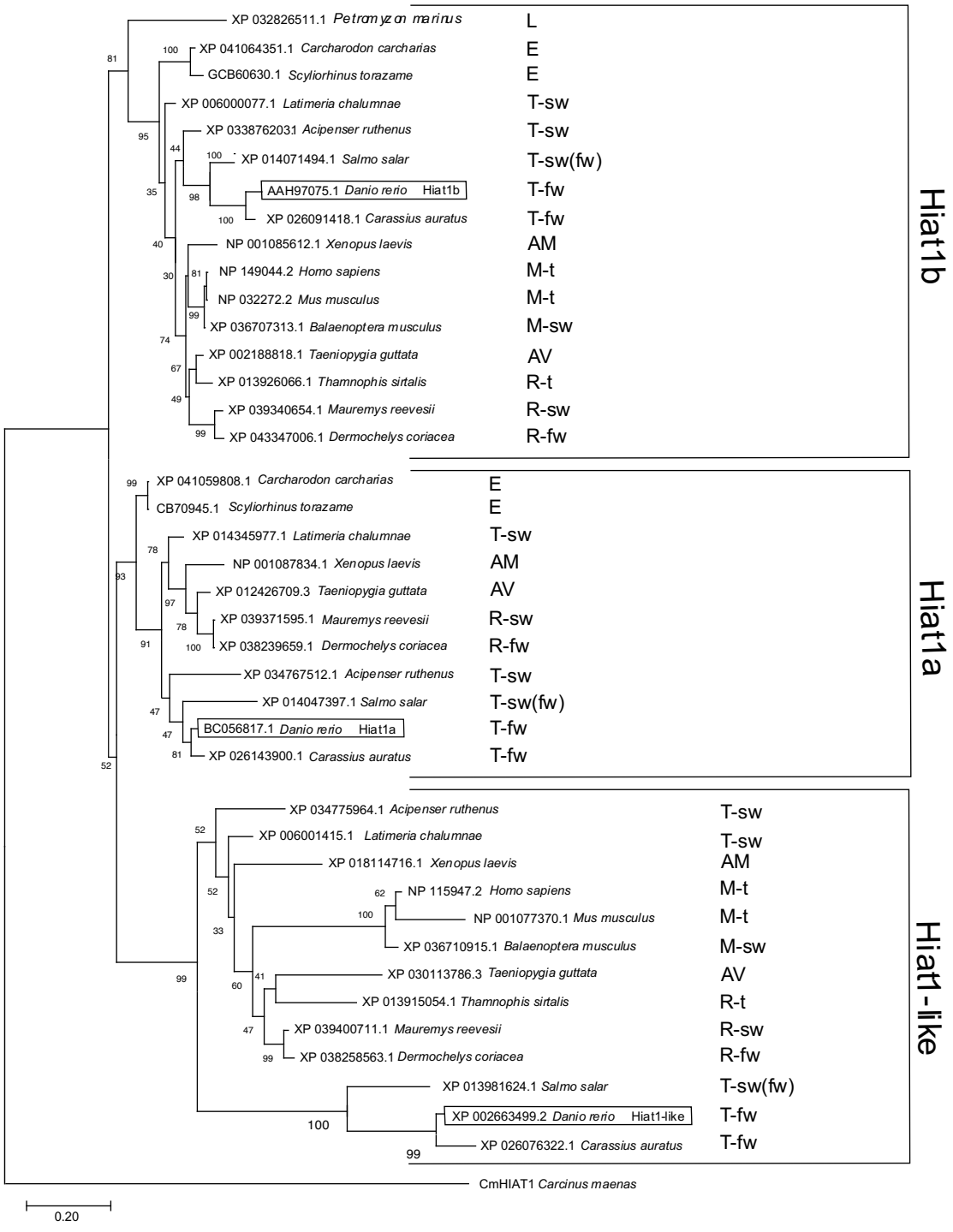

XP 032826511.1 Petromyzon marinus
XP 041064351.1 Carcharodon carcharias
GCB60630.1 Scyliorhinus torazame
XP 006000077.1 Latimeria chalumnae
XP 033876203.1 Acipenser ruthenus
XP 014071494.1 Salmo salar
AAH97075.1 Danio rerio Hiat1b
XP 026091418.1 Carassius auratus
NP 001085612.1 Xenopus laevis
NP 149044.2 Homo sapiens
NP 032272.2 Mus musculus
XP 036707313.1 Balaenoptera musculus
XP 002188818.1 Taeniopygia guttata
XP 013926066.1 Thamnophis sirtalis
XP 039340654.1 Mauremys reevesii
XP 043347006.1 Dermochelys coriacea
L
E
E
T-sw
T-sw
T-sw(fw)
T-fw
T-fw
AM
M-t
M-t
M-sw
AV
R-t
R-sw
R-fw
Hiat1b
XP 041059808.1 Carcharodon carcharias
CB70945.1 Scyliorhinus torazame
XP 014345977.1 Latimeria chalumnae
NP 001087834.1 Xenopus laevis
XP 012426709.3 Taeniopygia guttata
XP 039371595.1 Mauremys reevesii
XP 038239659.1 Dermochelys coriacea
XP 034767512.1 Acipenser ruthenus
XP 014047397.1 Salmo salar
BC056817.1 Danio rerio Hiat1a
XP 026143900.1 Carassius auratus
E
E
T-sw
AM
AV
R-sw
R-fw
T-sw
T-sw(fw)
T-fw
T-fw
Hiat1a
XP 034775964.1 Acipenser ruthenus
XP 006001415.1 Latimeria chalumnae
XP 018114716.1 Xenopus laevis
NP 115947.2 Homo sapiens
NP 001077370.1 Mus musculus
XP 036710915.1 Balaenoptera musculus
XP 030113786.3 Taeniopygia guttata
XP 013915054.1 Thamnophis sirtalis
XP 039400711.1 Mauremys reevesii
XP 038258563.1 Dermochelys coriacea
XP 013981624.1 Salmo salar
XP 002663499.2 Danio rerio Hiat1-like
XP 026076322.1 Carassius auratus
T-sw
T-sw
AM
M-t
M-t
M-sw
AV
R-t
R-sw
R-fw
T-sw(fw)
T-fw
T-fw
Hiat1-like
CmHIAT1 Carcinus maenas
0.20

Figure 5.1. Phylogenetic analysis of Hiat1-like, Hiat1a and Hiat1b across vertebrates. Shown is the Maximum likelihood consensus tree of the MUSCLE alignment (Edgar, 2004) of Hiat1 proteins as performed with MEGA X (Kumar et al., 2018). Numbers beside branches represent bootstrap values (1000 replicates). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site (indicated by bar at bottom). AM, amphibia; AV, aves; E, elasmobranchii; L, lamprey; M-sw, marine mammalia; M-t, terrestrial mammalia; R-fw, freshwater reptilia; R-sw, seawater reptilia; R-t, terrestrial reptilia, T-fw, freshwater teleostei; T-sw, seawater teleostei. *Salmo salar* has been indicated as both marine and freshwater as it returns to freshwater for reproduction. The tree has been rooted by including the recently identified Hiat-protein in the green crab, *Carcinus maenas* (Fehsenfeld et al., 2022 under review).

5.4.2 DrHiat1-mediated methylamine (MA) transport in *Xenopus laevis* oocytes

Both DrHiat1a and DrHiat1b-expressing oocytes showed significantly higher (2.7-fold) MA uptake compared to the SHAM (oocytes injected with water, negative control). There was no difference in the uptake of MA between oocytes expressing DrHiat1a or DrHiat1b (**Figure 5.2 A**; ANOVA with Tukey's pairwise comparisons with $P < 0.001$, $N = 20$).

When exposed to 1 mmol l^{-1} non-labeled NH_4Cl in addition to MA in the medium, the uptake of MA mediated by DrHiat1a and DrHiat1b was reduced by 45% and 40%, respectively. The degree of inhibition of MA by NH_4Cl was not significantly different between oocytes expressing DrHiat1a or DrHiat1b (**Figure 5.2 B**).

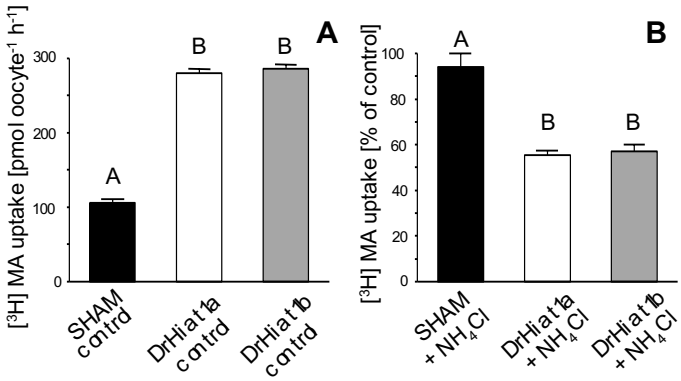


Figure 5.2 DrHiat1a and DrHiat1b-mediated methylamine/methylammonium (MA) uptake in *Xenopus laevis* oocytes. (A) Absolute [^{3}H]-methylamine/methylammonium (MA) uptake of sham-injected oocytes (black bar) and oocytes expressing DrHiat1a (white bar) or DrHiat1b (dark grey bar) under control conditions (medium pH 7.2). (B) Relative [^{3}H] MA uptake of sham-injected oocytes (black bar) and oocytes expressing DrHiat1a (white bar) or DrHiat1b (dark grey bar) in medium containing 1 mmol L^{-1} NH_4Cl. Values in (B) are depicted as % of the corresponding control values as shown in figure 2A. Data are shown as means ± s.e.m. Uppercase letters indicate significant differences as identified with one-way ANOVA and Tukey's post-hoc test ($P < 0.001$, $N = 19,20,20$).

Pre-loaded DrHiat1a- and DrHiat1b-expressing oocytes released significantly more ammonia compared to SHAM-injected oocytes (27% *vs.* 15%, **Figure 5.3**). There was no difference between the amounts of MA released in oocytes expressing Hiat1a compared to Hiat1b.

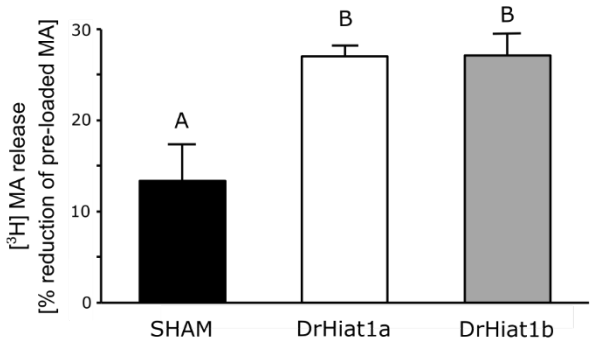


Figure 5.3 Relative methylamine/methylammonium (MA) release of pre-loaded *Xenopus laevis* oocytes expressing DrHiat1a or DrHiat1b. Oocytes were pre-loaded in MA-containing medium for 60 min after which internal oocyte MA concentration was measured (MA_{pre}). Subsequently, oocytes were transferred to MA-free medium and re-counted after additional 60 min (MA_{rel}). The graph shows the MA release relative to pre-loaded values calculated as MA release = 100% - (MA_{rel} / MA_{pre} * 100%). Data are shown as means ± s.e.m. Uppercase letters indicate significant differences as identified by Kruskal Wallis test with Mann-Whitney pairwise and Bonferroni corrected comparisons ($P < 0.03$, N = 20,20,20).

5.4.3 *In situ* hybridization of Hiat1 isoforms in zebrafish embryos and larvae

To determine Hiat1 expression in zebrafish, DrHiat1a and DrHiat1b antisense-RNA probes were generated and whole mount *in situ* hybridizations (WISH) was performed on zebrafish embryos and larvae of different developmental stages, beginning from the 4-cell stage until 4 days post-fertilization (dpf) (**Figure 5.4**). Maternal expression of both, DrHiat1a (**Figure 5.4 A-A**) and DrHiat1b (**Figure 5.4 B-A**) was observed, early onset of transcription during maternal-to-zygotic transition (**Figure 5.4 A-B and Figure 5.4 B-B**) and ubiquitous expression in all tissues during early segmentation stages (**Figure 5.4 A-C and C'; Figure 5.4 B-C and C'**). At 24 hours post-fertilization (hpf), towards the end of the segmentation period, expression of both genes is mainly localized to the head region, especially to the fore-, mid- and hindbrain including the midbrain-hindbrain-boundary, the lens and the trigeminal ganglion (**Figure 5.4 A-D, D' and figure 5.4 B-D, D'**). In addition, DrHiat1a shows distinct expression in the olfactory placode and the hatching gland in both examined segmentation stages (**Figure 5.4 A- C-D'**), while DrHiat1b shows additional expression in the optic vesicle, posterior lateral line ganglion and pectoral fin bud at 24 hpf (**see Supplementary Figure 5.11.1**). From 2 dpf to 4 dpf DrHiat1a and DrHiat1b expression can be detected in the brain, heart, gills, digestive system, pectoral fins and parts of the sensory nervous system including cranial and lateral line ganglia (DrHiat1a and DrHiat1b) and the retinal ganglion cell layer and lateral line neuromasts (DrHiat1b) (**Figure 5.4 A: E-F'; Figure 5.4 B: E-F' and Supplemental Figure 5.11.1**). For the digestive system, DrHiat1a appears to be more strongly expressed in the pancreas, whereas DrHiat1b is more present in the intestine (**Supplementary Figure 5.11.1, E-H**). Two observations stand out: First, beginning from maternal-to-zygotic transition, DrHiat1a required up to 5 times longer staining compared to DrHiat1b (while maternal staining developed simultaneously for both probes), which indicates a much lower expression level of DrHiat1a compared to DrHiat1b. Second, due to the proteinase K treatment, which was performed from 2 dpf onwards, staining of some of the more

external structures, like pectoral fins or lateral line neuromasts, was reduced or even completely abolished. For images without proteinase K treatment see **Supplementary Figure 5.11.1**.

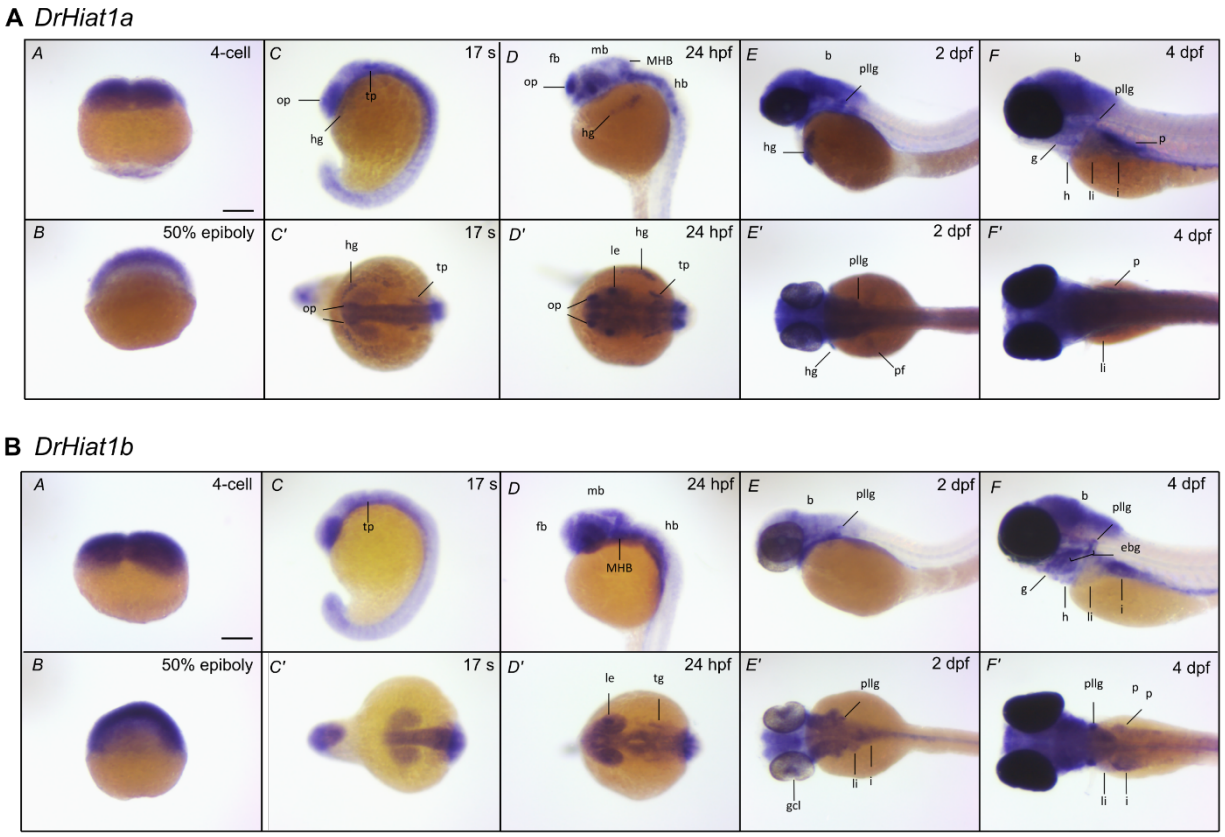


Figure 5.4. Expression pattern of *DrHiat1a* and *DrHiat1b* in zebrafish embryos and larvae during early development. Whole mount *in situ* hybridization of zebrafish, *Danio rerio*, DrHiat1a (A) and DrHiat1b (B) at the 4-cell stage, 50% maternal-to-zygotic transition (50% epiboly), 17 somite stage (17s), and 24 hours / 2 days / 4 days post-fertilization (24 hpf / 2 dpf / 4 dpf). *A, B*: lateral view, animal pole to the top. *C-F*: lateral view, anterior to the left. *C'-F'*: dorsal view, anterior to the left. b: brain; fb: forebrain; g: gills; h: heart; hb: hindbrain; hg: hatching gland; i: intestine; le: lens; li: liver; mb: midbrain; MHB: midbrain-hindbrain-boundary; op: olfactory placode; p: pancreas; pf: pectoral fin; pllg: posterior lateral line ganglion; tg: trigeminal ganglion; tp: trigeminal placode. Scale bar (for all images): 200 μm.

5.4.4 Ammonia excretion in zebrafish larvae with DrHiat1 knockdown

Successful knockdown of Hiat1a and Hiat1b using splice-blocking morpholinos was verified by PCR where reduced sizes of amplicons (from 450bp to 300bp) indicate successful splice blocking of the 4th exon in pre-mRNA DrHiat1a and DrHiat1b such that the respective exon was not included in mature RNA (**Figure 5.5 A**).

Significant reduction of ammonia excretion was found in larvae with DrHiat1b knockdown, but not with DrHiat1a knockdown (**Figure 5.5 B**). SIET was then used to measure the regional flux of NH_4^+ on larvae with DrHiat1b knockdown (**Figure 5.6 A**). NH_4^+ selective electrodes were placed at the yolk sac (**Figure 5.6 A, (1)**) and jaw (**Figure 5.6 A, (2)**) regions. Significant reduction of NH_4^+ flux was detected at both regions in DrHiat1b knockdown larvae compared to the SHAM (**Figure 5.6 B**).

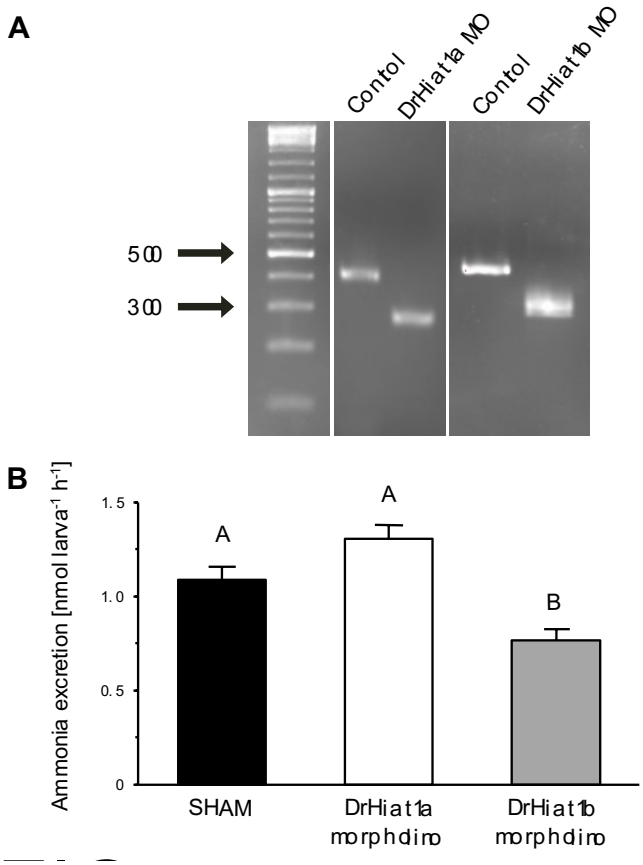


Figure 5.5 Whole animal ammonia excretion in *DrHiat1a* and *DrHiat1b* knock-down (morpholino) larvae (4dpf). (A) Confirmation of successful *DrHiat1a/DrHiat1b* knockdown (morpholinos, MO) indicated by a reduced size of PCR products compared to control, caused by blockage of pre-mRNA splicing in the morpholinos. (B) Ammonia excretion in sham-injected larva and *DrHiat1a/DrHiat1b* knockdown larvae (morpholinos). Data are shown as means ± s.e.m. Uppercase letters indicate significant differences as identified by Kruskal Wallis test with Mann-Whitney pairwise and Bonferroni corrected comparisons ($P < 0.004$, $N = 41,24,29$).

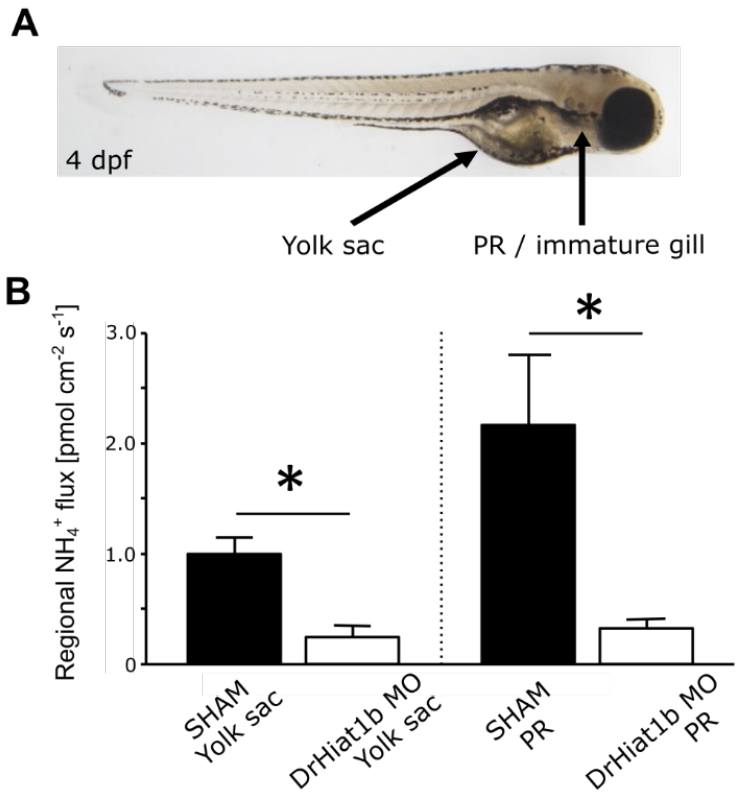


Figure 5.6 Regional NH_4^+ flux in *DrHiat1b* knock-down larvae. (A) Microelectrodes were placed at the yolk sac and pharyngeal region (PR; immature gill) of larvae 4 days post-fertilization (4 dpf). (B) The Scanning Ion-selective micro-Electrode technique (SIET) was used to measure NH_4^+ at these distinct regions in sham and DrHiat1b morpholinos (MO). Data are shown as means ± s.e.m. Asterisks denote significant differences as identified with Student's t-test ($P < 0.004$, $N = 4$-7).

5.5 DISCUSSION

Despite their high level of sequence conservation (86%), hippocampus abundant transporters 1a, 1b and 1-like (Hiat1a, Hiat1b, Hiat1-like; renamed mfsd14a/b in mammals) (Doran et al., 2016; Lekholm et al., 2017) can be clearly distinguished in the phylogenetic analysis. Unfortunately, likely due to the high level of amino acid similarity and the lack of research on the phylogeny and function of this/these transporter(s), there is currently no consistent nomenclature available in the literature and databases. Consequently, the reader needs to be aware that many proteins annotated as one or the other in the NCBI database did not hold true in our analysis (i.e., many transcripts annotated as Hiat1-like rather cluster as Hiat1b or -1a). It should also be noted that some sequences that were annotated as different isoforms turned out to be simply truncated/not properly assembled versions of the same gene. We therefore present here a potential reference for future annotations of this transporter based on the zebrafish names for each of the Hiat1 paralogues.

The fact that only one isoform of Hiat1 seems to be present in invertebrates (Fehsenfeld et al., 2022, under review), and Hiat1b being the only transcript identified in lampreys, indicates that this latter transcript is the most ancient of the three paralogues in vertebrates that was then duplicated in elasmobranchs (Hiat1a), followed by another gene duplication event in teleosts (Hiat1-like). This hierarchy is supported by the identification of the phosphoserine / threonine binding group Neck 6 and proline-dependent serine/threonine kinase group Cdk5 only in the ancient Hiat1b, but then being lost in the other two paralogues. Furthermore, the second duplication step seems to have resulted in the loss of two additional phosphoserine / threonine binding groups (Neck1 + 2), leaving the sugar binding motif as the only conserved domain in all three paralogues. Interestingly, Hiat1a seems to be lost again in terrestrial reptilians as well as terrestrial and aquatic mammals. Especially in the original Hiat1b, a clear distinction between fishes on the one hand and all other animal classes on the other hand can be made. This distinction seems to be less clear concerning Hiat1a, where the elasmobranch

isoform seems to stand at the base of this whole clade and *Latimeria* (as one of the most basic teleosts) seems to be the direct ancestor to Hiat1a in amphibians, birds and reptilians. All other fishes then cluster together as seen with Hiat1b with seawater fishes being more ancient than freshwater species. Finally, Hiat1-like seems to resemble Hiat1a more closely than Hiat1b in the sense that the more modern fishes form their own clade.

5.5.1 Zebrafish Hiat1a and Hiat1b as novel ammonia transporters

Despite possessing a sugar transporter-specific motif D-R/K-X-G-R-R/K between TMD2 and TMD3 (Matsuo et al., 1997), recent results clearly showed a lack of glucose transport in *Xenopus laevis* oocytes when expressing crustacean Hiat1 (Fehsenfeld et al., 2022, under review). Similarly, mammalian Mfsd14a was not capable of mediating glucose uptake into the oocyte (Zhouyao, unpublished observation). Alternatively, the authors showed that CmHiat1 promoted ammonium transport (likely NH_4^+), shown either directly with scanning ion-selective microelectrodes, or indirectly with radiolabelled MA as a proxy (Fehsenfeld et al., 2022). The present study revealed very similar results for zebrafish, in that both DrHiat1a and DrHiat1b could promote MA uptake when expressed in frog oocytes. Furthermore, MA release was enhanced in pre-loaded oocytes when expressing either of the Hiat1 paralogues. Adding the fact that external NH_4^+ was able to compete with MA uptake in transgenic oocytes (as seen by the immediate inhibition of MA uptake in the presence of NH_4Cl in the medium) allows us to conclude that DrHiat1a and DrHiat1b also function as ammonia transporters.

5.5.2 Hiat1b-mediated ammonia excretion in zebrafish larvae

Ammonia handling and excretion seem to be especially important in early life stages of fish for a variety of reasons. First, due to an amino acid-focussed metabolism, the ammonia load in fish embryos is substantial (Zimmer et al., 2017). Second, embryos experience physiological constraints; these include the presence of a chorion capsule as a barrier that impedes diffusion between the embryo

and its external environment, and the lack of fully functional gills which function as the main ammonia excretory organ in more mature fish (Zimmer et al., 2017).

Generally, ammonia excretion in larval zebrafish occurs in part *via* H^+-ATPase-rich (HR) cells in the yolk sac and gills that also contain apical Rhesus-glycoprotein 1 (Rhcg1 or Rhcgb) (Nakada et al., 2007), Na^+/H^+-exchanger 3 (Ito et al., 2013) and H^+-ATPase (Lin et al., 2006). Furthermore, with their unique expression patterns along the skin, gills, and yolk sac, Rhag, Rhbg, Rhcg1 (Rhcgb) and Rhcg2 (Rhcgl1) have all been shown to contribute to whole zebrafish larvae ammonia excretion to varying extents (Braun et al., 2009; Zimmer and Perry, 2020). Additionally, Rhbg and/or Rhcg1 (Rhcgb) seem to more specifically participate also in the excretion of ammonia *via* skin keratinocytes (Shih et al., 2008; Shih et al., 2013). In contrast, a Na^+/NH_4^+ metabolon as mentioned in the introduction, seems to rather contribute to ammonia excretion only under challenging conditions (i.e., low pH and/or low Na^+ environments) and less when fish were kept in unaltered laboratory conditions (i.e., circumneutral water with adequate salt concentrations) (Kumai and Perry, 2011; Shih et al., 2012).

Here, we provide first evidence that ammonia excretion in this early life stage appears to be additionally mediated by DrHiat1b and hence provides an alternative route for basic ammonia transport. Interestingly, Hiat1 is endogenous to frog oocytes (Fehsenfeld et al., 2022 under review) so it is ensured to being passed on with the maternal lineage, implying a crucial role for this transporter as early as development begins. The fact that both DrHiat1a and DrHita1b are ubiquitously expressed in the zebrafish embryo and early larvae emphasizes their importance as potential ammonia transporters in this crucial phase of life. They might provide a very general route for baseline ammonia excretion in every tissue, complemented by more differentially expressed and tissue-specific Rhesus-glycoproteins. It is not surprising that both Hiat1-transcripts can initially be found in the nervous system, resembling the first discovery of Hiat in marsupials' hippocampus (Matsuo et al., 1997).

DrHiat1b, as the most ancient isoform, being more abundantly expressed might indicate its more important role for basic ammonia maintenance and might also explain the subtle differences in tissue expression. For instance, while mRNA for DrHiat1b was present in the intestine, mRNA for DrHiat1a was not detectable.

In DrHiat1b morpholino knockdowns, whole larval ammonia excretion was reduced by ca. 30%. The fact that we observed even further reduction of ammonia excretion when we specifically screened the yolk sac (-75%) and pharyngeal region (immature gill; -85%) underlines the complexity of excretory mechanisms in this stage. Similar discrepancies have been observed in earlier studies on zebrafish larvae, when for example, Zimmer and Perry (2020) found that a knock-down of Rhcgb resulted in an increase of whole larvae ammonia excretion, whereas NH_4^+ flux at the yolk sac assessed by SIET was reduced in response to knockdown using the same morpholino (Shih et al., 2008, 2012). Further thorough investigations are needed to fully understand these respective mechanisms, including potential off-target effects and/or compensatory responses caused by the knock-down, i.e., the possibility that other ammonia transporters like Rhesus-glycoproteins take over for the loss of Hiat1 (Kok et al., 2015; Rossi et al., 2015).

5.6 CONCLUSION

The findings of the present study warrant indeed a reconsideration of ammonia-related transport processes in the gills and potentially other excretory organs of fish. We show here that especially early on in development, the involvement of Hiat1b and/or Hiat1a provides an alternate route for ammonia transport besides the involvement of Rhesus-glycoproteins and/or Na^+/NH_4^+ metabolon. It would be highly desirable for future studies to identify the transporters' exact cellular localization in the branchial membrane to allow for more detailed modelling of branchial ammonia transport. Clearly, Hiat1 transporters must be considered in future studies on ammonia transport in fish and other species, including the mammalian kidney.

5.7 ACKNOWLEDGEMENTS

We like to thank the Animal Care and Veterinary Services at the University of Ottawa, and in particular Christine Archer and Vishal Saxena for taking care of the zebrafish.

5.8 COMPETING INTERESTS

The authors declare no competing interests.

5.9 FUNDING

Financial support was provided by Natural Sciences and Engineering Research Council of Canada Discovery grants (NSERC RGPIN-5013-2018 to Dirk Weihrauch, RGPIN-2018-06027 to Peter Eck, and NSERC # G13017 to Steve Perry). The research was further supported by a University of Manitoba UCRP grant (2017-18/SUB 318995) to Dirk Weihrauch and Peter Eck. Haonan Zhouyao was awarded a Company of Biologists travel grant for this collaboration. Sandra Fehsenfeld and Alex Zimmer are/were the recipients of a NSERC postdoctoral fellowship.

5.10 REFERENCES

Adlimoghaddam, A., Boeckstaens, M., Marini, A. M., Treberg, J. R., Brassinga, A. K. C., & Weihrauch, D. (2015). Ammonia excretion in *Caenorhabditis elegans*: Mechanism and evidence of ammonia transport of the Rhesus protein CeRhr-1. *Journal of Experimental Biology*, *218*(5), 675–683. https://doi.org/10.1242/jeb.111856

Adlimoghaddam, A., O'Donnell, M. J., Kormish, J., Banh, S., Treberg, J. R., Merz, D., & Weihrauch, D. (2016). Ammonia excretion in *Caenorhabditis elegans*: Physiological and molecular characterization of the rhr-2 knock-out mutant. *Comparative Biochemistry and Physiology -Part A : Molecular and Integrative Physiology*, *195*(April), 46–54. https://doi.org/10.1016/j.cbpa.2016.02.003

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, *25*(17), 3389–3402. https://doi.org/10.1016/0031-9422(92)80418-E

Arillo, A., Margiocco, C., Melodia, F., Mesi, P., & Schenone, G. (1981). Ammonia toxicity mechanism in fish: Studies on rainbow trout (*Salmo gairdneri* Rich.). *Ecotoxicology and Environmental Safety*, *5*(3), 316–328. https://doi.org/10.1016/j.ecoenv.2005.12.006

Boo, M. V., Hiong, K. C., Goh, E. J. K., Choo, C. Y. L., Wong, W. P., Chew, S. F., & Ip, Y. K. (2018). The ctenidium of the giant clam, *Tridacna squamosa*, expresses an ammonium transporter 1 that displays light-suppressed gene and protein expression and may be involved in ammonia excretion. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, *188*(5), 765–777. https://doi.org/10.1007/s00360-018-1161-6

Braun, M. H., Steele, S. L., Ekker, M., & Perry, S. F. (2009). Nitrogen excretion in developing zebrafish (*Danio rerio*): A role for Rh proteins and urea transporters. *American Journal of Physiology - Renal Physiology*, *296*(5), F994–F1005. https://doi.org/10.1152/ajprenal.90656.2008

Bucking, C. (2017). A broader look at ammonia production, excretion, and transport in fish: a review

of impacts of feeding and the environment. *Journal of Comparative Physiology B*, *187*(1), 1–18. https://doi.org/10.1007/s00360-016-1026-9

Cameron, J. N., & Heisler, N. (1983). Studies of ammonia in the rainbow trout: physico-chemical parameters, acid-base behaviour and respiratory clearance. *Journal of Experimental Biology*, *105*(1), 107–125. Retrieved from http://jeb.biologists.org/content/105/1/107.short

Chasiotis, H., Ionescu, A., Misyura, L., Bui, P., Fazio, K., Wang, J., … Donini, A. (2016). An animal homolog of plant Mep/Amt transporters promotes ammonia excretion by the anal papillae of the disease vector mosquito *Aedes aegypti*. *Journal of Experimental Biology*, *219*(9), 1346–1355. https://doi.org/10.1242/jeb.134494

Diricx, M., Sinha, A. K., Liew, H. J., Mauro, N., Blust, R., & De Boeck, G. (2013). Compensatory responses in common carp (*Cyprinus carpio*) under ammonia exposure: Additional effects of feeding and exercise. *Aquatic Toxicology*, *142–143*, 123–137. https://doi.org/10.1016/j.aquatox.2013.08.007

Donini, A., & O'Donnell, M. J. (2005). Analysis of Na^+, Cl^-, K^+, H^+ and NH_4^+ concentration gradients adjacent to the surface of anal papillae of the mosquito *Aedes aegypti*: application of self-referencing ion-selective microelectrodes. *Journal of Experimental Biology*, *208*, 603–610. https://doi.org/10.1242/jeb.01422

Doran, J., Walters, C., Kyle, V., Wooding, P., Hammett-burke, R., & Colledge, W. H. (2016). *Mfsd14a* (*Hiat1*) gene disruption causes globozoospermia and infertility in male mice. *Reproduction*, *152*, 91–99. https://doi.org/10.1530/REP-15-0557

Dosdat, A., Ruyet, J. P.-L., Coves, D., Dutto, G., Gasset, E., Le Roux, A., & Lemarie, G. (2003). Effect of chronic exposure to ammonia on growth, food utilisation and metabolism of the European sea bass (*Dicentrarchus labrax*). *Aquatic Living Resources*, *16*(6), 509–520.

Durant, A. C., & Donini, A. (2018). Ammonia excretion in an osmoregulatory syncytium is facilitated

by AeAmt2, a novel ammonia transporter in *Aedes aegypti* larvae. *Frontiers in Physiology*, *9*, 1–16. https://doi.org/10.3389/fphys.2018.00339

Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, *5*, 113.

Fehsenfeld, S., Kiko, R., Appelhans, Y. S., Towle, D. W., Zimmer, M., & Melzner, F. (2011). Effects of elevated seawater pCO_2 on gene expression patterns in the gills of the green crab, *Carcinus maenas*. *BMC Genomics*, *12*(1), 488. https://doi.org/10.1186/1471-2164-12-488

Fehsenfeld, S., Quijada-Rodriguez, A. R., Zhouyao, H., Durant, A. C., Donini, A., Sachs, M., … Weihrauch, D. (2022). HIAT1 – a new highly conserved transporter involved in ammonia regulation. *FASEB Journal*, (under review).

Fehsenfeld, S., & Wood, C. M. (2018). Section-specific expression of acid-base and ammonia transporters in the kidney tubules of the goldfish *Carassius auratus* and their responses to feeding. *American Journal of Physiology - Renal Physiology*, *315*(6), F1565–F1582. https://doi.org/10.1152/ajprenal.00510.2017

Hammer, Ø., Harper, D. A. T., & Ryan, P. D. (2001). PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica*, *4*(1), 1–9. https://doi.org/10.1016/j.bcp.2008.05.025

Hans, S., Quijada-Rodriguez, A. R., Allen, G. J. P., Onken, H., Treberg, J. R., & Weihrauch, D. (2018). Ammonia excretion and acid-base regulation in the American horseshoe crab, *Limulus polyphemus*. *Journal of Experimental Biology*, *221*, jeb151894. https://doi.org/10.1242/jeb.151894

Hughes, M. C., Zimmer, A. M., & Perry, S. F. (2019). Role of internal convection in respiratory gas transfer and aerobic metabolism in larval zebrafish (*Danio rerio*). *American Journal of Physiology - Regulatory Integrative and Comparative Physiology*, *316*(3), R255–R264. https://doi.org/10.1152/ajpregu.00315.2018

Ito, Y., Kobayashi, S., Nakamura, N., Miyagi, H., Esaki, M., Hoshijima, K., & Hirose, S. (2013). Close association of carbonic anhydrase (CA2a and CA15a), Na^+/H^+ exchanger (Nhe3b), and ammonia transporter Rhcg1 in zebrafish ionocytes responsible for Na^+ uptake. *Frontiers in Physiology*, *4*, 1–17. https://doi.org/10.3389/fphys.2013.00059

Jones, D. T., Taylor, W. R., & Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Computer Applications in the Biosciences*, *8*(3), 275–282.

Kok, F. O., Shin, M., Ni, C. W., Gupta, A., Grosse, A. S., VanImpel, A., … Lawson, N. D. (2015). Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish. *Developmental Cell*, *32*, 97–108. https://doi.org/10.1016/j.devcel.2014.11.018

Kumai, Y., & Perry, S. F. (2011). Ammonia excretion *via* Rhcg1 facilitates Na^+ uptake in larval zebrafish, *Danio rerio*, in acidic water. *American Journal of Physiology - Regulatory Integrative and Comparative Physiology*, *301*(5), R1517–R1528. https://doi.org/10.1152/ajpregu.00282.2011

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Larsen, E. H., Deaton, L. E., Onken, H., O'Donnell, M., Grosell, M., Dantzler, W. H., & Weihrauch, D. (2014). Osmoregulation and excretion. *Comprehensive Physiology*, *4*(2), 405–573. https://doi.org/10.1002/cphy.c130004

Lekholm, E., Perland, E., Eriksson, M. M., Hellsten, S. V, Lindberg, F. A., Rostami, J., & Fredriksson, R. (2017). Putative membrane-bound transporters MFSD14A and MFSD14B are neuronal and affected by nutrient availability. *Frontiers in Molecular Neuroscience*, *10*, 1–13. https://doi.org/10.3389/fnmol.2017.00011

Lin, L. Y., Horng, J. L., Kunkel, J. G., & Hwang, P. P. (2006). Proton pump-rich cell secretes acid in

skin of zebrafish larvae. *American Journal of Physiology - Cell Physiology*, *290*(2), 371–378. https://doi.org/10.1152/ajpcell.00281.2005

Mallery, C. H. (1983). A carrier enzyme basis for ammonium excretion in teleost gill. NH_4^+-stimulated Na-dependent ATPase activity in *Opsanus beta*. *Comparative Biochemistry and Physiology -- Part A: Physiology*, *74*(4), 889–897. https://doi.org/10.1016/0300-9629(83)90364-X

Masui, D. C., Furriel, R. P. M., McNamara, J. C., Mantelatto, F. L. M., & Leone, F. A. (2002). Modulation by ammonium ions of gill microsomal (Na^+,K^+)-ATPase in the swimming crab *Callinectes danae*: A possible mechanism for regulation of ammonia excretion. *Comparative Biochemistry and Physiology - C Toxicology and Pharmacology*, *132*(4), 471–482. https://doi.org/10.1016/S1532-0456(02)00110-2

Matsuo, N., Kawamoto, S., Matsubara, K., & Okubo, K. (1997). Cloning of a cDNA encoding a novel sugar transporter expressed in the neonatal mouse hippocampus. *Biochemical and Biophysical Research Communications*, *238*, 126–129. https://doi.org/10.1006/bbrc.1997.7252

McKenzie, D. J., Shingles, A., Claireaux, G., & Domenici, P. (2009). Sublethal concentrations of ammonia impair performance of the teleost fast-start escape response. *Physiological and Biochemical Zoology*, *82*(4), 353–362. https://doi.org/10.1086/590218

McKenzie, D. J., Shingles, A., & Taylor, E. W. (2003). Sub-lethal plasma ammonia accumulation and the exercise performance of salmonids. *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology*, *135*(4), 515–526. https://doi.org/10.1016/S1095-6433(03)00116-8

Nakada, T., Hoshijima, K., Esaki, M., Nagayoshi, S., Kawakami, K., & Hirose, S. (2007). Localization of ammonia transporter Rhcg1 in mitochondrion-rich cells of yolk sac, gill, and kidney of zebrafish and its ionic strength-dependent expression. *American Journal of Physiology - Regulatory Integrative & Comparative Physiology*, (293), R1743–R1753. https://doi.org/10.1152/ajpregu.00248.2007.

Nakhoul, N. L., & Hamm, L. L. (2014). The challenge of determining the role of Rh glycoproteins in transport of NH_3 and NH_4^+. *WIREs Membrane Transport and Signaling*, *3*, 53–61. https://doi.org/10.1002/wmts.105

Nawata, C. M., Hirose, S., Nakada, T., Wood, C. M., & Kato, A. (2010). Rh glycoprotein expression is modulated in pufferfish (*Takifugu rubripes*) during high environmental ammonia exposure. *The Journal of Experimental Biology*, *213*, 3150–3160. https://doi.org/10.1242/jeb.044719

Obenauer, J. C., Cantley, L. C., & Yaffe, M. B. (2003). Scansite 2.0: Proteome-wide prediction of cell signalling interactions using short sequence motifs. *Nucleic Acids Research*, *31*(13), 3635–3641. https://doi.org/10.1093/nar/gkg584

Omasits, U., Ahrens, C. H., Mu, S., & Wollscheid, B. (2014). Sequence analysis Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics*, *30*(6), 884–886. https://doi.org/10.1093/bioinformatics/btt607

Pitts, R. J., Derryberry, S. L., Pulous, F. E., & Zwiebel, L. J. (2014). Antennal-expressed ammonium transporters in the malaria vector mosquito *Anopheles gambiae*. *PLoS ONE*, *9*(10), e111858. https://doi.org/10.1371/journal.pone.0111858

Quijada-Rodriguez, A. R., Schultz, A. G., Wilson, J. M., He, Y., Allen, G. J. P., Goss, G. G., & Weihrauch, D. (2017). Ammonia-independent sodium uptake mediated by Na^+ channels and NHEs in the freshwater ribbon leech *Nephelopsis obscura*. *Journal of Experimental Biology*, *220*(18), 3270–3279. https://doi.org/10.1242/jeb.159459

Rossi, A., Kontarakis, Z., Gerri, C., Nolte, H., Hölper, S., Krüger, M., & Stainier, D. Y. R. (2015). Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature*, *524*, 230–233. https://doi.org/10.1038/nature14580

Sayer, M. D. J., & Davenport, J. (1987). The relative importance of the gills to ammonia and urea excretion in five seawater and one freshwater teleost species. *Journal of Fish Biology*, *31*, 561–570.

https://doi.org/10.1111/j.1095-8649.1987.tb05258.x

Shih, T. H., Horng, J. L., Hwang, P. P., & Lin, L. Y. (2008). Ammonia excretion by the skin of zebrafish (*Danio rerio*) larvae. *American Journal of Physiology - Cell Physiology*, *295*(6), C1625–C1632. https://doi.org/10.1152/ajpcell.00255.2008

Shih, T. H., Horng, J. L., Lai, Y. T., & Lin, L. Y. (2013). Rhcg1 and Rhbg mediate ammonia excretion by ionocytes and keratinocytes in the skin of zebrafish larvae: H^+-ATPase-linked active ammonia excretion by ionocytes. *American Journal of Physiology - Regulatory Integrative and Comparative Physiology*, *304*(12), R1130–R1138. https://doi.org/10.1152/ajpregu.00550.2012

Shih, T. H., Horng, J. L., Liu, S. T., Hwang, P. P., & Lin, L. Y. (2012). Rhcg1 and NHE3b are involved in ammonium-dependent sodium uptake by zebrafish larvae acclimated to low-sodium water. *American Journal of Physiology - Regulatory Integrative and Comparative Physiology*, *302*, R84–R93. https://doi.org/10.1152/ajpregu.00318.2011

Shingles, A., McKenzie, D. J., Taylor, E. W., Moretti, A., Butler, P. J., & Ceradini, S. (2001). Effects of sublethal ammonia exposure on swimming performance in rainbow trout (*Oncorhynchus mykiss*). *Journal of Experimental Biology*, *204*(15), 2691–2698. https://doi.org/10.1242/jeb.204.15.2691

Sinha, A. K., Liew, H. J., Diricx, M., Blust, R., & De Boeck, G. (2012a). The interactive effects of ammonia exposure, nutritional status and exercise on metabolic and physiological responses in gold fish (*Carassius auratus* L.). *Aquatic Toxicology*, *109*, 33–46. https://doi.org/10.1016/j.aquatox.2011.11.002

Sinha, A. K., Liew, H. J., Diricx, M., Kumar, V., Darras, V. M., Blust, R., & De Boeck, G. (2012b). Combined effects of high environmental ammonia, starvation and exercise on hormonal and ion-regulatory response in goldfish (*Carassius auratus* L.). *Aquatic Toxicology*, *114–115*, 153–164. https://doi.org/10.1016/j.aquatox.2012.02.027

Smith, A. A., Zimmer, A. M., & Wood, C. M. (2012). Branchial and extra-branchial ammonia excretion

in goldfish (*Carassius auratus*) following thermally induced gill remodeling. *Comparative Biochemistry and Physiology A*, *162*, 185–192. https://doi.org/10.1016/j.cbpa.2012.02.019

Smith, H. W. (1929). The excretion of ammonia and urea by the gills of fish. *Journal of Biological Chemistry*, *81*(3), 727–742. https://doi.org/10.1016/s0021-9258(18)63725-1

Thiel, D., Hugenschutt, M., Meyer, H., Paululat, A., Quijada-Rodriguez, A. R., Purschke, G., & Weihrauch, D. (2017). Ammonia excretion in the marine polychaete *Eurythoe complanata* (Annelida). *Journal of Experimental Biology*, *220*(3), 425–436. https://doi.org/10.1242/jeb.145615

Thisse, C., & Thisse, B. (2008). High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nature Protocols*, *3*(1), 59–69. https://doi.org/10.1038/nprot.2007.514

Verdouw, H., Van Echteld, C. J. A., & Dekkers, E. M. J. (1978). Ammonia determination based on indophenol formation with sodium salicylate. *Water Research*, *12*, 399–402. https://doi.org/10.1016/0043-1354(78)90107-0

Wall, S. M., & Koger, L. M. (1994). NH_4^+ transport mediated by Na^+-K^+-ATPase in rat inner medullary collecting duct. *American Journal of Physiology - Renal Physiology*, *267*(4 Pt 2), F660–F670. https://doi.org/10.1152/ajprenal.1994.267.4.f660

Weihrauch, D., Chan, A. C., Meyer, H., Döring, C., Sourial, M., & O'Donnel, M. J. (2012). Ammonia excretion in the freshwater planarian *Schmidtea mediterranea*. *Journal of Experimental Biology*, *215*(18), 3242–3253. https://doi.org/10.1242/jeb.067942

White, R. M., Sessa, A., Burke, C., Bowman, T., LeBlanc, J., Ceol, C., … Zon, L. I. (2008). Transparent adult zebrafish as a tool for in vivo transplantation analysis. *Cell Stem Cell*, *2*(2), 183–189. https://doi.org/10.1016/j.stem.2007.11.002

Wilkie, M. P. (1997). Mechanisms of ammonia excretion across fish gills. *Comparative Biochemistry and Physiology - A Physiology*, *118A*(1), 39–50. https://doi.org/10.1016/S0300-9629(96)00407-0

Wood, C. M., & Nawata, C. M. (2011). A nose-to-nose comparison of the physiological and molecular

responses of rainbow trout to high environmental ammonia in seawater versus freshwater. *Journal of Experimental Biology*, *214*(21), 3557–3569. https://doi.org/10.1242/jeb.057802

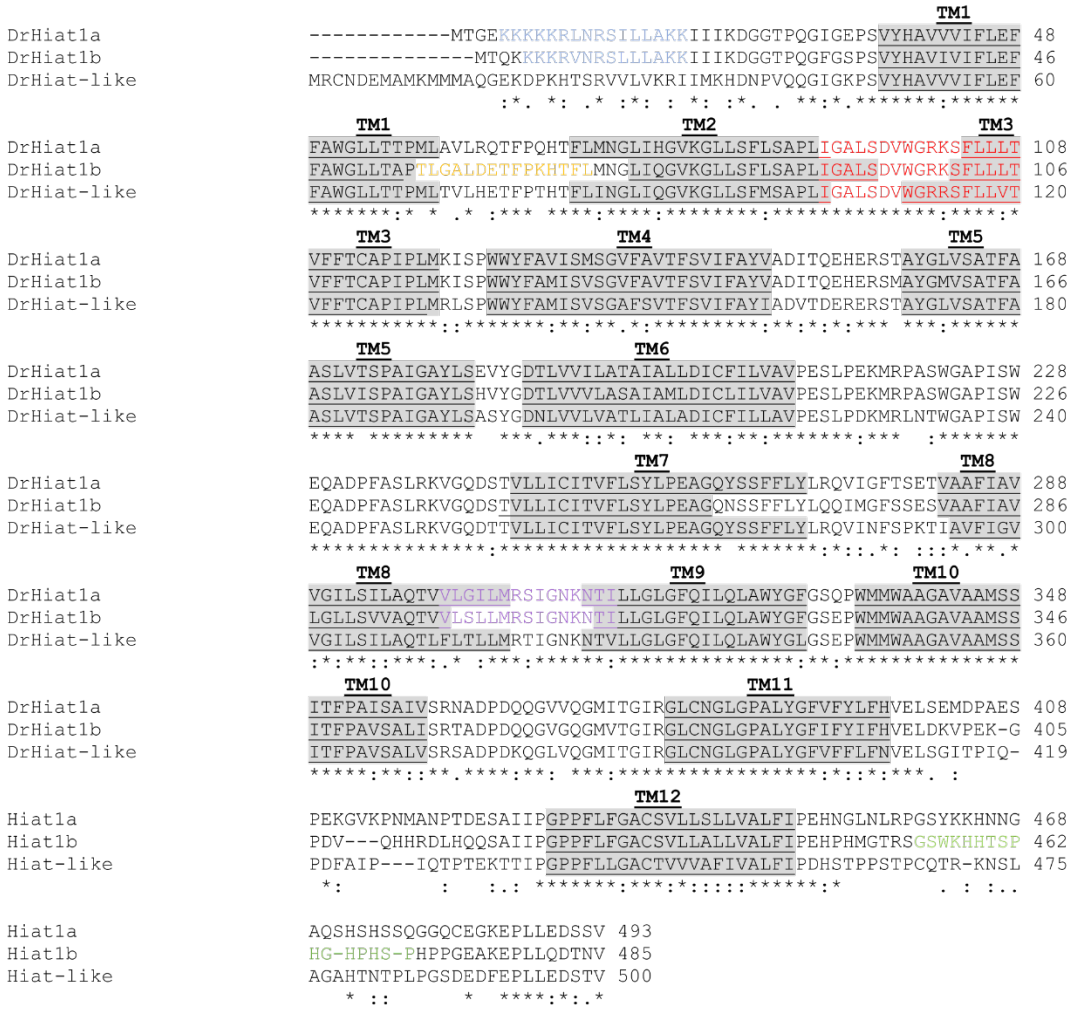
Wright, P. A., & Wood, C. M. (2009). A new paradigm for ammonia excretion in aquatic animals: role of Rhesus (Rh) glycoproteins. *The Journal of Experimental Biology*, *212*, 2303–2312. https://doi.org/10.1242/jeb.023085

Zimmer, A. M., Brauner, C. J., & Wood, C. M. (2014). Ammonia transport across the skin of adult rainbow trout (*Oncorhynchus mykiss*) exposed to high environmental ammonia (HEA). *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, *184*, 77–90. https://doi.org/10.1007/s00360-013-0784-x

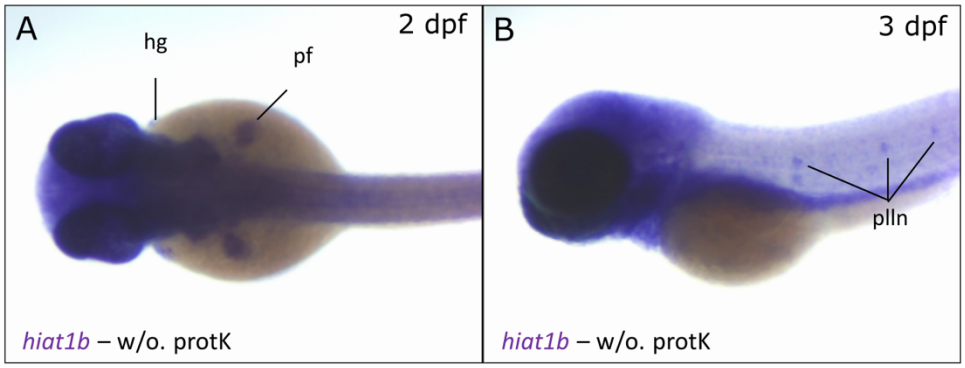
Zimmer, A. M., & Perry, S. F. (2020). The Rhesus glycoprotein Rhcgb is expendable for ammonia excretion and Na^+ uptake in zebrafish (*Danio rerio*). *Comparative Biochemistry and Physiology -Part A : Molecular and Integrative Physiology*, *247*, 110722. https://doi.org/10.1016/j.cbpa.2020.110722

Zimmer, A. M., Wright, P. A., & Wood, C. M. (2017). Ammonia and urea handling by early life stages of fishes. *Journal of Experimental Biology*, *220*(21), 3843–3855. https://doi.org/10.1242/jeb.140210

5.11 SUPPLEMENTARY INFORMATION



Supplementary Figure 5.11.1. Alignment of the three *Danio rerio* Hiat1 isoforms. Sequences for DrHiat1-like (GenBank accession number XP_002663499.2), DrHiat1a (Genbank acc. no. AAH97075.1) and DrHiat1b (GenBank acc. no. BC056817) were aligned with Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/). Shown are the sugar biding motif in red letters, transmembrane domains in bold and underlined (TMs), Phosphoserine/threonine binding groups in light blue (Neck1) / purple (Neck2) / yellow (Neck6) and a proline-dependent serine/threonine kinase group in green (Cdk5). Asterisks denote conserved sites, ":" denote conserved strong groups and "." denote conservation of weak groups.



Supplementary Figure 5.11.2. *In situ* hybridization of DrHiat1b without proteinase K treatment. Additional staining for DrHiat1b can be detected in (A) the pectoral fin (pf) and weakly in the hatching gland (hg) at 2 days post fertilization (dpf), and in (B) the posterior lateral line neuromasts (plln) at 3 dpf. B, dorsal view, anterior to the left; C: lateral view, anterior to the left.

GenBank accession no.	Clusters as	Species	Common name	Class
NP 001085612.1	Hiat1b	*Xenopus laevis*	African clawed frog	Amphibia
XP 018114716.1	Hiat1-like	*Xenopus laevis*	African clawed frog	Amphibia
NP 001087834.1	Hiat1a	*Xenopus laevis*	African clawed frog	Amphibia
XP 002188818.1	Hiat1b	*Taeniopygia guttata*	Zebrafinch	Aves
XP 030113786.3	Hiat1-like	*Taeniopygia guttata*	Zebrafinch	Aves
XP 012426709.3	Hiat1a	*Taeniopygia guttata*	Zebrafinch	Aves
XP_041064351.1	Hiat1b	*Carcharodon carcharias*	Great white shark	Chondrichthye
XP 041059808.1	Hiat1a	*Carcharodon carcharias*	Great white shark	Chondrichthye
GCB60630.1	Hiat1b	*Scyliorhinus torazame*	Cloudy catshark	Chondrichthye
GCB70945.1	Hiat1a	*Scyliorhinus torazame*	Cloudy catshark	Chondrichthye
XP 033876203.1	Hiat1b	*Acipenser ruthenus*	Sterlet	Chondrostei
XP 034775964.1	Hiat1-like	*Acipenser ruthenus*	Sterlet	Chondrostei
XP 034767512.1	Hiat1a	*Acipenser ruthenus*	Sterlet	Chondrostei
DW250260.1	Hiat1	*Carcinus maenas*	Green crab	Crustacea
XP_032826511.1	Hiat1b	*Petromycon marinus*	Sea lamprey	Hyperoartia (Agnatha)
NP 149044.2	Hiat1b	*Homo sapiens*	Human	Mammalia
NP 115947.2	Hiat1-like	*Homo sapiens*	Human	Mammalia
NP 032272.2	Hiat1b	*Mus musculus*	House mouse	Mammalia

NP 001077370.1	Hiat1-like	*Mus musculus*	House mouse	Mammalia
XP 036707313.1	Hiat1b	*Balaenoptera musculus*	Blue whale	Mammalia (SW
XP 036710915.1	Hiat1-like	*Balaenoptera musculus*	Blue whale	Mammalia (SW
XP 039340654.1	Hiat1b	*Mauremys reevesii*	Chinese pond turtle	Reptilia (FW)
XP 039400711.1	Hiat1-like	*Mauremys reevesii*	Chinese pond turtle	Reptilia (FW)
XP 039371595.1	Hiat1a	*Mauremys reevesii*	Chinese pond turtle	Reptilia (FW)
XP 043347006.1	Hiat1b	*Dermochelys coriacea*	Leatherback sea turtle	Reptilia (SW)
XP 038258563.1	Hiat1-like	*Dermochelys coriacea*	Leatherback sea turtle	Reptilia (SW)
XP 038239659.1	Hiat1a	*Dermochelys coriacea*	Leatherback sea turtle	Reptilia (SW)
XP 013926066.1	Hiat1b	*Thamnophis sirtalis*	Common garter snake	Reptilia (T)
XP 013915054.1	Hiat1-like	*Thamnophis sirtalis*	Common garter snake	Reptilia (T)
XP 006000077.1	Hiat1b	*Latimeria chalumnae*	West Indian ocean coelacanth	Sarcopterygii
XP 006001415.1	Hiat1-like	*Latimeria chalumnae*	West Indian ocean coelacanth	Sarcopterygii
XP 014345977.1	Hiat1a	*Latimeria chalumnae*	West Indian ocean coelacanth	Sarcopterygii
XP 026091418.1	Hiat1b	*Carassius auratus*	Goldfish	Teleostei (FW
XP 026076322.1	Hiat1-like	*Carassius auratus*	Goldfish	Teleostei (FW
XP 026143900.1	Hiat1a	*Carassius auratus*	Goldfish	Teleostei (FW
AAH97075.1	Hiat1b	*Danio rerio*	Zebrafish	Teleostei (FW
XP 002663499.2	Hiat1-like	*Danio rerio*	Zebrafish	Teleostei (FW
BC056817.1	Hiat1a	*Danio rerio*	Zebrafish	Teleostei (FW
XP 014071494.1	Hiat1b	*Salmo salar*	Atlantik salmon	Teleostei (SW

XP 013981624.1	Hiat1-like	*Salmo salar*	Atlantik salmon	Teleostei (SW)
XP 014047397.1	Hiat1a	*Salmo salar*	Atlantik salmon	Teleostei (SW)

Supplementary Table 5.11.1. Protein sequences for phylogenetic tree as shown in figure 1.

Accession numbers are according to GenBank.

Application	Name	Sequence (5'-3')	A s
ORF	DrHiat1a ORF forward	ACCATGACTGGAGAGAAAAAGAAGAAA	
	DrHiat1a ORF reverse	TTATACGCTGTCCTCCAGTAA	
	DrHiat1b ORF forward	ACCATGACCCAGAAAAAGAAGAAGCGA	
	DrHiat1b ORF reverse	TCAGACATTAGTGTCCTGAAGG	
Oocytes	DrHiat1a Xmal forward	TCCCCCCGGGACCATGACTGGAGAGAAAAAGAAGAAAA	
	DrHiat1a XbaI reverse	GCTCTAGATTATACGCTGCTGTCCTCCAGTAAA	
	DrHiat1b Xmal forward	TCCCCCCGGGACCATGACCCAGAAAAAGAAGAAGCGAG	
	DrHiat1b XbaI reverse	GCTCTAGATCAGACATTAGTGTCCTGTAGAAGG	
In situ	DrHiat1a insitu forward	AAAATCTAGAAACCATGACTGGAGAGAAAAAGAAGAAAA	
	DrHiat1a insitu reverse	TTTTGGGCCCTACGCTGCTGTCCTCCAGTAAAGGC	
	DrHiat1b insitu forward	AAAATCTAGAACCATGACCCAGAAAAAGAAGAAGCGAG	
	DrHiat1b insitu reverse	TTTTGGGCCCGACATTAGTGTCCTGTAGAAGGGGC	
Morpholinos	DrHiat1a MO forward	AAAAGAAGAAAAAGCGGCTGAAC	
	DrHiat1a MO reverse	ACAGCAAACACTCCAGACATGG	

DrHiat1b MO forward	ATTTCTGGAGTTCTTCGCTTGG
DrHiat1b MO reverse	GACTGAGATACGCACCGATGG

Supplementary Information Table 5.11.2 Primer sequences. Primers have been designed based on GenBank accession numbers XP_002663499.2 (DrHiat1-like), AAH97075.1 (DrHiat1a), and BC056817 (DrHiat1b).

CHAPTER 6:

DISCUSSION AND CONCLUSION

6.1 SUMMARY OF THE MAJOR FINDINGS

The focus of our studies was to functionally characterize two unknown genes: *SLC2A14* and *MFSD14A*, and to identify the substrate/substrates mediated by their encoded proteins, GLUT14 and MFSD14A. By heterologously expressing GLUT14 in the *Xenopus laevis* oocytes, we have confirmed that glucose and dehydroascorbic acid are transported by GLUT14. We have identified three additional single nucleotide polymorphisms within the *SLC2A14* gene that are associated to an increased risk of developing inflammatory bowel disease.

By utilizing a mixed approach where available tissue expression data from databases and *in situ* hybridization in zebrafish larvae, MFSD14A is highly conserved and is found to be ubiquitously expressed in all the tested tissues and organisms. Results from the transport studies rejected the hypothesis of MFSD14A being a novel glucose transporter since no glucose uptake was found in oocytes expressing MFSD14A. Instead, the MFSD14A-expressing oocytes mediated a bidirectional transport of radiolabeled methylamine. The transport of methylamine can be competitively inhibited by the presence of regular, non-labeled ammonia (in the form of NH_4Cl), indicating that MFSD14A is an ammonia transporter. This finding was further tested in zebrafish, a fish model commonly used in the

study of biological processes and human diseases. In Mfsd14a-knockdown zebrafish larvae, the total and regional ammonia excretion was both reduced, confirming our prior findings.

6.2 THE DISCOVERY OF THE SUBSTRATES OF GLUT14

Glucose as a vital source of energy for all living cells and requires special membrane transport proteins to be absorbed into the cells' cytoplasm. The major facilitator superfamily glucose facilitator (GLUT) is one of the three families of secondary glucose transporters in human. The GLUTs are found ubiquitously distributed across the body mediating facilitative diffusion of glucose down its concentration gradients (Holman, 2020; Mueckler & Thorens, 2013). There are 14 identified members in the GLUT family. Historically, the last member, GLUT14, encoded by gene SLC2A14, was thought to be a duplicon of GLUT3 that is exclusively expressed in the testis (Holman, 2020; Wu & Freeze, 2002). Later genetic association studies, however, linked *SLC2A14*/GLUT14 to non-testicular diseases such as Alzheimer's disease (Shulman et al., 2011; Wang et al., 2012), open angle glaucoma (Nag et al., 2013) and rheumatoid arthritis (Veal et al., 2014), suggesting the possibility of *SLC2A14*/GLUT14 expression in tissues besides the testis. Evidence presented by Amir Shaghaghi et al., 2016 supported these disease associations since the mRNA expression of *SLC2A14* was found in the colon, kidney, liver, ovary, lymph node, skeletal muscle, and ovary. Even though 94.5% DNA sequence identity was shared between *SLC2A14* and the well-characterized glucose transporter GLUT3, encoded by gene *SLC2A3*, GLUT14 appeared to not be a duplicon of GLUT3, requiring functional characterization to confirm the substrate/substrates. By heterologous expression experiments investigating the two identified GLUT14 isoforms (the long isoform and the short isoform), we have confirmed for the first time that GLUT14 is mediating the transport of glucose and dehydroascorbic acid. Our findings are consistent with the research on other members from the GLUT family, since they are often multi-substrate specific, transporting primarily hexoses (Augustin,

2010; Joost et al., 2002; Mueckler & Thorens, 2013; Uldry & Thorens, 2004), myo-inositol (Uldry et al., 2001), urate (Bibert et al., 2009; So & Thorens, 2010), glucosamine (Maher & Harrison, 1990), and ascorbate (Lee et al., 2010). Hence, it is possible that GLUT14 can mediate the transport of other substrates in addition to glucose and dehydroascorbic acid. Due to the physiological importance of glucose and the glucose transporters, future studies are needed to gain a more complete view on the substrate specificity of GLUT14, in order to improve our understandings on the existing disease associations, which will be discussed below.

6.2.1 *SLC2A14*/GLUT14 and implications for Alzheimer's disease

Alzheimer's disease is the most common type of dementia. For people of the age of 65 and above, at least 80% of people with dementia have been diagnosed with Alzheimer's disease (Weller & Budson, 2018). Alzheimer's disease is a progressive neurodegenerative disease with no described cure known to date. As Alzheimer's disease develops, patient's behavior and cognitive functions such as memory, comprehension, language, and judgement are gradually impaired, making Alzheimer's disease the sixth leading cause of death in the United States (Kumar et al., 2022; Weller & Budson, 2018).

The cause of Alzheimer's disease is multifactorial, and can be a combination of genetic and environmental factors that affect the brain health over time (Kumar et al., 2022). One factor that appears to be a cause of Alzheimer's disease is the impaired uptake and metabolism of glucose in the affected brain (Liu et al., 2008), disrupting the function of Tau (Li et al., 2006; Liu et al., 2004). Tau is a small microtubule-associated protein with a primary role of maintaining the stability of microtubules in the axons (Avila et al., 2004). In patients with Alzheimer's disease, Tau is abnormally hyperphosphorylated, forming pathological inclusions known as neurofibrillary tangles. To date, the cause of the impairment of glucose uptake/metabolism in the brains affected by Alzheimer's disease remain unclear. Since neurons and brain tissues are not capable of storing or producing glucose,

glucose transporters are required to facilitate the passage of glucose across the blood-brain barrier (Liu et al., 2008). The autopsy studies in Alzheimer's disease patients revealed significant reduction in GLUT3, and possibly GLUT1 in brain microvessels, cerebral neocortex and hippocampus. (Kyrtata et al., 2021; Liu et al., 2008). Since GLUT3 and GLUT1 are the two glucose transporters predominately expressed in the brain, it is highly likely that reduction in their expressions impair the glucose uptake in the brain (Kyrtata et al., 2021; Liu et al., 2008). Interestingly, the expression of GLUT2 and GLUT12, two glucose transporters found in the brain but to a lesser extent, are increased in the Alzheimer's disease-affected brains, suggesting a compensatory role of GLUT2 and GLUT12 to cope with the reduced function of GLUT3 (Kyrtata et al., 2021). With the recently-available single cell RNA sequencing data, the expression of *SLC2A14* is also detected in the brain, particularly in microglial cells and neuronal cells including the astrocytes, inhibitory neuron, excitatory neuron and oligodendrocytes (Human Protein Atlas, 2022a). Therefore, it is highly likely that *SLC2A14* is also playing a protective role against Alzheimer's disease by providing potentially additional glucose uptake into the brain cells (Amir Shaghaghi et al., 2017, Chapter 2). The presented results (Amir Shaghaghi et al., 2017, Chapter 2) concur with *SLC2A14* associations with Alzheimer's disease. With such findings, we have laid the foundation for future studies determining the pathology of Alzheimer's disease.

6.2.2 *SLC2A14*/GLUT14 and implications for rheumatoid arthritis

Rheumatoid arthritis is a complex and debilitating inflammatory autoimmune disease (Bullock et al., 2018). Although the exact cause and pathogenesis of Rheumatoid arthritis can be a combination of genetics, hormonal and environmental factors, increased glucose uptake as well as abnormal glucose metabolism due to insulin insensitivity has been associated with accelerated disease progression (Zezina et al., 2020). Patients with Rheumatoid arthritis suffer from severe pain due to the progressive

destruction of bone and cartilage (Bullock et al., 2018). One of the key characteristics of Rheumatoid arthritis is the poly-articular synovitis, which is a type of synovial inflammation where the synovium of a joint/joints becomes/become inflamed and swollen (Bullock et al., 2018; Pi et al., 2017). After characterizing the biology of synovial fluid in Rheumatoid arthritis patients, the immune cells in the synovial membrane, including both resident and infiltrating immune cells, are found to play a critical role in the occurrence and the development of Rheumatoid arthritis (Tran et al., 2005). The numbers of the T follicular helper cells, memory activated $CD4^+$ T cells, and plasma cells were significantly higher in Rheumatoid arthritis synovial tissues (Li et al., 2019; Tran et al., 2005). The T follicular helper cells are a specialized subset of $CD4^+$ T cells which assists the B cells to produce antibodies. The memory activated $CD4^+$ T cells are immune cells that are re-activated by antigen exposures, providing protective immunity by either producing early effector cytokines to direct other cells or kill the infected cells directly. The plasma cells are immune cells specializing in producing one specific type of antibody in large quantity.

Since the immune cells are rapidly adapting to stressors, environmental changes and proinflammatory stimuli, their energy demand is strongly elevated (Zezina et al., 2020). In order to provide fuel for the immune response, the glucose transporters expressed in the immune cells, particularly GLUT1 and GLUT3 are up-regulated (Zezina et al., 2020). Since Rheumatoid arthritis patients are already suffering from the poly-articular synovitis where an abnormal immune response is a known contributor, increased glucose uptake mediated by the up-regulated glucose transporters in the immune cells could potentially worsen the condition. This is supported by the finding of Veal and colleagues (Veal et al., 2014), who described protective effects against Rheumatoid arthritis occurring due to a 129-kb deletion on chromosome 12, eliminating the *SLC2A3* gene. In addition to the deletion of *SLC2A3*, *SLC2A14* was also partially disrupted. Veal and colleagues (Veal et al., 2014) dismissed *SLC2A14* as a potential candidate gene, since in the year of their study, expression of

SLC2A14 was only reported for testis (Wu & Freeze, 2002). However, since then, additional tissues expressing *SLC2A14* have been identified (Amir Shaghaghi et al., 2016), and we have confirmed that GLUT14 is mediating the transport of glucose, which is the same as GLUT3 (Amir Shaghaghi et al, 2017, Chapter 2). Moreover, single cell RNA-sequencing data also confirmed expression in immune cells (Human Protein Atlas, 2022a). Taken together, this strengthens the case that glucose dysregulation may indeed play a role in rheumatoid arthritis and implicates *SLC2A14* as a candidate gene.

6.2.3 *SLC2A14*/GLUT14 and implications for inflammatory bowel disease

The term inflammatory bowel disease is an umbrella term used for two conditions: Crohn's disease and ulcerative colitis (Ferrante, 2007). The symptoms for both Crohn's disease and ulcerative colitis include chronic inflammation of the digestive track (Fakhoury et al., 2014). For patients with Crohn's disease, inflammation can be found in any part of the digestive track from the mouth to the anus, with the distal portions, the small intestine (illum) and the large intestine, being the most susceptible regions (Gohil & Carramusa, 2014). For patients with ulcerative colitis, the affected intestinal regions include the inner linings of the large bowel, which are the colon and the rectum (Gohil & Carramusa, 2014). Previously, inflammatory bowel disease was characterized as an autoimmune disease where the inflammation in the bowel of patients was thought to be caused by the patient's immune cells attacking their body cells (Wen & Fiocchi, 2004). However, recent findings suggest that inflammatory bowel disease is rather an immune-mediated, autoinflammatory disease, where the inflammation in the bowel is a consequence of impaired innate immunity (Sartor & Wu, 2017; Wen & Fiocchi, 2004). The innate immunity is strong, fast-acting, and non-specific (Turvey & Broide, 2010). When the innate immunity is compromised, it leads to recurrent episodes of inflammation, which is a key symptom in patients with inflammatory bowel disease (Ferrante, 2007;

Mogensen, 2009). The innate immune system includes physical barriers such as the skin and the mucous membrane, as well as specialized immune cells such as the neutrophils. When the integrity of the physical barriers is compromised, and/or when the activity of immune cells is reduced, the innate immune system is impaired, causing inflammation (Mogensen, 2009; Turvey & Broide, 2010). Recent single cell RNA sequencing data revealed that the expression of *SLC2A14* is confined to the intestinal Paneth cells as well as the granulocytes, and we have found variations in SLC2A14 being associated to inflammatory bowel disease (Amir Shaghaghi et al., 2017). Moreover, we have confirmed that GLUT14 mediates the transport of glucose and dehydroascorbic acid (Amir Shaghaghi et al., 2017). Taken together, our finding implies a possible, direct role of GLUT14 in the maintenance of the intestinal innate immunity (Alhashim, 2022).

The Paneth cells are crucial members of the intestine epithelia since they are highly specialized secretory cells (Dayton & Clevers, 2017). They are primarily found in the crypts of Lieberkühn in the small intestine and are showed to be induced by inflammation in the colon (Dayton & Clevers, 2017). The key role of the Paneth cells is to produce and transport a collection of antimicrobial proteins and peptides such as α-defensins, angiogenin-4, lysozyme and secretory phospholipase A2 when bacteria are present (Wehkamp et al., 2006). In addition, also immunomodulating proteins that regulate the function and composition of the intestinal flora are produced by the Paneth cells, making them the basis of the intestinal antibacterial defense (Wehkamp & Stange, 2020).

There are many factors that negatively impact the function of the Paneth cells, such as smoking, antibiotic exposure, impaired intracellular bacterial recognition, and vitamin D deficiency; these are also known risk factors for the development of Crohn's disease (Piovani et al., 2019). Hence, multiple lines of evidence indicate that impaired Paneth cells immune function contributes to the development and the severity of the inflammatory bowel disease. The Paneth cells are highly glycolytic, requiring an abundant supply of glucose to fuel their metabolic activities (Dayton & Clevers,

2017). Since glucose requires transporters to gain entry into the cells, and we have shown that GLUT14 is a confirmed glucose transporter (Amir Shaghaghi et al., 2017) but also expressed in the Paneth cells, alteration in the GLUT14 function could impair the glucose absorption in the Paneth cells, disrupting its function, and affecting the development of inflammatory bowel disease negatively.

Another key member of the innate immune system are granulocytes, including the neutrophils. They constitute up to 70% of all white blood cells in humans and are recruited swiftly to sites of injury with ongoing acute inflammation (Rosales, 2018). In the digestive tract, the primary role of the neutrophils is to eliminate luminal pathogens to prevent them from crossing the intestinal mucosa (Fournier & Parkos, 2012). In addition to the production of antimicrobial chemicals similar to the Paneth cells, neutrophils additionally exhibit a so-called "oxidative burst", which critically depends on the metabolism of dehydroascorbic acid (Chen & Junger, 2012). During the oxidative burst, neutrophils produce and release reactive oxygen species and use them to directly harm extracellular pathogens (Mills et al., 2016). When this mechanism is abolished, severely delayed clearance of bacterial infections in chronic granulomatous disease was observed, indicating its importance in the functions of the neutrophils (Goldschmidt, 1991; Shilotri, 1977a, 1977b). When the reactive oxygen species are released into the extracellular environment, they oxidize ascorbic acid and convert it into dehydroascorbic acid, which becomes a substrate for GLUT14. A following process called "ascorbic acid recycling" will then take place where large quantities of dehydroascorbic acid are transported into the neutrophils and surrounding cells (Levine et al., 1993; Washko et al., 1993). Once the dehydroascorbic acid enters a cell, it is reduced to ascorbic acid, increasing the intracellular ascorbic acid level up to 20 folds (Washko et al., 1993). This increased level of intracellular ascorbic acid may serve as an important antioxidant, providing protective effects for the neutrophils against the oxidative burst and prolong their survivability or enhance their functioning. This theory was tested on scorbutic animals and their neutrophils showed decreased chemotaxis, phagocytosis, the reactive oxygen species

production and microbial eliminating compared to the control groups. When ascorbic acid is supplemented, the activity of neutrophils is restored, indicating the availability of vitamin C, as well as the cell's ability to utilize vitamin C are crucial in the preservation of the immunity (Goldschmidt, 1991; Shilotri, 1977a, 1977b). Since GLUT14 has been proven to transport dehydroascorbic acid, it could play a direct role in the uptake of dehydroascorbic acid into the neutrophils. With properly functioning neutrophils, local, acute inflammation in the digestive tract could be ameliorated to prevent it from progressing into chronic and severe inflammation, improving the prognosis in patients with inflammatory bowel disease.

6.3 DISCOVERY OF MFSD14A AS A NOVEL AMMONIUM TRANSPORTER

Ammonia is a byproduct produced during the catabolism of nitrogen-containing compounds such as amino acids and nucleic acids (Adeva et al., 2012; Veauvy et al., 2005; Veauvy et al., 2009). Ammonia is highly toxic, and as the production of ammonia is an ongoing process in all living cells, an efficient detoxification and excretion is required to avoid the deleterious neurotoxicity effects as the consequences of ammonia buildup (Adlimoghaddam et al., 2016). In humans, ammonia is detoxified in the liver through the urea cycle where toxic ammonia is being converted into urea prior to excretion (Adeva et al., 2012; Weiner & Verlander, 2017). Some ammonia, on the other hand, is generated in the kidneys during the synthesis of sodium bicarbonate, an important buffering agent in the maintenance of overall acid-base homeostasis (Weiner & Verlander, 2017).

Regardless of the fate of the ammonia, whether to undergo detoxification in the liver, or to be excreted from the kidney, transportation of ammonia is required by all living cells and must be tightly regulated. Our understanding of ammonia transport mechanisms, however, is far from being complete and ongoing process, tightly linked to the discoveries of new ammonia transporters. Prior to our knowing that the ammonia transporters do indeed exist, ammonia was thought to cross the cell

membrane exclusively by simple diffusion down the partial pressure gradient of NH_3. This theory was later proven to not be entirely correct, since the NH_3 permeability of biological membranes is relatively low (Antonenko et al., 1997) and membrane transporter proteins, such as members of the Rh-protein family have been discovered to be expressed in mammalian and invertebrate systems (Marini et al., 2000; Weihrauch et al., 2004). To date, a handful of transporters are found to accept ammonia as a substrate, and several ammonia transport models have been developed accordingly. These models, however, can only explain the overall ammonia transport mechanism partially due to their limitations, which will be discussed in detail. Since we report the discovery of a novel ammonium transporter, MFSD14A, which has not yet been included in any of the existing ammonia transport model, our results warrant the update and refinement on the current models accordingly.

In humans, the kidney is strongly involved in the production as well as the excretion of ammonia during the maintenance of acid-base homeostasis. Several detailed reviews on the human renal ammonia transport have been published by Weiner & Verlander (Weiner & Verlander, 2013, 2017, 2019). Key pathways will be discussed below. In the kidney, the proximal tubule accounts for up to 80% of the total renal ammonia production (Weiner & Verlander, 2013). When glutamine is catabolized in the mitochondria of the proximal tubule cells, it is believed to be excreted into the lumen of the proximal tubule via apically localized Na^+/H^+ antiporter 3, NHE3, encoded by gene *SLC9A3* (Ambühl et al., 1996; Amemiya et al., 1999). However, a recent study on the proximal tubule specific NHE3 knockout mice found reduced fluid and bicarbonate reabsorption but an unaltered NH_4^+ excretion (Li et al., 2013). These findings suggests that the key role of NHE3 is rather linked to fluid and bicarbonate reabsorption in the proximal tubule, instead of NH_4^+ excretion (Li et al., 2013). Our findings provided proof that MFSD14A functions as an NH_4^+ transporter and recent observations by single cell RNA-seq data confirming the presence of *MFSD14A* transcripts in

proximal tubule cells (Human Protein Atlas, 2022b), thus, it can be speculated that MFSD14A is candidate and involved in the proximal tubule's ammonia excretion mechanism.

The Rhesus (Rh) glycoproteins-mediated NH_3 excretion via an acid-trapping mechanism is another key ammonia transport model established by investigating epithelial cells from the collecting duct (Planelles, 2007; Weiner, 2004; Weiner & Verlander, 2013). Briefly, the basolateral RhBG mediate the transport of NH_3 down its partial pressure gradient (P_{NH3}) from the interstitium into the epithelial cells (Weiner, 2004). The cellular ammonia will then be excreted through the apical RhCG into the lumen of the collecting duct also along the P_{NH3}. This is created by an active proton excretion via a co-expressed V-ATPase, converting NH_3 in NH_4^+, lowering thereby the NH_3 abundance in the lumen (Weiner, 2004). The involvement of RhCG and ammonia excretion has been supported by experimental evidence as Rhcg-knockout mice developed metabolic acidosis and had reduced renal ammonia output (Biver et al., 2008). However, Rh-protein mediated ammonia excretion might not be the sole pathway in epithelial cells. For instance, when the branchial apical localized Rhcg protein in zebrafish was eliminated in a knock-out study, the total ammonia excretion was unaltered (Zimmer & Perry, 2020), supporting the possibility of an existing undiscovered ammonia excretion pathways. Since the mRNA expression of MFSD14A is also found in collection duct epithelial cells (Human Protein Atlas, 2022b), one can speculate that MFSD14A mediates NH_4^+ efflux into the lumen. This assumption needs to be verified by demonstrating apical localization of MFSD14A in collecting duct cells and perfusion experiments, ideally using a genetically-modified mouse model with the *Mfsd14a* gene being eliminated.

The transcellular ammonia transport has been described for many excretory tissues and epithelia particularly for the gills of ammonotelic aquatic invertebrates and fish and, of course the mammalian kidney, as reviewed by Wilkie on the ammonia and urea handling of the fish gills (Wilkie, 2002), by Fehsenfeld & Weihrauch in the European green crabs (Fehsenfeld & Weihrauch, 2013), and

by Smith and colleagues on goldfish (Smith et al., 2012). In general ammonia is either entering in the form of NH_3 via the basolateral Rhbg transporter or as NH_4^+ by means of the Na^+/K^+-ATPase, substituting K^+ (Larsen et al., 2016). Cellular release occurs then via an aforementioned apical ammonia trapping mechanism and Rhcg. On the other hand, so far, the mechanism of how regular tissue cells are getting rid of metabolically produced and potentially very toxic ammonia, has not been investigated in any detail. It is noteworthy that the abundance the Rh-protein in non-excretory tissues is indeed very low (Martin et al., 2011), while in some cells these ammonia transporters have not been detected at all (see table 6.2.1). By contrast, the newly discovered ammonia transporter, *MFSD14*A, was found in every cell/tissue so far investigated. Therefore, we assume that *MFSD14A* plays an important role in keeping cellular ammonia levels in a tolerable range.

Cell type	**Tissue**
Alpha cells	Pancreas
Ascending loop of Henle cells	Kidney
Cardiomyocytes	Heart
Cholangiocytes	Liver
Colon enterocytes	Colon
Colon enteroendocrine cells	Colon
Heart mitotic cells	Heart
Lung natural killer cells	Lung
Thyroid glandular cells	Thyroid

Table 6.3.1 A list of cells with confirmed *MFSD14A* mRNA expression and without the expression of *RhBG* and *RhCG*. The list was generated by cross-comparison between *MFSD14A*, *RhBG* and *RhCG* tissue cell RNA sequencing data (Human Protein Atlas, 2022c; Human Protein Atlas, 2022d; Human Protein Atlas, 2022e).

6.3.1 *MFSD14A*/MFSD14A and implications for globozoospermia

Globozoospermia is a condition responsible for 0.1% of the total male infertility and is characterized by abnormally formed sperm cells lacking the acrosome (Modarres et al., 2019). The acrosome produces digestive enzymes which break down the outer membrane of an egg cell (Modarres et al., 2019). When the acrosome is aberrant or missing, the sperms are no longer able to penetrate the zona pellucida, resulting in infertility (Modarres et al., 2019). The exact cause of globozoospermia is unknown, however, mutations in several genes have been associated with this disease. The *DPY19L2* gene has been the most frequently associated gene since over 70% of the patients with globozoospermia are found to carry mutations in *DPY19L2* (Dam et al., 2007; Pierre et al., 2012; Xiao et al., 2009; Yassine et al., 2015). When the *Mfsd14*a gene was knocked-down in mice, adult male homozygous mice (*Mfsd14a*$^{-/-}$) mutants became infertile with sperms lacking the acrosome, similar to the phenotype of globozoospermia (Doran et al., 2016). This finding indicates that MFSD14A could play a similar role as *DPY19L2* in the process of acrosome formation. Unfortunately, the function of *DPY19L2* has not yet been confirmed and the exact mechanism of how globozoospermia is linked to the mutations in *DPY19L2* is unknown (Modarres et al., 2019; Pierre et al., 2012; Yassine et al., 2015). In any case, MFSD14A seems to be important for proper spermatogenesis, likely by keeping the presence of toxic ammonia low in the gonads. Since the formation of the acrosome depends on transporting proteins to the developing head of the spermatid by the manchette, one can predict that protein catabolism increases during this process. This would result in an elevated production of ammonia, that needs to be excreted from the cells In fact, single-cell RNA sequencing data (Human Protein Atlas, 2022f) revealed that *MFSD14A* is highly expressed in the early and late spermatids, possibly playing a critical role eliminating cytotoxic ammonia build-up by excretion, allowing thereby normal acrosome formation. Clearly, more detailed studies are needed to confirm this hypothesis.

6.4 STRENGTHS

One of the key strengths for our studies is the utilization of multiple information sources as well as experimental model systems to contest the same research question. By doing so, the comprehensiveness and the accuracy of our results was maximized. For instance, when determining the genomic organization, expression profiles, secondary and tertiary structures for *MFSD14A*/MFSD14A, multiple databases and programs were used followed by manual annotation to ensure the validity of the information. Since most of the data deposited in the databases have not yet been manually curated, and the synchronization between databases are often delayed or lacking, pooling information from all available databases and manually annotate them is crucial to ensure that the results are error-free and all the details are being captured as much as possible.

When exploring the functions of GLUT14 and MFSD14A, we chose to use the *Xenopus laevis* oocyte system as a well-established model. There are several advantages of the frog oocyte expression system. First, the oocyte system is mature and well-established. It has been used to study the heterologously expressed proteins for more than four decades (Gurdon, 1973). Second, the natural, unmodified oocytes are very inactive and do not transport many molecules, including glucose, making it an ideal model to study membrane transporters (Bröer, 2010). Since glucose is not transported by unmodified oocytes, the glucose transport mediated by oocytes expressing GLUT14 or MFSD14A could be determined easily. Moreover, thousands of oocytes can be harvested from one female *Xenopus laevis*, providing sufficient samples for all experiments to ensure the experimental and statistical accuracy. During the experiments, we also took extra precaution by using 15 or more oocytes in each experimental

group, quantifying the substrate concentration in the oocytes individually rather than homogenizing them together. Moreover, each experiment was repeated at least twice with oocytes harvested from two different frogs to further rule out any possible error caused by the biological variables naturally found in the frogs. Last but not least, radiolabeled isotopes can be used in the oocyte transport studies. By tracking the movement of the H^3 and C^{14} labels from the incubation buffer to the oocyte's cytoplasm, each oocyte's transport data can be calculated without the influence of background or the possible internal difference within the oocytes.

Besides testing the function of MFSD14A by heterologously expressing it in the frog oocytes, we additionally employed the zebrafish model system to further verify the function of MFSD14A. The zebrafish model is commonly used for the study of biological processes and human diseases (Bradford et al., 2017). One of the key advantages of the zebrafish model is that it can be easily genetically modified to study the physiological function of a gene. After confirming that ammonia is the substrate of MFSD14A, we have generated Mfsd14a-knockdown zebrafish larvae using morpholino oligomers and measured their regional and total ammonia excretion. In contrast to the measurement of radiolabeled isotopes in oocyte-expression studies, we used a colorimetric assay to determine the total ammonia excretion and the Scanning Ion-Selective Electrode Technique (SIET) to measure in real time NH_4^+ movements right at the surface of the tissues. These approaches not only confirmed our data generated from the previous oocyte experiments, but also allowed us to gain insights on the role of MFSD14A in ammonia excretion.

6.5 LIMITATIONS

There are several limitations that can be found in our studies. First, the *Xenopus laevis* oocyte system we chose for the substrate verification had a complex, endogenous ammonia transport mechanism. Some ammonia transporters, such as the Rhesus protein Rhag, are expressed endogenously in the oocyte, even though to very a low extent, likely to protect the cell from toxic intracellular ammonia build-up (Fehsenfeld person. communication). Second, ammonia itself cannot be radiolabeled. Even though radiolabled methylamine is an acceptable proxy for ammonia, the structural difference between methylamine and ammonia is distinct and transport kinetic have been reported to be different (Nawata et al., 2010).

Taken together, the oocyte system is less ideal for studying ammonia transporters compared to glucose or dehydroascorbic acid transporters, but still the best system to be used.

Another limitation of our studies was that the localization of MFSD14A could not be determined by the experimental approaches we took. The *Xenopus* oocyte system allow the heterologous expression of MFSD14A, however, the oocytes are individual, non-polarized cells. Since oocytes are not epithelial cells and thereby do not possess apical and basolateral membranes, the cellular localization of MFSD14A in ammonia transporting epithelial cells could not be determined in this system. Since the cellular localization of MFSD14A is lacking, we were unable to precisely incorporate MFSD14A into the existing ammonia transport models, e.g. the freshwater gill or the mammalian nephron, since most of them are established in tissues containing epithelial cells.

6.6 CONCLUSION AND FUTURE DIRECTION

Since we have confirmed the substrates of GLUT14 (glucose and dehydroascorbic acid), the role of GLUT14 and its gene, SLC2A14 can be re-evaluated in the diseases that have been previously associated with them. Results regarding the function of GLUT14 from our studies will be useful in the development of future clinical trials, where the impact of the GLUT14 substrates in the disease interventions can be tested.

For MFSD14A, a newly discovered ammonia transporter, more work needs to be done. First, the transport kinetics of MFSD14A need to be determined. The transport module of MFSD14A should also be assessed to determine whether or not the transport of ammonia mediated by MFSD14A is sodium dependent. This can be done either using different radiolabeled isotopes or through electrophysiological approaches such as the patch clamp techniques. The localization of MFSD14A needs to be determined in not only the epithelial cells, but also internal cells since the expression of *MFSD14A* is found to be ubiquitous. Immunohistochemistry utilizing reliable antibodies, ideally custom-made monoclonal antibodies could be considered as a method to determine the localization of MFSD14A. Once this is determined, the existing ammonia transport models, such as the human renal ammonia transport model, can be updated accordingly. Novel ammonia transport models can also be developed for the internal, non-epithelial cells expressing *MFSD14A* since there is currently no ammonia transport mechanism established for them.

Additional studies utilizing the existing Mfsd14a-knockout mouse model can also be done by challenging the Mfsd14a-mutant mice with increased ammonia load and measure the ammonia levels from different regions of the testis. This will reveal whether or not MFSD14A plays a direct, protective role during the development of spermatocytes. Existing cohort studies focusing on globozoospermia in humans can also check the expression status of MFSD14A in their participants, to further understand the relationship between MFSD14A and globozoospermia in humans. Moreover, the

kidneys from the Mfsd14a-knockout mouse model can also be used in perfusion studies where the ammonia transport and filtration in the Mfsd14a-eliminated mouse kidney can be compared to kidneys harvested from the control mice, to further elucidate the role of Mfsd14a in the renal ammonia transport. Importantly, when conducting experiments using the Mfsd14a-knockout mouse model, one must consider the potential gene compensatory effects and measure the expression profiles of the other ammonia transporters expressed in the tissues, such as the Rhesus glycoproteins, to ensure the results generated from the Mfsd14a-knockout mouse model are not affected by the possible compensatory effects. Alternatively, conditional knockout mouse models can be generated using the Cre-LoxP Recombination technique, to further reduce the potential compensatory effects found in the global knockout models. A detailed review on this specific technique can be found in Tsien, 2016.

6.7 REFERENCES

Adamson, S. R., & Schmidli, B. (1986). Industrial mammalian cell culture. *The Canadian Journal of Chemical Engineering*, *64*(4), 531-539. https://doi.org/https://doi.org/10.1002/cjce.5450640402

Adeva, M. M., Souto, G., Blanco, N., & Donapetry, C. (2012). Ammonium metabolism in humans. *Metabolism*, *61*(11), 1495-1511. https://doi.org/10.1016/j.metabol.2012.07.007

Adlimoghaddam, A., Sabbir, M. G., & Albensi, B. C. (2016). Ammonia as a Potential Neurotoxic Factor in Alzheimer's Disease. *Front Mol Neurosci*, *9*, 57. https://doi.org/10.3389/fnmol.2016.00057

Alhashim, A. (2022). *The homo-sapiens solute carrier family 2 member 14 (SLC2A14) – further insight into the genomic organization, protein isoforms, substrates and kinetics*

Ambühl, P. M., Amemiya, M., Danczkay, M., Lötscher, M., Kaissling, B., Moe, O. W., . . . Alpern, R. J. (1996). Chronic metabolic acidosis increases NHE3 protein abundance in rat kidney. *Am J Physiol*, *271*(4 Pt 2), F917-925. https://doi.org/10.1152/ajprenal.1996.271.4.F917

Amemiya, M., Tabei, K., Kusano, E., Asano, Y., & Alpern, R. J. (1999). Incubation of OKP cells in low-K+ media increases NHE3 activity after early decrease in intracellular pH. *Am J Physiol*, *276*(3), C711-716. https://doi.org/10.1152/ajpcell.1999.276.3.C711

Amir Shaghaghi, M., Murphy, B., & Eck, P. (2016). The SLC2A14 gene: genomic locus, tissue expression, splice variants, and subcellular localization of the protein. *Biochem Cell Biol*, *94*(4), 331-335. https://doi.org/10.1139/bcb-2015-0089

Amir Shaghaghi, M., Zhouyao, H., Tu, H., El-Gabalawy, H., Crow, G. H., Levine, M., . . . Eck, P. (2017). The SLC2A14 gene, encoding the novel glucose/dehydroascorbate transporter GLUT14, is associated with inflammatory bowel disease. *Am J Clin Nutr*, *106*(6), 1508-1513. https://doi.org/10.3945/ajcn.116.147603

Antonenko, Y. N., Pohl, P., & Denisov, G. A. (1997). Permeation of ammonia across bilayer lipid membranes studied by ammonium ion selective microelectrodes. *Biophys J*, *72*(5), 2187-2195. https://doi.org/10.1016/s0006-3495(97)78862-3

Arathoon, W. R., & Birch, J. R. (1986). Large-scale cell culture in biotechnology. *Science*, *232*(4756), 1390-1395. https://doi.org/10.1126/science.2424083

Augustin, R. (2010). The protein family of glucose transport facilitators: It's not only about glucose after all. *IUBMB Life*, *62*(5), 315-333. https://doi.org/10.1002/iub.315

Avila, J., Lucas, J. J., Perez, M., & Hernandez, F. (2004). Role of tau protein in both physiological and pathological conditions. *Physiol Rev*, *84*(2), 361-384. https://doi.org/10.1152/physrev.00024.2003

Backer, M. P., Metzger, L. S., Slaber, P. L., Nevitt, K. L., & Boder, G. B. (1988). Large-scale production of monoclonal antibodies in suspension culture. *Biotechnol Bioeng*, *32*(8), 993-1000. https://doi.org/10.1002/bit.260320807

Bibert, S., Hess, S. K., Firsov, D., Thorens, B., Geering, K., Horisberger, J. D., & Bonny, O. (2009). Mouse GLUT9: evidences for a urate uniporter. *Am J Physiol Renal Physiol*, *297*(3), F612-619. https://doi.org/10.1152/ajprenal.00139.2009

Biver, S., Belge, H., Bourgeois, S., Van Vooren, P., Nowik, M., Scohy, S., . . . Marini, A. M. (2008). A role for Rhesus factor Rhcg in renal ammonium excretion and male fertility. *Nature*, *456*(7220), 339-343. https://doi.org/10.1038/nature07518

Bradford, Y. M., Toro, S., Ramachandran, S., Ruzicka, L., Howe, D. G., Eagle, A., . . . Westerfield, M. (2017). Zebrafish Models of Human Disease: Gaining Insight into Human Disease at ZFIN. *Ilar j*, *58*(1), 4-16. https://doi.org/10.1093/ilar/ilw040

Bröer, S. (2010). Xenopus laevis Oocytes. *Methods Mol Biol*, *637*, 295-310. https://doi.org/10.1007/978-1-60761-700-6_16

Bullock, J., Rizvi, S. A. A., Saleh, A. M., Ahmed, S. S., Do, D. P., Ansari, R. A., & Ahmed, J. (2018). Rheumatoid Arthritis: A Brief Overview of the Treatment. *Med Princ Pract*, *27*(6), 501-507. https://doi.org/10.1159/000493390

Chen, Y., & Junger, W. G. (2012). Measurement of oxidative burst in neutrophils. *Methods Mol Biol*, *844*, 115-124. https://doi.org/10.1007/978-1-61779-527-5_8

Dam, A. H., Koscinski, I., Kremer, J. A., Moutou, C., Jaeger, A. S., Oudakker, A. R., . . . Viville, S. (2007). Homozygous mutation in SPATA16 is associated with male infertility in human globozoospermia. *Am J Hum Genet*, *81*(4), 813-820. https://doi.org/10.1086/521314

Dayton, T. L., & Clevers, H. (2017). Beyond growth signaling: Paneth cells metabolically support ISCs. *Cell Res*, *27*(7), 851-852. https://doi.org/10.1038/cr.2017.59

Doran, J., Walters, C., Kyle, V., Wooding, P., Hammett-Burke, R., & Colledge, W. H. (2016). Mfsd14a (Hiat1) gene disruption causes globozoospermia and infertility in male mice. *Reproduction*, *152*(1), 91-99. https://doi.org/10.1530/REP-15-0557

Fakhoury, M., Negrulj, R., Mooranian, A., & Al-Salami, H. (2014). Inflammatory bowel disease: clinical aspects and treatments. *J Inflamm Res*, *7*, 113-120. https://doi.org/10.2147/jir.S65979

Fehsenfeld, S., & Weihrauch, D. (2013). Differential acid-base regulation in various gills of the green crab Carcinus maenas: Effects of elevated environmental pCO2. *Comp Biochem Physiol A Mol Integr Physiol*, *164*(1), 54-65. https://doi.org/10.1016/j.cbpa.2012.09.016

Ferrante, M. (2007). Fast Facts: Inflammatory Bowel Disease, 2nd ed. *Gut*, *56*(9), 1332. https://doi.org/10.1136/gut.2007.121400

Fournier, B. M., & Parkos, C. A. (2012). The role of neutrophils during intestinal inflammation. *Mucosal Immunol*, *5*(4), 354-366. https://doi.org/10.1038/mi.2012.24

Gohil, K., & Carramusa, B. (2014). Ulcerative colitis and Crohn's disease. *P t*, *39*(8), 576-577.

Goldschmidt, M. C. (1991). Reduced bactericidal activity in neutrophils from scorbutic animals and the effect of ascorbic acid on these target bacteria in vivo and in vitro. *Am J Clin Nutr*, *54*(6 Suppl), 1214S-1220S. https://doi.org/10.1093/ajcn/54.6.1214s

Gurdon, J. B. (1973). The translation of messenger RNA injected in living oocytes of Xenopus laevis. *Acta Endocrinol Suppl (Copenh)*, *180*, 225-243. https://doi.org/10.1530/acta.0.074s225

Holman, G. D. (2020). Structure, function and regulation of mammalian glucose transporters of the SLC2 family. *Pflugers Arch*, *472*(9), 1155-1175. https://doi.org/10.1007/s00424-020-02411-3

Joost, H. G., Bell, G. I., Best, J. D., Birnbaum, M. J., Charron, M. J., Chen, Y. T., . . . Thorens, B. (2002). Nomenclature of the GLUT/SLC2A family of sugar/polyol transport facilitators. *Am J Physiol Endocrinol Metab*, *282*(4), E974-976 https://doi.org/10.1152/ajpendo.00407.2001

Kumar, A., Sidhu, J., Goyal, A., & Tsao, J. W. (2022). Alzheimer Disease. In *StatPearls*. StatPearls Publishing, Copyright © 2022, StatPearls Publishing LLC.

Kyrtata, N., Dickie, B., Emsley, H., & Parkes, L. (2021). Glucose transporters in Alzheimer's disease. *BJPsych Open*, *7*(S1), S265-S266. https://doi.org/10.1192/bjo.2021.707

Larsen, B. R., Stoica, A., & MacAulay, N. (2016). Managing Brain Extracellular K+ during Neuronal Activity: The Physiological Role of the Na+/K+-ATPase Subunit Isoforms [Review]. *Frontiers in Physiology*, *7*. https://doi.org/10.3389/fphys.2016.00141

Lee, Y. C., Huang, H. Y., Chang, C. J., Cheng, C. H., & Chen, Y. T. (2010). Mitochondrial GLUT10 facilitates dehydroascorbic acid import and protects cells against oxidative stress: mechanistic insight into arterial tortuosity syndrome. *Hum Mol Genet*, *19*(19), 3721-3733. https://doi.org/10.1093/hmg/ddq286

Levine, M., Dhariwal, K. R., Washko, P. W., Welch, R. W., & Wang, Y. (1993). Cellular functions of ascorbic acid: a means to determine vitamin C requirements. *Asia Pac J Clin Nutr*, *2 Suppl 1*, 5-13.

Li, H. C., Du, Z., Barone, S., Rubera, I., McDonough, A. A., Tauc, M., . . . Soleimani, M. (2013). Proximal tubule specific knockout of the Na(+)/H(+) exchanger NHE3: effects on bicarbonate absorption and ammonium excretion. *J Mol Med (Berl)*, *91*(8), 951-963. https://doi.org/10.1007/s00109-013-1015-3

Li, W., Qu, G., Choi, S.-C., Cornaby, C., Titov, A., Kanda, N., . . . Morel, L. (2019). Targeting T Cell Activation and Lupus Autoimmune Phenotypes by Inhibiting Glucose Transporters [Original Research]. *Frontiers in Immunology*, *10*. https://doi.org/10.3389/fimmu.2019.00833

Li, X., Lu, F., Wang, J. Z., & Gong, C. X. (2006). Concurrent alterations of O-GlcNAcylation and phosphorylation of tau in mouse brains during fasting. *Eur J Neurosci*, *23*(8), 2078-2086. https://doi.org/10.1111/j.1460-9568.2006.04735.x

Liu, F., Iqbal, K., Grundke-Iqbal, I., Hart, G. W., & Gong, C. X. (2004). O-GlcNAcylation regulates phosphorylation of tau: a mechanism involved in Alzheimer's disease. *Proc Natl Acad Sci U S A*, *101*(29), 10804-10809. https://doi.org/10.1073/pnas.0400348101

Liu, Y., Liu, F., Iqbal, K., Grundke-Iqbal, I., & Gong, C. X. (2008). Decreased glucose transporters correlate to abnormal hyperphosphorylation of tau in Alzheimer disease. *FEBS Lett*, *582*(2), 359-364. https://doi.org/10.1016/j.febslet.2007.12.035

Maher, F., & Harrison, L. C. (1990). Hexose specificity for downregulation of HepG2/brain-type glucose transporter gene expression in L6 myocytes. *Diabetologia*, *33*(11), 641-648. https://doi.org/10.1007/bf00400564

Marini, A. M., Matassi, G., Raynal, V., André, B., Cartron, J. P., & Chérif-Zahar, B. (2000). The human Rhesus-associated RhAG protein and a kidney homologue promote ammonium transport in yeast. *Nat Genet*, *26*(3), 341-344. https://doi.org/10.1038/81656

Martin, M., Fehsenfeld, S., Sourial, M. M., & Weihrauch, D. (2011). Effects of high environmental ammonia on branchial ammonia excretion rates and tissue Rh-protein mRNA expression levels in seawater acclimated Dungeness crab Metacarcinus magister. *Comp Biochem Physiol A Mol Integr Physiol*, *160*(2), 267-277. https://doi.org/10.1016/j.cbpa.2011.06.012

Mills, E. L., Kelly, B., Logan, A., Costa, A. S. H., Varma, M., Bryant, C. E., . . . O'Neill, L. A. (2016). Succinate Dehydrogenase Supports Metabolic Repurposing of Mitochondria to Drive

Inflammatory Macrophages. *Cell*, *167*(2), 457-470 e413. https://doi.org/10.1016/j.cell.2016.08.064

Modarres, P., Tavalaee, M., Ghaedi, K., & Nasr-Esfahani, M. H. (2019). An Overview of The Globozoospermia as A Multigenic Identified Syndrome. *Int J Fertil Steril*, *12*(4), 273-277. https://doi.org/10.22074/ijfs.2019.5561

Mogensen, T. H. (2009). Pathogen recognition and inflammatory signaling in innate immune defenses. *Clin Microbiol Rev*, *22*(2), 240-273, Table of Contents. https://doi.org/10.1128/cmr.00046-08

Mueckler, M., & Thorens, B. (2013). The SLC2 (GLUT) family of membrane transporters. *Mol Aspects Med*, *34*(2-3), 121-138. https://doi.org/10.1016/j.mam.2012.07.001

Nag, A., Venturini, C., Hysi, P. G., Arno, M., Aldecoa-Otalora Astarloa, E., Macgregor, S., . . . Hammond, C. J. (2013). Copy number variation at chromosome 5q21.2 is associated with intraocular pressure. *Invest Ophthalmol Vis Sci*, *54*(5), 3607-3612. https://doi.org/10.1167/iovs.13-11952

Bradford, Y. M., Toro, S., Ramachandran, S., Ruzicka, L., Howe, D. G., Eagle, A., . . . Westerfield, M. (2017). Zebrafish Models of Human Disease: Gaining Insight into Human Disease at ZFIN. *Ilar j*, *58*(1), 4-16. https://doi.org/10.1093/ilar/ilw040

Bröer, S. (2010). Xenopus laevis Oocytes. *Methods Mol Biol*, *637*, 295-310. https://doi.org/10.1007/978-1-60761-700-6_16

Gurdon, J. B. (1973). The translation of messenger RNA injected in living oocytes of Xenopus laevis. *Acta Endocrinol Suppl (Copenh)*, *180*, 225-243. https://doi.org/10.1530/acta.0.074s225

Nawata, C. M., Wood, C. M., & O'Donnell, M. J. (2010). Functional characterization of Rhesus glycoproteins from an ammoniotelic teleost, the rainbow trout, using oocyte expression and SIET analysis. *J Exp Biol*, *213*(Pt 7), 1049-1059. https://doi.org/10.1242/jeb.038752

Pi, H., Zhou, H., Jin, H., Ning, Y., & Wang, Y. (2017). Abnormal Glucose Metabolism in Rheumatoid Arthritis. *BioMed Research International*, *2017*, 9670434. https://doi.org/10.1155/2017/9670434

Pierre, V., Martinez, G., Coutton, C., Delaroche, J., Yassine, S., Novella, C., & Arnoult, C. (2012). Absence of Dpy19l2, a new inner nuclear membrane protein, causes globozoospermia in mice by preventing the anchoring of the acrosome to the nucleus. *Development*, *139*(16), 2955-2965. https://doi.org/10.1242/dev.077982

Piovani, D., Danese, S., Peyrin-Biroulet, L., Nikolopoulos, G. K., Lytras, T., & Bonovas, S. (2019). Environmental Risk Factors for Inflammatory Bowel Diseases: An Umbrella Review of Meta-analyses. *Gastroenterology*, *157*(3), 647-659 e644. https://doi.org/10.1053/j.gastro.2019.04.016

Planelles, G. (2007). Ammonium Homeostasis and Human Rhesus Glycoproteins. *Nephron Physiology*, *105*(1), p11-p17. https://doi.org/10.1159/000096979

Rosales, C. (2018). Neutrophil: A Cell with Many Roles in Inflammation or Several Cell Types? [Review]. *Frontiers in Physiology*, *9*. https://doi.org/10.3389/fphys.2018.00113

Sartor, R. B., & Wu, G. D. (2017). Roles for Intestinal Bacteria, Viruses, and Fungi in Pathogenesis of Inflammatory Bowel Diseases and Therapeutic Approaches. *Gastroenterology*, *152*(2), 327-339 e324. https://doi.org/10.1053/j.gastro.2016.10.012

Schneider, M., Marison, I. W., & von Stockar, U. (1996). The importance of ammonia in mammalian cell culture. *J Biotechnol*, *46*(3), 161-185. https://doi.org/10.1016/0168-1656(95)00196-4

Shilotri, P. G. (1977a). Glycolytic, hexose monophosphate shunt and bactericidal activities of leukocytes in ascorbic acid deficient guinea pigs. *J Nutr*, *107*(8), 1507-1512. https://doi.org/10.1093/jn/107.8.1507

Shilotri, P. G. (1977b). Phagocytosis and leukocyte enzymes in ascorbic acid deficient guinea pigs. *J Nutr*, *107*(8), 1513-1516. https://doi.org/10.1093/jn/107.8.1513

Shulman, J. M., Chipendo, P., Chibnik, L. B., Aubin, C., Tran, D., Keenan, B. T., . . . De Jager, P. L. (2011). Functional screening of Alzheimer pathology genome-wide association signals in Drosophila. *Am J Hum Genet*, *88*(2), 232-238. https://doi.org/10.1016/j.ajhg.2011.01.006

Smith, A. A., Zimmer, A. M., & Wood, C. M. (2012). Branchial and extra-branchial ammonia excretion in goldfish (Carassius auratus) following thermally induced gill remodeling. *Comp Biochem Physiol A Mol Integr Physiol*, *162*(3), 185-192. https://doi.org/10.1016/j.cbpa.2012.02.019

So, A., & Thorens, B. (2010). Uric acid transport and disease. *J Clin Invest*, *120*(6), 1791-1799. https://doi.org/10.1172/jci42344

Tsien, J. Z. (2016). Cre-Lox Neurogenetics: 20 Years of Versatile Applications in Brain Research and Counting… [Perspective]. *Frontiers in Genetics*, *7*. https://doi.org/10.3389/fgene.2016.00019

Tran, C. N., Lundy, S. K., & Fox, D. A. (2005). Synovial biology and T cells in rheumatoid arthritis. *Pathophysiology*, *12*(3), 183-189. https://doi.org/10.1016/j.pathophys.2005.07.005

Turvey, S. E., & Broide, D. H. (2010). Innate immunity. *J Allergy Clin Immunol*, *125*(2 Suppl 2), S24-32. https://doi.org/10.1016/j.jaci.2009.07.016

Uldry, M., Ibberson, M., Horisberger, J. D., Chatton, J. Y., Riederer, B. M., & Thorens, B. (2001). Identification of a mammalian H(+)-myo-inositol symporter expressed predominantly in the brain. *Embo j*, *20*(16), 4467-4477. https://doi.org/10.1093/emboj/20.16.4467

Uldry, M., & Thorens, B. (2004). The SLC2 family of facilitated hexose and polyol transporters. *Pflugers Arch*, *447*(5), 480-489. https://doi.org/10.1007/s00424-003-1085-0

Veal, C. D., Reekie, K. E., Lorentzen, J. C., Gregersen, P. K., Padyukov, L., & Brookes, A. J. (2014). A 129-kb deletion on chromosome 12 confers substantial protection against rheumatoid arthritis, implicating the gene SLC2A3. *Hum Mutat*, *35*(2), 248-256. https://doi.org/10.1002/humu.22471

Veauvy, C. M., McDonald, M. D., Van Audekerke, J., Vanhoutte, G., Van Camp, N., Van der Linden, A., & Walsh, P. J. (2005). Ammonia affects brain nitrogen metabolism but not hydration status in the Gulf toadfish (Opsanus beta). *Aquat Toxicol*, *74*(1), 32-46. https://doi.org/10.1016/j.aquatox.2005.05.003

Veauvy, C. M., Walsh, P. J., & McDonald, M. D. (2009). Effect of elevated ammonia on tissue nitrogen metabolites in the ureotelic gulf toadfish (Opsanus beta) and the ammoniotelic midshipman (Porichthys notatus). *Physiol Biochem Zool*, *82*(4), 345-352. https://doi.org/10.1086/588829

Wang, W., Yu, J. T., Zhang, W., Cui, W. Z., Wu, Z. C., Zhang, Q., & Tan, L. (2012). Genetic association of SLC2A14 polymorphism with Alzheimer's disease in a Han Chinese population. *J Mol Neurosci*, *47*(3), 481-484. https://doi.org/10.1007/s12031-012-9748-y

Washko, P. W., Wang, Y., & Levine, M. (1993). Ascorbic acid recycling in human neutrophils. *J Biol Chem*, *268*(21), 15531-15535.

Wehkamp, J., Chu, H., Shen, B., Feathers, R. W., Kays, R. J., Lee, S. K., & Bevins, C. L. (2006). Paneth cell antimicrobial peptides: topographical distribution and quantification in human gastrointestinal tissues. *FEBS Lett*, *580*(22), 5344-5350. https://doi.org/10.1016/j.febslet.2006.08.083

Wehkamp, J., & Stange, E. F. (2020). An Update Review on the Paneth Cell as Key to Ileal Crohn's Disease. *Front Immunol*, *11*, 646. https://doi.org/10.3389/fimmu.2020.00646

Weihrauch, D., Morris, S., & Towle, D. W. (2004). Ammonia excretion in aquatic and terrestrial crabs. *J Exp Biol*, *207*(Pt 26), 4491-4504. https://doi.org/10.1242/jeb.01308

Weiner, I. D. (2004). The Rh gene family and renal ammonium transport. *Curr Opin Nephrol Hypertens*, *13*(5), 533-540. https://doi.org/10.1097/00041552-200409000-00009

Weiner, I. D., & Verlander, J. W. (2013). Renal ammonia metabolism and transport. *Compr Physiol*, *3*(1), 201-220. https://doi.org/10.1002/cphy.c120010

Weiner, I. D., & Verlander, J. W. (2017). Ammonia Transporters and Their Role in Acid-Base Balance. *Physiol Rev*, *97*(2), 465-494. https://doi.org/10.1152/physrev.00011.2016

Weiner, I. D., & Verlander, J. W. (2019). Emerging Features of Ammonia Metabolism and Transport in Acid-Base Balance. *Semin Nephrol*, *39*(4), 394-405. https://doi.org/10.1016/j.semnephrol.2019.04.008

Weller, J., & Budson, A. (2018). Current understanding of Alzheimer's disease diagnosis and treatment. *F1000Res*, *7*. https://doi.org/10.12688/f1000research.14506.1

Wen, Z., & Fiocchi, C. (2004). Inflammatory bowel disease: autoimmune or immune-mediated pathogenesis? *Clin Dev Immunol*, *11*(3-4), 195-204. https://doi.org/10.1080/17402520400004201

Wilkie, M. P. (2002). Ammonia excretion and urea handling by fish gills: present understanding and future research challenges. *J Exp Zool*, *293*(3), 284-301. https://doi.org/10.1002/jez.10123

Wu, X., & Freeze, H. H. (2002). GLUT14, a duplicon of GLUT3, is specifically expressed in testis as alternative splice forms. *Genomics*, *80*(6), 553-557. https://doi.org/10.1006/geno.2002.7010

Xiao, N., Kam, C., Shen, C., Jin, W., Wang, J., Lee, K. M., . . . Xia, J. (2009). PICK1 deficiency causes male infertility in mice by disrupting acrosome formation. *J Clin Invest*, *119*(4), 802-812. https://doi.org/10.1172/JCI36230

Yassine, S., Escoffier, J., Martinez, G., Coutton, C., Karaouzene, T., Zouari, R., . . . Arnoult, C. (2015). Dpy19l2-deficient globozoospermic sperm display altered genome packaging and DNA damage that compromises the initiation of embryo development. *Mol Hum Reprod*, *21*(2), 169-185. https://doi.org/10.1093/molehr/gau099

Zezina, E., Sercan-Alp, O., Herrmann, M., & Biesemann, N. (2020). Glucose transporter 1 in rheumatoid arthritis and autoimmunity. *Wiley Interdiscip Rev Syst Biol Med*, *12*(4), e1483. https://doi.org/10.1002/wsbm.1483

Zimmer, A. M., & Perry, S. F. (2020). The Rhesus glycoprotein Rhcgb is expendable for ammonia excretion and Na(+) uptake in zebrafish (Danio rerio). *Comp Biochem Physiol A Mol Integr Physiol*, *247*, 110722. https://doi.org/10.1016/j.cbpa.2020.110722

Zirkin, B. R., & Goldberg, E. (2018). Spermatids. In M. K. Skinner (Ed.), *Encyclopedia of Reproduction (Second Edition)* (pp. 42-46). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-801238-3.64426-4