SLC11A1 PROMOTER POLYMORPHISMS, GENE EXPRESSION AND ASSOCIATION WITH AUTOIMMUNE AND INFECTIOUS DISEASES

A Thesis Submitted for the Degree

of

Doctor of Philosophy

by

Nicholas Steven Archer
B. Sc.(Hons1)

School of Medical and Molecular Biosciences, Faculty of Science,
University of Technology Sydney, Australia.

2012

CERTIFICATE OF AUTHORSHIP/ORIGINALITY

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

_____

Nicholas Steven Archer

2012

ACKNOWLEDGMENTS

I am grateful to many people who have supported and helped me throughout the completion of my postgraduate studies. In particular, my supervisors Dr Bronwyn O'Brien and Dr Najah Nassif, both of whom have provided immeasurable guidance, support and wisdom throughout the completion of my postgraduate work. I thank you for the effort and enthusiasm you have shown to my work.

I wish to extend a special thanks to Stephanie Dowdell for her friendship and support through the completion of my PhD and assistance with reverse-transcriptase real-time PCR. I also appreciate and acknowledge the support provided to me from the numerous postgraduate students, postdoctoral and support staff that I have had the privilege and honour to work alongside of and socialise with.

Futhermore, I would like to thank the following people for their technical guidance for the experimental work completed in this project. Paul Held and Sharon Guffogg for assistance with the Biotek fluorescent plate reader. Dr Lisa Sedger for assistance with flow cytometery and Dr Mike Johnson for confocal microscopy analysis. Narelle Woodland, Gilian Rozenberg and the Prince of Wales Hospital for assistance with staining of THP-1 cells and positive controls. Additionally, a big thankyou to David Hyatt and Lonza for the loan of the Nucleofector instrument.

I would like to also acknowledge Jenefer M. Blackwell, Anna Dubaniewicz, A Graham, Leonardo A Sechi, Lee E Sieswerda, Maria Gazouli, Margje Haverkamp, Linda Wicker, Jennie Yang, Eileen Hoal, Timothy Sterling and Alison Motsinger-Reif for supplying additional population data for the completion of meta-analyses.

Finally, I would like to thank the support of my family, my partner, Sarah, my parents, Lynette and Warwick and sister and brother-in law, Kerri and Daniel. While you never did understand exactly what I was doing, you still tried to show an interest!

ABSTRACT

Solute Carrier Family 11A Member 1 (SLC11A1) is a member of a highly conserved group of ion transporters and has restricted localisation to the phagosomal membrane of monocytes/macrophages. SLC11A1 plays an immunomodulatory role in influencing macrophage activation status and the T helper 1/T helper 2 bias. As such it modulates susceptibility to infectious/autoimmune diseases. A polymorphic $(GT)_n$ promoter microsatellite repeat is known to alter *SLC11A1* promoter activity. Of the nine $(GT)_n$ alleles identified, alleles 3 and 2, which account for a combined allele frequency of greater than 95%, drive high and low *SLC11A1* expression, respectively. The increased *SLC11A1* expression, driven by $(GT)_n$ allele 3 is hypothesised to result in a heightened activation status of classically activated macrophages, affording resistance to infectious disease, but conferring susceptibility to pro-inflammatory autoimmune diseases. Conversely, decreased *SLC11A1* expression in the presence of allele 2 would confer susceptibility to infectious disease, but resistance to autoimmune disease.

A large number of studies assessing the association between the presence of specific $(GT)_n$ promoter alleles with the incidence of infectious and autoimmune disease have produced inconsistent associations. Meta-analyses are powerful analytical tools which combine individual association studies to estimate the strength of an association, therefore, meta-analyses of case control association studies (from 1991-2006) analysing the association of *SLC11A1* promoter $(GT)_n$ alleles 2 and 3 with the incidence of autoimmune disease were performed. The meta-analyses found a weak predominance of disease in the absence of allele 2, with a fixed effects pooled OR of 0.80 (95% CI = 0.22), however, a random effects pooled odds ratio (OR) of 0.88 (95% CI = 0.66) for allele 3 suggested no association with the incidence of autoimmune disease.

The publication of additional case control studies between 2006 and the present allowed a more comprehensive meta-analysis to be completed. This analysis, which included additional *SLC11A1* polymorphisms, represents the largest study assessing the association of *SLC11A1* polymorphisms with disease occurrence to date. Allele 2 of the $(GT)_n$ microsatellite was associated with increased and reduced incidence of infectious [OR=1.32 (1.20-1.46)] and autoimmune diseases [OR=0.90 (0.81-1.00)], respectively. Allele 3 was significantly associated with reduced incidence of infectious disease

[OR=0.82 (0.76-0.88)], however, the association with susceptibility to autoimmune disease occurrence did not reach statistical significance [OR=1.11 (0.98-1.26)]. The findings of the meta-analysis challenges the hypothesis that allele 3 is the disease causing variant at the $(GT)_n$ microsatellite repeat.

The results of these meta-analyses highlight small sample sizes as a major limitation of case control association studies. Completion of large-scale studies has been impractical because conventional *SLC11A1* $(GT)_n$ genotyping methodologies are time consuming and cannot differentiate all $(GT)_n$ variants. A high resolution melt curve methodology has been designed and optimised to genotype two *SLC11A1* polymorphisms, the $(GT)_n$ and $(CAAA)_n$ microsatellite repeats. Assay validation yielded a 100% success rate for genotyping of the $(GT)_n$ and $(CAAA)_n$ microsatellites. The designed methodology is the first to enable accurate, sensitive and high-throughput genotyping of these microsatellites and will enable the completion of sufficiently large association studies required to determine the association between the *SLC11A1* $(GT)_n$ and $(CAAA)_n$ polymorphisms and disease occurrence.

In addition to the $(GT)_n$ microsatellite, the -237C/T polymorphism has also been shown to modulate *SLC11A1* expression, with the T variant driving low expression in the presence of $(GT)_n$ allele 3. Little is known about *SLC11A1* transcription or the mechanism by which the $(GT)_n$ and -237C/T promoter polymorphisms modulate *SLC11A1* expression. Bioinformatic studies were completed to identify putative regulatory elements involved in transcription and promoter constructs, containing different lengths of the *SLC11A1* promoter, were prepared and used to assess promoter function. A 581bp promoter region (-532 to +49) that controlled *SLC11A1* expression in monocytes was identified. Within this region was identified a 148bp minimal promoter region (-99 to +49) containing the core elements for the formation of the basal transcriptional complex. The greatest transcriptional enhancement was identified within a 170bp region (-532 to -362) containing a novel IRF-Ets composite sequence for the recruitment of transcription factors IRF-8 and PU.1. Additionally, the promoter constructs suggested that the *SLC11A1* promoter may mediate bidirectional transcription. It was further determined that, in monocytic cells, the ability of $(GT)_n$ alleles 2 and 3 to differentially modulate *SLC11A1* expression was not due to their differing abilities to form Z-DNA, but to monocyte-specific factor(s) binding to a 165bp

region (-362 to -197) of the *SLC11A1* promoter. Additional bioinformatic and functional assays suggested that the T variant of the -237C/T polymorphism reduced *SLC11A1* promoter activity independently of the $(GT)_n$ microsatellite repeat.

Infectious and autoimmune diseases are major contributors to morbidity and mortality. *SLC11A1* is instrumental in regulating macrophage function and hence susceptibility to infectious and autoimmune diseases. This study has provided insight into the association of *SLC11A1* with disease incidence, has developed a novel genotyping methodology to allow the completion of large association studies and has elucidated mechanisms of transcriptional regulation of *SLC11A1* and the influence of polymorphisms on *SLC11A1* expression.

PUBLICATIONS ARISING FROM THE WORK DESCRIBED IN THIS THESIS

(A) PUBLICATIONS IN PEER-REVIEWED JOURNALS

Nicholas S. Archer, Najah Nassif & Bronwyn A. O'Brien (2012) "The *SLC11A1* $(GT)_n$ promoter polymorphism modulates expression through monocyte specific factor(s) to alter susceptibility to infectious and autoimmune diseases", (Manuscript in preparation).

Nicholas S. Archer, Najah Nassif & Bronwyn A. O'Brien (2012) "Meta-analysis of *SLC11A1* polymorphisms: $(GT)_n$ allele 2 exerts selective pressure in infectious and autoimmune disease", (Manuscript in preparation).

Nicholas S. Archer, Melinda Sirmias, Stephanie Dowdell, Najah Nassif & Bronwyn A. O'Brien (2012) "Genotyping disease-associated *SLC11A1* microsatellite repeats by high resolution melt analysis", (Submitted).

Nicholas S. Archer, Najah Nassif & Bronwyn A. O'Brien (2010) "Discrimination of microsatellite repeat polymorphisms of the *SLC11A1* promoter by melting curve analysis using the Eppendorf Mastercycler ep *realplex*", Eppendorf Technical Application Note 206.

Bronwyn O'Brien, **Nicholas S. Archer**, Fraser Torpy & Najah Nassif (2008) "Association of *SLC11A1* Promoter Polymorphisms with the incidence of autoimmune and Inflammatory Diseases: A Meta-Analysis", *Journal of Autoimmunity*, 31(1): 42-51.

(B) CONFERENCE ABSTRACTS

Nicholas Archer, Najah Nassif & Bronwyn O'Brien (2011) Poster entitled: "Macrophage specific factors differentially regulate allele specific *SLC11A1* expression and consequent susceptibility to infectious and autoimmune disease", 32nd Lorne Genome Conference.

Nicholas Archer, Najah Nassif & Bronwyn O'Brien (2010) Presentation entitled: "The *SLC11A1* $(GT)_n$ promoter polymorphism modulates susceptibility to infectious and autoimmune disease", 27th Combined RNSH/UTS/USyd/KIMR Scientific Research Meeting – Winner of the John Hambly award for best UTS presentation.

Nicholas Archer, Najah Nassif & Bronwyn O'Brien (2010) Poster entitled: "Elucidation of the essential promoter region of *SLC11A1*", 31st Lorne Genome Conference.

Nicholas Archer, Melinda Sirmias, Stephanie Dowdell, Najah Nassif & Bronwyn O'Brien (2009) Poster entitled: "Genotyping of functional SLC11A1 polymorphisms associated with infectious and autoimmune diseases", 26th Combined RNSH/UTS/USyd/KIMR Scientific Research Meeting.

Nicholas Archer, Najah Nassif & Bronwyn O'Brien (2008) Presentation in the Young Investigator Category "Elucidation of the Essential Promoter Region of *SLC11A1*", 25th Combined RNSH/UTS/USyd/KIMR Scientific Research Meeting.

Nicholas Archer, Najah Nassif & Bronwyn O'Brien (2007) Presentation entitled: "Genotyping SLC11A1 promoter polymorphisms by high resolution melt (HRM) analysis", 24th Combined RNSH/UTS/USyd/KIMR Scientific Research Meeting.

(C) AWARDS

Awarded the John Hambly Award for the best UTS presentation at the Combined RNSH/UTS/USyd/KIMR Scientific Research Meeting 2010.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

γ-IRE	interferon-γ response element
ALL	acute lymphocytic leukaemia
AML	acute myeloid leukaemia
AMML	acute myelomonocytic leukaemia
AP1	Activator protein 1
ARNT	aryl hydrocarbon receptor nuclear translocator
bp	base pairs
BRE	TFIIB-recognition element
BSA	bovine serum albumin
CCF2-AM	coumarin cephalosporin fluorescein
C/EBP	CCAAT/enhancer binding protein
CI	confidence interval
Ct	cycle threshold
DCE	downstream core element
DMEM	Dulbecco's modified eagle medium
DNA	deoxyribonucleic acid
DPE	downstream promoter element
EDTA	ethylenediaminetetraacetic acid
EMSA	electrophoretic mobility shift assays
FBS	fetal bovine serum
GM-CSF	granulocyte macrophage colony-stimulating factor
h	hours
HBSS	Hanks buffered salt solution
HIF-1	Hypoxia inducible factor 1
HIV	Human immunodeficiency virus
Idd	insulin dependant diabetes (murine)
IDDM	insulin dependant diabetes mellitus (human)
IECS	IRF-Ets composite sequence
IFN-γ	interferon-gamma
IL	interleukin
iNOS	inducible nitric oxide synthase
Inr	Initiator element
IRF	interferon regulatory factors
ISRE	IFN-stimulated response element
kb	kilobase
KLF	kruppel-like factor
l	litre
Lamp1	lysosome-associated membrane protein 1
LB	Luria Bertani
LD	linkage disequilibrium
LPS	lipopolysaccharide
MHC	major histocompatibility complex
min	minutes
MTE	motif ten element
NF-IL6	nuclear factor IL-6
NF-κB	nuclear factor kappa-light-chain-enhancer of activated B cells
NO	nitric oxide

Nramp	natural resistance-associated macrophage protein
NTC	no template control
Oct-1	octamer binding protein 1
OR	odds ratio
PAS	periodic acid-schiff
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PMA	phorbol myristate acetate
PMN	polymorphonuclear
pol II	RNA polymerase II
PU.1	protein encoded by *SPI-1* gene
RES	reticuloendothelial system
RNA	ribonucleic acid
RNase	ribonuclease
RPMI	Roswell Park Memorial Institute
RT	room temperature
s	seconds
SBB	Sudan black B
SLC11A1	Solute carrier family 11A member 1 (Human protein)
SLC11A1	Solute carrier family 11A member 1 (Human gene)
Slc11a1	Solute carrier family 11A member 1 (non-human protein)
Slc11a1	Solute carrier family 11A member 1 (non-human gene)
SLC11A2	Solute carrier family 11A member 2 (Human protein)
SLC11A2	Solute carrier family 11A member 2 (Human gene)
SNP	single nucleotide polymorphism
Sp1	Specificity protein 1
SPI-1	spleen focus by forming virus proviral integration 1
TAF	TBP associated factor
TBP	TATA binding protein
TESS	Transcription Element Search Software
TFIID	transcription factor II D
TFBS	transcription factor binding site
Th1	T helper 1
Th2	T helper 2
T_m	melting temperature
TNF-α	tumour necrosis factor-alpha
TSS1	transcription start site 1
TSS2	transcription start site 2
UV	ultraviolet
XCPE1	X core promoter element 1
YY1	Ying-Yang 1
ZBP-1	Z-DNA binding protein 1