Mutational analysis of the solute carrier family 11 member 1 gene (*SLC11A1*) implicated in iron transport

MONIQUE G. ZAAHL

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Stellenbosch.

Supervisor: Dr M J Kotze

Co-supervisors: Prof L Warnich

Prof T A Winter

University of Stellenbosch

December 2003

Declaration

I, the undersigned, hereby declare that the work contained in this dissertation is my own original work and has not previously in its entirety or in part been submitted at any university for a degree.

Signature:

Date:

Summary

The solute carrier family 11 member 1 gene (*SLC11A1*) is a divalent metal ion transporter with various pleiotropic effects on macrophage function. This gene that regulates iron, and is also regulated by cellular iron levels, has previously been linked to many infectious and autoimmune diseases. In this analysis, *in vitro* studies using the luciferase reporter system as well as case-control association studies were applied to investigate the significance of *SLC11A1* allelic variation in patients with diverse disease phenotypes.

For *in vitro* studies, five different *SLC11A1* promoter constructs were generated, followed by transfection into U937 and THP-1 cells. The inserted fragments included two previously described alleles (alleles 2 and 3), two novel alleles identified in this study (alleles 8 and 9) and a C to T point mutation at nucleotide position –237 in the presence of allele 3. The most striking finding was the opposite effect observed for allele 3 in the presence of the -237C→T polymorphism, similar to that of allele 2. Although the *SLC11A1* gene has previously been implicated in iron transport, we have demonstrated, for the first time, that the various alleles investigated cause differential expression of the gene upon iron loading.

Association studies were performed by investigating diseases including oesophageal cancer (OC), inflammatory bowel disease (IBD) and hereditary haemochromatosis (HH) (or primary iron overload). Significant associations ($P<0.05$) were observed with allele 3 for all three conditions investigated only after stratification according to the presence of the –237C→T polymorphism. Re-assessment of the promoter alleles according to expression profiles determined by the *in vitro* studies, showed statistically significant associations for allele 3 with OC and primary iron overload, compared with the respective population-matched control groups. Additionally, several novel variants were identified in exon 2 (112G→A, 148delGACCAGCCC, 157insGACCAGCCCAG) and intron 1 (IVS1-28C→T), with variant IVS1-28C→T occurring at a significantly increased frequency in patients with OC compared with population-matched controls ($P<0.05$). Investigation of the *SLC11A1* gene in individuals presenting with iron overload in the absence of homozygosity for the *HFE* C282Y mutation, provided further support for the importance of sequence variation in

the promoter region of the *SLC11A1* gene in modified risk of iron-related disorders. Genes related to iron homeostasis, including *HFE*, *SLC11A3*, *HAMP* and *DCYTB*, were investigated in individuals with similar criteria and potential disease-causing mutations were identified in 11% White and 45% Black South African patients. The possible significance of the *SLC11A3* and *DCYTB* genes in iron overload in the Black South African population, and the possible involvement of the *DCYTB* gene in iron overload in general, are demonstrated for the first time.

This study contributed to a better understanding of the function of the *SLC11A1* gene in relation to iron metabolism. The involvement of *SLC11A1* in a range of disease phenotypes including cancer and inflammatory conditions that may involve iron dysregulation, can probably be explained by interaction with external factors such as infectious agents that may affect cellular iron status. Our findings provide both *in vivo* and *in vitro* evidence that iron dysregulation mediated by allelic effects of *SLC11A1* may underlie disease susceptibility to infectious and autoimmune conditions.

Opsomming

Die opgeloste stof draer familie 11 deel 1 geen (*SLC11A1*) is 'n divalente metaal ioon vervoerder met verskeie pleiotropiese effekte op makrofaagfunksie. Die geen, wat yster reguleer en ook deur sellulêre ystervlakke gereguleer word, is voorheen verbind met verskeie infektiewe en outo-immune siektes. In hierdie studie is *in vitro* analises, deur middel van die lusiferase verklikker sisteem, asook gevalle-kontrole assosiasie studies gebruik om die rol van *SLC11A1* alleel variasie in pasiënte met diverse siektefenotipes te ondersoek.

Vyf verskillende *SLC11A1* promotor variante is geskep vir *in vitro* studies en gevolg deur transfeksie in U937 en THP-1 sellyne. Die ingevoegde fragmente het twee voorheen beskryfde allele (allele 2 en 3), twee nuwe allele wat in hierdie studie geïdentifiseer is (allele 8 en 9) en 'n C na T puntmutasie by nukleotied posisie –237 in die teenwoordigheid van alleel 3 ingesluit. Die opvallendste bevinding was die teenoorgestelde effek wat waargeneem is wanneer alleel 3 in die teenwoordigheid van die -237C→T polimorfisme voorkom, soortgelyk aan alleel 2 uitdrukking. Alhoewel die *SLC11A1* geen voorheen geïmpliseer is in yster vervoer, is daar vir die eerste keer aangetoon dat na yster lading, die verskillende allele differensiële uitdrukking van die geen veroorsaak.

Verskeie siektes, insluitend slukderm kanker (OC), inflammatoriese dermsiekte (IBD) en oorerflike hemochromatose (HH) (of primêre ysteroorlading), is ondersoek deur middel van assosiasie studies. Betekenisvolle verskille (*P*<0.05) is waargeneem vir alleel 3 tussen die kontrole- en pasiëntgroepe in al drie siektes wat ondersoek is, maar slegs na stratifikasie volgens die teenwoordigheid van die –237C→T polimorfisme. Na hersiening van die promotor allele volgens ekspressie profiele verkry met *in vitro* studies is statisties betekenisvolle assosiasie ook verkry vir alleel 3 met OC en primêre ysteroorlading in vergelyking met die onderskeie populasie kontrolegroepe. Verder is verskeie nuwe variante ook geïdentifiseer in ekson 2 (112G→A, 148delGACCAGCCC, 157insGACCAGCCCAG) en intron 1 (IVS1-28C→T) en 'n statisties betekenisvolle verhoogde frekwensie van variant IVS1-28C→T is waargeneem in pasiënte met OC in vergelyking met die populasie kontrolegroep (*P*<0.05). Die belangrikheid van variasie in die promotor area van die

iv

SLC11A1 geen as 'n modifiserende faktor in ysterverwante siektes, is verder ondersteun deur die *SLC11A1* geen in individue met ysteroorlading in die afwesigheid van homosigositeit vir die *HFE* C282Y mutasie te ondersoek. Ander gene geassosieerd met yster homeostase, insluitend *HFE*, *SLC11A3*, *HAMP* and *DCYTB,* is ondersoek in individue met soortgelyke seleksie kriteria en potensiële siekte-verwante mutasies is geïdentifiseer in 11% Wit en 45% Swart Suid-Afrikaanse pasiënte. Die moontlike belang van die *SLC11A3* en *DCYTB* gene in ysteroorlading in die Swart Suid-Afrikaanse populasie en die moontlike betrokkenheid van die *DCYTB* geen in yster oorlading oor die algemeen, is vir die eerste keer aangetoon.

Hierdie studie dra by tot 'n beter insig in die funksie van die *SLC11A1* geen ten opsigte van ystermetabolisme. Die betrokkenheid van *SLC11A1* in 'n reeks siekte fenotipes, wat insluit kanker en inflammatoriese toestande wat verband kan hou met 'n yster wanbalans, kan moontlik verklaar word deur interaksie met eksterne faktore soos infektiewe agente wat die sellulêre yster status kan beïnvloed. Ons bevindinge verskaf beide *in vivo* en *in vitro* getuienis dat yster wanbalans, wat bemiddel word deur alleliese effekte van *SLC11A1*, verantwoordelik mag wees vir vatbaarheid vir infektiewe en outoimmune siekte toestande.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION

CHAPTER 2. DETAILED EXPERIMENTAL PROCEDURES

CHAPTER 3. METHODS, RESULTS AND DISCUSSION

LIST OF TABLES

CHAPTER 3. METHODS, RESULTS AND DISCUSSION

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

CHAPTER 2. DETAILED EXPERIMENTAL PROCEDURES

CHAPTER 3. METHODS, RESULTS AND DISCUSSION

LIST OF ABBREVIATIONS

%	percentage
°C	degrees Celcius
5'	5-prime
3'	3-prime
χ^2	chi-square
5'-(GT)n	*SLC11A1* promoter GT repeat
β-ME	beta-mercaptoethanol
µg/ml	microgram per millilitre
µl	microlitre
µl/ml	microlitre per millilitre
x g	times gravity
A	adenosine
A (Ala)	alanine
AA	acrylamide
ADC	adenomacarcinoma
Ann	annealing temperature
AP-1	activator protein 1
Apaf-1	apoptotic protease activating factor 1
APC	adenomatous polyposis of the colon gene
APS	ammonium persulphate
ATCC	American Type Culture Collection
BAA	bisacrylamide
BanI	*Bacillus aneurinolyticus*, 1st enzyme
Bcg	*Mycobacterium bovis*
BglII	*Bacillus globigii*, 2nd enzyme
bp	base-pair
BSA	bovine serum albumin
C	cytidine
C	crosslinking
C (Cys)	cysteine
CARDs	caspase recruitment domains
CARD15	caspase recruitment domain 15 gene

CCND1	G1/S-specific cyclin D1
CD	Crohn's disease
CDAI	Crohn's disease activity index
Cd^{2+}	cadmium
cDNA	complementary deoxyribonucleic acid
CGH	comparative genomic hybridisation
CH_3COOK	potassium acetate
CH_3COONa	sodium acetate
cm	centimeter
C-MYC	myc protein
CO_2	carbon dioxide
CO^{2+}	cobalt
Cu^{2+}	copper cupric
D (Asp)	aspartic acid
dATP	2'-deoxy-adenosine-5'-triphosphate
DCC	deleted in colorectal carcinoma gene
dCTP	2'-deoxy-cytidine-5'-triphosphate
DCYTB	duodenal cytochrome B gene
ddH_2O	double distilled water
del	deletion
df	degrees of freedom
dGTP	2'-deoxy-guanosine-5'-triphosphate
dH_2O	distilled water
dHPLC	denaturing high performance liquid chromatography
DMT1	divalent metal transporter 1 gene
DNA	deoxyribonucleic acid
dNTP	2'-deoxy-nucleotide-5'-triphosphate
DTE	dithioerythritol
DTT	dithiothreitol
dTTP	2'-deoxy-thymidine-5'-triphosphate
E	extracellular loop
EDTA	ethylenediaminetetraacetic acid
EtOH	ethanol
F	forward primer

FAC	ferric ammonium citrate
Fe	iron
Fe^{2+}	ferrous iron
Fe^{3+}	ferric iron
FPN1	ferroportin-1 gene
Fre1	ferric reductase transmembrane component 1
fs	frameshift
Ft	ferritin
g	gram
G	guanosine
G (Gly)	glycine
GMCSF	granulocyte-macrophage colony-stimulating factor
H^+	hydrogen
H (His)	histidine
HAMP	hepcidin antimicrobial peptide gene
HEPC	hepcidin gene
HEX-SSCP	heteroduplex single-strand conformation polymorphism analysis
HFE	haemochromatosis gene
HFE1	haemochromatosis type 1
HFE2	haemochromatosis type 2
HFE3	haemochromatosis type 3
HFE4	haemochromatosis type 4
HH	hereditary haemochromatosis
HhaI	*Haemophilus haemolyticus*, 1st enzyme
HinfI	*Haemophilus influenzae Rf*, 1st enzyme
HLA-A2	major histocompatibility complex class I A2
HLA-G	major histocompatibility complex class I G
I	intracellular loop
IBD	inflammatory bowel disease
IFN-γ	interferon-gamma
IL-1β	interleukin-1 beta
iNOS	inducible nitric oxide synthase

INT-2	INT-2 proto-oncogene protein
IPTG	isopropyl-β-D-thio-galactopyranoside
IRE(s)	iron response element(s)
IREG1	iron-regulated transporter 1 gene
Ity	*Salmonella typhimurium*
IVS	intervening sequence
JH	juvenile haemochromatosis
kb	kilobase
KCl	potassium chloride
kD	kilo Dalton
$KHCO_3$	potassium hydrogen carbonate
KH_2PO_4	potassium phosphate dibasic
LB	Luria-Bertani medium
LEAP	liver–expressed antimicrobial peptide gene
LOH	loss of heterozygosity
LPS	bacterial lypopolysaccharides
LRR	leucine-rich repeat
Lsh	*Leishmania donovani*
M	moles per litre
MIIC	major histocompatibility complex class II
MCC	mutated in colorectal cancers gene
mg	milligram
$MgCl_2$	magnesium chloride
mg/ml	milligram per millilitre
Mg_2SO_4	magnesium sulphate
MHC	major histocompatibility complex
ml	millilitre
ml/min	millilitre per minute
mm	millimetre
mM	milli-moles per litre
Mn	manganese
MOPS	3-morpholinopropanesulfonic acid
MTP1	metal transporter 1 gene
N (Asn)	asparagine

NaCl	sodium chloride
NaClO$_4$	sodium perchlorate
Na$_2$CO$_3$	sodium carbonate
Na$_2$HPO$_4$	di-sodium hydrogen phosphate
NaOH	sodium hydroxide
NBD	nucleotide-binding domain
NF-κB	nuclear factor κB
ng	nanogram
ng/µl	nanogram per microlitre
NH$_4$Cl	ammonium chloride
NheI	*Neisseria mucosa heidelbergensis*, 1st enzyme
Ni^{2+}	nickel
nm	nanometre
NO	nitric oxide
NOD2	NOD2 protein gene
Nos2A	nitric oxide synthase 2A
NRAMP1	natural resistance-associated macrophage protein 1 gene
NRAMP2	natural resistance-associated macrophage protein 2 gene
OC	oesophageal cancer
OH	hydroxyl
ONPG	o-nitrophenyl-β-δ galactopyranoside
p	short arm of chromosome
PAA	polyacrylamide
Pb^{2+}	lead plumbous
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PKC	protein kinase C
pmol	picomole
pmol/µl	picomole per microlitre
Pwo	*Pyrococcus woesei*
q	long arm of chromosome
R	reverse primer
R (Arg)	arginine
RACE	rapid amplification of cDNA ends

RB1	retinoblastoma gene
RFLP	restriction fragment length polymorphism
RLU	relative light units
RNA	ribonucleic acid
RsaI	*Rhodopseudomonas sphaeroides*, 1st enzyme
RT-PCR	reverse transcriptase polymerase chain reaction
SAP	shrimp alkaline phosphatase
Sau96I	*Staphylococcus aureus PS96*, 1st enzyme
SCC	squamous cell carcinoma
SDS	sodium dodecyl sulphate
SH3	src homology 3 domain
SLC11A1	solute carrier family 11 member 1 gene
SLC11A2	solute carrier family 11 member 2 gene
SLC11A3	solute carrier family 11 member 3 gene
SMF1	transporter protein SMF1
SMF2	transporter protein SMF2
SSCP	single-strand conformation polymorphism
T	thymidine
Taq	*Thermus aquaticus*
TBE	tris-borate/EDTA
TE	tris-EDTA
TEAA	triethylamine acetate
TEMED	N, N, N' N',-tetramethylethylenediamine
Temp	temperature
Tf	transferrin
TFR1	transferrin receptor 1 gene
TFR2	transferrin receptor 2 gene
THP-1	ATCC monocyte cell line
Tm	melting temperature
TM	transmembrane domain
TNF-α	tumor necrosis factor-alpha
TP53	tumor protein p53
Tris-HCl	tris hydrochloride [2-Amino-2-(hydroxymethyl)-1,3-propanediol-hydrochloride]

TTR	transthyretin
U	units
U/µl	units per microlitre
U937	ATCC macrophage cell line
UC	ulcerative colitis
Usf2	upstream stimulatory factor 2 gene
UTR	untranslated region
V	volts
V (Val)	valine
v/v	volume per volume
W (Trp)	tryptophan
w/v	weight per volume
X-gal	5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside
Y (Tyr)	tyrosine
YAC	yeast artificial chromosome
Zn^{2+}	zinc

CHAPTER 1

INTRODUCTION

1. THE SOLUTE CARRIER FAMILY 11 (PROTON-COUPLED DIVALENT METAL ION TRANSPORTER) MEMBER 1 GENE (*SLC11A1*) AND DISEASE

Preface

The *SLC11A1* gene has previously been known as the natural resistance-associated macrophage protein 1 gene (*NRAMP1*) and is associated with susceptibility to various autoimmune and infectious diseases (reviewed by Blackwell *et al* 2001, 2003). The results of various studies support the hypothesis that specific alleles of the *SLC11A1* gene 5'-(GT)n repeat polymorphism contribute to autoimmune (allele 3) and infectious (allele 2) disease susceptibility. However, the direct contribution of the many pleiotropic effects of the *SLC11A1* gene on macrophage function and disease susceptibility is still unclear.

In this study we focussed on mutational analysis of the *SLC11A1* gene implicated in iron transport within the context of other genes that may be involved in iron homeostasis and/or disease phenotype. We hypothesised that the involvement of *SLC11A1* in a range of disease phenotypes may be explained by its involvement in iron metabolism. The potential role of *SLC11A1* (as a modifier locus) was investigated in patients with different disease types, including oesophageal cancer (OC), inflammatory bowel disease (IBD) and hereditary haemochromatosis (HH). These diseases and other genes implicated in their pathogenesis are discussed as an introduction to the study.

1.1. LITERATURE REVIEW

1.1.1. Identification of the *SLC11A1* gene

In mice, natural resistance to infection by unrelated intracellular parasites is controlled by the expression of a single dominant gene on mouse chromosome 1. This gene was designated *Leishmania donovani* (*Lsh*) (Bradley 1974), *Salmonella typhimurium* (*Ity*) (Plant and Glynn 1974) and *Mycobacterium bovis* (*Bcg*) (Gros *et al* 1981) and affects the capacity of macrophages to destroy ingested intracellular parasites early during infection (Skamene *et al* 1982).

The murine solute carrier family 11 member 1 gene (*Slc11a1*) was identified as the *Bcg/Ity/Lsh* gene by positional cloning (Vidal *et al* 1993). The identification of the *Slc11a1* gene as *Bcg/Ity/Lsh* was confirmed by using transfected macrophages *in vitro* (Barton *et al* 1995) and gene disruption *in vivo* (Vidal *et al* 1995). The full-length cDNA sequence of the gene was described by Barton *et al* (1994). *Slc11a1* encodes a novel macrophage-specific polypeptide, predicted to act as an integral membrane protein that plays a role early in the macrophage activation pathway (Vidal *et al* 1993, Blackwell *et al* 2000). Malo *et al* (1994) analysed the *Slc11a1* gene in inbred mice with either resistance or susceptibility phenotypes and identified a nonconservative amino acid substitution [glycine (resistant/dominant) to aspartic acid (susceptible/recessive) at amino acid residue 169], within the fourth transmembrane domain of the protein, to be associated with a susceptibility phenotype. This mutation has been predicted to affect either proper targeting and folding of the protein in the membrane or to directly interfere with the enzymatic activity of *Slc11a1*. These mice are as susceptible to *Leishmania donovani*,

3

Salmonella typhimurium and *Mycobacterium bovis* as gene-disrupted mice (Vidal *et al* 1995).

Susceptibility to a range of other pathogens in mice has also been identified, including *Mycobacterium lepraemurium* (Brown *et al* 1982, Skamene *et al* 1984), *Mycobacterium intracellulare* (Goto *et al* 1989), *Toxoplasma gondii* (Blackwell *et al* 1994), *Candida albicans* (Puliti *et al* 1995) and *Leishmania infantum* (Leclercq *et al* 1996). The human *SLC11A1* gene was subsequently cloned and characterised by using cDNA clones (Cellier *et al* 1994, Kishi 1994) and localised to chromosome region 2q35 by linkage analysis (Blackwell *et al* 1995) as well as PCR analysis of somatic cell hybrids and YAC cloning (Liu *et al* 1995).

1.1.2. Structure and protein of the *SLC11A1* gene

The *SLC11A1* gene comprises 15 exons (Table 1.1), and an alternatively spliced exon encoded by an Alu element present within intron 4, spanning a region of 12 kb (Cellier *et al* 1994, Blackwell *et al* 1995). The human *SLC11A1* gene demonstrated 92% similarity (85% identity) to the mouse *Slc11a1* gene, 73% (55% identity) to *Drosophila melanogaster*, 58% (40% identity) to *Oryza sativa* (rice) and 46% (28% identity) to *Saccharomyces cereviviae* homologues (Cellier *et al* 1995). The nucleotide sequence of the 5' flanking sequence of the *SLC11A1* gene has been characterised up to position – 415 bp by Blackwell *et al* (1995), up to position –3456 bp by Kishi *et al* (1996) and up to position –948 by Searle and Blackwell (1999). The promoter region of *SLC11A1* lacks any conventional TATA, GC and CCAAT boxes (Blackwell *et al* 1995).

4

Table 1.1. Exon/Intron boundaries of the *SLC11A1* gene are tabulated with the sizes of the various exons and introns (Blackwell *et al* 1995, Marquet *et al* 2000). The percentage amino acid (AA) identity with mouse *Slc11a1* is also indicated.

Exon	5' end of exon	Exon length (bp)	3' end of exon	Intron length (bp)	%AA identity (Mouse *Slc11a1*)
1	ATGACAG	182	TGACAGgtga	584	50
2	acagGTGACA	143	AAACCGgtgg	1140	68
3	acagGGCACC	123	TTCAAAgtaa	780	95
4	acagCTTCTC	120	CCTAAGgtga	1368	100
*4a	tcagAATCCA	74	CCCAGgtac	1149	U
5	atagGTGCCC	107	TGGACGgtacc	418	91
6	tcagAATCCC	71	ACTACGgtggg	334	100
7	gtagGGCTGC	68	TATGAGgtag	167	100
8	gcagTATGTG	156	GTCAAGgtga	1915	88
9	gtagTCTCGA	159	GCTGCGgtga	1171	87
10	acagTTCAAC	90	CAGGGGgtga	88	80
11	gcagGGCGTG	120	ATGGAGgtag	1485	98
12	ccagGGCTTC	150	CTGCTGgtga	987	94
13	ccagCTCCCG	74	GGCCTgtga	438	84
14	ccagGCTGAA	154	TACCTGgtac	139	73
15	ccagGTCTGG	108	TAG		67

Exon sequences are in uppercase letters and intron sequences are in lowercase letters. *Alternatively spliced exon, U, unknown

Several regulatory elements have been identified, including a PU box [the binding site for the transcription factor PU.1 (Klemsz *et al* 1990)], LPS (NF-IL6, AP-1, GMCSF) and IFN-γ (Isshiki *et al* 1990, Yang *et al* 1990) sequence motifs. A Z-DNA forming (GT)n repeat was also identified in the promoter region (Liu *et al* 1995, Blackwell *et al* 1995). In eukaryotes, a distinct class of binding proteins exist which interact exclusively with DNA in the Z-conformation, and roles in both positive and negative regulatory signalling have been attributed to this form of DNA (Rich *et al* 1984). The consensus sequences identified in the promoter region are shown in Figure 1.1. Several discrepancies were observed between the published promoter sequences of Kishi *et al* (1996) and Searle and Blackwell (1999) and have been highlighted on the sequence insert (Figure 1.1).

Sequence analysis of the human SLC11A1 protein, with a calculated mass of 60 kD, indicated that the polypeptide consists of 550-amino acids with 10 to 12 putative transmembrane (TM) domains. Additionally, 2 protein kinase C (PKC) phosphorylation sites within the first intracytoplasmic loop, 2 N-linked glycosylation sites within the segment separating TM domains 5 and 6, and an evolutionary conserved consensus transport motif within the segment separating TM domains 6 and 7 were also identified (Figure 1.2) (Cellier *et al* 1994, Blackwell *et al* 1995). The null mutation in murine *Slc11a1* occurs in TM domain 4 (Vidal *et al* 1993), corresponding to TM domain 2 (exon 6) of the human *SLC11A1* gene.

                                                                    1
-1083 AATTAGCCAGGTGTGGTTGGCAGGTGCCTGTAATCCCAGCTACTCAGGAGGCTGAGGAAGAAGAATTACTTGAA

-1009 CCCAAGAGGCAGAGGTTTCGGTGAGCCAAGGTCGCGCCATTGCACTTCAGCCTGGGTGACAGAGCAAGACGCC

 -936 ATCTCAAAGTAAATAAATAAATACATACATACATACATACAAAAAGTAGCCGGGTGTGGTGGTGCATGCCTGTAGT
                                                                    1
 -860 CCCAGCTACTCAGGAGGCTGAGGCAGGAGAATTGCTTAAACATGGAGGACAGAGGTTGCAGTGAGCTGAGATT

 -787 ATGCCACTGCACTCCAGTCTGGGCAACAGAGTAAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAGGA

 -712 GTTTGAGGCTGCAGTAAAACTAAGATGGCACCACTGCACTTCACACCTCACTCTAGCCTGGGTGACAGAGCGAGA

 -638 CCTTGTCTCTAAAAAAAAAAACAAAAACGAAAACCAAAAAAAAAAAAGGGTGGGGGTGGAATAGAATAAACAAC

 -562 TCTGAGAAGGGACATGATCTGGTGACAATCTCAAGTGAATCAGTGGTTGAACCAGGACCAGATCCCAGTGCCCC
                                                                    1
 -488 ATCTTCTGGGGACGTCCTGGGTCTGAATCTTCAATGCATGTCCCTTCTGCAGTGCCTTCCTCTGTGGCCCTCAAA

 -413 GGGAAACTGAAGCCTTTGAGGACATGAAGACTCGCATTAGGCCAACGAGGGGTCTTGGAACTCCAGATCAAAGA

 -339 GAATAAGAAAGACCTGACTCTGTGTGTGTGTACGTGTGTGTGTACGTGTGTGTGTGTGTGTGTGGCAGAGGGGG

 -265 GTGTGGTCATGGGGTATTGACATGAATACGCAAGGGGCAGGAAGCATCTGAAATCAGAGCTAACTTGGGAGGC

 -192 ACAGAACACGGGGTGCCTGGAAGGGGAACAGATGTGTTGTGGGGCACAGGGCAGGCTGGGAGGGGAACAAAG

                                                                +1 Transcription start site
 -120 GTCCACTCCATGGGTAACCAGACCCTTCCGCCAGGGCTGGCCACTTCTGCCTTTGGAAAATGTTTCACAACGCC

  -46 CCATGTTGTGTGTGTGTGTGAATCGGCCGATGTGAACCGAATGTTGATGTAAGAGG

Figure 1.1.

▬	**PU.1**
▬	**Interferon-γ (γ-IRE)**
▬	**NFκB**
▬	**AP1**
▬	**Z-DNA forming repeat polymorphism**
▬	**NF-IL6**
ATTAG/AATG	**GMCSF**

Legend to Figure 1.1.

Sequence of the promoter region of the *SLC11A1* gene with 1083 bp of sequence 5' of the transcription start site (adapted from Kishi *et al* 1996 and Searle and Blackwell 1999). Sequence up to position –937 and response elements are as published by Searle and Blackwell 1999. [1]Denoted only by Kishi *et al* 1996. Discrepancies observed: Kishi *et al* 1996, transcription start site at position –27 (doubly underlined), putative TATA box (red), additional G residues at position –127 and –148, T instead of C at position –237.

Figure 1.2.
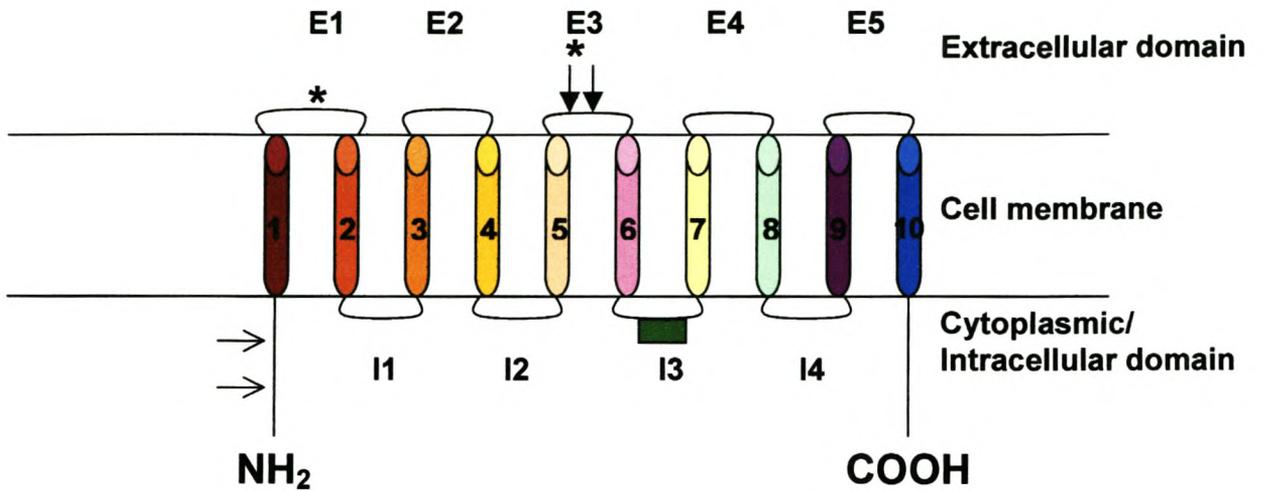
*   Cysteine residues conserved within 1st and 3rd extracellular loop domain

→   Predicted N-linked glycosylation sites

⇢   Consensus sites for protein kinase C (PKC) phosphorylation

▬   Consensus transport motif

▌   Transmembrane domains 1 to 10 (numbered 1 to 10)

E1 – E5   Extracellular loop 1 to 5

I1 – I4   Intracellular loop 1 to 4

Legend to Figure 1.2.
Schematic representation of the putative structure of the SLC11A1 protein and its orientation in the cell membrane (modified from Vidal *et al* 1995).

9

The protein is localised to late endosomal and lysosomal compartments (Atkinson *et al* 1997, Gruenheid *et al* 1997, Searle *et al* 1998) but not to early endosomes (Gruenheid *et al* 1997). More than half of the polypeptide composition (53%) consist of highly hydrophobic amino acids such as leucine (15%), isoleucine (6%), valine (7%), alanine (10%), phenylalanine (6%) and the apolar residue glycine (9%) (Cellier *et al* 1994). Charged residues account for only 12% of the polypeptide composition with arginine/lysine and aspartic/glutamic acid each accounting for 6%. The structure of the *Slc11a1* protein was shown to relate to that of transmembrane ion transporters (Cellier *et al* 1995, Brown *et al* 1995, Barton *et al* 1995, Atkinson *et al* 1997, Brown *et al* 1997, Gruenheid *et al* 1997, Searle *et al* 1998).

1.1.3. Polymorphisms identified in the *SLC11A1* gene

No mutations were identified in the *SLC11A1* gene that were comparable to those identified in mice although several polymorphic sites have been described (White *et al* 1994, Blackwell *et al* 1995, Liu *et al* 1995, Buu *et al* 1995, Lewis *et al* 1996, Searle and Blackwell 1999, Graham *et al* 2000, Kojima *et al* 2001, Lee *et al* 2002). These polymorphisms included variants in i) the coding region of the gene, including one deletion (a 9 bp deletion in exon 2), three missense mutations (A318V, V350M and D543N) and four silent nucleotide substitutions (L39, G65, F66 and G249), ii) intronic variants (IVS4+14G→C, IVS5-18G→A, IVS7+22G→A), and iii) variants in the untranslated regions (UTR), including two variants in the 5'-UTR [5'-(GT)n repeat and -236C→T] and one variant in the 3' UTR (1729+55del4).

Seven alleles (Table 1.2) have previously been identified for the Z-DNA-forming (GT)n repeat polymorphism in the promoter region of the *SLC11A1* gene (Blackwell *et al* 1995, Liu *et al* 1995, Blackwell 1996, Shaw *et al* 1996, 1997, Searle and Blackwell 1999, Graham *et al* 2000, Kojima *et al* 2001) that may affect levels of *SLC11A1* gene expression (Blackwell 1996, Searle and Blackwell 1999). *In vitro* promoter studies of this functional repeat polymorphism suggested a direct role of specific alleles in autoimmune and infectious disease susceptibility (Searle and Blackwell 1999). Allele 3, the high promoter activity sequence, is associated with susceptibility to autoimmune diseases whereas allele 2, the low promoter activity sequence, is associated with infectious disease susceptibility (and conversely protects against autoimmune diseases).

To investigate the significance of the 5'-(GT)n repeat, functional studies were performed in the presence and absence of exogeneous stimuli including IFN-γ and LPS (Searle and Blackwell 1999). The 5'-(GT)n polymorphism regulated gene expression in the absence of an exogeneous stimulus and therefore acts as an enhancer element. The addition of IFN-γ caused up regulation and LPS caused differential expression for the various polymorphic alleles studied. Blackwell *et al* (2000) suggested that the allelic association with autoimmune and infectious disease susceptibility are maintained in human populations by some form of balancing selection.

Table 1.2. Previously-documented *SLC11A1* Z-DNA-forming 5'-(GT)n repeat polymorphic variants

Allele	Repeat polymorphism	*Allele frequency	Reference
Allele 1	t(gt)$_5$ac(gt)$_5$ac(gt)$_{11}$ggcaga(g)$_6$	0.001	Blackwell *et al* 1995
Allele 2	t(gt)$_5$ac(gt)$_5$ac(gt)$_{10}$ggcaga(g)$_6$	0.20 to 0.25	Blackwell *et al* 1995
Allele 3	t(gt)$_5$ac(gt)$_5$ac(gt)$_9$ggcaga(g)$_6$	0.75 to 0.80	Blackwell *et al* 1995
Allele 4	t(gt)$_5$ac(gt)$_9$ggcaga(g)$_6$ [$]	0.001	Searle and Blackwell 1999
Allele 5	t(gt)$_4$ac(gt)$_5$ac(gt)$_{10}$ggcaga(g)$_6$	0.02	Graham *et al* 2000
Allele 6	t(gt)$_5$ac(gt)$_5$ac(gt)$_4$at(gt)$_4$ggcaga(g)$_7$	Unknown[^]	Graham *et al* 2000
Allele 7	t(gt)$_5$ac(gt)$_5$at(gt)$_{11}$ggcaga(g)$_6$	0.045[#]	Kojima *et al* 2001

*Allele frequencies in the general population, [#]Identified only in the Japanese population, [^]Allele frequency not denoted, [$]Previously identified as t(gt)$_5$ac(gt)$_{10}$ggcaga(g)$_6$ (Blackwell *et al* 1995, Shaw *et al* 1996, 1997) and t(gt)$_5$ac(gt)$_5$ac(gt)$_4$ggcaga(g)$_6$ (Blackwell *et al* 1995).

A potentially functional polymorphism was also identified in exon 2 of the *SLC11A1* gene and comprises deletion of a nine-nucleotide repeat in the N-terminal proline/serine-rich putative SH3 binding domain (Barton *et al* 1994, White *et al* 1994, Blackwell *et al* 2000). The spacing of prolines were shown to be important in binding SH3 domains (Booker *et al* 1993) and this polymorphic area may influence the conformation and hence function of the protein. An additional 3-amino-acid insert in human exon 2 compared to mouse in this area was reported by White *et al* (1994) and this was shown to be the normal allele in humans. This deletion was observed only in a heterozygous state at a frequency of 0.02 in Brazilians and was shown to be segregating within Brazilian leprosy families, but not with the disease (White *et al* 1994). A frequency of less than 0.001 was observed by

Blackwell *et al* (2000) in the general population. *SLC11A1* is presumably an essential gene and deleterious coding region mutations will therefore be rare.

1.1.4. Expression of the *SLC11A1* gene

Expression of the *SLC11A1* gene was initially thought to be restricted to myeloid cell lineages (Vidal *et al* 1993, Cellier *et al* 1997). However, recent evidence indicated expression of the gene in vesicles of neurons (Evans *et al* 2001), cortical pyramidal neurons (layers III and V), a subset of striatal neurons, cerebellar Purkinje cell bodies, the anterior pituitary, pancreatic islets and adrenal medulla (Blackwell *et al* 2001).

1.1.5. The proposed function of the *SLC11A1* gene

The exact function of the *SLC11A1* gene is still to be elucidated but several hypotheses have been described.

1.1.5.1. Pleiotropic effect of *SLC11A1* on macrophage function

Murine *Slc11a1* has been shown to have many pleiotropic effects on macrophage activation pathways, including regulation of the CXC chemokine KC, interleukin-1β (IL-1β), inducible nitric oxide synthase (iNOS), major histocompatibility complex (MHC) class II molecules, tumour necrosis factor α (TNF-α), nitric oxide (NO) release, L-arginine flux, oxidative burst and tumoricidal as well as antimicrobial activity (Blackwell *et al* 1988, Blackwell 1989, Blackwell *et al* 1991, Schurr *et al* 1991, Searle and Blackwell 1999, Blackwell *et al* 2000).

13

Slc11a1 enhances expression of iNOS (encoded by *Nos2A*) and generates toxic NO in wildtype (*Slc11a1^{Gly169}*) versus mutant (*Slc11a1^{Asp169}*) macrophages (Roach *et al* 1991, Arias *et al* 1997). It was assumed that NO combines with OH$^{\bullet}$ to produce the more toxic peroxynitrite and is therefore crucial to *Slc11a1*-mediated resistance *in vivo*. Recent evidence suggest that iNOS is not important in early *Slc11a1*-mediated Kupffer cell regulation of infection but may be important later in granuloma-dependent control of the parasite (Blackwell *et al* 2003). Other pleiotropic effects mediated by *SLC11A1* entailed up-regulation of various functions including class II I-A antigen expression, protein kinase C and antigen presentation (Blackwell *et al* 1991). Individuals homozygous for allele 3 of the *SLC11A1* gene were shown to produce higher levels of TNF-α (Blackwell *et al* 2003). However, the exact role for TNF-α, IL-1β, chemokines or their receptors in *Slc11a1*-regulated autoimmune and infectious disease phenotypes are not yet known and needs further investigation.

1.5.1.2. *SLC11A1* functions as a metal ion transporter

Several lines of evidence support the hypothesis of *SLC11A1* being a metal ion transporter. Firstly, a consensus transport motif was identified in the third intracellular/cytoplasmic loop of the gene (exon 11) (Vidal *et al* 1993). This evolutionary conserved consensus motif, known as the "binding protein-dependent transport system inner membrane component signature", is conserved from prokaryotes to eukaryotes (Cellier *et al* 1994) and was previously identified in a series of bacterial periplasmic transport proteins (Kerppola and Ames 1992). The motif is predicted to be localised in cytoplasmic loops of periplasmic proteins, possibly mediating interaction of the

hydrophobic membrane anchors with peripheral ATP binding subunits to energise these transporters (Kerppola and Ames 1992).

Secondly, the polypeptide localises to intracellular vesicular membranes in resting cells, suggesting that *Slc11a1* could influence the micro-environment of the invading pathogen within the phagolysosome by either directing the efflux or influx of some essential substrate or toxin respectively (Atkinson and Barton 1998).

Thirdly, the *Slc11a* family is defined by a highly conserved hydrophobic core composed of ten TM domains, including several invariant charged residues in these domains (Cellier *et al* 1995). Helical periodicity of sequence conservation which predicts a helical bundle within the membrane, occurring with a conserved charged interior and a semi-conserved hydrophobic exterior, is typical of families of ion transporters and channels. *SMF1* and *SMF2*, the yeast homologues of *Slc11a1*, were shown to be involved in high and low affinity Mn^{2+} uptake respectively, inhibitable by Zn^{2+} (Supek *et al* 1996). Similarly, the *Slc11a2* gene (78% identity to *Slc11a1*) was identified as the gene controlling microcytic anaemia in mice (Gruenheid *et al* 1995), having broad specificity as a metal ion transporter (Gunshin *et al* 1997). By using a *Xenopus oocyte* transfection system, *Slc11a2* was shown to transport Fe^{2+}, Zn^{2+}, Mn^{2+}, Co^{2+}, Cd^{2+}, Cu^{2+}, Ni^{2+} and Pb^{2+}. Fe^{2+}, Zn^{2+} and Mn^{2+} are of specific interest in regulating macrophage activation as candidate metal ions contributing to the many pleiotropic effects ascribed to *Slc11a1* (Blackwell *et al* 2000). In particular, Fe^{2+} is of interest for its potential role in the direct generation of toxic antimicrobial hydroxyl radicals within the phagolysosomes of infected macrophages (Zwilling *et al* 1999). Zn^{2+} has also been shown to be important in

15

directing endosomal fusion events (Aballay *et al* 1995, 1997) and Mn^{2+} may be important for metalloprotease activity in the late endosomal MIIC compartment (Lang *et al* 1997). It appears that removal of divalent cations from the phagosome could both restrain pathogen growth and result in the pleiotropic effects of macrophage activation (Buschman and Skamene 2001).

Cellier *et al* (1995) suggested that the *SLC11A1* polypeptides form part of a group of transporters or channels. A 50% reduction in cellular iron content was observed in interferon-γ stimulated macrophages from *Slc11a1* wildtype compared with *Slc11a1* mutant mice (Zwilling *et al* 1999). *Slc11a1* was subsequently shown to function as a metal ion transporter that is regulated by (Atkinson *et al* 1997) and regulates cellular iron levels (Atkinson and Barton 1998). Atkinson *et al* (1997) demonstrated that chelation of iron with deferoxamine influences *Slc11a1* protein expression in interferon-γ and lipopolysaccharide activated macrophages. Atkinson and Barton (1998) suggested that *Slc11a1* could be involved in the recycling of iron and one essential function of phagocytes is to degrade and recycle dead or dying cells including erythrocytes. By directing iron efflux, the gene could therefore participate in the iron salvage pathway in the resting macrophage.

Differences in iron uptake into phagosomes isolated from *Slc11a1* wildtype versus *Slc11a1* mutant macrophages was also demonstrated. Goswami *et al* (2001) showed, by using *Xenopus oocytes*, that *Slc11a1* is a divalent cation (Fe^{2+}, Zn^{2+} and Mn^{2+}) transporter like *Slc11a2* (Gunshin *et al* 1997). However, *Slc11a2* is a symporter of H^+

and metal ions whereas *Slc11a1* is an antiporter, dependent on the pH, that can flux divalent cations in either direction against a proton gradient (Goswami *et al* 2001). Divalent cations (like Mn^{2+} and Fe^{2+}) are known to play a major role in the survival of microbes within phagolysosomal vesicles since many pathogenic organisms secrete defensive Mn- and Fe-dependent superoxide enzymes for scavenging reactive oxygen metabolites. Iron depletion was also suggested as a mechanism by which the growth of intracellular pathogens may be regulated. These results suggested a new model for metal ion homeostasis in macrophages as illustrated in Figure 1.3 (adapted from Blackwell *et al* 2001).

More recently, Blackwell *et al* (2003) investigated the role of *Slc11a1* in the metabolism of phagocytically acquired iron from phagocytosed effete red cells. They found that iron uptake and release, following direct acquisition of transferrin (Tf)-bound iron via the Tf receptor, were equivalent in wildtype (*Slc11a1^{Gly169}*) and mutant (*Slc11a1^{Asp169}*) macrophages and were not influenced by interferon-γ/lipopolysaccharide activation. Iron uptake was equivalent and up-regulated similarly with activation, following phagocytosis of [^{59}Fe]Tf-anti-Tf immune complexes, but intracellular distribution and release of iron was different. In wildtype macrophages 60% of iron was in the soluble form and 12% in the insoluble fraction with 28% ferritin (Ft)-bound. The soluble component increased to 82% with activation. In mutant macrophages 50-60% of the iron was in the soluble form, 40-50% was in the insoluble form and less than 5% was ferritin (Ft)-bound. Western blotting confirmed these results and implied failure in the degradative process in mature phagosomes in mutant macrophages. Nitric oxide (NO)-mediated iron release was also

17

higher (2.4-fold) in activated wildtype macrophages compared with mutant macrophages suggesting that iron, acquired by phagocytosis and degradation, is retained within the phagosomal compartment in wildtype macrophages and that NO triggers iron release by direct secretion of phagosomal contents rather than via the cytoplasm.
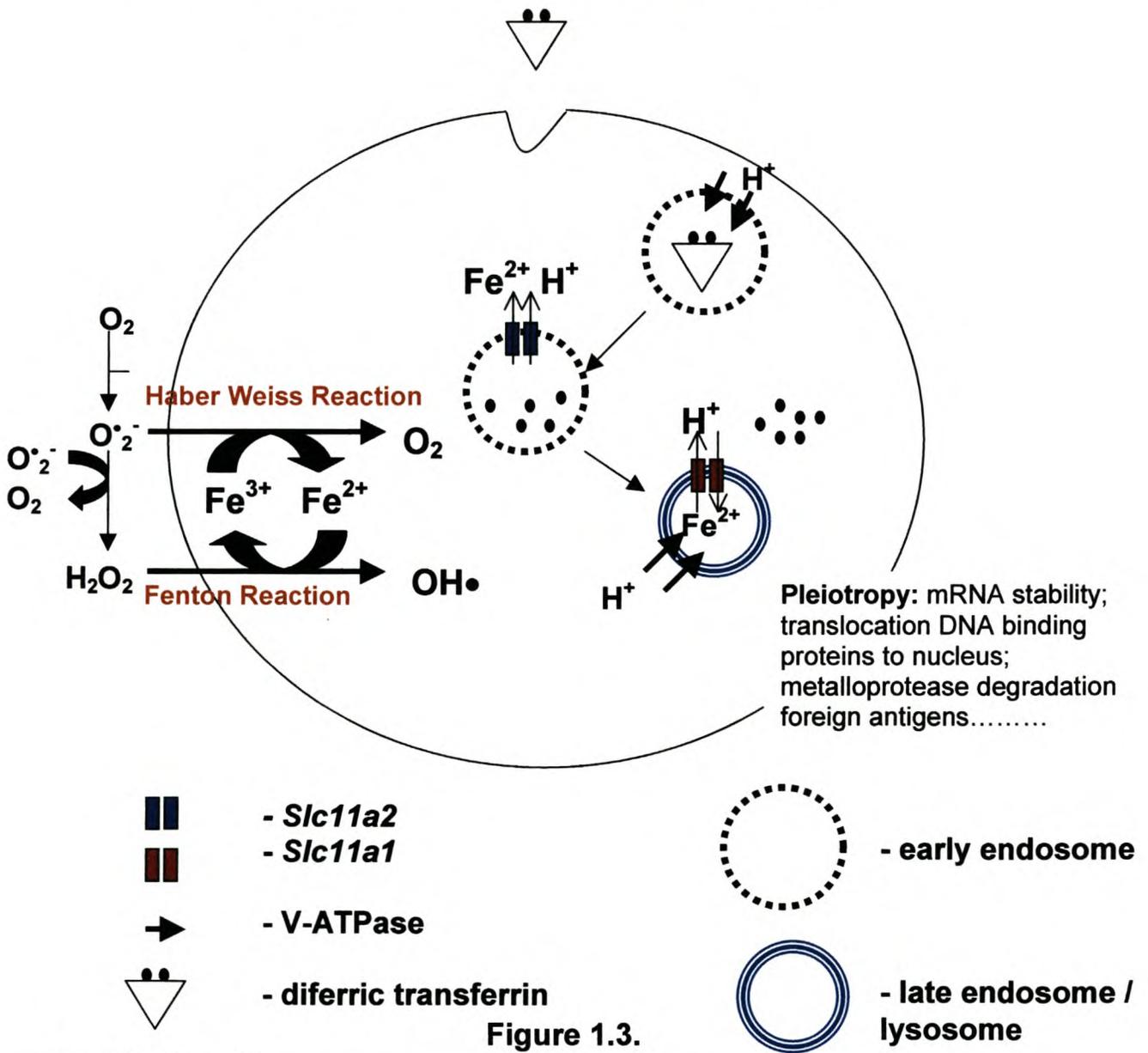
Figure 1.3.

Legend to Figure 1.3.

Schematic representation of divalent cation homeostasis in macrophages and its relationship to oxygen- and nitrogen-dependent antimicrobial activity. $Slc11a2$ delivers extracellularly acquired divalent cations into the cytosol in early endosomes after recruitment of V-ATPase and acidification of the vacuole. $Slc11a1$ delivers divalent cations from the cytosol to late endosomes/lysosomes (an acidic compartment). The Fenton reaction then uses ferrous iron to generate toxic hydroxyl (OH') radicals which may react with NO to produce toxic peroxynitrate (adapted from Blackwell et al 2001).

1.1.6. The role of *SLC11A1* in autoimmune and infectious disease susceptibility

Susceptibility to both autoimmune and infectious diseases has been shown to be associated with variants of the *SLC11A1* gene. Autoimmune diseases include rheumatoid arthritis (Shaw *et al* 1996, Singal *et al* 2000), juvenile rheumatoid arthritis (Shaw *et al* 1996, 1997, Sanjeevi *et al* 2000), diabetes (Esposito *et al* 1998), sarcoidosis (Maliarik *et al* 2000), Crohn's disease (Hofmeister *et al* 1997, Kojima *et al* 2001) and multiple sclerosis (Kotze *et al* 2001). Infectious diseases include tuberculosis (Bellamy *et al* 1998, Bellamy 1999, Cervino *et al* 2000, Gao *et al* 2000, Greenwood *et al* 2000, Ryu *et al* 2000), leprosy (Abel *et al* 1998) and human immunodeficiency virus (Marquet *et al* 1999). Blackwell (2001) summarised the approach in determining susceptibility to infectious diseases by host-pathogen interactions (Figure 1.4). The various studies support the hypothesis that the *SLC11A1* Z-DNA forming promoter repeat polymorphism may contribute directly to the allelic associations with infectious versus autoimmune disease susceptibilities as shown in Table 1.3 (adapted from Blackwell *et al* 2003).

Infectious disease as a complex trait

Host genetic factors

Genetic epidemiology
- Careful selection of disease and disease-related phenotypes
- Comparison of phenotypes in monozygous and dizygous twins, and/or examine segregation of phenotypes in families

Genetic studies in man
- Collect DNAs: family-based (extended multicase pedigrees; dizygous twins or affected sib-pairs plus parents; affecteds plus parents), and/or population-based (ethnically matched case-control pairs)

Linkage and association studies of function-related candidate genes

Genome scan with microsatellite or SNP markers to identify new candidates on basis of map location ('positional cloning')

- Linkage disequilibrium/allelic association mapping to pinpoint and identify disease-related mutation/polymorphism

Analysis of gene function
- *In vitro* and *in vivo* analysis of polymorphism and functional phenotype
- Functional studies using mouse models

Pathogen genetic factors

Genetic studies in pathogen
- Collect pathogen isolates from same cases used in human genetic studies
- Multilocus sequence typing or DNA fingerprinting
- Look for association between disease phenotype/ virulence and pathogen genes

Host–pathogen genetic interaction

Genetic studies in pathogen
- Stratify by pathogen genotype to study host genetics
- Stratify by host genotype to study pathogen genetics
- Look for functional interaction between host and pathogen genotype

TRENDS in Molecular Medicine

Figure 1.4.

Legend to Figure 1.4.
A summary for the approach in determining susceptibility to infectious diseases by host-pathogen interactions (adapted from Blackwell 2001).

Table 1.3. A summary of the various *SLC11A1* Z-DNA-forming 5'-(GT)n repeats associated with autoimmune, infectious disease and immune response phenotypes (adapted from Blackwell *et al* 2003).

Phenotype/disease	Population	(GT)n allele association	Reference
Autoimmune disease			
Rheumatoid arthritis	United Kingdom	Allele 3	Shaw *et al* 1996
Rheumatoid arthritis	Canada	(GT)n not associated	Singal *et al* 2000
Rheumatoid arthritis in HLA shared epitope negative patients	Spain	Allele 2	Rodriguez *et al* 2002
Crohn's disease	USA	Not done	Hofmeister *et al* 1997
Type 1 diabetes in sibs of RA patients	United Kingdom	Allele 3	Esposito *et al* 1998
Early onset type 1 diabetes	Japan	Allele 2 lower in patients	Bassuny *et al* 2002
Juvenile rheumatoid arthritis	Latvia	Allele 3	Sanjeevi *et al* 2000
Primary biliary cirrhosis	United Kingdom	Allele 5	Graham *et al* 2000
Sarcoidosis	African Americans	Allele 3	Maliarik *et al* 2000
Multiple Sclerosis	South Africa	Alleles 3 and 5	Kotze *et al* 2001
Inflammatory bowel disease (Crohn's disease and ulcerative colitis)	Japan	Allele 7	Kojima *et al* 2001

21

Table 1.3. A summary of the various *SLC11A1* Z-DNA-forming 5'-(GT)n repeats associated with autoimmune, infectious disease and immune response phenotypes (adapted from Blackwell *et al* 2003) (Continued).

Phenotype/disease	Population	(GT)n allele association	Reference
Infectious disease			
Pulmonary tuberculosis	Brazil	Allele 2	Shaw *et al* 1997
Pulmonary tuberculosis	The Gambia	Allele 2	Bellamy *et al* 1998
Pulmonary tuberculosis	Japan	Allele 2	Gao *et al* 2000
Pulmonary tuberculosis	Korea	Not done	Ryu *et al* 2000
Pulmonary tuberculosis	Guniea-Conakry	(GT)n not associated	Cervino *et al* 2000
Pulmonary tuberculosis	Aboriginal Canadian	Not done	Greenwood *et al* 2000
Pulmonary tuberculosis	Taiwan	Not done	Liaw *et al* 2002
Leprosy	Vietnam	Not done	Abel *et al* 1998
Leprosy type	West Africa	Not done	Meisner *et al* 2001
HIV	Columbia	Allele 3	Marquet *et al* 1999
Visceral leishmaniasis	Sudan	Allele 3	Mohamed *et al* unpublished data
Meningcoccal meningitis	United Kingdom	Allele 3	Hibberd *et al* unpublished data

Table 1.3. A summary of the various *SLC11A1* Z-DNA-forming 5'-(GT)n repeats associated with autoimmune, infectious disease and immune response phenotypes (adapted from Blackwell *et al* 2003) (Continued).

Phenotype/disease	Population	(GT)n allele association	Reference
Immune response			
Granulomatous mitsuda reaction to lepromin	Vietnam	Not done	Alcais *et al* 2000
IL10 responses to mycobacterial PPDs	Malawi	(GT)n not associated	Blackwell *et al* unpublished data
IL10 responses to LPS and LPS+IFN-	Gambia	Allele 2	Awomoyi *et al* in press

"Not done" indicates that the 5'-(GT)n polymorphism was not investigated but association with other polymorphisms in the *SLC11A1* gene was demonstrated, Allele 2, $t(gt)_5ac(gt)_5ac(gt)_{10}ggcaga(g)_6$, Allele 3, $t(gt)_5ac(gt)_5ac(gt)_9ggcaga(g)_6$, Allele 5, $t(gt)_4ac(gt)_5ac(gt)_{10}ggcaga(g)_6$, Allele 7, $t(gt)_5ac(gt)_5at(gt)_{11}ggcaga(g)_6$.

In this study we focussed on mutation analysis of the *SLC11A1* gene implicated in both autoimmune and infectious disease susceptibility as well as iron transport, within the context of other genes that may be involved in iron homeostasis and/or disease phenotype. These diseases include oesophageal cancer, inflammatory bowel disease and hereditary haemochromatosis and are discussed briefly.

1.1.7. OESOPHAGEAL CANCER (OC)

1.1.7.1. Background

Oesophageal cancer (OC) shows marked geographic variation occurring at very high frequencies in certain areas of China, Iran, Uruguay, France, Italy and South Africa (Parkin *et al* 1988), and displays unique epidemiological features. The disease mainly occurs in adults (mid to late adulthood) and is rarely seen in persons younger than 25, with mortality rates increasing steadily with age (Blot 1994). OC occurs more frequently in males with an estimated frequency of two to four times higher in males than in females.

Two main forms of OC, squamous cell carcinoma (SCC) and adenocarcinoma (ADC), with different aetiological and pathological characteristics have been identified (reviewed by Lu 2000). SCC (more prevalent in developing countries) occurs at a higher frequency than ADC (more common in developed countries), with the highest incidence of ADC in Northern America and in some parts of Europe. Differences have been found between various racial backgrounds, with SCC occurring more often in Blacks and ADC being more prevalent in Caucasians. The occurrence of OC among Blacks is higher at

younger ages compared to Whites; with the Black/White ratio for mortality exceeding six fold for those younger than age 55 (Blot and Fraumeni 1987).

1.1.7.2. Iron homeostasis and OC

Iron overload was previously identified as a possible risk factor in the development of OC and evidence for the involvement of iron in cancers was demonstrated with a clear correlation between dietary iron overload and the development of hepatocellular carcinoma in the South African Black population (Mandishona *et al* 1998). Iron overload associated with OC in the Black South African population was mainly thought to be the result of a high consumption of home-brewed alcoholic beverages contaminated with iron (MacPhail *et al* 1979, Isaacson *et al* 1985). It was speculated that iron loading might be related to the enhanced absorption of iron or ironbound compounds from traditionally brewed maize and sorghum beer (Derman *et al* 1980). The involvement of iron in the development of OC had also been implicated in other populations (Amer *et al* 1990, Rogers *et al* 1993, Hsing *et al* 1995). This observation was further illustrated in a rat model where iron supplementation was indicated as a cause of developing OC (Goldstein *et al* 1998, Chen *et al* 1999, 2000).

Normal anti-tumour activity occurs through the loss of iron from the target cell and reduced anti-tumour activity is observed in the presence of excess iron (Green *et al* 1988, Huot *et al* 1990). Iron therefore acts as an indirect carcinogen. Excess iron also interferes with the growth of tumouricidal-activated macrophages and leads to enhanced tumor growth (Weiss *et al* 1992). The generation of oxygen free radicals due to iron

25

induced mutagenesis can also contribute to the development of cancer, leading to DNA strand damage and disruption of DNA structure (Togokuni 1996).

1.1.7.3. Factors involved in OC pathogenesis

i) Environmental involvement

Of the environmental risk factors contributing to the development of SCC, cigarette smoking and alcoholic beverage consumption are the major determinants (International Agency for Research on Cancer 1986, 1988, Blot 1994). The risk of developing OC is five times higher in cigarette smokers (tenfold among heavy cigarette smokers) than non-smokers and increased alcohol consumption among smokers further increases the risk of developing OC in a synergistic manner. The risk of developing OC also increases with the amount of alcohol consumed (Blot 1992), regardless of the smoking category, with a 20 to 50-fold increase (La Vecchia and Negri 1989, Blot 1994). Diet and nutrition may also play a major role with diets high in fresh fruit and vegetables consistently associated with reduced risk (Van Rensburg 1981, Steinmetz and Potter 1991, Blot 1994). Other factors contributing to a lesser extent to disease pathogenesis were occupational exposure (Selikoff *et al* 1979, Norell *et al* 1983, Gustavisson *et al* 1993), ionising radiation (Smith 1984), drinking of hot beverages (Yang and Wang 1993) and viral agents (papilloma viruses) (Hille *et al* 1985).

ii) Genetic involvement

Molecular studies have revealed frequent chromosomal abnormalities (Whang-Peng *et al* 1990, Rosenblum-Vos *et al* 1993, Rao *et al* 1995) and gene alterations. Gene alterations include amplification of the cellular oncogenes *C-MYC* (Lu *et al* 1988) and

the *INT-2* gene (Lu 2000), and amplification and expression of CYCLIN-D1 *(CCND1)* (Liang and Lu 1991, Adelaide *et al* 1995). Frequent mutations and deletions were identified in genes such as the retinoblastome (*RB1*), tumor protein p53 (*TP53*), adenomatous polyposis of the colon (*APC*), mutated in colorectal cancers (*MCC*) and deleted in colorectal carcinoma (*DCC*) genes (reviewed by Lu 2000). Several other genes involved in cell growth and regulatory pathways have also been found to be altered in OC (reviewed by Klimstra 1994, Rosen 1994, Mandard *et al* 2000). Rosen (1994) stated that most tumors develop as a result of multiple mutations in genes involved in regulating growth and differentiation. Studies involving loss of heterozygosity (LOH) (Tarmin *et al* 1994, Barrett *et al* 1996, Shimada *et al* 1996) and comparative genomic hybridisation (CGH) (Moskaluk *et al* 1998, Sarlomo-Rikala *et al* 1998, du Plessis *et al* 1999) have also identified several chromosomal regions of potential importance implicated in OC. These studies suggest the involvement of many oncogenes and tumor suppressor genes in the initiation and development of OC.

1.1.8. INFLAMMATORY BOWEL DISEASE (IBD)

1.1.8.1. Background

Inflammatory bowel disease (IBD) occurs due to chronic inflammation of the intestines, resulting in swelling, ulcerations and loss of function, and is often referred to as an autoimmune disease. The disease is generally subdivided into two phenotypes known as Crohn's disease (CD) and ulcerative colitis (UC) (Sanderson 1986, Griffiths 1995). The combined prevalence of CD and UC is estimated to be 100 to 200 per 100 000 in the Western world (Kyle 1992, Loftus *et al* 1998).

27

UC involves continuous inflammation of the rectum and colonic mucosal layers, without fistulas and granulomas, with inflammation occurring rarely in the last section of the ileum. In contrast, in CD, discontinuous inflammation occurs in any part of the gastrointestinal tract, most frequently affecting the terminal ileum and colon, with fistulas and granulomas. Two distinct clinical forms of CD, known as perforating and non-perorating CD, have also been documented (Greenstein *et al* 1988, Gilberts *et al* 1994). Nonperforating CD is the more benign form, whereas perforating CD, with the presence of absesses and/or free perforations, is the more aggressive form. CD or UC cannot be determined in approximately 10% of cases, referred to as indeterminate colitis, when inflammation is confined to the rectum and colon. CD is suggested to be the outcome of three interactive factors, including host susceptibility, enteric microflora and the mucosal immune system (Fiocchi 1998).

Patients with IBD present with symptoms such as diarrhea, abdominal pain and rectal bleeding (Sanderson 1986, Griffiths 1995). In some cases constipation, fever, weight loss and dehydration also occur. Severe constipation causes perforation of the intestine and results in toxic dilation of the colon, referred to as toxic megacolon. The presence of CD or UC is based on clinical presentation and confirmed by radiology, or by endoscopy and biopsy (Heikenen *et al* 1999). Extraintestinal inflammation of the skin, eyes or joints can also occur and the prevelance of IBD increases when associated with other autoimmune diseases such as ankylosing spondylitis, psoriasis, sclerosing cholangitis and multiple sclerosis (Snook *et al* 1989, Yang and Rotter 1993). Several genetic syndromes, such as Turner syndrome, glycogen storage disease type Ib and

Hermansky-Pudlak syndrome, have also been reported to be associated with IBD (Shanahan *et al* 1988, Yang and Rotter 1993).

1.1.8.2. Iron homeostasis and IBD

Anaemia of iron deficiency, due to the inadequate intake or loss of iron, and anaemia of chronic diseases, resulting from decreased erythropoiesis, frequently occur in IBD. Iron homeostasis is usually altered in conditions of chronic inflammation, predominantly resulting from chronic blood loss in the intestine (Oldenburg *et al* 2001). Iron deficiency can occur due to iron malabsorption because of inflammatory activity in the small intestine and resection or impaired dietary intake. Iron is essential for bacterial growth and it is speculated that anaemia may result from an attempt by the host to deprive infiltrative microorganisms of iron and thereby hindering them from growing. This is supported by findings that high body iron leads to increase susceptibility to a variety of bacteria (Calver *et al* 1979, Wright *et al* 1981, Van Asbeck *et al* 1982, Martinez *et al* 1990) and a low level of iron appear to protect against various pathogens (Murray and Murray 1977, Murray *et al* 1980, Oppenheimer *et al* 1986).

However, in chronic inflammatory conditions such as rheumatoid arthritis where no microorganisms are involved, inflammation is ascribed to the generation of reactive oxygen species (Oldenburg *et al* 2001). These species give rise to the formation of OH in the Haber-Weiss reaction (Kehrer 2000) and the availability of chelated iron is essential for the formation of the OH radical in the Fenton reaction (Cohen 1985). A close relationship exists between iron and immunity but the exact influence of iron metabolism on cell-mediated immunity is not well established yet. Iron overload has

been shown to impair intracellular killing of pathogens, through IFNγ-mediated pathways, by macrophages (Alford *et al* 1991). The control of iron metabolism is interconnected with other cellular pathways, and further analysis of these aspects are likely to provide a better understanding of pathological processes in which iron may participate.

1.1.8.3. Factors involved in IBD pathogenesis

Research involving both cellular and molecular biology revealed the possible involvement of several factors in the aetiology of IBD. These included both genetic and environmental factors (Sartor 1995, Shanahan 1995, Targan 1995). By reviewing potential animal models for CD as well as results of studies on the mechanisms of its pathogenesis and human molecular genetics, Targan and Murphy (1995) concluded that CD heterogeneity is at the primary genetic level and expression of genetic susceptibility requires environmental triggers.

i) Environmental involvement

The role of environmental factors is thought to be one of the less understood and most difficult factors to interpret in the progression of IBD, and possibly as important as genetic factors in the risk of developing IBD (Fiocchi 1998). Potential environmental factors include prenatal events, breastfeeding, childhood infections, microbial agents, smoking, oral contraceptives, diet, hygiene, occupation, education, climate, pollution, stress, and miscellaneous components such as toothpaste, appendectomy, tonsillectomy, blood transfusions, contact with animals, and physical activity. The most

established association and intriguing factor is smoking, with non-smoking being associated with UC and smoking with CD (Harries *et al* 1982a, 1982b, Tysk *et al* 1988, Somerville *et al* 1994, Reif *et al* 1995).

ii) Genetic involvement

The involvement of genes in the pathogenesis of IBD was confirmed by familial studies (Orholm *et al* 1991, Yang *et al* 1993, Binder 1998), difference in concordance rates in twin pairs (Tysk *et al* 1988, Binder 1998) and linkage studies (Hugot *et al* 1996, Satsangi *et al* 1996a, Cho *et al* 1998, 2000, Rioux *et al* 2000, IBD International Genetics Consortium 2001, Rioux *et al* 2001, Hugot *et al* 2001, Ogura *et al* 2001a). Linkage studies provided evidence for linkage between IBD and loci on several chromosomes associated with CD and/or UC (Cho *et al* 1998, 2000, Satsangi *et al* 1996a, Duerr *et al* 1998, Parkes *et al* 2000, Rioux *et al* 2001).

Prevalence of UC in first-degree relatives is estimated to be between 4 and 16% (Lewkonia and McConnell 1976, Farmer *et al* 1980) with the frequency being higher in CD than in UC (Satsangi *et al* 1994). A positive family history is more common in patients with CD than in UC, with relatives of patients with CD developing CD or UC, but most likely CD (Satsangi *et al* 1997). Satsangi *et al* (1996b) studied the clinical characteristics of IBD and found statistically significant differences in the median age of onset, being higher in the patients than in their offspring, possibly reflecting genetic anticipation between generations. Inheritance in families did not follow simple mendelian models and therefore IBD is considered as a complex genetic trait largely due to genetic heterogeneity.

31

Concordance rates in twin pairs provided strong evidence of genetic predisposition in IBD, especially CD (Tysk *et al* 1988, Binder 1998). A difference in concordance rate was also identified between CD and UC in these twin studies, suggesting differences in the genetic basis of these diseases. Genetic factors seem to be relatively more important in CD than in UC and are considered to be related polygenic disorders (Satsangi *et al* 1997). Satsangi *et al* (1996b) suggested that CD and UC are closely related, but distinct polygenic disorders, sharing some susceptibility genes.

The involvement of the solute carrier family 11 member 1 gene (*SLC11A1*) in IBD pathogenesis was first investigated by Hofmeister *et al* (1997). The SLC11A1 protein is believed to regulate macrophage function and is a good candidate gene because hyperreactive immune response toward the intestinal flora was previously implicated in the pathogenesis of IBD (Duchmann *et al* 1995). Hofmeister *et al* (1997) identified significant differences in haplotype distributions in patients with CD and reported the first association between the *SLC11A1* gene and CD. However, analysis of the *SLC11A1* gene polymorphisms (274C→T and 823C→T) in a study performed by Stokkers *et al* (1999), showed no association of the *SLC11A1* gene with IBD susceptibility. A recent study by Kojima *et al* (2001) confirmed association of a novel allele $[t(gt)_5ac(gt)_5at(gt)_{11}ggcaga(g)_6$ - Allele 7] of the *SLC11A1* promoter repeat polymorphism identified in the Japanese population with IBD.

Recently, independent studies performed by Hugot *et al* (2001) and Ogura *et al* (2001a), using positional-cloning strategies and linkage association studies, respectively, identified disease-predisposing alleles of the caspase recruitment domain-containing protein 15 gene (*CARD15*), formerly known as the NOD 2 protein gene (*NOD2*). The gene is located on chromosome region 16q12 and is comprised of 12 exons (Ogura *et al* 2001a). This gene showed strong linkage to CD susceptibility. Thirteen sequence variants were identified and significant association of three of these polymorphisms (R702W, G908R and 1007fs) with CD was demonstrated. Association of the frameshift variant with CD was confirmed in a German and British population (Hampe *et al* 2001). Mutations of *CARD15* were identified in 40 to 50% of CD patients (Vermeire *et al* 2000, Ogura *et al* 2001a, Hugot *et al* 2001, Hampe *et al* 2001, Lesage *et al* 2002, Van Heel *et al* 2002), with a causative genetic defect in about 15% of the patients. In a recent study by Yamazaki *et al* (2002), significant association with *CARD15* could not be confirmed in the Japanese population. The authors suggested that the prevalent *CARD15* mutations are independent risk factors for CD in Caucasian patients.

The gene encodes a member of the Apaf-1/Ced4 superfamily of apoptosis regulators that is expressed in monocytes. The CARD15 protein is homologous to a class of cytosolic plant disease resistance genes. These genes recognize bacterial components and induce appropriate pathogen specific responses (Weitzman 2001). The gene comprises two N-terminal capase-recruitment domains (CARDs) fused to a central nucleotide-binding domain (NBD) and ten leucine-rich repeats (LRR) (Ogura *et al* 2001a).

The variants identified are predicted to alter structure of either the LRR domain of the protein, possibly affecting host-interactions with bacterial LPS and inducing the nuclear factor κB (NF-κB) signaling pathway, or the adjacent region (Weitzman 2001). NF-κB is a target of anti-inflammatory steroids and sulfasalazine and is activated by *CARD15* (Ogura *et al* 2001b). The function of the protein encoded by *CARD15* is only partly known and is predicted to detect bacteria in the gut and to help control inflammatory response to them (Ogura *et al* 2001b). The role of the CARD15 protein in IBD pathogenesis is speculated to: (1) regulate the induction of cytokines (via NF-κB) that controls inflammatory response and (2) alter or change pathogen recognition patterns due to mutations in the LRR (Weitzman 2001, Ogura *et al* 2001b).

1.1.9. HEREDITARY HAEMOCHROMATOSIS (HH)

1.1.9.1. Background

Hereditary haemochromatosis (HH) is generally an autosomal recessive condition (Edwards *et al* 1988). Features of haemochromatosis include cirrhosis of the liver, diabetes mellitus, athropathy, cardiomyopathy, endocrine abnormalities and hepatocellular carcinoma, which is the result of tissue damage from the progressive accumulation of iron (Bothwell *et al* 1995). Men usually present with clinical symptoms by the fourth to fifth decade and women presenting about a decade later. Women may be protected against iron accumulation due to menstruation and pregnancy (Bothwell and MacPhail 1998).

HH is characterised by the inappropriate absorption of iron from the gut, resulting in iron overload (Bothwell *et al* 1995). Excessive iron absorption causes iron overload primarily in parenchymal cells. Muir *et al* (1984) identified four groups of disease expression, possibly indicating the involvement of other genes in iron metabolism leading to HH. These groups include i) a classic form with elevated transferrin saturation, serum ferritin levels and liver iron content (group I); ii) severe iron overload, accelerated disease manifesting at an early age (group II); iii) elevated total body iron stores, normal transferrin saturation and serum ferritin levels (group III) and iv) markedly elevated findings on serum biochemical tests, i.e., transferrin saturation and serum ferritin, with minimal elevation in total body iron stores (group IV).

1.1.9.3. Factors involved in HH pathogenesis

i) Environmental involvement

Non-genetic factors influencing penetrance of different *HFE* genotypes include physiological and pathological blood loss, dietary intake of iron and alcohol intake (Adams *et al* 2000).

ii) Genetic involvement

Type 1 haemochromatosis, also designated *HFE1*, is associated with mutations in the *HFE* gene that maps to 6p21.3 and mutations C282Y and H63D have been identified as the most common cause of the disorder (Feder *et al* 1996) [see section 1.1.9.4(i)]. More than 90% of individuals presenting with genetic haemochromatosis in the United Kingdom will have an *HFE* related mutation (The UK Haemochromatosis Consortium 1997). Non-*HFE* related forms of HH, which are less common, have also been reported.

35

The gene involved in type 2 haemochromatosis, also known as juvenile haemochromatosis (JH) or *HFE2*, has been mapped to chromosome region 1q but has not yet been characterised (Camaschella *et al* 1997). JH is a rare and severe disorder of iron overload with liver and heart disease being evident by the third decade. Recently, mutations identified in the hepcidin antimicrobial peptide gene (*HAMP*), that maps to 19q13, were associated with JH (Roetto *et al* 2003). JH linked to the unidentified gene on chromosome region 1q has been designated *HFE2A* and *HFE2B* refers to JH associated with mutations in *HAMP*. Mutations in the transferrin receptor 2 gene (*TFR2*), that maps to 7q22, contribute to haemochromatosis type 3 (*HFE3*) (Camaschella *et al* 2000, Roetto *et al* 2001) and haemochromatosis type 4 (*HFE4*), associated with an autosomal dominant form of inheritance and maps to 2q32, is associated with mutations in the solute carrier family 11 member 3 gene (*SLC11A3*) (Njajou *et al* 2001, Montosi *et al* 2001).

1.1.9.4. Genes involved in iron homeostasis

Several genes have been implicated in regulating iron homeostasis and includes the transferrin receptor 2 gene (*TFR2*) (Camaschella *et al* 2000, Roetto *et al* 2001), the solute carrier family 11 (proton-coupled divalent metal ion transporter) member 2 gene [*SLC11A2*, also known as the natural resistance-associated macrophage protein 2 gene (*NRAMP2*) or divalent metal transporter 1 gene (*DMT1*)] (Gruenheid *et al* 1995, Gunshin *et al* 1997), hepcidin antimicrobial peptide gene [*HAMP*, also known as the liver–expressed antimicrobial peptide gene (*LEAP1*) or hepcidin (*HEPC*)] (Krause *et al* 2000, Park *et al* 2001, Pigeon *et al* 2001, Nicolas *et al* 2001), solute carrier family 11 (proton-coupled divalent metal ion transporter) member 3 gene [*SLC11A3*, also known as the

ferroportin 1 gene (*FPN1*), iron-regulated transporter 1 gene (*IREG1*) or metal transporter 1 gene (*MTP1*)] (Donovan *et al* 2000, McKie *et al* 2000, Abboud and Haile 2000), duodenal cytochrome b gene (*DCYTB*) (McKie *et al* 2001), ceruloplasmin (Cairo *et al* 2001) and hephaestin (Vulpe *et al* 1999, Kaplan and Kushner 2000). The function and localisation of the proteins encoding the various genes are illustrated in figure 1.5 (with permission adapted from Pointon 2001)
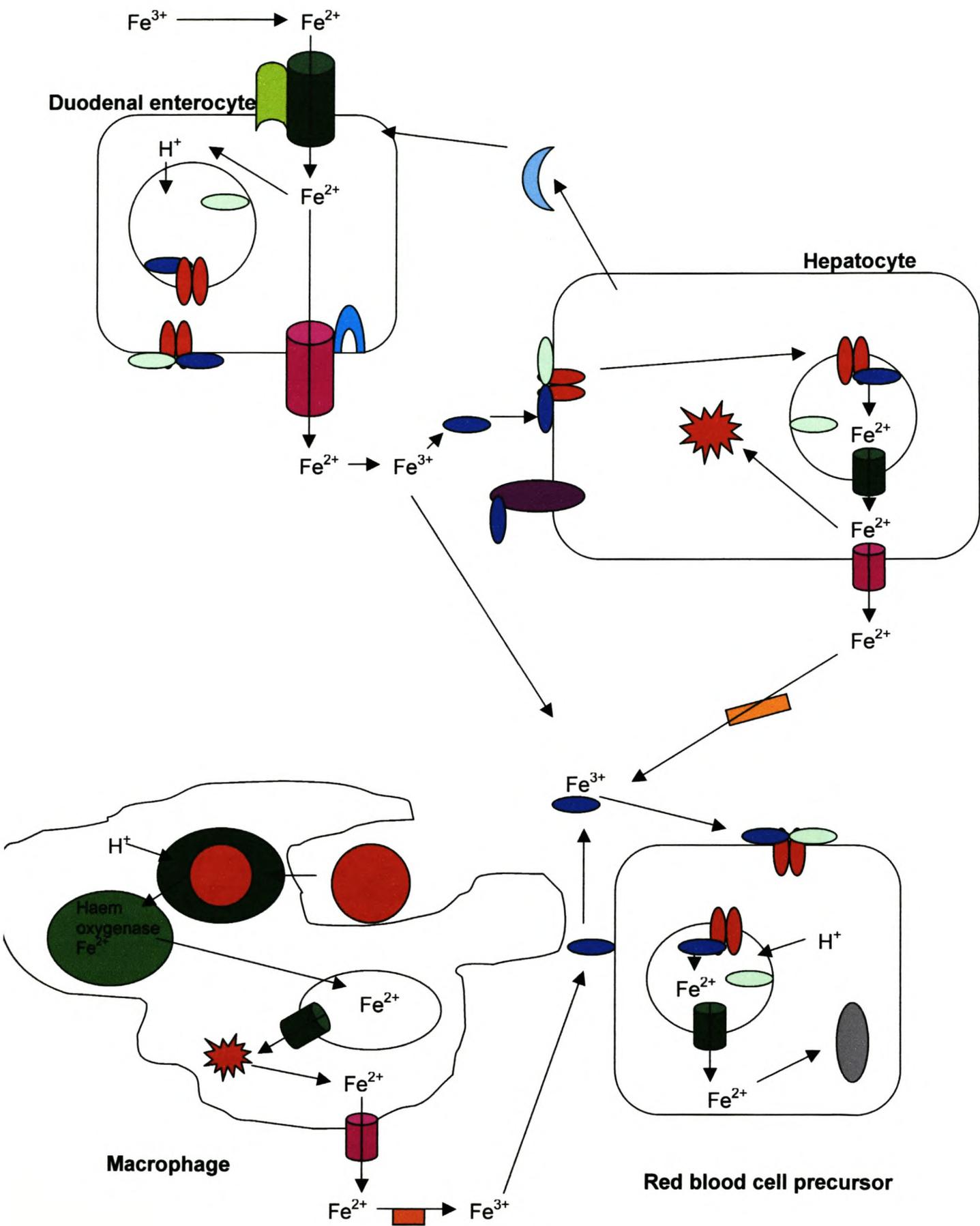
Figure 1.5.
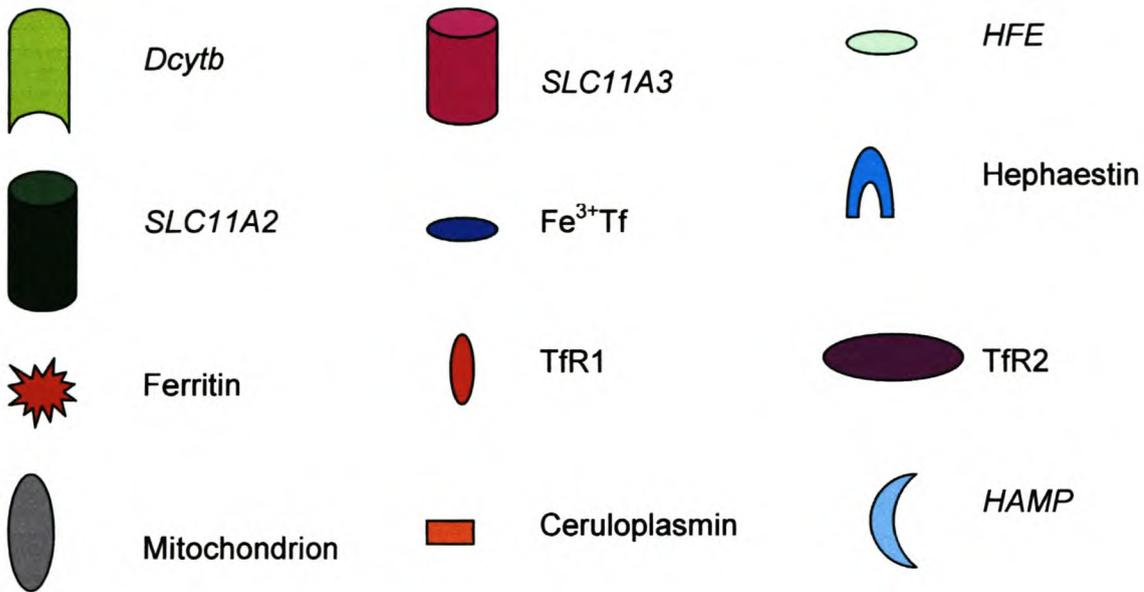
Legend to Figure 1.5.

A schematic representation of the movement of iron in cell types having a role in iron metabolism. Dietary iron is absorbed by duodenal enterocytes where *DCYTB* reduces ferric iron to ferrous iron at the luminal surface. Ferrous iron is transported across the apical membrane by *SLC11A2*, followed by the export of iron at the basal membrane by *SLC11A3*. Ferrous iron is then oxidised to ferric iron by hephaestin. In hepatocytes, iron is taken up via either the Tf-Tf1 endocytic pathway possibly influenced by *HFE* or as Tf bound iron via TfR2. Iron is stored bound to ferritin and can be exported to the circulation via *SLC11A3*, if required, and reduced to ferric iron by ceruloplasmin. Iron is taken up in the red blood cell precursor by TfR1 endocytosis and used for the synthesis of haemoglobin. Macrophages recycle iron from phagocytosed effete red blood cells where haem iron is released by haem oxidase.

In this study, we investigated the involvement of *HFE, SLC11A3, HAMP* and *DCYTB* in patients presenting with hereditary haemochromatosis and only these genes will be discussed further.

i) HFE

HFE has been identified as the major gene underlying HH. The *HFE* gene is localised on chromosome 6p (Simon *et al* 1975, Feder *et al* 1996). The protein consists of 343 amino acids that is homologous to the MHC class I protein, HLA-A2 and the non-classical class I protein, HLA-G. The HFE protein consists of three extracellular domains (α1-domain, α2-domain and α3-domain), a transmembrane domain and a small cytoplasmic tail, and forms a heterodimer with β_2-microglobulin for cell surface expression. *HFE* forms a complex with transferrin receptor 1 (TfR1) (Parkkila *et al* 1997, Feder *et al* 1998), reducing the affinity of TfR1 for transferrin (Tf) by five to ten-fold (Feder *et al* 1998, Gross *et al* 1998, Ikuta *et al* 2000). Evidence of *HFE* being involved in iron loading was observed in mice being deficient in *Hfe* (Zhou *et al* 1998, Fleming *et al* 2001).

Homozygosity for a mutation in the *HFE* gene, causing a cysteine to tyrosine substitution at amino acid 282 (C282Y), has been shown to be the most common cause of HH in up to 90% of individuals from European descent (The UK Haemochromatosis Consortium 1997). This mutation, located in the α3-domain, disrupts a disulphide bond that is essential for the interaction with β_2-microglobulin and is predicted to result in the lack of cell surface expression of the gene (Feder *et al* 1996). One-third of deaths in affected individuals being homozygous for the C282Y mutation occur due to primary hepatocellular carcinoma

40

A second missense mutation that results in a histidine to aspartic acid substitution at amino acid 63 (H63D) is also associated with HH. Homozygosity occurs in approximately 4.5% of HH patients (Merryweather-Clarke *et al* 2000) and up to 13.3% is heterozygous for the H63D mutation (Piperno *et al* 1998). The H63D mutation is much more common in European populations, approximately 30.4% (Merryweather-Clarke 2000), than in patients with HH. Approximately 10% of HH patients are C282Y and H63D compound heterozygotes (Piperno *et al* 1998, Merryweather-Clarke *et al* 2000). The H63D mutation is located in the α1-domain of the gene. Other mutations in *HFE* accounts for up to 2 to 9.6% of HH individuals in different populations (Feder *et al* 1996, Henz *et al* 1997, Douabin *et al* 1999, De Villiers *et al* 1999, Barton *et al* 1999a, Rochette *et al* 1999, Wallace *et al* 1999, Worwood *et al* 1999, Rosmorduc *et al* 2000, Pointon *et al* 2000, Piperno *et al* 2000).

ii) Solute carrier family 11 (proton-coupled divalent metal ion transporter) member 3 gene (SLC11A3)

The *SLC11A3* gene, localised to chromosome 2q32 (Haile 2000), encodes a transmembrane protein involved in releasing iron from cells (Donovan *et al* 2000, McKie *et al* 2000, Abboud and Haile 2000). The gene was first identified by positional cloning, in *weissherbst zebrafish*, as the gene responsible for hypochromic anaemia (Donovan *et al* 2000). Mouse and human cDNA was subsequently isolated by RT-PCR of liver and placenta, respectively (Donovan *et al* 2000), and by using a subtractive cloning strategy and PCR analysis (McKie *et al* 2000).

41

The *SLC11A3* gene consists of 8 exons and spans a region of 20 kb (Njajou *et al* 2001). The human *SLC11A3* protein consists of 571 amino acids and contains 10 transmembrane domains (McKie *et al* 2000). The protein is localised to the basolateral membrane of polarised epithelial cells. A functional iron responsive element (IRE), predicted to form a hairpin-loop, was identified in the 5' untranslated region. The highest level of expression of the gene in humans is in placenta, liver, spleen and kidney. Expression of *SLC11A3* in *Xenopus oocytes* stimulated iron efflux (McKie *et al* 2000). Recently, mutations in the *SLC11A3* gene have been shown to be associated with an autosomal dominant form of haemochromatosis, also termed haemochromatosis type 4 (*HFE4*) (Njajou *et al* 2001, Montosi *et al* 2001, Wallace *et al* 2002, Devalia *et al* 2002, Roetto *et al* 2002).

iii) Hepcidin antimicrobial peptide gene (HAMP)

Biochemical purification of blood ultrafiltrate using a cysteine alkylation assay and mass spectrometry, followed by micropeptide sequence and RT-PCR analysis as well as 5' and 3' RACE was used to isolate the cDNA encoding *HAMP* (Krause *et al* 2000). Park *et al* (2001) also isolated the cDNA for *HAMP* by biochemical purification and amino acid sequence analysis of hepcidin peaks in urine, followed by EST database searching and 5' RACE. The human gene was mapped to chromosome region 19q13 (Krause *et al* 2000, Park *et al* 2001, Pigeon *et al* 2001) and the mouse gene to chromosome 7 (Pigeon *et al* 2001) by genomic sequencing analysis.

The gene consists of 3 exons with the final exon encoding the active peptide (Krause *et al* 2000, Park *et al* 2001, Pigeon *et al* 2001). The protein consists of 84 amino acids that undergo enzymatic cleavage into mature 20, 22 and 25 amino acids. The protein contains a 24-residue N-terminal signal sequence and a pentaarginyl proteolysis site followed by the active C-terminal 25-amino acid peptide (Krause *et al* 2000). The active peptide contains a unique 17-residue stretch with 8 cysteines forming 4 disulfide bridges that stabilize the beta-sheet structure.

By using animal models, Nicolas *et al* (2001) found that *Hamp* gene expression was totally inhibited in mice exhibiting iron overload consequent to targeted disruption of the upstream stimulatory factor 2 gene (*Usf2*), which is in close proximity to *Hamp*. In these mice the development of iron overload was similar to that observed in human hereditary haemochromatosis and in mice with knockout of the *Hfe* gene (Zhou *et al* 1998, Bahram *et al* 1999, Levy *et al* 1999). These authors suggested that overexpression of hepcidin might result in phenotypic traits of iron deficiency. Transgenic mice expressing hepcidin under the control of the liver-specific transthyretin (TTR) promoter presented with decreased body iron levels and severe microcytic hypochromic anemia. These results support the proposed role of hepcidin as an iron-regulatory hormone. Mutations in this gene has also been identified and associated with juvenile haemochromatosis (Roetto *et al* 2003).

iv) Duodenal Cytochrome b gene (DCYTB)

The *DCYTB* gene encodes a cytochrome b-like molecule and was isolated by using a subtractive cloning strategy (McKie *et al* 2000). The gene consists of 4 exons, which code for 286 amino acids containing six predicted transmembrane domains and four conserved histidine residues.

DCYTB is a ferric reductase and has nucleotide sequence homology to the cytochrome b561 family and the yeast ferrireductase Fre1 (Shatwell *et al* 1996). The gene is highly expressed at the intestinal brush border and reduces ferric ions to ferrous ions in the gut lumen for transport across the apical membrane as the first step in intestinal iron absorption (McKie *et al* 2001). Dcytb expression, as for Slc11a2, increases when iron absorption is stimulated and may use ascorbate as a cofactor because it lacks the motifs for binding the more common cofactors (nicotinamide adenine dinucleotide, nicotinamide adenine dinucleotide phosphate and flavin adenine dinucleotide) (McKie *et al* 2001).

1.2. OBJECTIVES OF THIS STUDY

i) Mutational analysis of the *SLC11A1* gene 5'-(GT)n repeat polymorphism was performed to determine whether the effects mediated by the gene are related to its involvement in iron homeostasis.

ii) The functional polymorphisms of the *SLC11A1* gene were analysed in patients presenting with oesophageal cancer (OC), inflammatory bowel disease (IBD) and hereditary haemochromatosis (HH) respectively, and association studies were performed to investigate the *SLC11A1* gene as a potential modifier locus in these disease phenotypes.

iii) Further investigation of the *SLC11A1* gene as a potential modifier locus was exploited by investigating possible gene-gene interaction with the *CARD15* gene, which has been implicated in IBD susceptibility.

CHAPTER 2

DETAILED EXPERIMENTAL PROCEDURES

2.1. DNA EXTRACTION FROM WHOLE BLOOD

DNA was extracted from whole blood samples by a modification of the technique of Miller *et al* (1988). Cold lysis buffer (40 ml) (155 mM NH_4Cl, 10 mM $KHCO_3$, 0.1 mM EDTA – pH 7.4) was added to each whole blood sample in a 50 ml polypropylene tube (Falcon). The solution was kept on ice until the red blood cells had undergone lysis and the cell suspension was then centrifuged at 250 x g for 10 minutes. After the supernatant was removed, the pellet was washed twice with 5 ml cold PBS (27 mM KCl, 137 mM NaCl, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4). The intact pellet was resuspended in 3 ml nucleic lysis buffer (0.01M Tris-HCl (pH 8.2), 0.4 M NaCl, 2 mM EDTA - pH 8.2), 1% SDS (w/v) and 1.5 mg/ml proteinase K (Roche Diagnostics), and incubated overnight in a water bath at 55^oC.

Following overnight incubation, 1 ml of saturated 6 M NaCl was added and the solution shaken vigorously for 1 minute, followed by centrifugation at 950 x g for 15 minutes. The supernatant containing the DNA was transferred to a clean Falcon tube. Two times the volume ice-cold EtOH (99%-) (v/v) was added and the solution was left at room temperature for 30 minutes. The precipitated DNA was transferred to a fresh 1.5 ml tube (Eppendorf) and washed twice with 1 ml 70% EtOH (v/v) to remove any excess salt. After centrifugation at 800 x g for 15 minutes, the excess EtOH was carefully removed and the DNA was left to air dry at room temperature. The DNA pellet was dissolved in 500 µl ddH_2O overnight at room temperature and the DNA stored at 4^oC. DNA quality and quantity was measured by agarose gel electrophresis or spectrophotometry (µQuant Bio-Tek Instruments)

2.2. POLYMERASE CHAIN REACTION (PCR) AMPLIFICATION

2.2.1. Oligonucleotide Primers

The PCR primers used in this study for heteroduplex single-strand conformation polymorphism (HEX-SSCP) and denaturing high performance liquid chromatography (dHPLC) analysis are listed in Table 2.1. Primers used for PCR amplification of products subjected to restriction fragment length polymorphism (RFLP) analysis are listed in Table 2.2.

48

Table 2.1. Oligonucleotide primers used for amplifying PCR products subjected to HEX-SSCP and dHPLC analysis

GENE	EXON	PRIMERS				PCR	ANN 1	ANN 2	ANALYSIS
		Forward (5'-3')	*Tm (°C)	Reverse (5'-3')	*Tm (°C)	FRAGMENT SIZE (bp)	(°C)	(°C)	
SLC11A1	[a]5'-UTR	GGGGTCTTGGAACTCCAGAT	62	GGCAGCTCCTCAGCCTGCAC	68	467	65	60	HEX-SSCP
	[b]2	AGGAGGGAAAGGATCAGG	56	GGATCCTGTCAATCTTGCAGGC	68	143	65		HEX-SSCP
[c]*SLC11A3*	1	CCCCGACTCGGTATAAGA	56	TTCCTCCAGAACTCGTGTAG	60	489	54		dHPLC
	2	TGGATAAGCATTCTGCCCTC	60	AAAGCATGTGTACTTGGATG	56	275	54		dHPLC
	3	AATGTAGCCAGGAAGTGCC	58	AGAGGTGGTGCCATCTAAG	58	416	58		dHPLC
	4	GGATAAGAACAGTCTCACTG	58	TTCATCCTTTACCACTACCAG	60	243	54		dHPLC
	5	TTAAACTGCCTTGTTTAGTG	54	GCCTCATTTATCACCACCG	58	279	58		dHPLC
	6	TTGTGTAAATGGGCAGTCTC	58	CCTCGTCTACCAAAGCGATA	60	460	54		dHPLC
	7i	GCTTTTATTTCTACATGTCC	54	GCTGTGCCAATCCTGAGATC	62	435	54		dHPLC
	7ii	GAGCATCAGCTATAACTGG	56	TAATGGATTCTCTGAACCTAC	58	435	54		dHPLC
	8i	TTGAAATGTATGCCTGTAAAC	56	TTTCCATGCCTCAACATAAGG	60	567	54		dHPLC
	8ii	GTTTTTACCACAGCTGTGCC	60	ATACCTTAAGATCAATAGGAT	54	411	54		dHPLC

49

Table 2.1. Oligonucleotide primers used for amplifying PCR products subjected to HEX-SSCP and dHPLC analysis (Continued)

GENE	EXON	PRIMERS				PCR	ANN 1	ANN 2	ANALYSIS
		Forward (5'-3')	*Tm (°C)	Reverse (5'-3')	*Tm (°C)	FRAGMENT SIZE (bp)	(°C)	(°C)	
ᵈDCYTB	1	AGACAGCCCCAAGAAGTC	56	TCCTCCAGCACTTCCTCAG	60	407	62.3		dHPLC
	2	GAGGGGAGAAGCAAAAGCCAAG	56	AAGATCGCAGCACTGCACTC	62	476	63.1		dHPLC
	3	GTAGTGGAACTGGAGCGAG	56	TGTCAGACTTTTTGGTCCTTC	60	496	62.1		dHPLC
	4i	CTGTTTTGCATGTTGCTGTATC	62	ATCTCTGCCCAGCCTCATC	60	352	58		dHPLC
	4ii	AGGTTCCATGCCAGCCTACTC	66	ATCTATTCTGCACAAGGCACCG	64	406	58		dHPLC
ᵈHAMP	1	AGCAAAGGGGAGGGGGCTCAGACC	80	TCCCATCCCTGCTGCCCTGCTAAGGA	84	292	60		dHPLC
	2+3	TTGCCGGGAGCCAGTCTAGAGGTCCA	60	TGCAAGGCAGGGTCAGGACAAGCT	76	497	63.1		dHPLC

50

Table 2.1. Oligonucleotide primers used for amplifying PCR products subjected to HEX-SSCP and dHPLC analysis (Continued)

GENE	EXON	PRIMERS				PCR	ANN 1	ANN 2	ANALYSIS
		Forward (5'-3')	*Tm (°C)	Reverse (5'-3')	*Tm (°C)	FRAGMENT SIZE (bp)	(°C)	(°C)	
[e]HFE	4	TGGCAAGGGTAAACAGATCC	60	TACCTCCTCAGGCACTCCTC	64	396	65	60	HEX-SSCP
	1	GGACACTGGATCACCTAGTGTT	66	CTAGTTTCGATTTTTCCACCCC	64	355	58		dHPLC
	2	ACATGGTTAAGGCCTGTTGC	60	CAGCTGTTTCCTTCAAGATGCA	64	373	54		dHPLC
	3	AAATAGGGACCTATTCCTTTGGT	64	GTGCCCTGCAACCTCCTCCA	66	393	58		dHPLC
	4	TGGCAAGGGTAAACAGATCC	60	CTCAGGCACTCCTCTCAACC	64	390	58		dHPLC
	5	GTATGTGACTGATGAGAGCCA	62	CAGAGGTACTAAGAGACTTC	58	310	54		dHPLC
	6	TAGTGCCCAGGTCTAAATTG	58	TGAGTCTCTAGTTTTGTCTCC	60	201	58		dHPLC
[f]CARD15/	4	ACCTTCAGATCACAGCAGCC	62	GCTCCCCCATACCTGAAC	58	494	65	60	HEX-SSCP
NOD2	11	GAATGTCAGAATCAGAAGGG	58	CTCACCATTGTATCTTCTTTTC	56	230	60	55	HEX-SSCP

*Tm = 2(nA+nT) + 4(nG+nC), Abbreviations: Tm – melting temperature, Ann – annealing temperature, UTR, untranslated region
References: [a]Kotze *et al* 2001, [b]This study, [c]Njajou *et al* 2001, Primer sequences supplied by [d]Dr A Merryweather-Clarke and [e]Prof C Camaschella, [f]Lesage *et al* 2002

Table 2.2. Oligonucleotide primers used for amplifying PCR products subjected to RFLP analysis

| GENE | EXON/INTRON (Variant) | PRIMERS | | | | PCR | ANN 1 | ANN 2 | #ENZYME |
		Forward (5'-3')	*Tm (°C)	Reverse (5'-3')	*Tm (°C)	FRAGMENT SIZE (bp)	(°C)	(°C)	
[a]SLC11A1	5'-UTR [5'-(GT)n]	GGGGTCTTGGAACTCCAGAT	62	TACCCATGACCACACCC	54	116	65		[1]Rsal
[b]HFE	2 (S65C)	TCTGTCTAATCATGAGTATT	48	CTTGCTGTGGTTGTGATTTTC	58	513	52		[2]Hinfl
	2 (IVS2+4T/C)	ACATGGTTAAGGCCTGTTG	56	TGCCACTAGAGTATAGGGGC	58	870	56		[1]Rsal
	4 (IVS4-44T/C)	TGTCTCTCCTGTAGCTTGTT	58	TTCTGTCTCCCAGGGTGCT	56	825	55		[3]Sau96l
	5 (IVS5-47G/A)	CCTGAACATCTGTGGTGTAG	60	TGGGACTACAGGCGTCTGC	62	1038	55		[4]Banl
[c]CARD15/NOD2	8 (G908R)	CCCAGCTCCTCCCTCTTC	60	AAGTCTGTAATGTAAAGCCAC	58	380	65	60	[5]Hhal

*Tm = 2(nA+nT) + 4(nG+nC), #All enzymes were incubated at 37°C, Abbreviations: Tm – melting temperature, Ann – annealing temperature, References: [a]Kotze *et al* 2001, [b]Primer sequences supplied by Dr J Pointon, [c]Lesage *et al* 2002

Restriction endonucleases recognition sites on sense strand:

[1]GT♦AC; Recommended buffer – NEBuffer 1 [10 mM Tris-HCl, 10 mM MgCl$_2$, 1 mM DTT (pH 7.0)] (New England Biolabs)

[2]G♦ANT C; Recommended buffer – Buffer B [6 mM Tris-HCl, 6 mM MgCl$_2$, 50 mM NaCl (pH 7.5)] (Promega)

[3]G♦GNC C; Recommended buffer – Buffer C [10 mM Tris-HCl, 10 mM MgCl$_2$, 50 mM NaCl (pH 7.9)] (Promega)

[4]G♦G(T/C)(A/G)C G; Recommended buffer - Buffer G [50 mM Tris-HCl, 5 mM MgCl$_2$ (pH 8.2)] (Promega)

[5]G CG♦C; Recommended buffer – Buffer C [10 mM Tris-HCl, 10 mM MgCl$_2$, 50 mM NaCl (pH 7.9)] (Promega)

RFLP sizes

SLC11A1 5'(GT)n repeat (most common alleles): Allele 1 (5511): 57 bp, 51 bp, 5 bp, Allele 2 (5510): 57 bp, 49 bp, 5 bp; Allele 3 (559): 57 bp, 47 bp, 5 bp, Allele 5 (4510): 55 bp, 49 bp, 5 bp

HFE S65C: S, 367bp, 72 bp, 68 bp, 6 bp; C, 435 bp, 72 bp, 6 bp

HFE IVS2+4T/C: T, 557 bp, 245 bp, 19 bp; C, 297 bp, 260 bp, 245 bp, 19 bp

HFE IVS4-44T/C: T, 523 bp, 195 bp, 108 bp; C, 305 bp, 218 bp, 195 bp, 108 bp

HFE IVS5-47G/A: G, 681 bp, 405 bp, 307 bp, 150 bp; A, 988 bp, 405 bp, 150 bp

CARD15/NOD2 G908R: G, 380 bp; R, 242 bp, 138 bp

2.2.2. DNA amplification

General PCR

The DNA samples were amplified on an Applied Biosystems PCR cycler (GeneAmp®PCR system 2700). PCR amplification of the various exons were performed in 25 µl reactions. Each reaction consisted of 0.2 pmol/µl of each primer, 0.5 U *Taq* polimerase enzyme (Roche Diagnostics), 0.2 mM of each dNTP (dGTP, dCTP, dATP, dTTP) (Promega), 1 X buffer with $MgCl_2$ (Roche Diagnostics) and 20 ng/µl DNA.

The PCR conditions for the exons were as follows: an initial denaturation step for 2 min at 95°C, followed by 10 cycles of denaturation at 95°C for 30 seconds, annealing as specified for each exon in Tables 2.1 and 2.2 (Ann 1) for 45 seconds, and an extension at 72°C for 30 seconds. This was followed by 30 cycles of the same conditions but with an annealing temperature as specified in Tables 2.1 and 2.2 (Ann 2), followed by an extension step at 72°C for 10 minutes and a final step of 1 minute at 25°C. The exons with only one annealing temperature (Ann 1) specified were amplified for 35 cycles.

2.3. AGAROSE GEL ELECTROPHORESIS

Successful amplification of PCR products was tested on a 2% (w/v) horizontal agarose gel [consisting of 2g agarose in 100 ml 1 X TBE (90 mM Tris-HCl (pH 8.0), 90 mM boric acid and 1 mM EDTA)]. The PCR product (5 µl) was mixed with an equal volume of ficoll-orange G loading buffer [consisting of 2 mM orange G, 2 mM ficoll and 9 mM EDTA (pH 7.0)] and loaded on the gel. The products were resolved at 150 V for 30 minutes in 1 X TBE buffer solution and 0.01% (v/v) of ethidium bromide

for staining purposes. The DNA was visualised by ultraviolet light transillumination assisted by a UVP Image Store 5000 Ultra-Violet transilluminator system and photographed with the Sony Video Graphic Printer (UP 860 LE) system.

2.4. HETERODUPLEX SINGLE-STRAND CONFORMATION POLYMORPHISM (HEX-SSCP) ANALYSIS

HEX-SSCP analysis (Kotze *et al* 1995) was performed on a 30 cm vertical gel apparatus. These fragments were electrophoresed on a 10% polyacrylamide (PAA) gel supplemented with urea (gel consisting of 7.5% urea (w/v), 1.5 X TBE (135 mM Tris-Cl (pH 8.0), 135 mM boric acid and 2 mM EDTA), 10% PAA (w/v) (1%C of a 40% stock [99 acrylamide (AA):1 bisacrylamide (BAA)], 0.1% APS (w/v) and 0.1% (v/v) TEMED.

A 0.75 mm gel was cast and once polymerised, prepared for loading of the PCR products. Gels were placed in a Hoefer electrophoresis tank filled with fresh 0.5 X TBE buffer (45 mM Tris-Cl (pH 8.0), 45 mM boric acid and 0.9 mM EDTA). The upper buffer chamber was filled with 1.5 X TBE buffer. An equal volume of bromophenol blue loading buffer [consisting of 95% formamide (de-ionised) (v/v), 20 mM EDTA (disodiumsalt), 0.05% xylene cyanol (w/v) and 0.05% bromophenol blue (w/v)] was added to the PCR products, heat denatured for 5 minutes at 95°C and then kept on ice. Approximately 15 µl of the denatured PCR product was loaded on the gel. Electrophoresis of the gels was performed at 300 V at 4°C for 20 hours.

Following electrophoresis, the gels were dismantled and the DNA stained in 0.01% (v/v) ethidium bromide for 10 min and destained in ddH_2O for 3 min. The DNA

fragments were visualised by ultraviolet light transillumination assisted by the UVP Image Store 5000 Ultra-Violet transilluminator system and photographed with the Sony Video Graphic Printer (UP 860 LE) system. Automated DNA sequencing was performed on PCR products demonstrating HEX-SSCP mobility shifts (see section 2.8. for automated sequencing protocol).

2.5. DENATURING HIGH PERFORMANCE LIQUID CHROMATOGRAPHY (dHPLC) ANALYSIS

PCR amplification was performed, in a skirted 96-well PCR plate (Southern Cross Biotechnology) as previously described [see section 2.3.2. (i)] and an additional 0.05 U proofreading enzyme, *Pwo* DNA Polymerase (Roche Diagnostics), was added to the PCR reaction prior to amplification. Following PCR amplification, heteroduplex formation was induced by heat denaturation at 98 ^{o}C for 15 minutes followed by slow cooling (2°C/minute) and analysed by denaturing high-performance liquid chromatography (dHPLC) on the WAVE DNA Fragment Analysis System (Transgenomic). Five microliters of the heteroduplexed PCR products were automatically loaded on the column and were eluted on a linear acetonitrile (Sigma-Aldrich) gradient in a 0.1 M triethylamine acetate buffer (TEAA) (pH 7.0) with a constant flow rate of 0.9 ml/min. The start and end points of the gradient were adjusted according to the size of the PCR products. The denaturing run temperatures, as predicted by the Wavemaker 4.1.44 program, are listed in Table 2.3. Data were acquired using an ultraviolet detector at 260 nm and heterozygous profiles were identified by visual inspection of the chromatograms on the basis of the appearance of superimposed multiple eluting peaks, followed by sequencing analysis.

56

Table 2.3. Experimental run temperatures as used for dHPLC analysis

GENE	EXON	TEMPERATURES		
		*TEMP 1	*TEMP 2	*TEMP 3
SLC11A3	1	60 (shift-2.0)	62 (shift 0.0)	66 (shift+3.0)
	2	57 (shift-2.0)	59 (shift-0.5)	62 (shift+2.5)
	3	57 (shift-1.5)	59 (shift-0.5)	61 (shift+1.5)
	4	59 (shift-0.5)	61 (shift+2.0)	
	5	57 (shift-1.0)	59 (shift+1.0)	
	6	57 (shift-1.0)	60 (shift+1.5)	
	7i	59 (shift-1.0)	61 (shift+0.5)	
	7ii	58 (shift-0.5)	60 (shift+1.0)	
	8i	56 (shift-1.0)	59 (shift+1.0)	
	8ii	53 (shift+1.0)	56 (shift-1.0)	
DCYTB	1	65 (shift-0.5)		
	2	59 (shift+0.5)		
	3	51 (shift-0.5)	55 (shift+1.5)	58 (shift+2.5)
	4i	59 (shift-1.0)	61 (shift+1.0)	
	4ii	59 (shift+2.0)		
HAMP	1	64 (shift-0.5)		
	2+3	60 (shift-0.5)	63 (shift+1.0)	
HFE	1	59 (shift-0.5)	62 (shift+1.0)	66 (shift+3.0)
	2	61 (shift 0.0)		
	3	61 (shift-0.5)	63 (shift+0.5)	
	4	63 (shift+0.5)		
	5	58 (shift-0.5)	60 (shift+0.5)	
	6	59 (shift-1.0)	62 (shift+1.0)	

*Tm > oven temperature = negative shift (0.5 min/°C), Tm < oven temp = positive shift (0.5 min/°C)

57

2.6. RESTRICTION FRAGMENT LENGTH POLYMORPHISM (RFLP) ANALYSIS

PCR products subjected to RFLP analysis (as indicated in Table 2.2) were digested overnight in 20 μl reactions, consisting of 10 μl PCR product, 1 U of the relevant enzyme and 1 X buffer, in a 37°C water bath. Following digestion, an equal volume of bromophenol blue loading buffer was added to the PCR products.

Electrophoresis of the digested PCR products of *SLC11A1* and *CARD15/NOD2* was performed on a 9 cm vertical (Hoefer system) PAA gel [consisting of 12% PAA (w/v) [5%C of a 40% stock (95 acrylamide (AA):5 bisacrylamide (BAA)], 1 X TBE, 0.1% APS (w/v) and 0.1% TEMED (v/v)]. The gel was cast and once polymerised, placed in a Hoefer electrophoresis tank filled with fresh 1 X TBE buffer. Approximately 15 μl of the digested PCR product with loading buffer was loaded on the gel and electrophoresis performed at 250 V for 30 minutes at room temperature. Staining and visualization of the gels were performed as for HEX-SSCP analysis.

Electrophoresis of the digested PCR products of *HFE* was performed on a 3% (w/v) horizontal agarose gel [consisting of 3g agarose in 100 ml 1 X TBE (90 mM Tris-HCl (pH 8.0), 90 mM boric acid and 1 mM EDTA). Loading of the digested PCR product, electrophoresis and visualisation are similar as described in section 2.4.

2.7. CONSTRUCTION OF LUCIFERASE REPORTER GENE CONSTRUCTS AND TRANSFECTION

2.7.1. Cloning and transformation

DNA samples with the various alleles of the *SLC11A1* promoter region were PCR amplified using primers 5'- agt<u>gctagc</u>ctgcagtgccttcctctg-3' (Forward) and 5'-

tgaagatctgccgacttcaggtactct-3' (Reverse) (Searle and Blackwell 1999). The primers used were designed to incorporate *NheI* (restriction endonuclease recognition site: G♦CTAG C) and *BglII* (restriction endonuclease recognition site: A♦GATC T) restriction sites, generating a PCR product size of 584 bp (Searle and Blackwell 1999). Following amplification, the PCR products were purified by using a PCR purification kit (Amersham Biosciences) and dissolved in 10 μl dH$_2$O. Briefly, the GFX column (MicroSpin™ columns pre-packed with a glass fiber matrix) was placed in a collection tube (2 ml capless microcentrifuge tube) and 500 μl capture buffer (buffered solution containing acetate and chaotrope) added to the column. The PCR product (100 μl) was added to the capture buffer and mixed thoroughly by pipetting the sample. The sample was then centrifuged at 10000 x g for 30 seconds. Following centrifugation, the flow-through was discarded and 500 μl wash buffer (buffer containing 10 mM Tris-HCl (pH 8.0), 1 mM EDTA, 80% EtOH) was added to the column. The sample was, again, centrifuged at 10000 x g for 30 seconds and the flow-through discarded. The GFX column was transferred to a 1.5 ml microcentrifuge tube and 10 μl ddH$_2$O added to the column. The sample was incubated at room temperature for 1 minute and centrifuged at 10000 x g for 1 minute.

The purified PCR products and luciferase-reporter vector, the pGL2-Basic vector (Promega) (see Figure 2.1), were digested overnight in 20 μl reactions with *NheI* (New England Biolabs) and *BglII* (Roche Diagnostics) enzymes in a 37°C water bath. The reaction consisted of 1 U of each enzyme, 1 X buffer M [10 mM Tris-HCl (pH 7.5), 10 mM CH$_3$COOK, 50 mM NaCl, 1 mM DTE (Roche Diagnostics)] and 10 μl PCR product. A similar reaction was performed for the pGL2-Basic vector using 500 ng of the vector. After digestion, the pGL2-Basic vector was enzymatically treated by

adding 0.3 U/µl SAP (Roche Diagnostics) (storage buffer: 25 mM Tric-HCl (pH 7.6), 1

mM MgCl2, 50% glycerol) to the reaction and incubated for 15 minutes at 37°C,

followed by incubation at 65°C for 15 minutes to inactivate the enzyme. The vector

was purified by using a PCR purification kit (Amersham Biosciences) and dissolved

in 10 µl dH$_2$O.
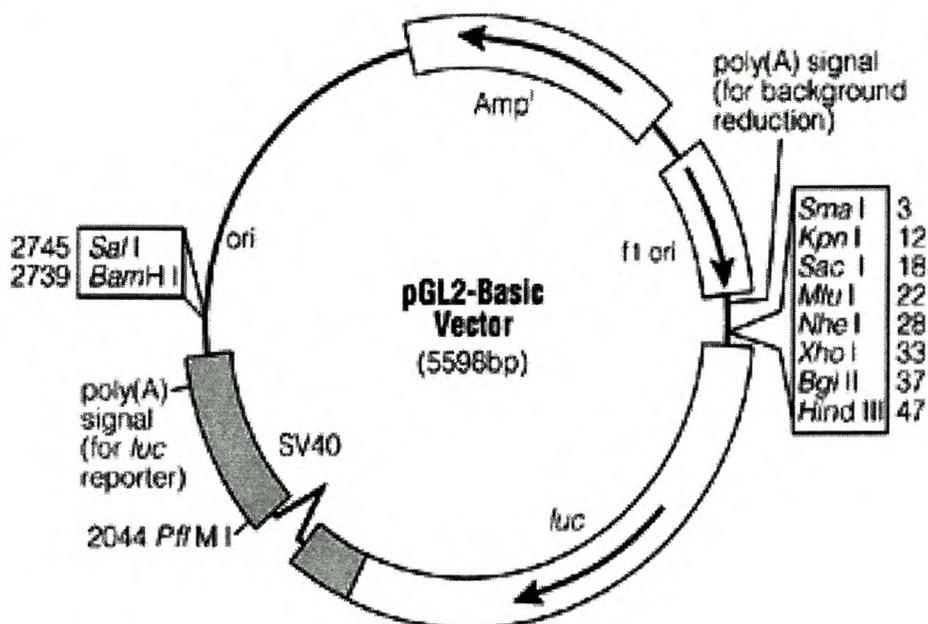
Figure 2.1.

Legend to Figure 2.1.
Vector circle map of pGL2-Basic vector [storage buffer: 10 mM Tris-HCl (pH 7.4), 1 mM EDTA (Promega)], *luc*, cDNA coding firefly luciferase, SV40, DNA derived from SV40 virus, poly(A) signal, mRNA polyadenylation signal, ori, plasmid replication origin in *Epicurian coli*, Amp^r, gene for ampicillin resistance, f1 ori, origin of replication derived from filamentous phage, arrow within *luc*, indicates direction of transcription of the luciferase gene, arrow within f1 ori, indicates direction of ssDNA strand synthesis.

The digested PCR products (10 µl) was mixed with an equal volume of ficoll-orange

G loading buffer and loaded on a 1% (w/v) agarose gel. The products were resolved

at 80 V for 30 minutes in a 1 X TBE buffer and 0.01% (v/v) ethidium bromide solution

for staining purposes. The DNA fragment was visualised by ultraviolet light transillumination and excised from the gel with a clean, sharp scalpel.

The fragment was purified by using a GFX gel band purification kit (Amersham Biosciences). Briefly, the gel slice was weighed in a 1.5 ml microcentrifuge tube and 10 µl of capture buffer was added for each 10 mg of gel slice. The amount of gel slice should not exceed 300 mg for one GFX column. The mixture was incubated at 60 °C for 10 minutes (or until the slice has completely dissolved), with vortexing every 2 to 3 minutes during the incubation. Following incubation, the sample was centrifuged briefly to collect the sample at the bottom of the tube. To bind the DNA, the sample was applied to the GFX column, placed in a 2 ml collection tube, incubated for 1 minute at room temperature and centrifuged at 10000 x g for 30 seconds. The flow-through was discarded and 500 µl wash buffer added to the column, followed by centrifugation at 10000 x g for 30 seconds. The flow-through was discarded and the column centrifuged at 10000 x g for 1 minute to remove residual EtOH. The GFX column was placed into a clean 1.5 ml microcentrifuge tube and the DNA eluted by adding 10 µl ddH2O to the center of the membrane. The column was incubated at room temperature for 1 minute followed by centrifugation at 10000 x g for 1 minute.

Ligation of the pGL2-Basic vector and PCR fragments, with a ratio of 1:1 (insert:plasmid), were performed by using a Rapid ligation kit (Roche Diagnostics). Briefly, the insert DNA and vector DNA (plasmid) was dissolved in 1 X DNA dilution buffer [20 mM Tris-HCl (pH 7.4), 1 mM EDTA, 5 mM DTE, 60 mM KCl, 50% glycerol (v/v)] in a total volume of 10 µl. T4 DNA ligation buffer (10 µl) was added and thoroughly mixed by pipetting. T4 DNA ligase [storage buffer: 20 mM Tris-HCl (pH

7.4), 1 mM EDTA, 5 mM DTE, 60 mM KCl, 50% glycerol (v/v)] was added to the mixture, thoroughly mixed and incubated for 5 minutes at room temperature. The competent cells were transformed by adding 2 µl of the ligation solution to 10 µl of XL1-Blue supercompetent cells (Stratagene's *Epicurian coli* competent cells) and then incubated on ice for 30 minutes. The cells were heat-shocked at 54°C for 60 seconds and incubated on ice for 2 minutes. Following incubation, 90 µl pre-warmed SOC medium (2% tryptone (w/v), 0.5% yeast extract (w/v), 8 mM NaCl, 2.5 mM KCl, 10 mM $MgSO_4$, 0.5 mM NaOH and 0.02 M glucose) was added to the cells and incubated for 60 min at 37°C with shaking at 250 rpm [Innova™ 2100 platform shaker (New Brunswick Scientific)]. The transformed solution was plated onto a LB agar plate [0.17 M NaCl, 1% tryptone (w/v), 0.5% yeast extract (w/v), 2% agar (w/v)], supplemented with 50 µg/ml ampicillin (Roche Diagnostics), and incubated overnight at 37°C.

After overnight incubation, colonies were picked from the plate, spread onto a master plate (LB-agar plate with grid) and the needle placed into a 0.2 ml PCR tube containing reagents for PCR amplification, using plasmid primers adjacent to the insert (Forward – 5'tgtatcttatggtactgtaactg3', Reverse – 5'ctttatgtttttggcgtcttcca3'). The plates were then incubated overnight at 37°C. Successful PCR amplification was tested on a 1% agarose gel. Colonies showing amplification was picked from the master plate and inoculated into 50 ml LB medium [0.17 M NaCl, 1% tryptone (w/v), 0.5% yeast extract (w/v)] supplemented with 50 µg/ml ampicillin. This solution was incubated overnight at 37°C with shaking at 250 rpm [Innova™ 2100 platform shaker (New Brunswick Scientific)].

Following overnight incubation and shaking, plasmid extraction was performed by using an EndoFree plasmid maxi extraction kit (Qiagen). Briefly, the bacterial cells were harvested by centrifugation at 6000 x g for 30 minutes at 4°C. Culture volumes should not exceed 100 ml for high-copy plasmids and 250 ml for low-copy plasmids. The bacterial pellet was completely resuspended in 10 ml buffer P1 (50 mM Tris-Cl pH 8.0, 10 mM EDTA, 100 μg/ml RNase A) by vortexing. Subsequently, 10 ml buffer P2 [200 mM NaOH, 1% SDS (w/v)] was added to the solution and mixed by inversion, followed by incubation at room temperature for 5 minutes. A QIAfilter cartridge was prepared by screwing the cap onto the outlet nozzle of the QIAfilter Maxi cartridge and placed into a 50 ml Falcon tube. Following incubation of the lysate, 10 ml chilled buffer P3 (3 M CH$_3$COOK, pH 5.5) was added and mixed by inverting the tube. The lysate was immediately poured into the barrel of the QIAfilter cartridge and incubated at room temperature for 10 minutes. Following incubation, the lysate was filtered into a 50 ml Falcon tube by removing the cap from the outlet nozzle and the plunger gently inserted into the QIAfilter cartridge.

To the filtered lysate, 2.5 ml buffer ER (containing isopropanol and polyethylene glycol octylphenyl ether) was added and mixed by inverting the tube. The lysate was then incubated on ice for 30 minutes. The QIAGEN-tip 500 was equilibrated by applying 10 ml buffer QBT [750 mM NaCl, 50 mM MOPS (pH 7.0), 15% isopropanol, 0.15% Triton-X 100 (v/v)] and the column was allowed to empty by gravity flow. The filtered lysate was added to the QIAGEN-tip and allowed to enter the resin by gravity flow. The QIAGEN-tip was washed twice with 30 ml buffer QC [1 M NaCl, 50 mM MOPS (pH 7.0), 15% isopropanol (v/v)] and the DNA eluted by adding 15 ml buffer QN [1.6 M NaCl, 50 mM MOPS (pH 7.0), 15% isopropanol (v/v)]. The DNA was

precipitated by adding 10.5 ml room temperature isopropanol, mixed by inverting the tube and centrifuged at 5000 x g for 45 minutes at 4°C. The precipitated DNA was washed with 70% (v/v) endotoxin-free room temperature EtOH by centrifuging at 5000 x g for 60 minutes at 4°C. The supernatant was carefully decanted and the pellet left to air dry. The DNA was dissolved in 50 ml endotoxin-free buffer TE (10 mM Tris-Cl (pH 8.0), 1 mM EDTA) and the yield determined by spectrophotometry and/or agarose gel electrophoresis. Using the Endofree Plasmid Maxi protocol a yield of 300 to 500 μg for high-copy plasmids and 50 to 250 μg for low-copy plasmids are expected. The DNA was concentrated at a concentration of 1 μg/ml.

Proper orientation and insertion of the alleles were verified by automated sequencing analysis (see section 2.8. for automated sequencing protocol), using primers Forward – 5'tgtatcttatggtactgtaactg3', and Reverse - 5'ctttatgtttttggcgtcttcca3'.

2.7.2. Transfection

Two macrophage cell lines, THP-1 and U937, were used to investigate transcriptional activity of the various alleles of the *SLC11A1* promoter region. These cells were cultured at 37°C in an atmosphere of 5% CO_2 and 95% air in the flask, in RPMI 1640 medium (Sigma) supplemented with 10% foetal calf serum (v/v) (Sigma), 2 mM L-glutamine (Sigma), 100 U/ml Penicillin and Streptomycin (Sigma). Transient transfections were performed by using FUGENE 6 (Roche Diagnostics) as transfection reagent. The various cell lines were resuspended at $2X10^6$ cells in 2 ml fully supplemented medium with 2 μg of the reporter construct and 1 μg of a control β-galactosidase (Promega) construct. Viability of the cells were measured by staining with a 0.4% Trypan Blue solution (v/v) (Sigma) and cells counted using a

haemocytometer. Transfected cells were incubated at 37°C for 24 hours with or without the addition of interferon-γ [IFN-γ, 0.1 μg/μl (Sigma)], bacterial lipopolysaccharide [LPS, 10 ng/ml (Sigma)], interferon-γ (0.1 μg/μl) and bacterial lipopolysaccharide (10 ng/ml), and ferric ammonium citrate (FAC, 200 μg/ml), respectively.

Following incubation, the cells were harvested and lysed using a luciferase assay system (Promega). Briefly, the cells were collected in a 2 ml microcentrifuge tube and centrifuged at 200 x g for 5 minutes at room temperature. The growth medium was removed and the cells rinsed with 1 X PBS [137 mM NaCl, 2.7 mM KCl, 4.3 mM Na$_2$HPO$_4$, 1.4 mM KH$_2$PO$_4$ (pH 7.3)], followed by centrifugation at 200 x g for 5 minutes. The supernatant was removed and the cells resuspended in 1 X reporter lysis buffer [25 mM Tris-phosphate (pH 7.8), 2 mM DTT, 2 mM 1,2-diaminocyclohexane-N,N,N',N'-tetraacetic acid, 10% glycerol (v/v), 1% Triton X-100 (v/v)]. The cells were briefly centrifuged and the supernatant transferred to a new 2 ml microcentrifuge tube. Luciferase activity was measured using a luciferase assay reagent (Promega) where 20 μl of cell lysate was added to 100 μl assay lysis buffer and relative light units (RLU) read in a luminometer (Turner Designs Luminometer Model TD-20/20).

The β-galactosidase assay was performed on 100 μl cell lysate in 1650 μl buffer containing 20 mM Na$_2$HPO$_4$ (pH 7.8), 2 mM KCl, 0.2 mM Mg$_2$SO$_4$, 0.54 ml/l β-ME, 4 mg/ml ONPG and 285 mM Na$_2$CO$_3$. The cell suspension was incubated at 37°C for 30 minutes followed by spectophotometry (absorbance measured at 420 nm). Protein assays were performed with the protein assay ESL (exact, sensitive and low

interference) kit (Roche Diagnostics). Briefly, a series of standards were prepared to cover an appropriate concentration range by diluting the BSA protein standard in the same diluent as the unknown sample. Reagent A (alkaline copper tartrate solution) (100 µl) was added to 50 µl of sample in a 1.5 ml microcentrifuge tube and mixed. The tubes were incubated for at least 5 minutes at 15 to 25 °C. Reagent B (ascorbic acid, bathocuproine disulfonic acid solution) (1000 µl) was subsequently added, mixed briefly and the absorbance was measured at 485 nm after exactly 30 seconds. The concentration was calculated against a standard calibration curve. Relative light units of luciferase activity were normalized against the protein concentration and β-galactosidase readings for each construct within an experiment. These experiments were performed three times with triplicate transfections for each construct and set of conditions within each experiment.

2.8. AUTOMATED SEQUENCING

Automated DNA sequencing was performed using the same primers as for HEX-SSCP, dHPLC and RFLP analysis as indicated in Tables 2.1 and 2.2. The PCR products were purified by using a PCR purification kit (Amersham Biosciences). The purified product (3 µl) was added to a reaction mix consisting of 4 µl termination ready reaction mix [BigDye® Terminator v3.1 cycle sequencing kit (Applied Biosystems)] and 3.3 pmol primer. This was followed by cycle sequencing on a 2400 Perkin Elmer PCR cycler. The cycle program was as follows: an initial denaturation step at 96°C for 2 minutes, followed by 25 cycles of denaturation at 96 °C for 20 seconds, annealing at 55 °C for 20 seconds and an extension at 60 °C for 4 minutes.

Following cycle sequencing, the products were cleaned by spun-column chromatography using Centri-sep columns (Applied Biosystems). After spun-column chromatography, 2 X the volume $MgCl_2$ (20 µg/ml) and 60% EtOH (v/v) were added to the products. The solution was left on the bench for 10 to 15 minutes, followed by centrifugation for 15 to 20 minutes at 10000 x g. The supernatant was removed and the pellet left to air-dry. HiDi formamide (Applied Biosystems) was added to the dry pellet and the products were loaded on an ABI Prism 3100 genetic analyser (Applied Biosystems) automated sequencer. Following electrophoresis, analysis of electropherograms was performed by using sequencing analysis software supplied by Applied Biosystems.

Table 2.4. Alphabetic list of generally used chemicals/reagents and their suppliers

Chemical/Reagent	Supplier
AA	Merck
Agar	Sigma
agarose	Promega
APS	Promega
β-ME	BDH
BAA	Promega
Boric acid	Promega
Bromophenol blue	Sigma
Chloroform	Merck
dNTPs (dATP, dTTP, dCTP, dGTP)	Promega
EDTA	Sigma
EtOH	Merck
Ficoll	Sigma
Formamide (De-ionised)	Merck
Glucose	Merck
Isopropanol	Merck
KCl	BDH
$KHCO_3$	Merck
KH_2PO_4	Merck
Mg_2SO_4	Sigma
NaCl	Sigma

Table 2.4. Alphabetic list of generally used chemicals/reagents and their suppliers (Continued)

Chemical/Reagent	Supplier
Na_2CO_3	Sigma
Na_2HPO_4	BDH
NaOH	Merck
NH_4Cl	BDH
ONPG	Sigma
Orange G	Merck
SDS	Merck
TEAA	Transgenomics
TEMED	Promega
Tris-HCl	Promega
tryptone	Sigma
Urea	Promega
Xylene cyanol	Sigma
Yeast extract	Sigma

CHAPTER 3

METHODS, RESULTS AND DISCUSSION

3.1. Differential expression of allelic variants of the 5'-(GT)n repeat in the promoter region of the *SLC11A1* gene upon iron loading: Detection of an opposite allelic effect in the presence of the –237C→T polymorphism

3.1.1. ABSTRACT

Association of various autoimmune and infectious diseases with genetic variation in the solute carrier family 11 member 1 (*SLC11A1)* gene is in accordance with its role in iron metabolism and immune function. In this investigation *in vitro* studies were performed to determine whether allelic variants in the promoter region of the gene is affected by iron loading, thereby leading to differential expression of *SLC11A1*. Constructs containing five different *SLC11A1* 5'-(GT)n polymorphic alleles identified in the South African population (alleles 2, 3, 5, 8 and 9) and a C to T point mutation at nucleotide position – 237, both in the absence and presence of allele 3, were cloned into the pGL2-Basic luciferase-reporter vector and transfected into U937 and THP-1 cells. Addition of exogenous stimuli, including interferon-γ, bacterial lipopolysaccharide and ferric ammonium citrate, demonstrated significant differences in the ability of these alleles to regulate gene expression. Striking differences were obtained upon iron loading, with allele 3 showing opposite effects in the presence or absence of promoter polymorphism -237C→T. Our findings provide direct evidence that this promoter polymorphism is functional and support the hypothesis that iron dysregulation mediated by allelic effects of *SLC11A1* underlies disease susceptibility linked to infectious and autoimmune conditions.

71

3.1.2. INTRODUCTION

The solute carrier family 11 member 1 gene (*SLC11A1*) has been linked to various autoimmune and infectious diseases (reviewed by Blackwell *et al* 2003). The multiple pleiotropic effects of *SLC11A1* associated with macrophage activation (Blackwell 1996) are potentially important for induction and maintenance of autoimmune diseases and crucially important in resistance to intramacrophage pathogens. The effects on macrophage function include enhanced chemokine KC, tumour necrosis factor α, interleukin 1β, inducible nitric oxide synthase, and MHC class II expression.

Blackwell and co-workers (2000) recently speculated that regulation of iron by *SLC11A1* might underlie the allelic association with various autoimmune and infectious diseases. Conclusive evidence has been provided that the Z-DNA forming repeat in the promoter region of the human *SLC11A1* represents a functional polymorphism that correlates with infectious and autoimmune disease susceptibility (Searle and Blackwell 1999). The two most common variants designated alleles 2 [t(gt)$_5$ac(gt)$_5$ac(gt)$_{10}$ggcaga(g)$_6$] and 3 [t(gt)$_5$ac(gt)$_5$ac(gt)$_9$ggcaga(g)$_6$], showed a similar degree of enhancement of reporter gene expression in the presence of interferon-γ (IFN), while addition of bacterial lipopolysaccharide (LPS) caused significant reduction in expression driven by allele 2 and enhanced expression driven by allele 3. These results were in accordance with the hypothesis that chronic hyperactivation of macrophages associated with *SLC11A1* allele 3 is functionally linked to autoimmune disease susceptibility, while the poor level of allele 2 expression contributes to infectious disease susceptibility. The likelihood that allele 3 would conversely protect against infectious disease and allele 2 against autoimmune

disease, was supported by data from Kotze *et al* (2001), who compared the allele frequencies of the promoter variant between different age groups. *SLC11A1* alleles considered to be detrimental in relation to autoimmune disease susceptibility appear to be maintained in the population as a consequence of improved survival to reproductive age following infectious disease challenge, as previously suggested by Searle and Blackwell (1999). A single base-pair substitution has also been reported at position −236 in the promoter region of the *SLC11A1* gene (Lewis *et al* 1996) but to our knowledge, *in vitro* studies involving this polymorphism have thus far not been investigated. Sequencing data obtained in our laboratory consistently identified an additional G residue between position −167 and −169 (numbered according to the sequence published by Lewis *et al* 1996) and therefore this polymorphism occurs at position −237 (du Plessis 2000).

The involvement of *SLC11A1* in many clinically distinct autoimmune and infectious diseases may suggest a mechanistic continuum of the same disease process involving iron dysregulation. This notion was substantiated by reports on abnormalities of iron metabolism (Weber *et al* 1988, Nielsen *et al* 1994) in arthritic conditions shown to be associated with genetic variation in the *SLC11A1* gene (Shaw *et al* 1996, Sanjeevi *et al* 2000, Singal *et al* 2000, Yang *et al* 2000), which may be linked to defective iron supply for erythropoiesis (Cazzola *et al* 1996).

In an attempt to demonstrate a possible direct relation of *SLC11A1* alleles with cellular iron status, functional studies were performed in this study to test for differential

expression of allelic variants of the Z-DNA forming promoter polymorphism upon iron loading. The potential effect of the $-237C{\rightarrow}T$ polymorphism was also investigated.

3.1.3. MATERIALS AND METHODS

The study protocol was approved by the Ethics Review Committee of the University of Stellenbosch.

3.1.3.1. Methods

i) Plasmid construction

DNA samples of individuals harboring five different allelic variants (alleles 2, 3, 5, 8 and 9) (Table 3.1.1) of the 5'-(GT)n repeat polymorphism in the *SLC11A1* gene, and an individual with the -237C→T polymorphism occurring *in cis* with allele 3, were available for this study. These were used as templates to generate six different luciferase reporter gene constructs cloned into the pGL2-Basic vector (Promega), using oligonucleotide primers as previously described (Searle and Blackwell 1999). All the constructs constitute alleles of the *SLC11A1* Z-DNA forming repeat promoter polymorphism, with plasmid construct allele 3/-237T incorporating the T-allele at nucleotide position −237 in association with allele 3. Alleles 2, 3 and 5 have been described previously and are designated as published (Blackwell *et al* 1995, Graham *et al* 2000), with numbering of new variants detected in this study following consecutively. The alleles were cloned into the vector using standard procedures and the DNA purified by using an endotoxin-free plasmid DNA extraction kit (EndoFree Plasmid Maxi Kit - Qiagen). Proper orientation and insertion of the various alleles were verified by sequencing analysis.

Table 3.1.1. Variants incorporated into the pGL2-Basic vector

ALLELE	VARIANTS	REFERENCE
Allele 2	t(gt)$_5$ac(gt)$_5$ac(gt)$_{10}$ggcaga(g)$_6$	Blackwell *et al* 1995
Allele 3	t(gt)$_5$ac(gt)$_5$ac(gt)$_9$ggcaga(g)$_6$	Blackwell *et al* 1995
Allele 5	t(gt)$_4$ac(gt)$_5$ac(gt)$_{10}$ggcaga(g)$_6$	Graham *et al* 2000
Allele 8	t(gt)$_5$ac(gt)$_5$ac(gt)$_6$ggcaga(g)$_6$	M. Zaahl, unpublished data
Allele 9	t(gt)$_5$ac(gt)$_5$ac(gt)$_8$ggcaga(g)$_6$	M. Zaahl, unpublished data
Allele 3/-237T[#]	t(gt)$_5$ac(gt)$_5$ac(gt)$_9$ggcaga(g)$_6$/-237T[#]	Blackwell *et al* 1995, Lewis *et al* 1996

[#]Previously designated –236C→T (Lewis *et al* 1996)

ii) Transfection and luciferase activity

Myeloid lineage cell lines U937 and THP-1 were maintained in RPMI-1640 (Sigma) supplemented with 10% fetal calf serum (Sigma) and 2 mmol/l L-glutamine (Gibco). These cells were cultured at 37°C in an atmosphere of 5% CO_2 and 95% air in the flask. Transient transfections were performed by using FUGENE 6 (Roche Diagnostics) as transfection reagent. The various cell lines were resuspended at 2X10^6 cells in 2 ml fully supplemented medium with 2 μg of the reporter construct and 1 μg of a control β-galactosidase (Promega) construct. Viability of the cells were measured by staining with a 0.4% Trypan Blue solution (Sigma) and cells counted using a haemocytometer. Transfected cells were incubated at 37°C for 24 hours with or without the addition of interferon-γ (IFN-γ, 0.1 μg/μl) (Sigma), bacterial lipopolysaccharide (LPS, 10 ng/ml)

(Sigma), interferon-γ (0.1 μg/μl) and bacterial lipopolysaccharide (10 ng/ml), and ferric ammonium citrate (FAC, 200 μg/ml) (Sigma), respectively.

Following incubation, the cells were harvested and lysed using a reporter lysis buffer (Promega). Luciferase activity was measured using a luciferase assay reagent (Promega) where 20 μl of cell lysate was added to 100 μl of assay buffer and relative light units (RLU) read in a luminometer. The β-galactosidase assay was performed on 100 μl of cell lysate in 1650 μl of buffer containing 20 mM Na_2HPO_4 (pH 7.8) (BDH), 2 mM KCl (BDH), 0.2 mM Mg_2SO_4 (Sigma), 0.54 ml/l β-mercaptoethanol (BDH), 4 mg/ml ONPG (o-nitrophenyl-β-δ galactopyranoside) (Sigma) and 285 mM Na_2CO_3 (Sigma). The cell suspension was incubated at 37°C for 30 minutes followed by spectrophotometry (absorbance measured at 420 nm). Protein assays were performed with the protein assay ESL (exact, sensitive and low interference) kit (Roche Diagnostics). Absorbance was measured at 485 nm and the concentration calculated against a standard calibration curve. Relative light units of luciferase activity were normalised against the protein concentration and β-galactosidase readings for each construct within an experiment. These experiments were performed three times with triplicate transfections for each construct and set of conditions within each experiment.

3.1.3.2. Statistical analysis

Statistically significant differences between constructs and treatments were tested by the two tailed unpaired Student's t test. A probability value smaller than 0.05 was regarded as statistically significant.

3.1.4. RESULTS

Various transcription factor binding sites and response elements have been identified in the promoter region of the *SLC11A1* gene (Blackwell *et al* 1995) and incorrect functioning of these elements and enhancers can result in altered gene expression. In order to investigate the effect of the various Z-DNA forming repeats and the -237C→T polymorphism on promoter activity, different constructs (Table 3.1.1) were generated in the pGL2-Basic vector. The ability of various alleles of the *SLC11A1* Z-DNA forming promoter repeat polymorphism to drive reporter gene expression was measured with and without the addition of exogeneous stimuli. The mean and standard error for triplicate points obtained over three independent transfection experiments is shown in Figure 3.1.1.

In the absence of exogeneous stimuli: Expression was observed for all of the constructs. Higher levels of reporter gene expression were observed for allele 3 compared to the remaining alleles. Statistically significant differences were observed for allele 3 compared to allele 2 ($p<0.05$). Alleles 9 ($p<0.05$) and 3/–237T ($p<0.05$) also differed significantly compared with allele 2.

Addition of interferon-γ (IFN-γ): Enhanced luciferase expression with the addition of IFN-γ as an exogeneous stimulus caused an increase in gene expression for all six constructs. No statistically significant differences were observed between constructs.

77

However, increases in reporter gene activity within the constructs were significant for alleles 2 ($p<0.05$) and 3/–237T ($p<0.04$) compared with no stimuli.
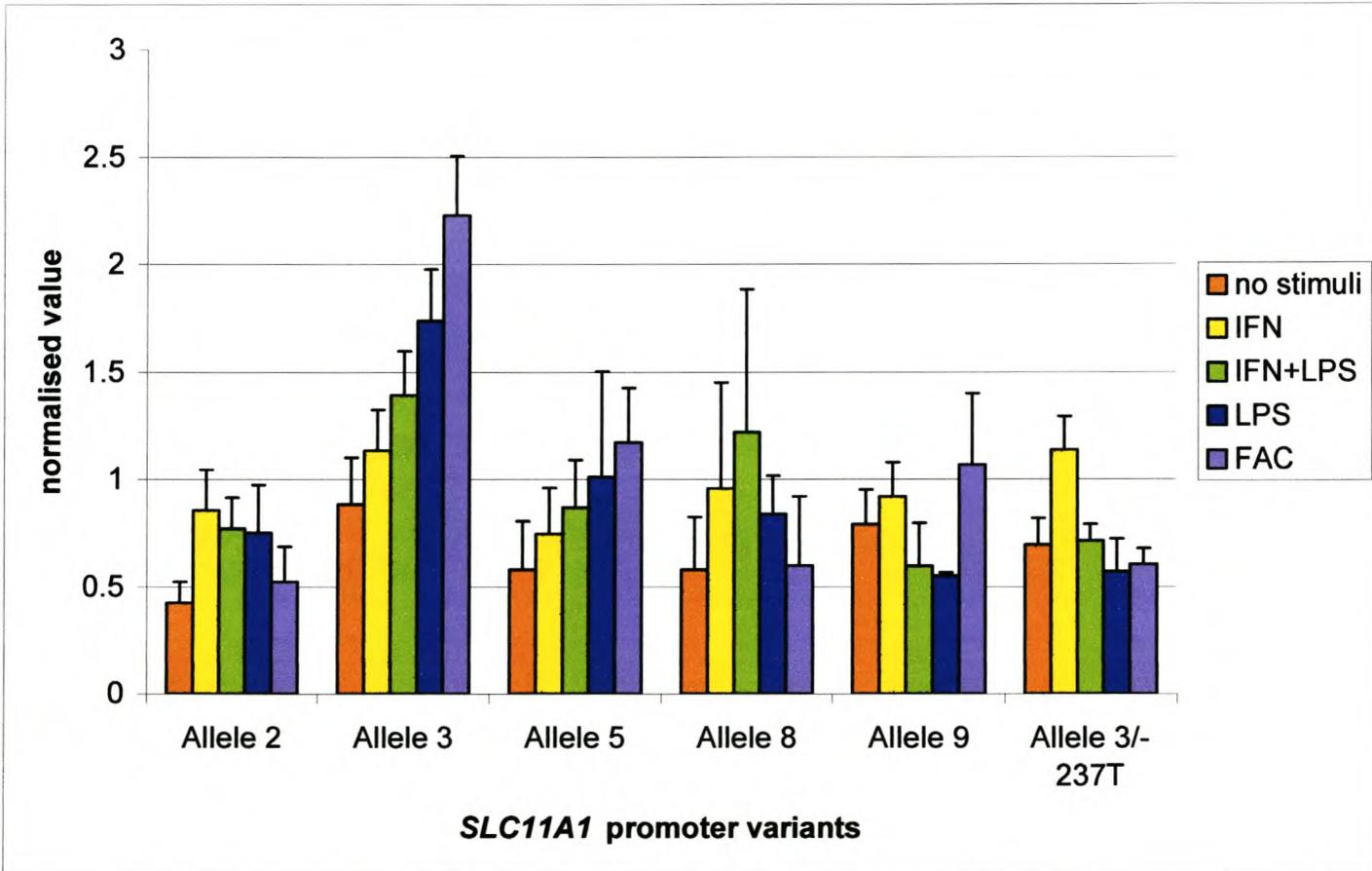
Figure 3.1.1.

Legend to Figure 3.1.1.

SLC11A1 promoter variants transfected into U937 cells. The effect of ferric ammonium citrate (FAC) on levels of luciferase reporter gene activity driven by three polymorphic alleles for *SLC11A1* detected, in relation to interferon-γ (IFN) and bacterial lipopolysaccharide (LPS) enhanced expression, is illustrated. The mean (SD) normalised values obtained over three independent experiments, including triplicate transfections for each experiment, are graphed for *SLC11A1* constructs. Similar results were obtained for these variants transfected into THP-1 cells.

Addition of interferon-γ (IFN-γ) and bacterial lipopolysaccharide (LPS): The addition of LPS as a second exogeneous stimulus with IFN-γ results in further enhancement of gene expression for alleles 3, 5 and 8, and causes a reduction in alleles 2, 9 and 3/–237T allele driven response, compared to the addition of IFN-γ alone. Compared to allele 3 activity, statistically significant differences were observed for alleles 2 ($p < 0.03$), 5 ($p < 0.05$), 9 ($p < 0.02$) and 3/–237T ($p < 0.02$). Statistically significant differences were also observed for alleles 2 (compared to no stimuli, $p < 0.05$) and 3/–237T (compared to the addition of IFN-γ alone, $p < 0.04$) within the constructs.

Addition of bacterial lipopolysaccharide (LPS): In the presence of LPS alone, an increase in luciferase activity was observed for alleles 3 and 5, and a decrease for alleles 2, 8, 9 and 3/–237T, compared to the addition of IFN-γ and LPS. Compared with allele 3, statistically significant differences were observed for alleles 2 ($p < 0.04$), 8 ($p < 0.04$), 9 ($p < 0.05$) and 3/–237T ($p < 0.02$). Within the various constructs, significant differences were observed for alleles 3 (addition of IFN-γ and LPS compared to LPS alone, $p < 0.03$), 9 (addition of IFN-γ compared with LPS alone, $p < 0.04$) and 3/–237T (addition of IFN-γ compared with LPS alone, $p < 0.05$).

Addition of ferric ammonium citrate (FAC): The addition of FAC caused an increase in luciferase activity for alleles 2, 3, 5 and 9, and a decrease in activity for allele 3/–237T, compared to no stimuli. Activity observed for allele 8 with the addition of FAC was similar compared to no stimuli. Significant differences were observed for all constructs compared to allele 3 FAC activity with $p < 0.006$ for allele 2, $p < 0.02$ for allele 5, $p < 0.009$

79

for allele 8, p<0.02 for allele 9 and p<0.01 for allele 3/–237T. Allele 5 FAC activity was also statistically significant compared to alleles 2 (p<0.05) and 3/–237T (p<0.05). Significant differences within the various constructs were observed for alleles 3 (compared with no stimuli, p<0.05; compared with IFN-γ activity, p<0.01), 5 (compared with no stimuli, p<0.05), 9 (compared with LPS activity, p<0.05) and 3/–237T (compared with IFN-γ activity, p<0.02).

Compared to no stimulus, an increase for all the treatments were observed for the various alleles except for alleles 9, with the addition of IFN-γ and LPS, and LPS alone, and 3/-237T with the addition of LPS and FAC respectively. Similar expression was observed for alleles 8, with the addition of FAC, and 3/-237T, with the addition of IFN-γ and LPS, compared to no stimulus. Similar results were observed when transfecting the various plasmid constructs into THP-1 cells. The results are consistent with findings previously described for alleles 2 and 3 (Searle and Blackwell 1999). However, these authors have not determined the effect of FAC on luciferase gene expression.

3.1.5. DISCUSSION

The various statistical comparisons performed are shown in Table 3.1.2 and 3.1.3. The most striking observation made in this study was the opposite effects mediated by allele 3 of the *SLC11A1* 5'-(GT)n repeat in the presence or absence of the T-allele of the polymorphism at nucleotide position -237. High expression associated with allele 3 of the *SLC11A1* gene in the absence of this sequence variant is in accordance with its role in autoimmune disease susceptibility.

Table 3.1.2. Statistical comparisons following addition of different stimuli to the various luciferase reporter constructs investigated

Associations within constructs	*p-value					
Stimuli	Allele 2	Allele 3	Allele 5	Allele 8	Allele 9	Allele 3/-237T
No stimuli vs IFN-γ	0.05	0.22	0.31	0.27	0.30	0.04
No stimuli vs IFN-γ+LPS	0.05	0.08	0.20	0.22	0.25	0.45
No stimuli vs LPS	0.19	0.10	0.30	0.39	0.13	0.19
No stimuli vs FAC	0.32	0.05	0.05	0.48	0.26	0.28
IFN-γ vs IFN-γ+LPS	0.36	0.20	0.35	0.38	0.14	0.04
IFN-γ vs LPS	0.27	0.14	0.36	0.31	0.04	0.05
IFN-γ vs FAC	0.12	0.01	0.13	0.29	0.36	0.02
IFN-γ+LPS vs LPS	0.37	0.03	0.45	0.25	0.07	0.20
IFN-γ+LPS vs FAC	0.16	0.31	0.21	0.23	0.16	0.17
FAC vs LPS	0.34	0.10	0.37	0.46	0.05	0.40

*Statistical analysis performed using the Student's unpaired *t* test; statistically significant associations are indicated in red

Table 3.1.3. Statistical comparisons between the luciferase reporter constructs following addition of different stimuli

Associations between constructs	*p-value				
Treatment	**Allele 2 vs Allele 3**	**Allele 2 vs Allele 5**	**Allele 2 vs Allele 8**	**Allele 2 vs Allele 9**	**Allele 2 vs Allele 3/-237T**
No stimuli	0.05	0.28	0.29	0.05	0.05
IFN-γ	0.17	0.35	0.43	0.40	0.15
IFN-γ+LPS	0.03	0.36	0.28	0.25	0.37
LPS	0.04	0.35	0.39	0.26	0.28
FAC	0.006	0.05	0.42	0.12	0.05
	Allele 3 vs Allele 5	**Allele 3 vs Allele 8**	**Allele 3 vs Allele 9**	**Allele 3 vs Allele 3/-237T**	
No stimuli	0.19	0.20	0.37	0.25	
IFN-γ	0.12	0.37	0.21	0.49	
IFN-γ+LPS	0.05	0.41	0.02	0.02	
LPS	0.20	0.04	0.05	0.02	
FAC	0.02	0.009	0.02	0.01	
	Allele 5 vs Allele 8	**Allele 5 vs Allele 9**	**Allele 5 vs Allele 3/-237T**		
No stimuli	0.49	0.24	0.33		
IFN-γ	0.36	0.27	0.10		
IFN-γ+LPS	0.33	0.20	0.28		
LPS	0.39	0.26	0.27		
FAC	0.11	0.40	0.08		
	Allele 8 vs Allele 9	**Allele 8 vs Allele 3/-237T**			
No stimuli	0.26	0.35			
IFN-γ	0.47	0.38			
IFN-γ+LPS	0.23	0.27			
LPS	0.17	0.19			
FAC	0.18	0.49			
	Allele 9 vs Allele 3/-237T				
No stimuli	0.33				
IFN-γ	0.19				
IFN-γ+LPS	0.31				
LPS	0.46				
FAC	0.16				

*Statistical analysis performed using the Student's unpaired *t* test; statistically significant associations are indicated in red

However, in the presence of the −237C→T variant low gene expression of allele 3 is obtained comparable to that of allele 2, known to have a protective effect against autoimmunity while conversely predisposing to infectious diseases. Compared to allele 2 gene expression, even lower expression was observed for allele 3 in association with the −237C→T polymorphism for treatments with IFN-γ and LPS, and LPS alone. Several IFN-γ and LPS-related response elements are in close proximity of the −237C→T polymorphism, suggesting that this polymorphism directly influence these response elements. Further studies are warranted to assess the *in vivo* significance of this allelic effect in relation to disease association, since occurrence of linkage disequilibrium may lead to failure to detect a true disease association with the *SLC11A1* gene.

Differential expression was observed without the addition of an exogeneous stimuli indicating that the promoter repeat polymorphism acts as an enhancer element, as previously suggested (Searle and Blackwell 1999). Enhanced expression for all the constructs studied were observed with the addition of IFN-γ and is consistent with the occurrence of multiple interferon response elements both 5' and 3' of the repeat polymorphism, therefore being important in up regulating gene expression. The addition of LPS caused differential expression of the gene with an up regulation in gene expression for alleles 3 and 5 and down regulation of alleles 2, 8, 9 and 3 in association with the −237C→T polymorphism (3/-237T), compared to the addition of IFN-γ and LPS, suggesting that the LPS related response elements (NFκB, AP-1 and NF-IL6) are differentially affected by the various alleles. *Slc11a1* expression was also shown to be

regulated by GMCSF (Brown *et al* 1995), IFN-γ (Brown *et al* 1995, Govoni *et al* 1996, Brown *et al* 1997) and LPS (Govoni *et al* 1996) in mice.

The addition of FAC caused an up-regulation in gene expression for alleles 2, 3, 5 and 9 and a down regulation in gene expression for allele 3 in association with the $-237C{\rightarrow}T$ polymorphism, compared with no stimulus. This is the first study demonstrating a direct influence of iron loading on the naturally occurring alleles of the *SLC11A1* 5'-(GT)n promoter repeat polymorphism. It thus appears that the *SLC11A1* promoter is sensitive to changes in iron levels thereby affecting expression of the *SLC11A1* gene, but it remains unclear whether variation in the promoter region causes the imbalance in iron locally by altering iron recycling/homeostasis in macrophages, or whether the imbalance of iron in the environment influences *SLC11A1* expression and therefore any associated pleiotropic effects. No conventional iron responsive elements (IRE) were thus far identified in the promoter region of *SLC11A1*. The effect of the $-237C{\rightarrow}T$ polymorphism on allele 3 expression raises the possibility that this sequence change affects a putative transcription factor-binding site. Unrecognised IREs could also be in close proximity of the $-237C{\rightarrow}T$ polymorphism, thereby influencing allele 3 expression and mRNA translation or stability.

Although the association with disease susceptibility may be related to the many pleiotropic effects of this gene on macrophage function (Searle and Blackwell 1999, Marquet *et al* 1999), our data provide additional evidence that its involvement in iron transport may contribute directly to the disease phenotype, possibly via the effect of iron

on pathogen survival, intracellular signalling pathways and/or mRNA stability (Searle and Blackwell 1999). Iron has to be immediately available for mounting a successful immunological defence, but at the same time be effectively unavailable to the invading pathogen (Weiss 2002). It therefore seems likely that the continuous battle for iron between the host and invading pathogens known to result in cytokine-mediated responses, could play an important role in the context of infectious and autoimmune disease susceptibility linked to the *SLC11A1* gene. Our findings provide direct support for the hypothesis that iron dysregulation may underlie the allelic association of the functional *SLC11A1* promoter polymorphism with infectious and autoimmune disease susceptibility.

3.1.6. ACKNOWLEDGEMENTS

3.2. Association of functional polymorphisms of *SLC11A1* with risk of oesophageal cancer in the Coloured population of South Africa

3.2.1. ABSTRACT

Several environmental factors, including smoking and alcohol, dietary factors, mycotoxins, nitrosamines, infection and chronic inflammation, have been implicated in the aetiology of oesophageal cancer (OC). Excess iron also leads to enhanced tumour growth. The purpose of this study was to assess the likelihood that variation in the *SLC11A1* gene contributes to OC susceptibility, possibly due to its role in inflammation and iron metabolism. The regions of the gene containing potential functional polymorphisms, including the promoter region and exon 2, were investigated. The study cohort included 105 OC patients with squamous cell carcinoma (SCC) and 110 population-matched controls. A significantly decreased frequency of the $-237C{\rightarrow}T$ promoter polymorphism was observed in the patient group with OC compared with the population-matched control group ($p<0.002$, χ^2 with Yates' correction $= 7.87$). A statistically significant disease association was also observed with allele 3 of the 5'-(GT)n promoter polymorphism ($p<0.0006$, χ^2 with Yates' correction $= 10.16$), but only in the absence of the T-allele at nucleotide position -237 following allelic stratification. Several novel variants were identified in intron 1 (IVS1-28C$\rightarrow$T) and exon 2 (112G$\rightarrow$A, 148delGACCAGCCC, 157insGACCAGCCCAG). A novel intronic polymorphism, IVS1-28C$\rightarrow$T, was also significantly associated with OC ($p<0.05$, χ^2 with Yates' correction $= 2.52$). We demonstrate association of genetic variation in both the promoter region and intron 1 of the *SLC11A1* gene with OC susceptibility.

86

3.2.2. INTRODUCTION

Oesophageal cancer (OC) occurs at high frequencies in certain areas of the world and is more prevalent in males than in females (Blot 1994). Differences have also been observed between racial backgrounds, with squamous cell carcinoma (SCC) occurring more often in Blacks and adenomacarcinoma (ADC) being more prevalent in Whites.

Several risk factors contribute to the development of squamous cell carcinoma and of these factors, cigarette smoking and alcoholic beverage consumption is the major determinants (IARC 1986, 1988, Blot 1994). Other factors contributing to a lesser extent includes diet and nutrition, mycotoxins, nitrosamines, infectious agents, repeated oesophageal injury and chronic inflammation. (Selikoff *et al* 1979, Van Rensburg 1981, Norell *et al* 1983, Hille *et al* 1985, Steinmetz and Potter 1991, Blot 1994, Gustavisson *et al* 1993, Orlando 2002). Iron overload has previously been implicated as a possible risk factor in the development of OC (Mandishona *et al* 1998) and evidence for the involvement of iron in OC was illustrated in a rat model where OC developed as a consequence of iron supplementation (Goldstein *et al* 1998, Chen *et al* 1999, 2000).

The *SLC11A1* gene is associated with susceptibility to various autoimmune and infectious diseases (reviewed by Blackwell *et al* 2003) and has been shown, in mice, to function as a metal ion transporter that is regulated by and regulates cellular iron levels (Atkinson *et al* 1997, Atkinson and Barton 1998, Goswami *et al* 2001). A 5'-(GT)n repeat polymorphism in the promoter region of the *SLC11A1* gene was shown to affect levels of gene expression (Searle and Blackwell 1999). *In vitro* promoter studies of this repeat suggested direct contribution of alleles to autoimmune (allele 3) and infectious (allele 2)

87

disease susceptibility (Blackwell *et al* 2000, Searle and Blackwell 1999). The 5'-(GT)n repeat was also differentially expressed in the presence of ferric ammonium citrate and these findings further supports the hypothesis that iron dysregulation may underlie the allelic association of the functional *SLC11A1* promoter polymorphism with infectious and autoimmune disease susceptibility (chapter 3.1). A potentially functional polymorphism was also identified in exon 2 of the *SLC11A1* gene and comprises deletion of a nine-nucleotide repeat in the N-terminal proline/serine-rich putative SH3 binding domain (Barton *et al* 1994, White *et al* 1994, Blackwell *et al* 2000).

Due to the role of *SLC11A1* in iron regulation and inflammatory diseases, we examined functional polymorphisms in the promoter region and exon 2 of the *SLC11A1* gene in the Coloured South African OC population. Association studies were performed to investigate the likelihood that the gene, as a potential susceptibility or modifier locus, may be involved in OC pathogenesis.

3.2.3. MATERIALS AND METHODS

The study protocol was approved by the Ethics Review Committees of the Universities of Stellenbosch and Cape Town.

3.2.3.1. Subjects

Blood samples were obtained from 105 unrelated patients with squamous cell carcinoma of the oesophagus, including 82 male and 23 female. These patients had been referred to clinics at the Grootte Schuur and Tygerberg Hospitals in Cape Town. Patients with OC present with symptoms such as difficulty to swallow and loss of weight.

These patients were referred for a barium swallow followed by biopsies in the theatre to confirm the presence of an ADC or SCC on the basis of the histology (Hermanek and Sobin 1987). In a questionnaire completed by the individuals included in the study, possible environmental risk factors of OC were denoted, including cigarette smoking, alcohol drinking, family history of cancer and occupation. Blood extracted DNA samples from 110 (43 male and 67 female) normal, unrelated population-matched individuals were included as controls. Control individuals were recruited at the Tygerberg Hospital. The study population included individuals from the South African "Coloured" population, referring to individuals of mixed ancestry (Loubser *et al* 1999), including San, Khoi, Madagascar, Javanese and European origin, from the Western Cape.

3.2.3.2. Experimental procedures

DNA extraction was performed on whole blood using standard techniques (Miller *et al* 1988). Polymerase chain reaction (PCR) amplification was performed on the promoter region (Kotze *et al* 2001) and exon 2 (forward primer - 5'gtgacaagggtccccaa3'; reverse primer - 5'atcccagacacaaaaccg3') of the *SLC11A1* gene. The PCR products of the promoter region (467-bp) and exon 2 (143-bp) were subjected to a combined heteroduplex single-strand conformation polymorphism (HEX-SSCP) method according to Kotze *et al* (1995). These PCR products were resolved on a 12% polyacrylamide gel supplemented with 7.5% urea at 4°C (350 V) for 18 hr. In order to better discriminate between the different alleles of the 5'-(GT)n repeat polymorphism in the promoter region of the gene, a 116-bp PCR product was generated (Kotze *et al* 2001) and subjected to restriction fragment length polymorphism (RFLP) analysis using *Rsa*I (Graham *et al* 2000). The digested PCR product was resolved on a 10% polyacrylamide gel at room

temperature (250V) for 30 minutes. Following electrophoresis, the gels were stained with ethidium bromide and the DNA visualised by ultraviolet light transillumination. Semi-automated DNA sequencing was performed on PCR products demonstrating unique mobility or conformational variants in the polyacrylamide gels, using an ABI 3100 PRISM automated sequencer.

3.2.3.3. Statistical analysis

Allele and genotype frequencies were estimated by allele counting and statistical differences between patient and control groups were tested for significance by the Fisher exact test and/or chi-squared (χ^2) analysis with Yates' correction. A probability value smaller than 0.05 was regarded as statistically significant. The Hardy-Weinberg equilibrium was performed to test equilibrium for the genetic traits investigated in the respective populations.

Allelic stratification applied in this study was based on previous findings (chapter 3.1), demonstrating an opposite allelic effect for allele 3 (similar to that of allele 2) upon iron loading when occurring in association with the $-237C \rightarrow T$ polymorphism. Therefore, when allele 3 occurred together with the T-allele of the $-237C \rightarrow T$ polymorphism, this allelic combination was grouped with allele 2 for further statistical analysis.

3.2.4. RESULTS

Our study population included 82 (78%) males and 23 (22%) females with a mean age of 61 years and 59 years respectively (Table 3.2.1), with the majority of patients being

90

older than 50 years (63% male and 65% female). Additional data was available for only 45 (54%) male and 15 (65%) female OC patients and all of these patients were cigarette smokers with an average of 10 cigarettes per day. Of these cigarette smokers, 9 of 60 (15%) (male: 6 of 45, 13.33%; female: 3 of 15, 20%) were only smokers, 31 of 60 (52%) (male: 27 of 45, 60%; female: 4 of 15, 26.67%) additionally consumed alcoholic beverages (including both beer (shop-bought) and wine or spirits), 15 of 60 (25%) (male: 8 of 45, 17.78%; female: 7 of 15, 46.67%) additionally consumed only beer and 5 of 60 (8%) (male: 4 of 45, 8.89%; female: 1 of 15 6.67%) additionally consumed only wine or spirits. Information concerning occupation was available for only 32 males and 15 females with the majority of the males being laborers (16 of 32, 50%), farm workers (10 of 32, 31%) and gardeners (6 of 32, 19%) and females being domestic workers (7 of 15, 47%) and housewifes (6 of 15, 40%). The OC patients were all from the Western Cape and are localised within the region as shown in Figure 3.2.1.

Table 3.2.1. Characteristics of OC patients

	Patients with OC	
	Male (n=82)	**Female (n=23)**
Mean Age ± *s.e. (range)	61 ± 1.41 (37-89)	59 ± 2.75(39-85)
Unknown	12 (14.6%)	3 (13%)
Alcoholic consumption (Beer)		
Weekly	18 (21.95%)	7 (30.43%)
Daily	8 (9.76%)	3 (13.04%)
Seldom	9 (10.98%)	1 (4.35%)
Never	10 (12.19%)	4 (17.39%)
Unknown	37 (45.12%)	8 (34.78%)
Alcoholic consumption (Wine/Spirits)		
Weekly	15 (18.29%)	2 (8.70%)
Daily	10 (12.20%)	2 (8.70%)
Seldom	6 (7.32%)	1 (4.35%)
Never	14 (17.07%)	10 (43.48%)
Unknown	37 (45.12%)	8 (34.78%)
Cigarette Smoking	45 (54.88%)	15 (65.22%)
Unknown	37 (45.12%)	8 (34.78%)

*s.e., standard of mean

Figure 3.2.1.

Legend to figure 3.2.1.
Area of high incidence of OC in the Western Cape are shown (green) and possible environmental factors contributing to OC should be considered within this region.

Mutation analysis of the promoter region of the *SLC11A1* gene using a combined HEX-SSCP method revealed two previously-documented polymorphic residues, including -237C→T (Lewis *et al* 1996) and −8G→A (Cashmore T, unpublished data). The −237C→T polymorphism was identified in both a heterozygous state [4 of 105 (3.81%) OC patients and 20 of 110 (18.18%) population-matched control individuals] and homozygous state [1 of 105 (0.95%) OC patients and 1 of 110 (0.91%) population-

93

matched control individuals]. The –8G→A variant was identified only in a heterozygous state in 1 of 82 (1.22%) male OC patients and in 2 of 43 (4.65%) male control individuals. By investigating the 5'-(GT)n promoter repeat polymorphism using RFLP analysis, we only identified three previously described alleles, including alleles 2, 3 (Blackwell *et al* 1995) and 5 (Graham *et al* 2000) in the study population.

Mutation analysis of intron 1/exon 2 of the *SLC11A1* gene using a combined HEX-SSCP method revealed four novel variants (112G→A, IVS1-28C→T, 148delGACCAGCCC and 157insGACCAGCCCAG) (Figure 3.2.2). Three of these variants (112G→A [1 of 110 (0.91%)], 148delGACCAGCCC [1 of 110 (0.91%)] and 157insGACCAGCCCAG [1 of 110 (0.91%)]) were identified only in the control individuals. The intronic variant IVS1-28C→T was identified in 6 of 105 (5.71%) OC patients and in 1 of 110 (0.91%) control individuals. The polymorphisms identified in exon 2 were only present in a heterozygous state. All the variants identified were verified by sequencing analysis. The numbering system for variants identified in exon 2 and intron 1 is according to the sequence published by Cellier *et al* (1994). HEX-SSCP analysis of the variants identified are shown in Figure 3.2.3. The variants were in Hardy Weinberg equilibrium in the population group studied.

Normal: TCT CCAGCCCGA CCAGCCCGA CCAGCCCAG GGC

9bp del: TCT ------------------CCAGCCCGA CCAGCCCAG GGC (White *et al* 1994)

9bp del: TCT CCAGCCCGA CCAGCCC ------------------AG GGC (A)

11bp ins: TCT CCAGCCCGA CCAGCCCGA CCAGCCG AG GGC (B)

GACCAGCCCAG

A

B

C

IVS1-28C→T

D
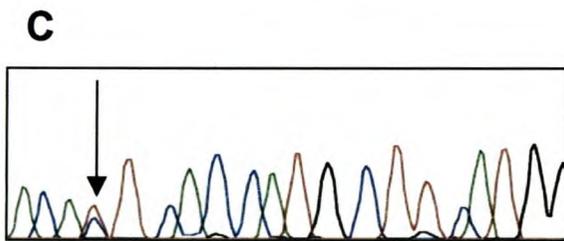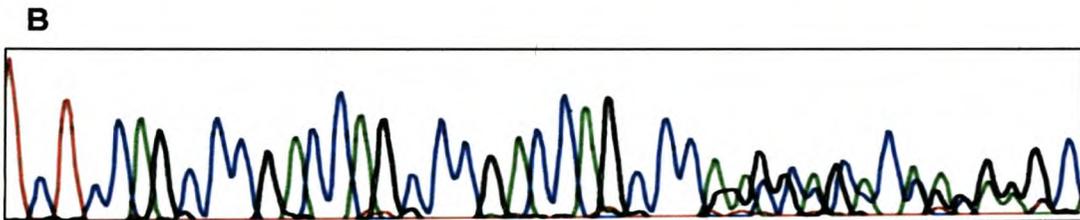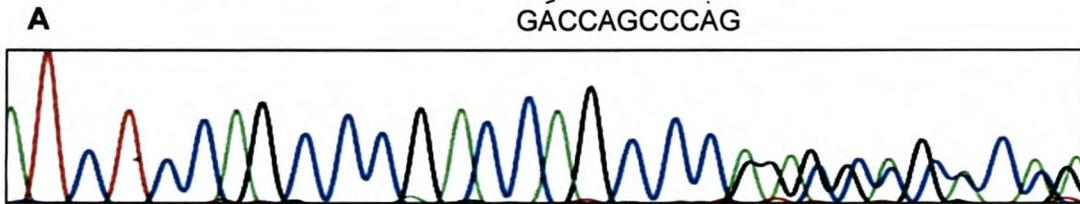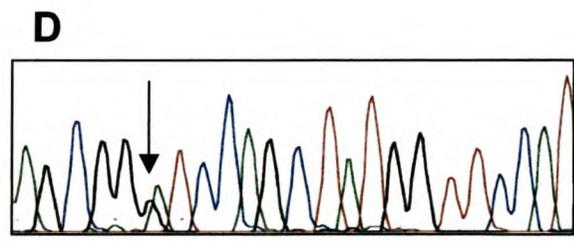
112G→A

Figure 3.2.2.

<u>**Legend to Figure 3.2.2.**</u>

Direct sequencing analysis of variants detected in exon 2 amplified DNA of the *SLC11A1* gene. (**A**) Novel 9-bp deletion, (**B**) novel 11-bp insersion, (**C**) splice variant (IVS1-28C/T) and (**D**) base pair substitution (112G/A), red, thymidine (T) blue, cytidine (C), green, adenosine (A), black, guanosine (G), arrows indicate point of variation.
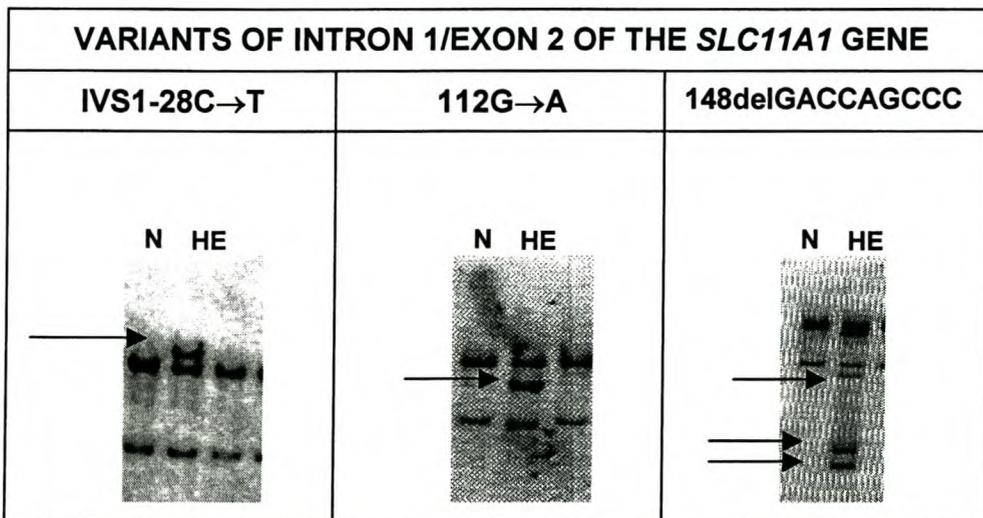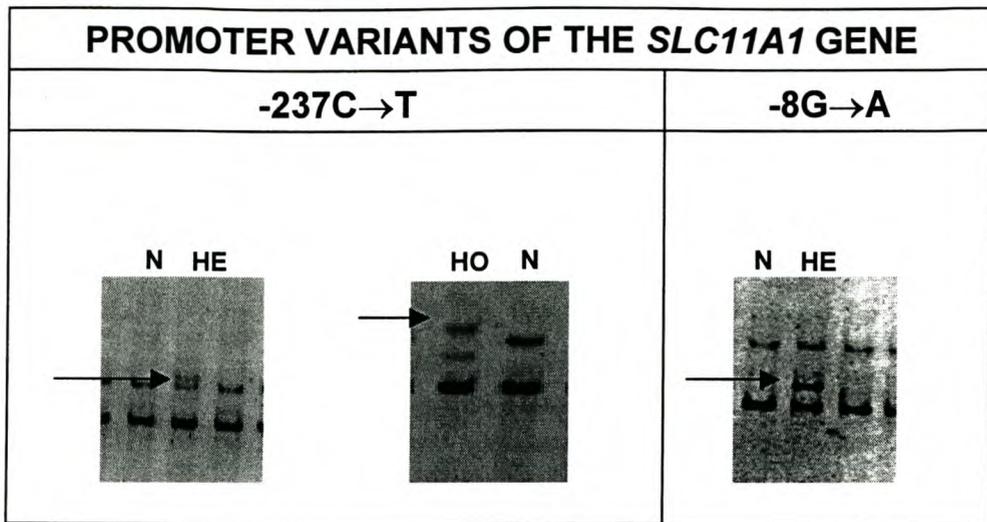
Figure 3.2.3.

Legend to Figure 3.2.3.
Ethidium bromide stained gel images of PCR-SSCP banding patterns of variants identified in the promoter region and intron 1/exon 2 of the *SLC11A1* gene in 12% polyacrylamide gels supplemented with 7.5% urea; N, banding pattern for homozygous wildtype variant; HE, banding pattern for heterozygous variant; HO, banding pattern for homozygous mutant variant; arrows indicate abnormal bands.

The allele and genotype frequencies of the different polymorphic *SLC11A1* variants detected are provided in Tables 3.2.2 and 3.2.3 for the patient and population-matched control groups. The T allele of the $-237C \rightarrow T$ polymorphism was under represented in the OC patient group. Statistically significant differences for the $-237C \rightarrow T$ polymorphism was observed when comparing allele ($p<0.002$, χ^2 with Yates correction = 7.87, df=1) and genotype frequencies ($p<0.003$, χ^2 = 11.14, df=2) between the patient and control groups. A significant difference was also observed when comparing male OC patients with the respective gender control group for the $-237C \rightarrow T$ polymorphism for both allele ($p<0.04$, χ^2 with Yates correction = 3.11, df=1) and genotype ($p<0.01$, χ^2 = 9.83, df=2) frequencies.

The $-237C \rightarrow T$ polymorphism was consistently detected in the presence of allele 3 (Table 3.2.3) and none of the patient or control individuals homozygous for allele 2 presented with the $-237C \rightarrow T$ polymorphism. Therefore, the $-237C \rightarrow T$ polymorphism appears to occur in *cis* with allele 3. No statistically significant differences were observed when comparing the variants of the 5'-(GT)n promoter repeat polymorphism of

Table 3.2.2. Allele frequencies of the various *SLC11A1* variants identified in the Coloured South African population

VARIANTS	[a]ALLELE FREQUENCIES					
	CC (2n=220)			OC (2n=210)		
Promoter variants	**Male (2n=86)**	**Female (2n=134)**	**Total (2n=220)**	**Male (2n=164)**	**Female (2n=46)**	**Total (2n=210)**
Allele 2	0.186	0.194	0.191	0.177	0.152	0.171
Allele 3	0.802	0.806	0.804	0.823	0.848	0.828
Allele 5	0.012	-	0.004	-	-	-
-8G/A	0.023	-	0.009	0.006	-	0.004
-237C→T	*[c]0.081*	0.112	*[b]0.10*	*[c]0.024*	0.043	*[b]0.029*
Exon 2 variants						
112G→A	0.012	-	0.004	-	-	-
IVS1-28C→T	-	0.007	*[d]0.004*	0.030	0.022	*[d]0.028*
148delGACCAGCCC	0.012	-	0.004	-	-	-
157insGACCAGCCCAG	0.012	-	0.004	-	-	-

[a]Allele frequencies of polymorphic alleles denoted, Statistically significant associations: [b]OC vs control (p<0.004), [c]OC males vs control males (p<0.04), [d]OC vs control (p<0.05), CC, Coloured control group, OC, oesophageal cancer patient group, Allele 2, $t(gt)_5ac(gt)_5ac(gt)_{10}ggcaga(g)_6$, Allele 3, $t(gt)_5ac(gt)_5ac(gt)_9ggcaga(g)_6$, Allele 5, $t(gt)_4ac(gt)_5ac(gt)_{10}ggcaga(g)_6$, IVS, intervening sequence

Table 3.2.3. Genotype frequencies of the various *SLC11A1* variants identified in the Coloured population of South Africa

Promoter alleles	GENOTYPE FREQUENCIES			
	CC (n=110)		OC (n=105)	
	Males (n=43)	Females (n=67)	Males (n=82)	Females (n=23)
Allele 2; Allele 2	0.05 (2)	0.01 (1)	0.05 (4)	0.09 (2)
Allele 2; Allele 3	0.28 (12)	0.36 (24)	0.26 (21)	0.13 (3)
Allele 3; Allele 3	0.65 (28)	0.63 (42)	0.70 (57)	0.78 (18)
Allele 3; Allele 5	0.02 (1)	-	-	-
-237CC	0.84 (36)	0.79 (53)	0.97 (79)	0.91 (21)
-237CT	0.16 (7)	0.19 (13)	0.02 (2)	0.09 (2)
-237TT	-	0.01 (1)	0.01 (1)	-
-8GG	0.95 (41)	1.0 (67)	0.99 (81)	1.0 (23)
-8GA	0.05 (2)	-	0.01 (1)	-
Allele 2; Allele 3; -237CT	0.02 (1)	0.03 (2)	0.01 (1)	-
Allele 3; Allele 3; -237CT	0.14 (6)	0.16 (11)	0.01 (1)	0.09 (2)
Allele 3; Allele 3; -237TT	-	0.01 (1)	0.01 (1)	-

CC, Coloured control group, OC, oesophageal cancer patient group, number of individuals with various genotypes are indicated in brackets, Allele 2, t(gt)$_5$ac(gt)$_5$ac(gt)$_{10}$ggcaga(g)$_6$, Allele 3, t(gt)$_5$ac(gt)$_5$ac(gt)$_9$ggcaga(g)$_6$, Allele 5, t(gt)$_4$ac(gt)$_5$ac(gt)$_{10}$ggcaga(g)$_6$

OC patients with the control group (data not shown), except when this was stratified according to the presence of the $-237C \rightarrow T$ polymorphism in association with allele 3. Allele 3, in the absence of the T-allele, was significantly associated with OC ($p<0.0006$, χ^2 with Yates correction = 10.16, df=1) and in particular with males ($p<0.01$, χ^2 with Yates correction = 5.13, df=1). In cases where allele 3 was observed in association with allele T, this genotype combination was also grouped with allele 2 for statistical analysis, based on the finding that the genotypes including these sequence changes exert a similar effect on transcriptional activity (chapter 3.1). Re-assessment of the 5'-(GT)n repeat in this manner confirmed the significant association of allele 3 with OC compared with the controls for both allele frequency ($p<0.01$, χ^2 with Yates' correction = 4.54, df=1) and genotype ($p<0.005$, $\chi^2 = 10.51$, df=2) distribution. Statistically significant differences for the intron 1 IVS1-28C$\rightarrow$T polymorphism were also observed when comparing the allele ($p<0.05$, χ^2 with Yates correction = 2.52, df=1) and genotype ($p<0.05$, χ^2 with Yates correction = 2.56, df=1) frequencies for the OC patient and control groups.

3.2.5. DISCUSSION

Both environmental and genetic factors appear to be involved in the development of OC. Of the environmental factors, cigarette smoking and consuming alcoholic beverages have been shown to be major determinants in developing OC. In our patient group the involvement of these environmental factors were confirmed since all the patients (of which we had data available) are cigarette smokers and/or consume large amounts of alcoholic beverages (including both beer and wine/spirits). Occupational exposures are also likely to play a minor role in the development of OC (reviewed by Blot 1994) and

further investigation into this aspect would be interesting. The majority of our OC patients are exposed to combustion products and a study performed in Sweden has shown that a nearly three-fold excess of developing OC are observed among individuals exposed to these products (Gustavisson *et al* 1993).

We assessed the likelihood that allele 3 of the 5'(GT)n promoter repeat polymorphism of *SLC11A1*, associated with autoimmune disease susceptibility, is involved in OC susceptibility due to its role in chronic hyperactivation of macrophages and the association of chronic inflammation with OC. A significant association was observed for allele 3 of *SLC11A1* with OC, and especially males with OC, when stratified according to the presence of the $-237C{\rightarrow}T$ polymorphism. This was done based on our findings obtained in a previous study (chapter 3.1) where *in vitro* studies of *SLC11A1* promoter variants using luciferase reporter gene constructs demonstrated opposite effects on promoter activity in relation to allele 3 in the presence and absence of the $-237C{\rightarrow}T$ polymorphism (similar to that of allele 2).

Significant association with OC was furthermore observed for the novel intronic IVS1-$28C{\rightarrow}T$ polymorphism identified in this study. A polymorphic area was previously identified in the adjacent coding region in exon 2, located within the proline-rich SH3 binding domain consensus sequence (Barton *et al* 1994). The 9-bp deletion previously identified (White *et al* 1994) was not similar to the 9-bp deletion identified in our study population and an additional 11-bp insertion was identified in this area (Figure 3.2.2). Although these variants were identified at very low frequencies, it could possibly

101

influence the conformation and function of the protein. The spacing of prolines have been shown to be important in binding SH3 domains (Booker *et al* 1993) and this polymorphic area could account for susceptibility to autoimmune and/or infectious diseases. Accurate splicing in eukaryotes are dependant on the 5' splice site, the 3' splice site and the branch site (Cartegni *et al* 2002). Motifs in the intron acting as a branch site have been identified with a conserved A residue in the branch site serving as a branchpoint. The IVS1-28C→T polymorphism lies within this region and could possibly affect splicing. Functional studies needs to be performed to elucidate the role of this variant. Alternatively, the observation of association of this polymorphism with OC could be ascribed to linkage of this polymorphism with an unidentified mutation contributing to OC pathogenesis. This polymorphism appears to be in linkage disequilibrium with allele 3 of the 5'-(GT)n promoter polymorphism since none of these individuals were homozygous for allele 2, while 5 of 6 (83%) were homozygous for allele 3 and 1 of 6 (17%) presenting with alleles 2 and 3. The −237C→T polymorphism was not detected in these six individuals.

The association of functional polymorphisms of the iron transporter *SLC11A1* with OC, and particularly with males in our study, makes this gene a possible candidate for explaining the higher frequency of OC in males compared to females. A previous study involving the Black South African population also demonstrated association between variants in the *SLC11A1* gene and OC susceptibility (du Plessis 2000), which provides considerable support for our findings. Direct mechanisms of iron induced mutagenesis may occur through generation of oxygen free radicals and in turn lead to DNA strand

damage and disruption in structure, contributing to the development of cancer (Togokuni 1996). Individuals homozygous for allele 3 of the *SLC11A1* gene have also been reported to produce higher levels of tumour necrosis factor α (TNF-α) (Blackwell *et al* 2003). The exact role for these pleiotrophic effects in *SLC11A1*-regulated autoimmune and infectious disease phenotypes are not yet known and needs further investigation.

Although two regions of the *SLC11A1* gene has been associated with OC in our study population, the possible influence of mutations in other genes closely linked to the *SLC11A1* gene cannot be excluded. A high frequency of loss on chromosome 2q was indicated in a study by du Plessis *et al* (1999), using comparative genomic hybridisation analysis, where various genes have been identified that could possibly be involved in OC pathogenesis. Although association of *SLC11A1* gene variants with susceptibility to various autoimmune and infectious diseases may be related to the many pleiotrophic effects of the gene on macrophage function (Searle and Blackwell 1999), we conclude that its involvement in inflammation and the generation of reactive oxygen species possibly contributes directly to OC pathogenesis

3.2.6. ACKNOWLEDGEMENTS

3.3. The $-237C{\to}T$ promoter polymorphism of the *SLC11A1* gene exerts a protective effect in relation to inflammatory bowel disease in the South African population

3.3.1. ABSTRACT

The purpose of this study was to assess the likelihood that variation in the promoter region of the solute carrier family 11 member 1 gene (*SLC11A1*) contributes to inflammatory bowel disease (IBD) susceptibility in the South African population. The study cohort included 102 IBD patients, 47 with Crohn's disease (CD) and 55 with ulcerative colitis (UC), and 192 population-matched controls. Mutation analysis revealed two novel alleles for the 5'-(GT)n repeat polymorphism, $t(gt)_5ac(gt)_5ac(gt)_6ggcaga(g)_6$ (allele 8) and $t(gt)_5ac(gt)_5ac(gt)_8ggcaga(g)_6$ (allele 9), and one previously-documented point mutation $-237C{\to}T$. A significantly decreased frequency of the $-237C{\to}T$ promoter polymorphism was observed in the patient group with IBD ($p<0.001$) and CD ($p<0.0006$) compared with the population-matched control group. These findings are in accordance with previous *in vitro* studies, which demonstrated that the point mutation at nucleotide position -237 represents a functional polymorphism that affects regulation of the upstream 5'-(GT)n repeat polymorphism differentially upon iron loading. Our findings raise the possibility that iron dysregulation mediated by allelic effects of *SLC11A1* may contribute to IBD susceptibility.

3.3.2. INTRODUCTION

Inflammatory bowel disease (IBD), generally classified as either Crohn's disease (CD) or ulcerative colitis (UC), is characterised by chronic inflammation of the intestines (Sanderson 1986, Griffiths 1995). It remains uncertain whether these subtypes of IBD are distinct entities or a mechanistic continuum of the same disease process. IBD frequently occurs in association with other autoimmune diseases, including ankylosing spondylitis, psoriasis, sclerosing cholangitis and multiple sclerosis (Snook *et al* 1989, Yang and Rotter 1993).

Both genetic and environmental (microbial) factors appear to be involved in the aetiology of IBD (Sartor 1995, Shanahan 1995, Targan 1995). The frequency of affected first-degree relatives in CD is higher than in UC (Satsangi *et al* 1994), and therefore genetic factors may be relatively more important in CD (Satsangi *et al* 1997). It has been suggested that CD might be triggered by intracellular pathogens in genetically predisposed individuals (Targan 1995). Support for this notion was obtained by the identification of a significant association with susceptibility alleles in the caspase recruitment domain-containing protein 15 gene (*CARD15*), which appear to cause inappropriate monocyte response to bacterial infection (Hugot *et al* 2001, Ogura *et al* 2001a, Hampe *et al* 2001, Lesage *et al* 2002). However, some studies failed to demonstrate association between mutations in the *CARD15* gene and CD, possibly due to population differences (Van Heel *et al* 2002). Other genes linked to IBD susceptibility (Podolsky 2002) that are also expressed in macrophages, include the solute carrier family 11 member 1 gene (*SLC11A1*) (Hofmeister *et al* 1997, Kojima *et al* 2001). This gene has been linked to various autoimmune and infectious diseases, which led to

speculation that regulation of iron by *SLC11A1* may be of major importance in this context (Blackwell *et al* 2000). This is substantiated by reports on abnormalities of iron metabolism (Weber *et al* 1988, Nielsen *et al* 1994) in arthritic conditions shown to be associated with genetic variation in the *SLC11A1* gene (Shaw *et al* 1996, Sanjeevi *et al* 2000, Singal *et al* 2000, Yang *et al* 2000), which may be linked to defective iron supply for erythropoiesis (Cazzola *et al* 1996).

Anaemia of iron deficiency and anaemia of chronic diseases frequently occurs in IBD (Oldenburg *et al* 2001). Iron deficiency in IBD patients are usually treated with oral iron supplements but this may lead to an increased inflammatory activity by generating reactive oxygen species. In an attempt to determine the possible role of the *SLC11A1* gene in the disease process, we assessed the likelihood that variation in the promoter region of the *SLC11A1* gene contributes to IBD susceptibility in the South African population. Previous studies of the functional Z-DNA forming repeat polymorphism in the promoter region of the *SLC11A1* gene suggested direct contribution of specific alleles to autoimmune (allele 3) and infectious (allele 2) disease susceptibility (Searle and Blackwell 1999, Blackwell *et al* 2000). These findings were considered in the current study in relation to recent *in vitro* evidence, demonstrating an opposite allelic effect for allele 3 (similar to that of allele 2) upon iron loading, in the event that allele 3 occurs in association with the $-237C \rightarrow T$ polymorphism (chapter 3.1).

3.3.3. MATERIALS AND METHODS

The study protocol was approved by the Ethics Review Committees of the Universities of Stellenbosch and Cape Town.

3.3.3.1 Subjects

Blood samples were obtained with informed consent from 102 unrelated patients diagnosed with IBD, including 55 with UC (24 male and 31 female) and 47 with CD (18 male and 29 female). All patients attended the Gastroenterology Clinic at Grootte Schuur Hospital. The presence of UC or CD was confirmed by radiology, or by endoscopy and biopsy (Heikenen *et al* 1999). The study population included 9 Black (9 females with UC), 77 Coloured (34 male – 18 UC and 16 CD patients, 43 female – 17 UC and 26 CD patients) and 16 Caucasian (8 male – 6 UC and 2 CD patients; 8 female – 5 UC and 3 CD patients) South African patients diagnosed with IBD and 192 control individuals from these respective population groups: 25 Black (12 male, 13 female), 110 Coloured (43 male, 67 female) and 57 Caucasian (33 males, 24 female). In this study "Caucasian" refers to an individual of European descent, mainly Dutch, French, German and British origin; "Coloured" refers to an individual of mixed ancestry, including San, Khoi, Madagascar, Javanese and European origin; "Black" refers to South Africans of central African descent.

3.3.3.2. Methods

DNA was extracted from whole blood using a standard method (Miller *et al* 1988). The promoter region of the *SLC11A1* gene was amplified by the polymerase chain reaction (PCR) using oligonucleotide primers previously described (Kotze *et al* 2001), generating

107

fragments of 116-bp and 467-bp, respectively. The 467-bp PCR products were subjected to combined heteroduplex single-strand conformation polymorphism (HEX-SSCP) analysis (Kotze *et al* 1995) and resolved on a 12% PAA gel supplemented with 7.5% urea at 4°C (350 V) for 18 hr. The 116-bp PCR product of the promoter region was subjected to restriction fragment length polymorphism (RFLP) analysis (Graham *et al* 2000) using *RsaI*, which allows discrimination between the various Z-DNA forming promoter repeat polymorphic alleles. The digested PCR products were resolved on a 10% PAA gel at room temperature (250V) for 30 minutes. Following electrophoresis, the gels were stained with ethidium bromide and the DNA visualised by ultraviolet light transillumination. Semi-automated DNA sequencing was performed on PCR products demonstrating mobility or conformational variants in the polyacrylamide gels, using an ABI 3100 PRISM automated sequencer.

3.3.3.3. Statistical analysis

Allele and genotype frequencies were estimated by allele counting and statistical differences between patient and control groups were tested for significance by the Fisher exact test and/or chi-squared (χ^2) analysis with Yates' correction. The Mantel Haenszel chi-squared test was performed to exclude the possibility of confounding factors due to population substructures. A probability value smaller than 0.05 was regarded as statistically significant. The Hardy-Weinberg equilibrium was performed to test equilibrium for the genetic traits investigated in the respective populations.

Allelic stratification applied in this study was based on previous findings (chapter 3.1), demonstrating an opposite allelic effect for allele 3 (similar to that of allele 2) upon iron loading when occurring in association with the $-237C \rightarrow T$ polymorphism. Therefore, when allele 3 occurred together with the T-allele of the $-237C \rightarrow T$ polymorphism, this allelic combination was grouped with allele 2 for further statistical analysis.

3.3.4. RESULTS

Mutation analysis of the promoter region of the *SLC11A1* gene was performed in South African patients with CD and IBD, using a combined HEX-SSCP method and RFLP analysis. This revealed several allelic variants which included the $-237C \rightarrow T$ polymorphism (Lewis *et al* 1996), and additional alleles for the 5'-(GT)n repeat polymorphism, designated alleles 8 $(t(gt)_5ac(gt)_5ac(gt)_6ggcaga(g)_6)$ and 9 $(t(gt)_5ac(gt)_5ac(gt)_8ggcaga(g)_6)$ (EMBL accession no. AJ535670). The variants identified are shown in Figure 3.3.1.

The $-237C \rightarrow T$ polymorphism was identified in both a heterozygous state (20 of 110 [18.2%] Coloured control individuals, 3 of 35 [8.57%] Coloured UC patients; 6 of 57 [10.5%] Caucasian control individuals, 2 of 11 [18.2%] Caucasian UC patients and 3 of 25 [12%] Black control individuals) and homozygous state (1 of 110 [0.91%] Coloured control individuals and 1 of 57 [1.75%] Caucasian control individuals). Allele 9 was identified in 2 of 42 CD patients (5%) in the Coloured population and was absent in the control individuals. Allele 8 was identified in 1 of the 110 Coloured control individuals (0.91%) and absent in the Black and Caucasian populations studied. All the different

109

variants identified in the promoter region were verified by sequencing analysis. The variants studied were in Hardy Weinberg equilibrium in the population group studied.
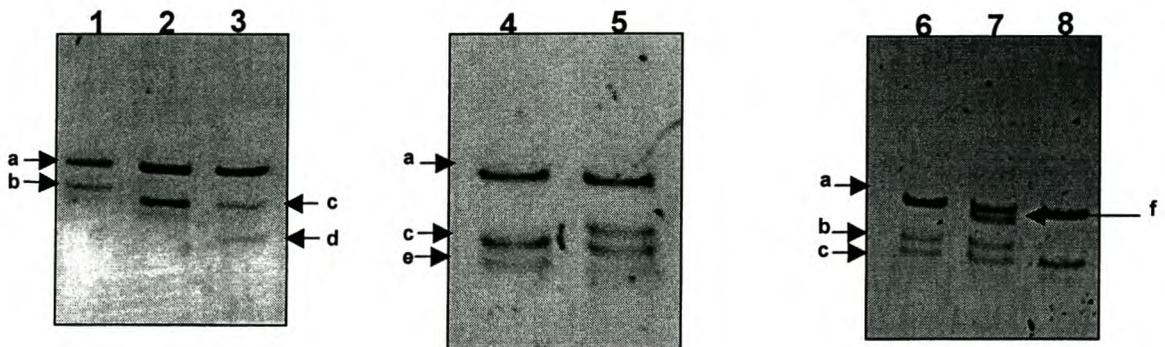
Figure 3.3.1.

Legend to Figure 3.3.1.
Restriction enzyme (*RsaI*) analysis of the 5'-(GT)n repeat of the promoter region. PCR products resolved on a 10% polyacrylamide gel and stained in 0.01% (v/v) ethidium bromide. Different fragment sizes observed are numbered from a to f with (a) 57 base pairs, (b) 49 base pairs (c) 47 base pairs (d) 41 base pairs (e) 45 base pairs and (f) 55 base pairs. Various genotypes observed are numbered from 1 to 8 with

lane 1: $t(gt)_5ac(gt)_5ac(gt)_{10}ggcaga(g)_6$ / $t(gt)_5ac(gt)_5ac(gt)_{10}ggcaga(g)_6$
lanes 2 and 8: $t(gt)_5ac(gt)_5ac(gt)_9ggcaga(g)_6$ / $t(gt)_5ac(gt)_5ac(gt)_9ggcaga(g)_6$
lane 3: $t(gt)_5ac(gt)_5ac(gt)_9ggcaga(g)_6$ / $t(gt)_5ac(gt)_5ac(gt)_6ggcaga(g)_6$
lane 4: $t(gt)_5ac(gt)_5ac(gt)_9ggcaga(g)_6$ / $t(gt)_5ac(gt)_5ac(gt)_8ggcaga(g)_6$
lanes 5 and 6: $t(gt)_5ac(gt)_5ac(gt)_9ggcaga(g)_6$ / $t(gt)_5ac(gt)_5ac(gt)_{10}ggcaga(g)_6$
lane 7: $t(gt)_4ac(gt)_5ac(gt)_{10}ggcaga(g)_6$ / $t(gt)_5ac(gt)_5ac(gt)_9ggcaga(g)_6$

The allele and genotype frequencies of the different polymorphic *SLC11A1* variants detected for the various populations studied are presented in Tables 3.3.1 and 3.3.2. The T allele of the -237C→T polymorphism is under represented in the patient groups and was not identified in any of the CD patients studied (Table 3.3.2). Statistically significant differences were observed in the Coloured population (Table 3.3.3) but not in the Caucasian and Black populations, most probably due to the small sample size for these population groups. The frequency of the -237C→T promoter polymorphism differed significantly in Coloured patients with IBD ($p<0.001$) and CD ($p<0.0006$), compared with the Coloured control group (Table 3.3.3). A significant difference was also observed when comparing the IBD female group (including CD and UC patients) and female CD Coloured patients with the respective gender Coloured control group ($p<0.005$). All the variants studied were in Hardy Weinberg equilibrium for the various population groups.

No statistically significant differences were observed when comparing the variants of the 5'-(GT)n repeat polymorphism in IBD patients with that in the control group (data not shown), except when this was stratified according to the presence of the −237C/T polymorphism in association with allele 3 (Table 3.3.3). Allele 3 was significantly associated with CD in the Coloured population in the absence of this sequence variant.

Table 3.3.1. Allele frequencies of the various *SLC11A1* promoter variants in three genetically distinct South African populations.

Promoter alleles	ALLELE FREQUENCIES							
	CC (2n=220)	CCD (2n=84)	CUC (2n=70)	WC (2n=114)	WCD (2n=10)	WUC (2n=22)	BC (2n=50)	[a]BUC (2n=18)
Allele 1	-	-	-	0.009	-	-	-	-
Allele 2	0.190	0.202	0.243	0.202	0.200	0.136	0.160	0.111
Allele 3	0.800	0.774	0.757	0.781	0.800	0.864	0.840	0.889
Allele 5	0.005	-	-	0.009	-	-	-	-
Allele 8	0.005	-	-	-	-	-	-	-
Allele 9	-	0.024	-	-	-	-	-	-
-237T	0.100	-	0.043	0.070	-	0.091	0.060	-

[a]No individuals in the Black population presented with CD in this study, CC, Coloured control group, CCD, Coloured Crohn's disease patient group; CUC, Coloured ulcerative colitis patient group, WC, Caucasian control group, WCD, Caucasian Crohn's disease patient group, WUC, Caucasian ulcerative patient group, BC, Black control group, BUC, Black ulcerative colitis patient group

Table 3.3.2. Genotype frequencies of the various *SLC11A1* promoter variants in three genetically distinct South African populations

Promoter alleles	GENOTYPE FREQUENCIES							
	CC (n=110)	CCD (n=42)	CUC (n=35)	WC (n=57)	WCD (n=5)	WUC (n=11)	BC (n=25)	[a]BUC (n=9)
Allele 2; Allele 2	0.03 (3)	0.05 (2)	0.03 (1)	0.02 (1)	-	-	-	-
Allele 2; Allele 3	0.33 (36)	0.31 (13)	0.43 (15)	0.37 (21)	0.40 (2)	0.27 (3)	0.32 (8)	0.22 (2)
Allele 3; Allele 3	0.63 (69)	0.60 (25)	0.54 (19)	0.58 (33)	0.60 (3)	0.73 (8)	0.68 (17)	0.78 (7)
Allele 3; Allele 5	0.009 (1)	-	-	0.02 (1)	-	-	-	-
Allele 3; Allele 8	0.009 (1)	-	-	-	-	-	-	-
Allele 3; Allele 9	-	0.05 (2)	-	-	-	-	-	-
[c]Allele 1; Allele 3	-	-	-	0.02 (1)	-	-	-	-
-237CC	0.81 (89)	1.0 (42)	0.91 (32)	0.88 (50)	1.0 (5)	0.82 (9)	0.88 (22)	1.0 (9)
-237CT	0.18 (20)	-	0.09 (3)	0.10 (6)	-	0.18 (2)	0.12 (3)	-
-237TT	0.009 (1)	-	-	0.02 (1)	-	-	-	-
Allele 2; Allele 3; -237CT	0.03 (3)	-	-	0.04 (2)	-	-	-	-
Allele 3; Allele 3; -237CT	0.15 (17)	-	0.09 (3)	0.07 (4)	-	0.18 (2)	0.12 (3)	-
Allele 3; Allele 3; -237TT	0.009 (1)	-	-	0.02 (1)	-	-	-	-

[a]No individuals in the Black patient group presented with CD in this study, CC, Coloured control group, CCD, Coloured Crohn's disease patient group; CUC, Coloured ulcerative colitis patient group, WC, Caucasian control group, WCD, Caucasian Crohn's disease patient group, WUC, Caucasian ulcerative patient group, BC, Black control group, BUC, Black ulcerative colitis patient group. The number of individuals for each genotype is indicated in brackets.

Table 3.3.3. Statistically significant differences of allele and genotype frequencies observed within the Coloured South African population

Variant	Association tested	GENOTYPE FREQUENCIES			ALLELE FREQUENCIES				
		χ^2	p-value	df	χ^2 with Yates' correction	p-value	Mantel-Haenszel χ^2 test	p-value	df
-237C→T	#IBD patient group vs control group	9.56	0.009	2	8.17	0.001	9.09	0.003	1
	#IBD female patient group vs female control group	7.68	0.02	2	6.40	0.003	5.73	0.02	1
	CD patient group vs control group	9.31	0.009	2	7.63	0.0006	8.35	0.005	1
	CD female patient group vs female control group	6.39	0.04	2	4.91	0.005	5.13	0.02	1
*3/3 without −237T	CD patient group vs control group	6.49	0.004	1	6.78	0.004	7.29	0.01	1

χ^2, chi-squared, df, degrees of freedom, p-value, probability value, *Allele 3/Allele 3 compared to Allele 3/Allele 3 associated with the −237T allele, #IBD patient group includes both CD and UC patients, refer to Table 3.3.2 for number of individuals used for statistical analysis

However, re-assessment of the data according to phenotypic effects determined *in vitro* (chapter 3.1), demonstrating an opposite allelic effect for allele 3 (similar to that of allele 2) upon iron loading in the event that allele 3 occurs in association with the $-237C \rightarrow T$ polymorphism, did not show a significant association of allele 3 with IBD when compared with the population-matched control group for the 5'-(GT)n repeat.

The likelihood that confounding factors influenced the statistical analysis was excluded by assessment of the Mantel Haenszel chi-squared test, which demonstrated that all the associations complied with the "rule of 5." The rare variants, alleles 5, 8 and 9, were insufficiently represented in both the patient and control groups and were not included for comparisons. The $-237C \rightarrow T$ polymorphism appears to be in *cis* with allele 3, since none of the patient and control groups shown to be homozygous for allele 2 presented with the T-allele.

3.3.5. DISCUSSION

Linkage studies previously implicated the involvement of the *SLC11A1* gene with CD susceptibility (Hofmeister *et al* 1997). However, investigation of intragenic polymorphisms ($274C \rightarrow T$ and $823C \rightarrow T$) could not confirm the association with IBD susceptibility (Stokkers *et al* 1999). In a more recent study performed in the Japanese population, a novel variant of the *SLC11A1* promoter repeat polymorphism (allele 7) was shown to be associated with IBD susceptibility (Kojima *et al* 2001). In this study we performed association studies to investigate the likelihood that genetic variation in the

115

promoter region of the *SLC11A1* gene is involved in the pathogenesis of IBD in the South African population.

No disease association was detected when the allelic distribution of the 5'-(GT)n repeat polymorphism was compared between patients and controls. However, statistically significant differences were observed when the frequency of the downstream $-237C \rightarrow T$ polymorphism was compared between patients with CD and control individuals. Under representation of the T-allele was observed and was identified in 3 of 35 UC patients (8.57%) and in none of the 42 CD patients compared with 21 of 110 control individuals (19.10%) in the Coloured population. Although the patient numbers were small, this allele was not detected in Caucasian patients with CD, while present in a considerable proportion of the population-matched control group (7 of 57, 12.28%). None of the Black patients presented with CD in this study. The $-237C \rightarrow T$ polymorphism was absent in the Black UC patient group, while present in the respective control group (3 of 25, 12.0%).

Allele 3 of the *SLC11A1* 5'-(GT)n promoter polymorphism has previously been linked to autoimmune disease susceptibility due to its role in chronic hyperactivation of macrophages. We therefore re-assessed the likelihood that this variant is involved in IBD susceptibility following allelic stratification. When the potential effect of allele 3 was considered in the absence of the $-237C \rightarrow T$ polymorphism, a significant association with CD was observed in the Coloured population. In cases where allele 3 was observed in association with allele T, this genotype combination was also grouped with allele 2 for

statistical analysis, based on the fact that the genotypes including these sequence changes exert a similar effect on transcriptional activity (chapter 3.1). No statistically significant association was observed in this analysis. This finding demonstrates the importance of assessment of allelic interaction in case-control association studies. The apparent protective effect of the −237C→T polymorphism in relation to IBD represents the most significant finding in this study.

Several other genes, including the *CARD15* (Hugot *et al* 2001, Ogura *et al* 2001) and *TNF-α* (Ruuls and Sedgwick 1999, Koss *et al* 2000, Sashio *et al* 2002, Van Heel *et al* 2002) genes, both expressed in macrophages, were shown to be implicated in CD pathogenesis. According to Van Montfrans *et al* (2002), *CARD15* mutations are responsible for disease in about 15% of CD patients. O'Callaghan *et al* (2003) have furthermore shown that association of a *TNF-α* promoter variant (-857C→T) with IBD was more prevalent in families with *CARD15* risk alleles, suggesting that both *TNF-α* and *CARD15* interact with NF-κB in macrophages. Individuals homozygous for allele 3 of *SLC11A1* express higher levels of TNF-α, possibly due to the pleiotropic effects of *SLC11A1* on *TNF-α*. LPS-related response elements, including NF-κB, in the promoter region of *SLC11A1* were also differentially affected by the various alleles of the 5'-(GT)n repeat. With *SLC11A1*, *CARD15* and *TNF-α* all being expressed in macrophages, we could speculate that these genes interact through a common pathway, with *SLC11A1* acting as a modifier locus, to cause IBD possibly due to its role in iron metabolism.

It should also be taken into consideration that CD activity [CD Activity Index – (CDAI)] could be influenced by an iron deficiency state due to the contribution of the haematocrit being affected by iron deficiency, although this activity is relatively small. Unfortunately iron status was not available for the patients included in this study. In conditions of chronic inflammation such as IBD, iron homeostasis is usually altered mainly due to chronic blood loss in the intestine (Oldenburg *et al* 2001). Iron deficiency can occur due to iron malabsorption because of inflammatory activity in the small intestine and resection or impaired dietary intake.

The significant decreased frequency of the $-237C{\rightarrow}T$ promoter polymorphism observed in this study in patients with Crohn's disease compared with population-matched controls confirm the role of the *SLC11A1* gene in IBD. It is clear that interaction between various genes and environmental factors contribute to the disease phenotype and future studies to assess the impact of *SLC11A1* promoter alleles on subsequent expression of disease activity and characterisation may prove valuable in patient management.

3.4. *CARD15* gene mutations in South African patients with inflammatory bowel disease patients: interaction with *SLC11A1* gene variants

3.4.1. ABSTRACT

The caspase recruitment domain gene (*CARD15*) was recently identified as an important susceptibility gene for Crohn's disease (CD). The purpose of this study was to determine whether the three most common *CARD15* mutations, R702W, G908R and 1007fs, contribute to inflammatory bowel disease (IBD) susceptibility in three genetically distinct South African populations. The study cohort included 101 IBD patients, 46 with CD and 55 with ulcerative colitis (UC), and 183 population-matched controls. Mutations R702W, G908R and 1007fs were present at relatively low frequencies (<20%) in our predominantly non-Caucasian study population. No statistically significant differences were observed in the mutation frequencies in UC and CD patients or when these groups were compared with healthy control individuals within each population group. Two additional mutations were identified, one novel (A661P) and one previously described (A725G), with the latter being identified mainly in the Coloured UC patients (4 of 35, 11%). In this population, statistically significant differences were obtained between UC and control individuals when comparing both allele (p<0.004, χ^2 with Yates' correction = 8.01) and genotype frequencies (p<0.004, χ^2 with Yates' correction = 8.14) for the A725G mutation, suggesting a possible role for this variant in disease expression. Since the same study population has previously been investigated for allelic association with the *SLC11A1* gene, possible combined allelic effects of both genes were finally considered in an attempt to explain our findings.

3.4.2. INTRODUCTION

Inflammatory bowel disease (IBD) is a disease of chronic inflammation in the intestines, and is often referred to as an autoimmune disease. The disease is generally subdivided into two phenotypes known as Crohn's disease (CD) and ulcerative colitis (UC) (Sanderson 1986, Griffiths 1995). Both genetic and environmental factors have been implicated in the development of IBD (Sartor 1995, Shanahan 1995, Targan 1995).

A susceptibility locus on chromosome 16 was first identified by Hugot *et al* (1996) and subsequently by other groups (reviewed in the IBD International Genetics Consortium 2001, Zouali *et al* 2001). However, several authors failed to replicate these findings (Rioux *et al* 1998, Vermeire *et al* 2000, Paavola *et al* 2001). Independent studies performed by Hugot *et al* (2001) and Ogura *et al* (2001a) identified disease-predisposing alleles of the caspase recruitment domain-containing protein 15 gene (*CARD15*). Several mutations and polymorphisms of the *CARD15* gene have been identified to date in IBD patients demonstrating significant association with three of these mutations (R702W, G908R and 1007fs), mainly with CD in the Caucasian population (Hugot *et al* 2001, Ogura *et al* 2001a, Hampe *et al* 2001, Lesage *et al* 2002). Mutations of *CARD15* are considered to be the causative genetic defect in about 15% of CD patients (van Montfrans *et al* 2002), altering either the structure of the LRR domain of the protein (Ogura *et al* 2001a) or the adjacent region (Weitzman 2001).

In this study we investigated the presence of mutations R702W, G908R and 1007fs in the *CARD15* gene in three genetically distinct South African populations and performed

case-control association studies to assess the significance of this gene in the pathogenesis of IBD in the South African population. The same study population has recently been subjected to analysis of the *SLC11A1* gene (chapter 3.3) and therefore mutation frequencies were also considered within the context of the *SLC11A1* genetic background and allelic effects of promoter variants. We specifically investigated the likelihood that the apparent protective effect associated with the *SLC11A1* –237C→T polymorphism in relation to IBD susceptibility has influenced our data. Similar to *CARD15*, the *SLC11A1* gene is also expressed in macrophages and gene variation is associated with various autoimmune and infectious diseases, including susceptibility to IBD (Hofmeister *et al* 1997, Kojima *et al* 2001, Podolsky 2002).

3.4.3. MATERIALS AND METHODS

The study was approved by the Ethics Review Committees of the Universities of Stellenbosch and Cape Town.

3.4.3.1. Subjects

Blood samples were obtained with informed consent from 101 unrelated patients diagnosed with IBD, including 55 with UC (23 male and 32 female) and 46 with CD (17 male and 29 female). All patients attended the Gastroenterology Clinic at Grootte Schuur Hospital. The presence of UC or CD was confirmed by radiology, or by endoscopy and biopsy (Heikenen *et al* 1999). The study population included 9 Black (9 females with UC), 76 Coloured (31 male – 16 UC and 15 CD patients, 45 female – 19 UC and 26 CD patients) and 16 Caucasian (9 male – 7 UC and 2 CD patients; 7 female

121

– 4 UC and 3 CD patients) South African patients diagnosed with IBD and 183 control individuals from these respective population groups: 25 Black (12 male, 13 female), 100 Coloured (39 male, 61 female) and 58 Caucasian (34 males, 24 female). In this study "Caucasian" refers to an individual of European descent, mainly Dutch, French, German and British origin; "Coloured" refers to an individual of mixed ancestry (Loubser *et al* 1999), including San, Khoi, Madagascar, Javanese and European origin; "Black" refers to South Africans of central African descent.

3.4.3.2. Methods

DNA was extracted from whole blood using a standard method (Miller *et al* 1988). Prior to inclusion in this study, all individuals have been genotyped for the 5'-(GT)n repeat polymorphism of the *SLC11A1* gene (chapter 3.3). Exons 4, 8 and 11 of the *CARD15* gene were amplified by polymerase chain reaction (PCR) amplification using oligonucleotide primers previously described (Lesage *et al* 2002). The PCR products of exons 4 (494 bp) and 11 (230 bp), incorporating polymorphisms R702W and 1007fs respectively, were subjected to a combined heteroduplex single-strand conformation polymorphism (HEX-SSCP) analysis (Kotze *et al* 1995) method and resolved on a 12% PAA gel supplemented with 7.5% urea at 4°C (350 V) for 18 hr. The PCR product of exon 8 (380 bp), incorporating polymorphism G908R, was subjected to restriction fragment length polymorphism (RFLP) analysis using *HhaI* (Promega). The digested PCR products were resolved on a 10% polyacrylamide gel at room temperature (250V) for 30 minutes. Following electrophoresis, the gels were stained with ethidium bromide and the DNA visualised by ultraviolet light transillumination. Semi-automated DNA

sequencing was performed on PCR products demonstrating mobility or conformational variants in the polyacrylamide gels, using an ABI 3100 PRISM automated sequencer.

3.4.3.3. Statistical analysis

Allele and genotype frequencies were estimated by allele counting and statistical differences between patient and control groups were tested for significance by the Fisher exact test and/or chi-squared (χ^2) analysis with Yates' correction. The total carrier frequency, including heterozygotes and homozygotes, was counted to estimate the prevalence of gene variants. The prevalence of combined mutation carriers in patient and control subjects were compared to elucidate possible gene-gene interaction, according to Butt *et al* (2003). A probability value smaller than 0.05 was regarded as statistically significant. The Hardy-Weinberg equilibrium was performed to test equilibrium for the genetic traits investigated in the respective populations.

Allelic stratification applied in this study was based on previous findings (chapter 3.1), demonstrating an opposite allelic effect for allele 3 (similar to that of allele 2) upon iron loading when occurring in association with the $-237C{\rightarrow}T$ polymorphism. Therefore, when allele 3 occurred together with the T-allele of the $-237C{\rightarrow}T$ polymorphism, this allelic combination was grouped with allele 2 for further statistical analysis.

3.4.4. RESULTS

The three most common *CARD15* mutations implicated in Crohn's disease were screened for by using a combined HEX-SSCP method (R702W and 1007fs) and RFLP

analysis (G908R). Mutations R702W and 1007fs were observed in both the heterozygous and homozygous state, whereas the G908R mutation was identified only in a heterozygous state (Table 3.4.1). A single previously described mutation (A725G) (Lesage *et al* 2002) and one novel variant (A661P) were also identified in exon 4 of the *CARD15* gene. Figure 3.4.1A and B shows the direct sequencing analysis of the five different variants detected in the *CARD15* gene.

The A661P variant caused a G to C transversion, changing alanine to proline, and was identified in a heterozygous state in 1 of 100 (0.01%) Coloured control individuals. Variant A725G caused a C to G transversion, changing alanine to glycine and was previously detected at a frequency of 0.3% in CD patients and was absent in the controls (Lesage *et al* 2002). These mutations would probably not have a major effect on the properties of the protein as both amino acids represent neutral hydrophobic amino acids. Mutation A725G was identified in 4 of 35 (11%) ulcerative colitis individuals of the Coloured population in our study and was observed only in a heterozygous state. The A725G variant was absent in the Coloured CD patient group and population-matched control group as well as in both the Black and Caucasian patient (including CD and UC) and control groups. One UC patient with the A725G variant was also heterozygous for the R702W mutation. Unfortunately, additional family members of this individual was not available for genotype-phenotype association studies.

124

Table 3.4.1. Allele and genotype frequencies of the variants identified in the *CARD15* gene in three genetically distinct South African populations

Phenotype	R702W					G908R					1007fs					A725G				
	Genotype			Allele		Genotype			Allele		Genotype			Allele		Genotype			Allele	
	CC	CT	TT	C	T	GG	GC	CC	G	C	CC	InsC (he)	InsC (ho)	C	insC	CC	CG	GG	C	G
CCD (n=41)	0.88 (36)	0.12 (5)	-	0.94	0.06	1.0 (41)	-	-	1.0	-	0.98 (40)	0.02 (1)	-	0.99	0.01	1.0 (41)	-	-	1.0	-
CUC (n=35)	0.97 (34)	0.03 (1)	-	0.99	0.01	1.0 (35)	-	-	1.0	-	0.91 (32)	0.09 (3)	-	0.96	0.04	0.89 (31)	ᵃ0.11 (4)	-	0.94	ᵃ0.06
CC (n=100)	0.95 (95)	0.05 (5)	-	0.97	0.03	0.99 (99)	0.01 (1)	-	0.995	0.005	0.93 (93)	0.07 (7)	-	0.96	0.04	1.0 (100)	-	-	1.0	-
WCD (n=5)	0.80 (4)	0.20 (1)	-	0.90	0.10	1.0 (5)	-	-	1.0	-	0.80 (4)	-	0.20 (1)	0.80	0.20	1.0 (5)	-	-	1.0	-
WUC (n=11)	0.82 (9)	0.09 (1)	0.09 (1)	0.86	0.14	1.0 (11)	-	-	1.0	-	1.0 (11)	-	-	1.0	-	1.0 (11)	-	-	1.0	-
WC (n=58)	0.90 (52)	0.08 (5)	0.02 (1)	0.94	0.06	1.0 (58)	-	-	1.0	-	0.90 (52)	0.07 (4)	0.03 (2)	0.93	0.07	1.0 (58)	-	-	1.0	-
BUC (n=9)	0.89 (8)	0.11 (1)	-	0.94	0.06	1.0 (9)	-	-	1.0	-	0.89 (8)	0.11 (1)	-	0.94	0.06	1.0 (9)	-	-	1.0	-
BC (n=25)	0.92 (23)	0.08 (2)	-	0.96	0.04	1.0 (25)	-	-	1.0	-	1.0 (25)	-	-	1.0	-	1.0 (25)	-	-	1.0	-

CCD, Coloured Crohn's disease patients; CUC, Coloured ulcerative colitis patients, CC, Coloured control individuals, WCD, Caucasian Crohn's disease patients; WUC, Caucasian ulcerative colitis patients, WC, Caucasian control individuals, BUC, Black ulcerative colitis patients, BC, Black control individuals, he, heterozygote; ho, homozygote, ᵃCUC vs CC (allele frequency and genotype frequencies: $p < 0.004$). The number of individuals for each genotype is indicated in brackets.
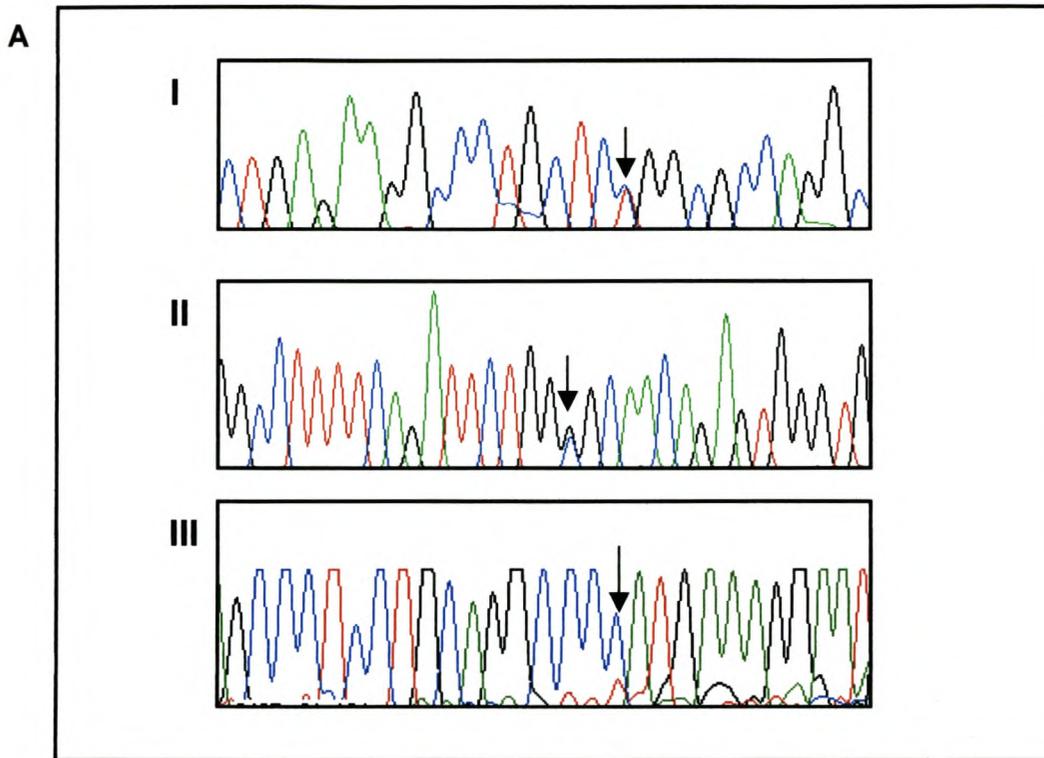
Figure 3.4.1.A

<u>**Legend to Figure 3.4.1.A**</u>
Direct sequence analysis of the three common *CARD15* gene mutations with i) sense strand of the R702W (2104 C→T) mutation, ii) sense strand of the G908R (2722 G→C) mutation and iii) sense strand of the 1007fs (3020insC) mutation (homozygous insertion); red, thymidine (T) blue, cytidine (C), green, adenosine (A), black, guanosine (G), arrows indicate point of variation.
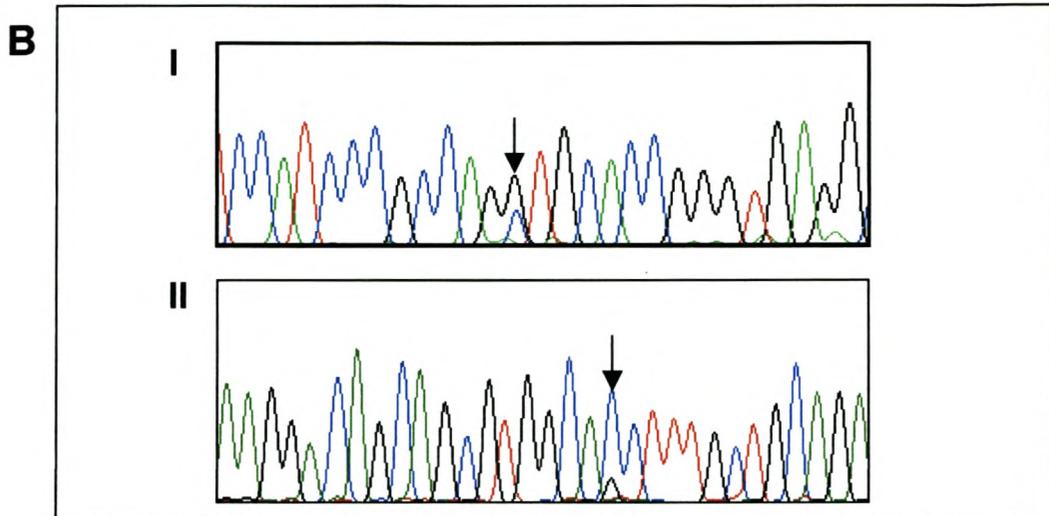
Figure 3.4.1.B

Legend to Figure 3.4.1.B
Direct sequence analysis of the two additional variants identified in exon 4 of the *CARD15* gene, with i) sense strand of variant A725G (2174 C→G) and ii) sense strand of variant A661P (1981 G→C), red, thymidine (T) blue, cytidine (C), green, adenosine (A), black, guanosine (G), arrows indicate point of variation.

127

The allele and genotype frequencies of the variants studied are provided in Table 3.4.1 for the different population groups. Mutations R702W, G908R and 1007fs were observed in 13%, 0% and 4% CD patients, 5%, 0% and 7% UC patients, and 7%, 0.5% and 7% of the control group, respectively. No statistically significant differences were observed for any of the common mutations (data not shown). However, when performing statistical analysis for the A725G variant, highly significant association was observed for both allele and genotype frequencies in the Coloured UC patient group compared with controls (allele frequency: $p < 0.004$, χ^2 with Yates' correction = 8.02; genotype frequency: $p < 0.004$, χ^2 with Yates' correction = 8.14). All the mutations studied were in Hardy Weinberg equilibrium in all the population groups, except for the 1007fs mutation in the healthy Caucasian control group ($p < 0.002$). Explanations for this finding are not clear since the control group was not selected for or against a particular genotype.

Observed and theoretical prevalence of combined mutation carriers in the patient and control groups are shown in Table 3.4.2. Analysis of the *CARD15* mutations was also investigated in combination with *SLC11A1* promoter variant genotypes (Table 3.4.3) as genetic background. No statistically significant associations were observed when possible combined effects of the *SLC11A1* promoter variants and three common *CARD15* mutations were investigated. A similar analysis was not performed for variant A725G since this mutation did not occur in the control population.

128

Assessment of the frequency of the −237C→T polymorphism in the patient group compared with controls having one or more of the three common mutations (R702W, G908R and 1007fs) or the A725G did not show any co-existence of these *CARD15* and *SLC11A1* variants in either the patient or control groups.

Table 3.4.2. Theoretical and observed prevalence for combined carriers of various mutations identified in the *CARD15* gene and *SLC11A1* promoter variants identified in three genetically distinct South African populations

Phenotype	Observed prevalence for combined genotype frequencies (%)						Theoretical prevalence for combined carriers (Observed prevalence for combined carriers)							
	CARD15				*SLC11A1*									
	R702W	G908R	1007fs	A725G	*Allele 2	*Allele 3	R702W/ Allele 2	R702W/ Allele 3	G908R/ Allele 2	G908R/ Allele 3	1007fs/ Allele 2	1007fs/ Allele 3	A725G/ Allele 2	A725G/ Allele 3
CCD (n=41)	12%	0%	2%	0%	37%	95%	4.4% (4.9%)	11.4% (12.2%)	0% (0%)	0% (0%)	0.7% (0%)	1.9% (2.4%)	0% (0%)	0% (0%)
CUC (n=35)	3%	0%	9%	11%	54%	97%	1.6% (2.9%)	2.9% (2.9%)	0% (0%)	0% (0%)	4.9% (2.9%)	8.7% (8.6%)	5.9% (5.7%)	10.7% (11.4%)
CC (n=100)	5%	1%	7%	0%	52%	92%	2.6% (1.0%)	4.6% (5.0%)	0.51% (1.0%)	0.91% (1.0%)	3.6% (2.0%)	6.4% (7.0%)	0% (0%)	0% (0%)
WCD (n=5)	20%	0%	20%	0%	40%	100%	8.0% (20.0%)	20.0% (20.0%)	0% (0%)	0% (0%)	8.0% (0%)	20.0% (20.0%)	0% (0%)	0% (0%)
WUC (n=11)	18%	0%	0%	0%	54%	100%	9.7% (9.1%)	18.0% (18.2%)	0% (0%)	0% (0%)	0% (0%)	0% (0%)	0% (0%)	0% (0%)
WC (n=58)	10%	0%	10%	0%	47%	90%	4.7% (1.7%)	9.0 (10.3%)	0% (0%)	0% (0%)	4.7% (6.9%)	9.0% (10.3%)	0% (0%)	0% (0%)
BUC (n=9)	11%	0%	11%	0%	22%	100%	2.4% (0%)	11.0% (11.1%)	0% (0%)	0% (0%)	2.4% (11.1%)	11.0% (11.1%)	0% (0%)	0% (0%)
BC (n=25)	8%	0%	0%	0%	44%	100%	3.5% (4.0%)	8.0% (8.0%)	0% (0%)	0% (0%)	0% (0%)	0% (0%)	0% (0%)	0% (0%)

*T-allele of the −237C→ T polymorphism in the presence of allele 3 was grouped with allele 2 based on *in vitro* allelic effect (chapter 3.1), CCD, Coloured Crohn's disease patients; CUC, Coloured ulcerative colitis patients, CC, Coloured control individuals, WCD, Caucasian Crohn's disease patients; WUC, Caucasian ulcerative colitis patients, WC, Caucasian control individuals, BUC, Black ulcerative colitis patients, BC, Black control individuals,

Table 3.4.3. Genotype frequencies of the various *SLC11A1* promoter variants in three genetically distinct South African populations.

Promoter alleles	GENOTYPE FREQUENCIES							
	CC (n=100)	CCD (n=41)	CUC (n=35)	WC (n=58)	WCD (n=5)	WUC (n=11)	BC (n=25)	*BUC (n=9)
Allele 2; Allele 2	0.03 (3)	0.05 (2)	0.03 (1)	0.02 (1)	-	-	-	-
Allele 2; Allele 3	0.34 (34)	0.32 (13)	0.43 (15)	0.36 (21)	0.4 (2)	0.36 (4)	0.32 (8)	0.22 (2)
Allele 3; Allele 3	0.62 (62)	0.63 (26)	0.54 (19)	0.58 (34)	0.6 (3)	0.64 (7)	0.68 (17)	0.78 (7)
Allele 1; Allele 3	-	-	-	0.02 (1)	-	-	-	-
Allele 5; Allele 3	0.01 (1)	-	-	0.02 (1)	-	-	-	-
-237CC	0.82 (82)	1.0 (41)	0.91 (32)	0.88 (51)	1.0 (5)	0.82 (9)	0.96 (22)	1.0 (9)
-237CT	0.17 (17)	-	0.09 (3)	0.10 (6)	-	0.18 (2)	0.04 (3)	-
-237TT	0.01 (1)	-	-	0.02 (1)	-	-	-	-
Allele 2; Allele 3; -237C/T	0.03 (3)	-	-	0.03 (2)	-	-	-	-
Allele 3; Allele 3; -237C/T	0.14 (14)	-	0.09 (3)	0.07 (4)	-	0.18 (2)	0.04 (3)	-
Allele 3; Allele 3; -237T/T	0.01 (1)	-	-	0.02 (1)	-	-	-	-

*No individuals in the black patient group presented with CD in this study, CC, coloured control group, CCD, coloured Crohn's disease patient group; CUC, coloured ulcerative colitis patient group, WC, Caucasian control group, WCD, Caucasian Crohn's disease patient group, WUC, Caucasian ulcerative patient group, BC, black control group, BUC, black ulcerative colitis patient group. The number of individuals for each genotype is indicated in brackets.

131

Table 3.4.4. Assessment of the various *SLC11A1* variants in relation to the common *CARD15* mutations

*SLC11A1 Promoter alleles	CARD15/NOD2 mutation	GENOTYPE FREQUENCIES							
	1007fs	CC (n=100)	CCD (n=41)	CUC (n=35)	WC (n=58)	WCD (n=5)	WUC (n=11)	BC (n=25)	*BUC (n=9)
Allele 2; Allele 3	C	0.43 (43)	0.32 (13)	0.49 (17)	0.33 (19)	0.40 (2)	0.55 (6)	0.44 (11)	0.11 (1)
	insC (he)	0.02 (2)	-	0.03 (1)	0.05 (3)	-	-	-	0.11 (1)
	insC (ho)	-	-		0.02 (1)	-	-	-	-
Allele 3; Allele 3	C	0.42 (42)	0.61 (25)	0.40 (14)	0.47 (27)	0.40 (2)	0.45 (5)	0.56 (14)	0.78 (7)
	insC (he)	0.05 (5)	0.02 (1)	0.06 (2)	0.02 (1)	-	-	-	-
	insC (ho)	-		-	0.02 (1)	0.20 (1)	-	-	-
	G908R								
Allele 2; Allele 3	GG	0.44 (44)	0.32 (13)	0.51 (18)	0.40 (23)	0.40 (2)	0.55 (6)	0.44 (11)	0.22 (2)
	GC	0.01 (1)	-	-	-	-	-	-	-
	R702W								
Allele 2; Allele 3	CC	0.44 (44)	0.27 (11)	0.49 (17)	0.38 (22)	0.20 (1)	0.45 (5)	0.40 (10)	0.22 (2)
	CT	0.01 (1)	0.05 (2)	0.03 (1)	0.02 (1)	0.20 (1)	-	0.04 (1)	-
	TT	-	-	-	-	-	0.09 (1)	-	-
Allele 3; Allele 3	CC	0.43 (43)	0.56 (23)	0.46 (16)	0.41 (24)	0.60 (3)	0.36 (4)	0.52 (13)	0.67 (6)
	CT	0.04 (4)	0.07 (3)	-	0.05 (3)	-	0.09 (1)	0.04 (1)	0.11 (1)
	TT	-	-	-	0.02 (1)	-	-	-	-
Allele 3; Allele 5	CC	0.01 (1)	-	-	-	-	-	-	-
	CT	-	-	-	0.02 (1)	-	-	-	-
	A725G								
Allele 2; Allele 3	CC	0.45 (45)	0.32 (13)	0.46 (16)	0.40 (23)	0.40 (2)	0.55 (6)	0.44 (11)	0.22 (2)
	CG	-	-	0.06 (2)	-	-	-	-	-
	GG	-	-	-	-	-	-	-	-
Allele 3; Allele 3	CC	0.47 (47)	0.63 (26)	0.40 (14)	0.50 (29)	0.60 (3)	0.45 (5)	0.56 (14)	0.78 (7)
	CG	-	-	0.06 (2)	-	-	-	-	-
	GG	-	-	-	-	-	-	-	-

*Only genotypes in association with the *CARD15* variants are denoted and T-allele of the $-237C{\rightarrow}T$ polymorphism in the presence of allele 3 was grouped with allele 2 based on *in vitro* allelic effect, [#]No individuals in the Black patient group presented with CD in this study, CC, Coloured control group, CCD, Coloured Crohn's disease patient group; CUC, Coloured ulcerative colitis patient group, WC, Caucasian control group, WCD, Caucasian Crohn's disease patient group, WUC, Caucasian ulcerative patient group, BC, Black control group, BUC, Black ulcerative colitis patient group. The number of individuals for each genotype is indicated in brackets. Allele 2, $t(gt)_5ac(gt)_5ac(gt)_{10}ggcaga(g)_6$, Allele 3, $t(gt)_5ac(gt)_5ac(gt)_9ggcaga(g)_6$, Allele 5, $t(gt)_4ac(gt)_5ac(gt)_{10}ggcaga(g)_6$

3.4.5. DISCUSSION

Association of the *CARD15* gene with CD susceptibility has previously been identified in several studies. By investigating the gene regions spanning the three most common *CARD15* mutations in the diverse South African population, we additionally identified one novel variant (A661P) and a previously described variant (A725G) (Lesage *et al* 2002). The common mutations were identified at low frequencies in our study cohort and was present in only 14.9% of the patient group [8 of 46 (17.4%) CD patients and 7 of 55 (12.7%) UC patients]. No statistically significant differences were observed for these mutations between patient and control groups. In a previous study, Yamazaki *et al* (2002) investigated the three common mutations in Japanese CD patients and also found no association with these mutations. A recent study also failed to demonstrate an association between these variants and CD susceptibility in the Dutch population (Stokkers *et al* 2003).

Yamazaki *et al* (2002) suggested that the prevalent *CARD15* mutations are independent risk factors for CD in Caucasian patients. The Caucasian contribution to the South African Coloured population is only approximately 33% (Loubser *et al* 1999), and therefore the incidence of mutations in this population is expected to be a third of the incidence estimated in the South African Caucasian population. Although no significant associations were achieved for the R702W, G908R and 1007fs mutations in our study cohort, with the majority of individuals from the Coloured population, association cannot be confirmed or excluded for the Caucasian and Black populations because of a small sample group. We would, however, expect these mutations to contribute to CD in the

133

Caucasian population but possibly not in the Black population. In future studies, these population groups need to be expanded to elucidate the involvement of the *CARD15* gene in CD susceptibility in the general South African population, within the context of the other candidate genes such as *SLC11A1* (Kojima *et al* 2002, chapter 3.3) and *TNF-α* (Ruuls and Sedgwick 1999).

The previously described variant, A725G, was observed mainly in the Coloured UC patients. In a previous study, this variant was identified in only 3 of 906 (0.3%) CD patients from European descent and was absent in 318 UC patients and 206 controls (Lesage *et al* 2002). In our study, the A725G variant was identified in 4 of 35 UC patients of the Coloured population and statistically significant differences were achieved (p<0.004, χ^2 with Yates' correction = 8.02). The likelihood that this variant may be in linkage disequilibrium with a functional mutation in the *CARD15* gene or other disease-causing loci should be considered, as this mutation was considered to be a benign polymorphism.

Although no linkage with the *CARD15* gene and UC has previously been reported, the involvement of this gene with UC remains a possibility. In a recent study by Van Heel *et al* (2002), linkage could not be confirmed by microsatellite linkage and linkage disequilibrium mapping to the *CARD15* gene with CD susceptibility even though association was observed with the disease-causing alleles of the gene in the population studied. These authors and others highlighted the difficulties inherited in fine-mapping disease genes for complex traits (Risch and Merikangas 1996, Lernmark and Ott 1998)

and this should be considered before excluding *CARD15* as a susceptibility gene for UC. Here we found significant association of the gene with UC susceptibility although no linkage could be confirmed for UC and *CARD15* in previous studies. It is, however, possible that unique precipitating factors present in the Coloured population of South Africa converts the disease state from a genetic predisposition for CD to UC, but the nature of this proposed mechanism is unknown. In light of the multifactorial nature of CD, it seems likely that the *CARD15* gene plays a less important role in relation to other genetic and environmental factors (e.g. smoking, infectious agents), which may be required to express the disease in South African patients.

The difficulty of identifying genes involved in complex conditions has also been highlighted in the earlier study of the *SLC11A1* gene performed in the same study population (chapter 3.3). We have shown that variation in the promoter region of the *SLC11A1* gene may modify risk of IBD, despite the fact that no linkage to this chromosomal region on 2q has previously been found in genome-wide screens.

Although no significant associations were observed when possible combined effects of the *SLC11A1* promoter variants and mutations of *CARD15* were investigated, marginal significant associations were observed for combined carriers of R702W and allele 2 of *SLC11A1* ($p < 0.08$) and combined carriers of R702W and allele 3 ($p < 0.08$) for CD in the Coloured population. Because of the low frequency observed for the three common *CARD15* mutations, the study population needs to be expanded to further investigate/confirm the possible interaction between *SLC11A1* and *CARD15* implicated

by this analysis. Sub classification of our study population for disease type/behaviour (including fistulising, stricturing and inflammation) and location of disease (including colonic, enterocolonic and small bowel) could also give a better understanding of the contribution of these genes to CD susceptibility. A recent study has shown that variant 1007fs is associated with stricturing, G908R with inflammation and R702W with fistulising (Sambuelli *et al* 2003).

The *CARD15* gene needs to be screened for other variants that may also be involved in IBD susceptibility because linkage to chromosome 16 is not completely explained by the three disease-associated mutations (Van Heel *et al* 2001, Lesage *et al* 2002). Linkage has also been found with other regions of chromosome 16 as demonstrated by Hampe *et al* (2002). These authors confirmed the importance of the *CARD15* locus in CD and identified another susceptible IBD locus on the proximal region of 16p. Different areas of linkage were also previously detected on chromosome 16 (Hugot *et al* 1996, Curran *et al* 1998, Brant *et al* 1998, Annese *et al* 1999, Cavanaugh 2001, Cavanaugh *et al* 1998, Cho *et al* 1998), suggesting association of other genes on chromosome 16.

The unexpected association of a *CARD15* mutation (A725G) with UC, that we have demonstrated, further contributes to the contradictory results obtained over various studies and highlights the complexity of IBD. One common denominator recurs in all these studies, i.e. that multiple genes interact with environmental factors to contribute to the IBD phenotype.

136

3.4.6. ACKNOWLEDGEMENTS

3.5. Analysis of genes implicated in iron regulation in individuals presenting with primary iron overload

3.5.1. ABSTRACT

Extensive investigation into the molecular basis of iron overload disorders have provided new insight into the complexity of iron metabolism, known to be interconnected with various cellular pathways. The possible involvement of genes related to iron homeostasis, including *HFE*, *SLC11A1*, *SLC11A3*, *HAMP* and *DCYTB*, were investigated in individuals who were referred for confirmation or exclusion of a molecular diagnosis of haemochromatosis, but who tested negative or were heterozygous for the most frequent *HFE* mutation, C282Y. Denaturing high performance liquid chromatography (dHPLC) analysis of the *HFE, SLC11A3, HAMP* and *DCYTB* genes revealed a unique spectrum of mutations in the South African study population, including 67 unrelated South African patients and 50 population-matched controls. A case-control study performed in 81 unrelated patient referrals and 91 control individuals from the United Kingdom furthermore confirmed the significance of sequence variation in the promoter region of the *SLC11A1* gene in iron-related disorders. Two novel *DCYTB* gene mutations, R226H and IVS1-4C→G, not detected in the control population, could possibly explain the non-*HFE* iron overload in 11% of South African Caucasian patient referrals. In Africans with iron overload not related to the *HFE* gene, the possible involvement of the *SLC11A3* and *DCYTB* genes were demonstrated for the first time. This study confirms the genetic heterogeneity of haemochromatosis and highlights the significance of *DCYTB* mutations in relation to iron overload.

3.5.2. INTRODUCTION

Hereditary haemochromatosis (HH) is generally an autosomal recessive condition and is characterised by iron overload, primarily in parenchymal cells (Bothwell *et al* 1995). Iron accumulation usually results in tissue damage and causes cirrhosis of the liver, diabetes mellitus, athropathy, cardiomyopathy, endocrine abnormalities and hepatocellular carcinoma which is associated with HH (Powell *et al* 1994, Bothwell *et al* 1995).

Mutations in the *HFE* gene, with homozygosity for the C282Y mutation being the most common disease genotype, have been identified as the cause of HH in over 80% of individuals of European descent with this disorder (Feder *et al* 1996, Worwood *et al* 1997). Other mutations in the *HFE* gene account for the disease phenotype in approximately 2-10% of cases (reviewed by Pointon *et al* 2000). Non-*HFE* related forms of HH are less common but have also been reported and include mutations in the *TFR2* (Camaschella *et al* 2000, Roetto *et al* 2001), *SLC11A3* (Njajou *et al* 2001, Montosi *et al* 2001) and *HAMP* (Roetto *et al* 2003) genes. A complex spectrum of phenotypic expression has been observed for individuals with HH (Adams *et al* 1997, Rhodes *et al* 1997, de Villiers *et al* 1999, Milani and Kotze 1999, Sachot *et al* 2001) and a recent study showed that only 1% of individuals present with clinical expression of the disease (Beutler 2003), which raised the possibility of other modifier genes contributing to the clinical variability observed in HH.

In this study genes involved in iron overload, including *HFE* and *HAMP*, and genes involved in iron storage and transport, including *SLC11A3* and *DCYTB*, were analysed in patients referred for a molecular diagnosis of HH in the South African population.

139

Possible involvement of the *SLC11A1* gene involved in iron transport was investigated in patients with similar criteria from the United Kingdom. These individuals comprised a subpopulation of individuals who either tested negative for the HFE mutation C282Y, or who were heterozygous for this common mutation underlying HH.

3.5.3. MATERIALS AND METHODS

3.5.3.1. Subjects

The South African study population initially included 161 apparently unrelated patients who were referred for HH mutation screening based on abnormal iron parameters (Bacon and Sadiq 1997) in the absence of secondary causes for elevated ferritin and transferrin saturation levels. After exclusion of all C282Y homozygous patients, a subpopulation including 58 (86.6%) individuals without mutation C282Y and 9 (13.4%) C282Y heterozygotes were selected for further analysis. The patients and controls were from the South African Black (11 patients and 20 control individuals) and Caucasian (56 patients and 50 control individuals) populations, with "Caucasian" referring to an individual of European descent, mainly Dutch, French, German and British origin and "Black" referring to South Africans of central African descent.

Blood samples were additionally obtained from 81 unrelated patients from the United Kingdom with elevated serum iron and/or characteristics of organ damage compatible with HH. These included 59 (73%) males and 22 (27%) females that were C282Y negative [61 of 81 (75.3%)] or C282Y heterozygous [20 of 81 (24.7%)], who were referred to the MRC Molecular Haematology Unit in the United Kingdom for genetic testing. DNA samples of 91 unrelated population-matched control individuals, including

41 (45%) males and 50 (55%) females were available for comparative analysis. Iron status was unknown in these healthy control individuals.

3.5.3.2. Methods

i) dHPLC analysis of the HFE, SLC11A3, HAMP and DCYTB genes

DNA was extracted from whole blood using a standard method (Miller *et al* 1988). The various exons of *SLC11A3* (Njajou *et al* 2001), *HFE*, *HAMP* and *DCYTB* (A Merryweather-Clarke, unpublished data) genes were amplified by polymerase chain reaction (PCR) amplification using intronic oligonucleotide primers. The samples were heteroduplexed by heating at 95°C followed by cooling at 2°C per minute and analysed by denaturing high-performance liquid chromatography (dHPLC) on the WAVE DNA Fragment Analysis System (Transgenomic). Semi-automated DNA sequencing was performed on PCR products demonstrating variation on the chromatograms using an ABI 3100 PRISM automated sequencer.

ii) Mutation analysis of the promoter region of the SLC11A1 gene

PCR amplification was performed using oligonucleotide primers spanning the 5'-(GT)n repeat polymorphism and a point mutation at base −237 in the promoter region of the *SLC11A1* gene, generating fragments of 116-bp and 467-bp, respectively (Kotze *et al* 2001). The 467-bp PCR product of the promoter region was subjected to a combined heteroduplex single-strand conformation polymorphism (HEX-SSCP) method (Kotze *et al* 1995) for mutation detection. These PCR products were resolved on a 12% polyacrylamide gel supplemented with 7.5% urea at 4°C (350 V) for 18 hr. The 116-bp

141

PCR product of the promoter region was subjected to restriction fragment length polymorphism (RFLP) analysis (Graham *et al* 2000). These products were digested with *Rsal* to discriminate between the various Z-DNA forming promoter repeat polymorphic alleles. The digested PCR product was resolved on a 10% polyacrylamide gel at room temperature (250V) for 30 minutes. Following electrophoresis, the gels were stained with ethidium bromide and the DNA visualised by ultraviolet light transillumination.

3.5.3.3. Statistical analysis

Allele and genotype frequencies were estimated by allele counting and statistical differences between patient and control groups were tested for significance by the Fisher exact test and/or chi-squared (χ^2) analysis with Yates' correction. A probability value smaller than 0.05 was regarded as statistically significant. The Hardy-Weinberg equilibrium was performed to test equilibrium for the genetic traits investigated in the respective populations.

3.5.4. RESULTS

Five genes previously implicated in iron homeostasis (*HFE* and *HAMP*) or iron transport and storage (*SLC11A1*, *SLC11A3* and *DCYTB*) were investigated in patients presenting with primary iron overload that could not be ascribed to homozygosity for the common *HFE* C282Y mutation. Mutation analysis of *HFE*, *SLC11A3*, *HAMP* and *DCYTB*, in the South African population revealed several mutations, including promoter, exonic and intronic variants. Chromatograms and sequencing analysis of the novel variants identified are shown in Figures 3.5.1, 3.5.2 and 3.5.3.

142

Figure 3.5.1.(A)

Figure 3.5.1.(B)

144

Direct sequencing analysis of the (CGG)n repeat

(CGG)$_7$

(CGG)$_8$

* (CGG)$_9$ *
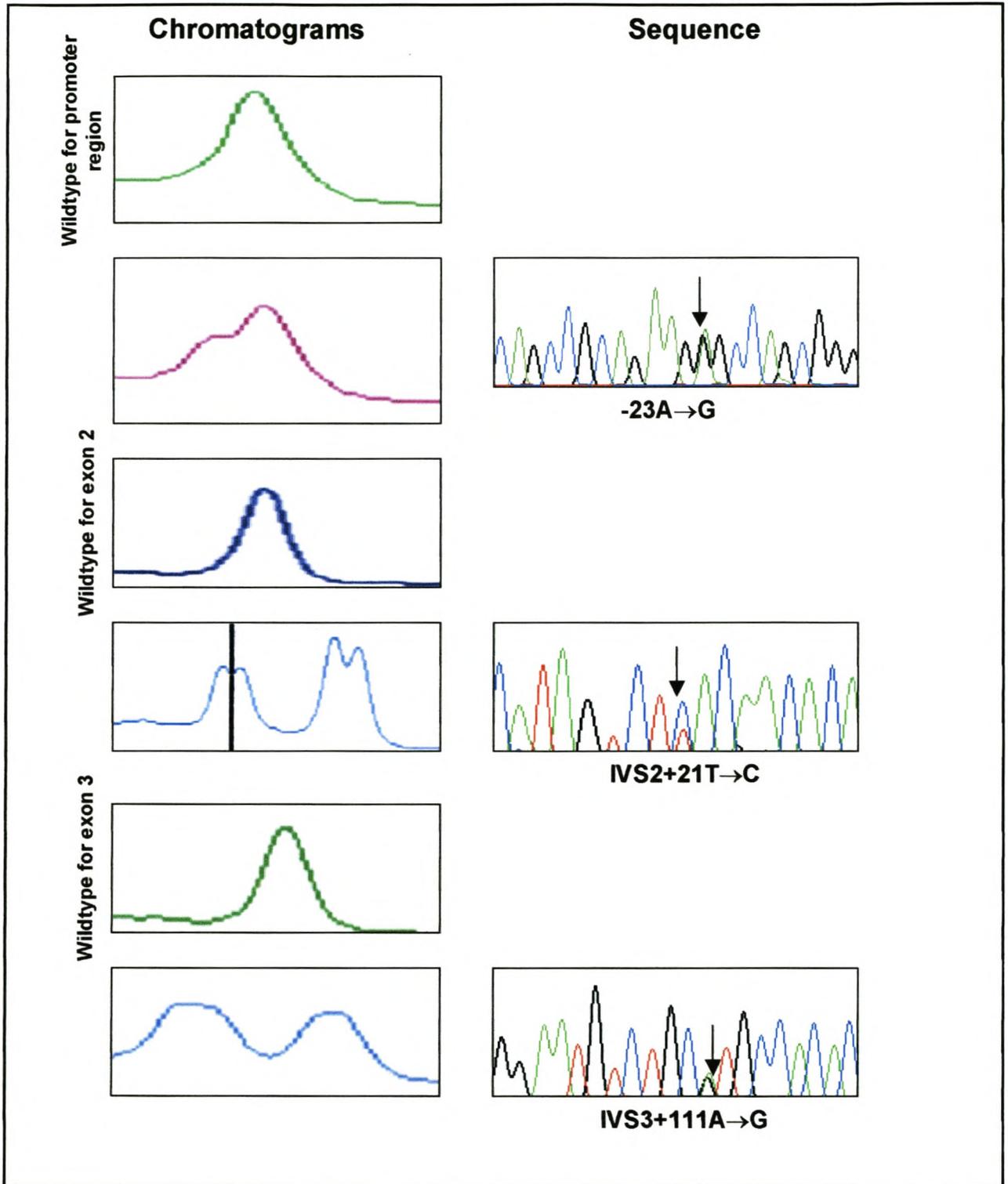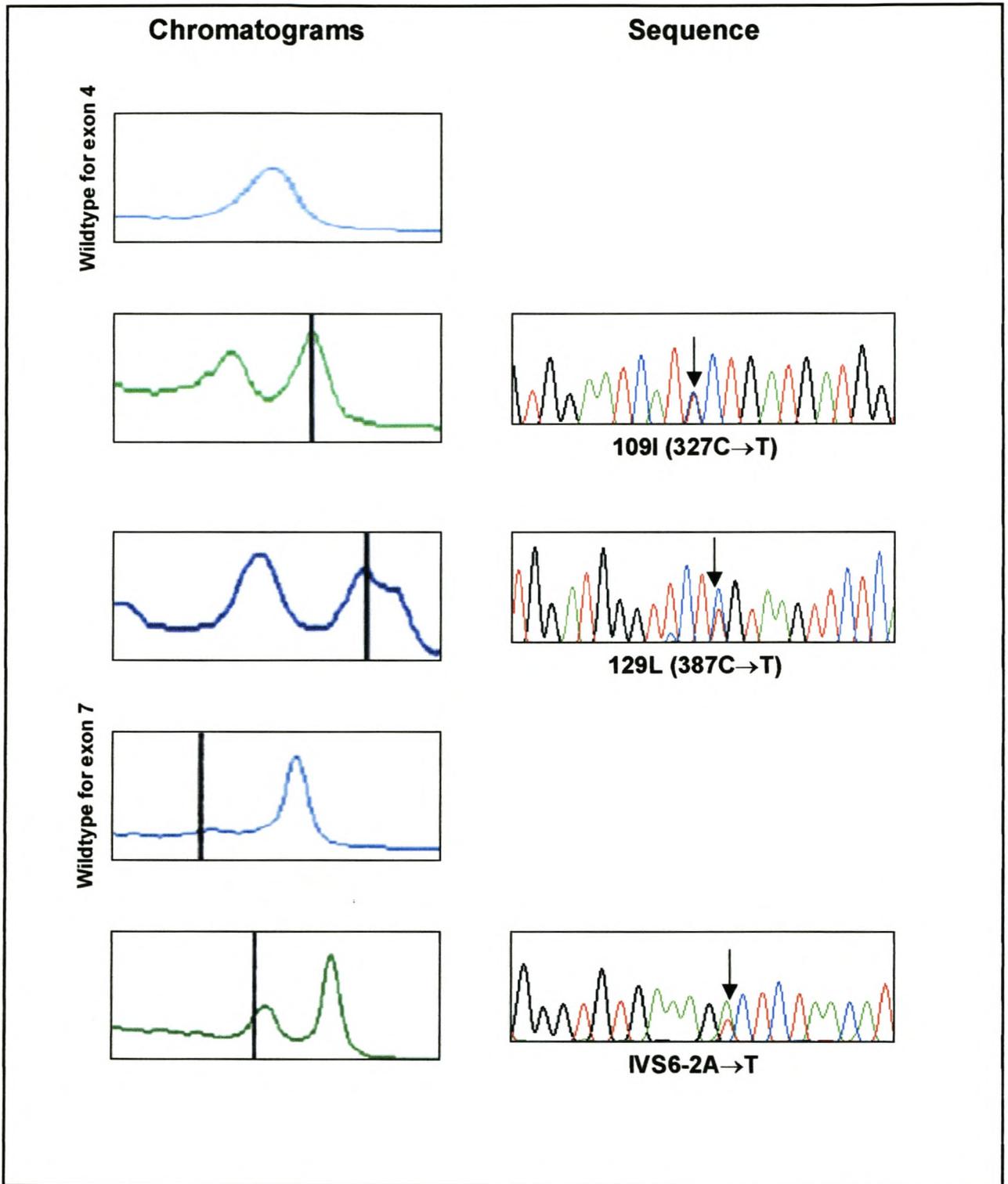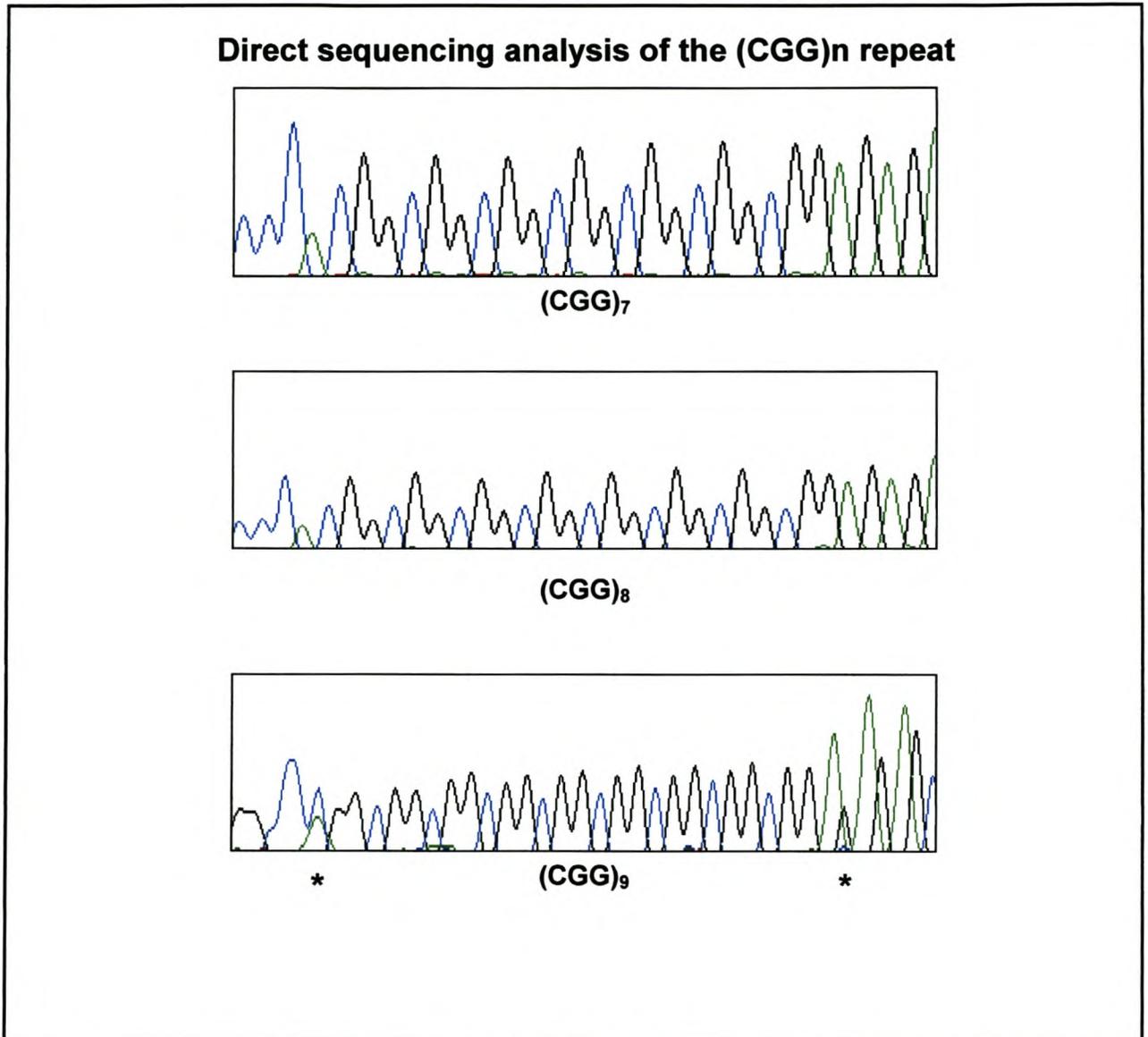
Figure 3.5.1.(C)

Legend to Figure 3.5.1.
A and B: Chromatograms and sequencing analysis of the novel variants identified in the *SLC11A3* gene, arrows indicate points of variation **C:** Sequencing analysis of the (CGG)n repeat in the promoter region of the *SLC11A3* gene, *Artefacts due to mobility shifts, red, thymidine (T) blue, cytidine (C), green, adenosine (A), black, guanosine (G).
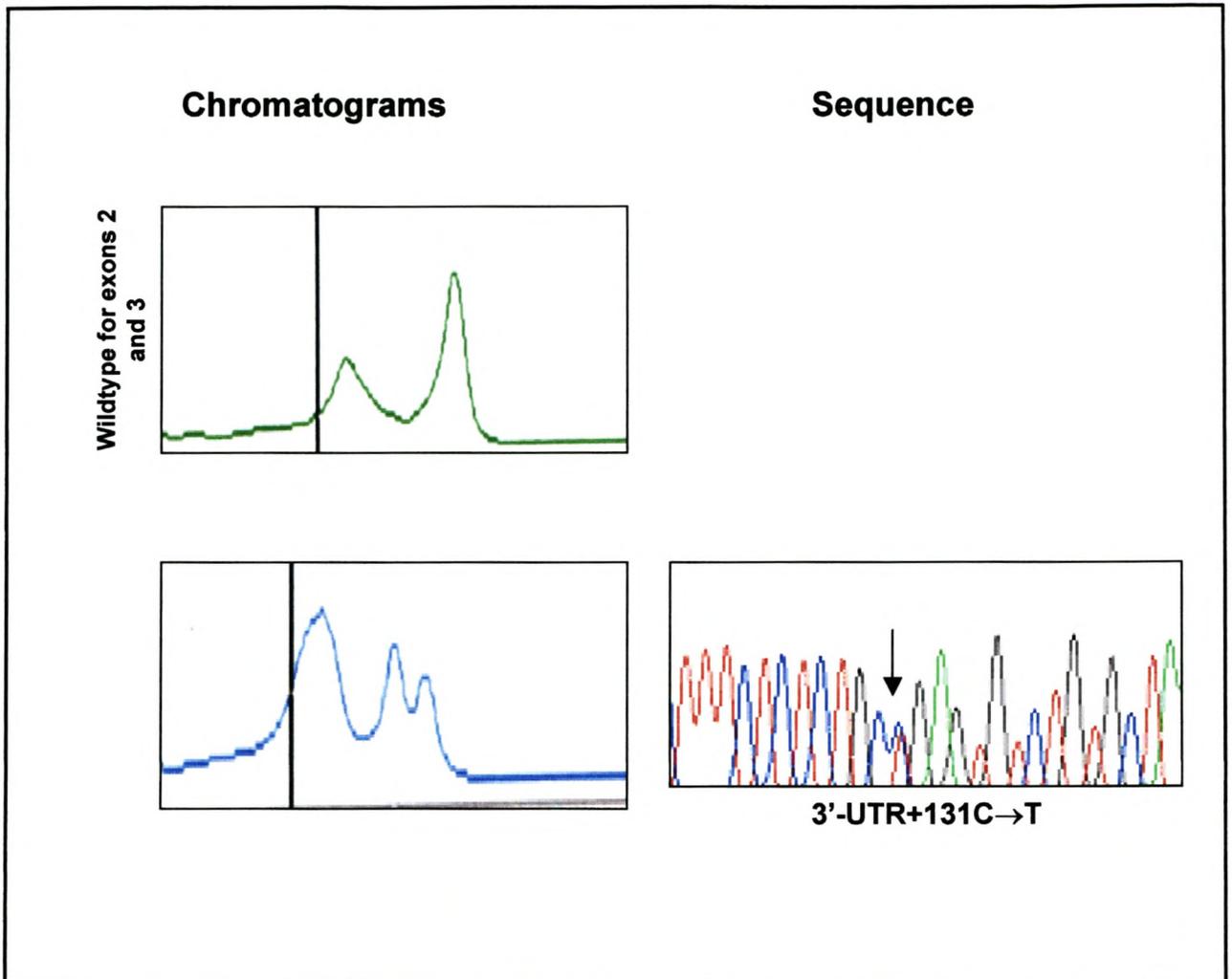
Figure 3.5.2

Legend to Figure 3.5.2.

Chromatograms and sequencing analysis of the novel variant identified in the *HAMP* gene. Arrow indicates point of variation, red, thymidine (T) blue, cytidine (C), green, adenosine (A), black, guanosine (G).

146
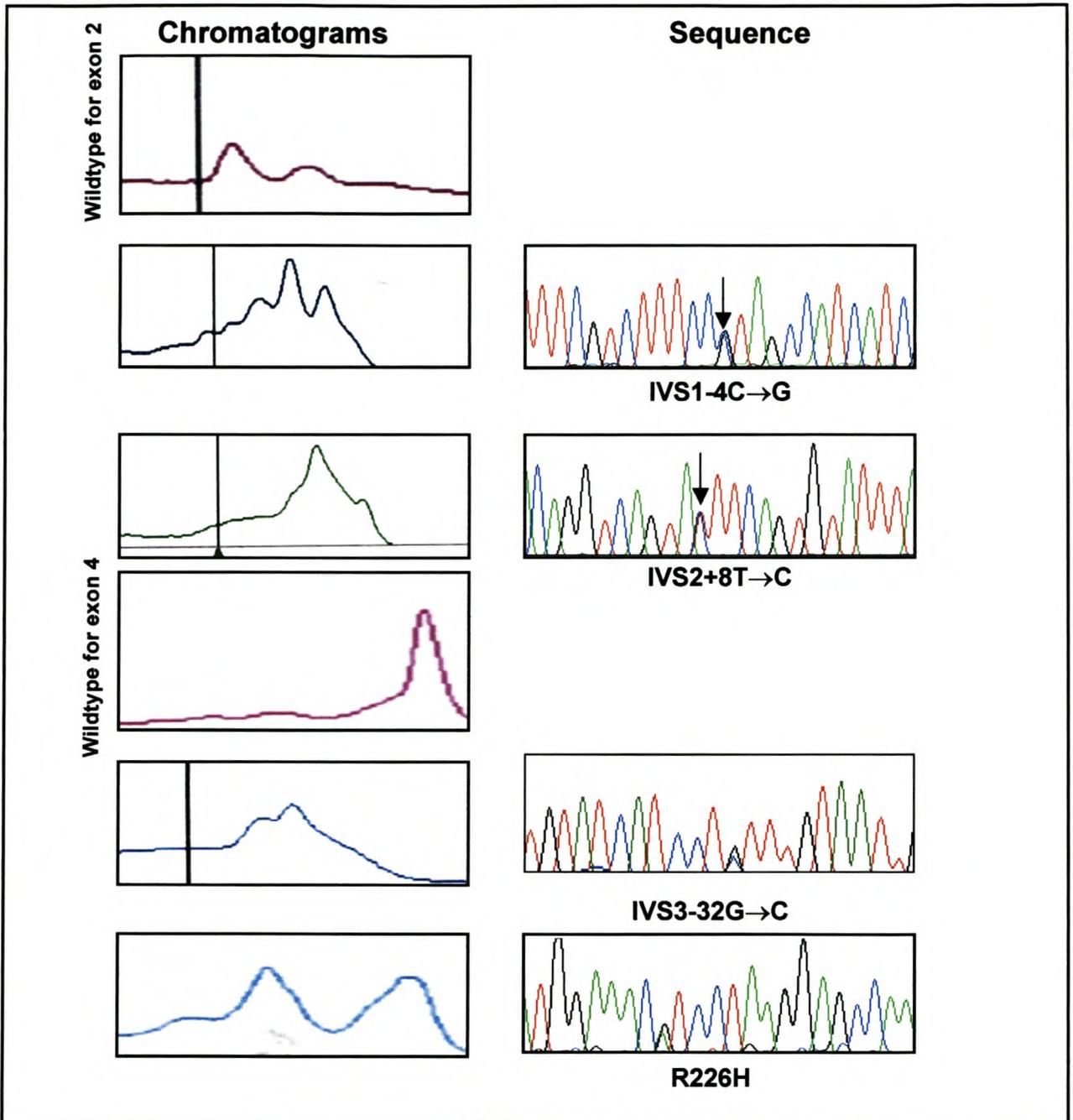
Figure 3.5.3.

Legend to Figure 3.5.3.

Chromatograms and sequencing analysis of the novel variants identified in the *DCYTB* gene. Arrows indicates points of variation, red, thymidine (T) blue, cytidine (C), green, adenosine (A), black, guanosine (G).

The variants identified in both the patient and control groups are shown in Table 3.5.1, together with their allele frequencies. Statistically significant differences detected between patient and control populations are indicated in the table. The Hardy-Weinberg equilibrium was determined within each patient subpopulation versus the respective control groups to assess possible effects of the selection criteria (exclusion of C282Y homozygotes) from the study population. The results obtained demonstrated that all the groups complied with the Hardy Weinberg equilibrium, except for the *DCYTB* IVS2+8T/C polymorphism ($p<0.02$) in the Caucasian patient group. This may be due to the fact that a subpopulation was studied, after exclusion of C282Y homozygotes.

HFE gene

Mutation analysis of the *HFE* gene revealed only previously described variants (Rochette *et al* 1999, Höhler *et al* 1999, de Villiers *et al* 1999, Beutler and West 1997) (Table 3.5.1). Heterozygosity for the C282Y mutation was identified in 9 of 56 (16.1%) Caucasian patients and 6 (10.7%) of these patients also had the H63D mutation. The S65C mutation was present in 3 of 56 (5.4%) Caucasian patients and was absent in the respective control group. One individual with the S65C mutation also had the H63D mutation.

Table 3.5.1. Allele frequencies of the variants identified in the South African population.

GENE	EXON/ INTRON	VARIANT	*ALLELE FREQUENCIES		P	*ALLELE FREQUENCIES		P
			WHH (2n=112)	WC (2n=100)		BHH (2n=22)	BC (2n=40)	
HFE	2	[a]IVS2+4T→C	0.39	0.31	ns	0.18	0.43	[e]*0.05*
	2	[b]H63D	0.17	0.15	ns	-	-	ns
	4	[a]IVS4-44T→C	0.10	0.06	ns	0.18	0.03	[d]*0.04*
	4	[a]IVS4-50A→G	0.06	0.15	ns	0.14	1.00	[e]*<0.0001*
	5	[a]IVS5-47G→A	0.50	0.46	ns	0.45	0.65	ns
SLC11A3	5'-UTR	[a](CGG)$_8$	0.65	0.78	ns	0.23	0.10	0.08
	5'-UTR	[a](CGG)$_9$	-	-	-	0.27	0.12	0.07
	5'-UTR	[a]-23A→G	-	-	-	0.36	0.03	[d]*0.0007*
	5'-UTR	[b]-8C→G	0.15	0.15	ns	-	0.08	ns
	5'-UTR	[b]-98G→C	0.20	0.15	ns	-	0.08	ns
	1	[a]IVS1-24G→C	0.84	0.88	ns	0.82	0.28	[d]*<0.0001*
	2	[b,c]IVS2+21T→C	-	-	-	0.04	-	ns
	3	[b,c]IVS3+111A→G	-	-	-	0.04	-	ns
	4	[b]I109	0.008	-	ns	-	0.03	ns
	4	[b]L129	-	-	-	0.23	0.05	[d]*0.04*
	6	[b]V221	0.21	0.08	[d]**0.03**	0.14	-	[d]*0.04*

Table 3.5.1. Allele frequencies of the variants identified in the South African population (Continued).

GENE	EXON/ INTRON	VARIANT	*ALLELE FREQUENCIES		P	*ALLELE FREQUENCIES		P
			WHH (2n=112)	WC (2n=100)		BHH (2n=22)	BC (2n=40)	
DCYTB	2	[a]IVS2+8T→C	0.67	0.64	ns	0.77	0.88	ns
	4	[a]S266N	0.76	0.86	ns	0.77	0.88	ns

*Allele frequency of polymorphic allele denoted, [a]Identified in both a heterozygous and homozygous state, [b]Identified only in a heterozygous state, [c]Rare variants, [d]Statistically significant over representation of the polymorphic allele in the patient compared to the control groups, [e]Statistically significant under representation of the polymorphic allele in the patient compared to the control groups, WHH, Caucasian haemochromatosis patient group, WC, Caucasian control group, BHH, Black haemochromatosis patient group, BC, Black control group, P, probability value, ns, not significant.

150

The C282Y, H63D and S65C mutations were absent in the Black population and only previously described polymorphisms were identified in this population group (Table 3.5.1). Statistically significant associations were, however, observed for various polymorphisms in the Black patient group compared with the population-matched control group. These associations were observed when comparing allele frequencies for IVS2+4T→C (p<0.05, χ^2 with Yates' correction = 2.59), and IVS4-44T→C (p<0.04, χ^2 with Yates' correction = 2.83). A statistically significant association was observed with IVS4-50A→G for both allele (p<0.0001, χ^2 with Yates' correction = 43.35) and genotype (p<0.0001, χ^2 with Yates' correction = 30.00) frequencies. The variants IVS2+4T→C and IVS4-50A→G were under represented and variant IVS4-44T→C was over represented in the Black patient group compared with the population-matched control group. This finding may be indicative of linkage disequilibrium with functional polymorphisms in the gene affecting expression of the *HFE* gene.

SLC11A3 gene

Several previously described polymorphisms were identified and are presented in Table 3.5.1. Seven novel variants were also identified: 1) an A to G substitution at position –23 in the 5'-UTR region of the gene, 2) nine CGG repeats [(CGG)$_9$] in the 5'-UTR region of the gene, 3) a T to C transition within the second intron (IVS2+21), 4) an A to G transition within the third intron (IVS3+111), 5) a C to T substitution at nucleotide position 327, involving amino acid 109 (I109), 6) a C to T substitution at nucleotide position 387, involving amino acid 129 (L129) and 7) a splice variant involving an A to T transversion within the sixth intron (IVS6-2). No clear correlation could be made

151

between the CGG repeat in the promoter region and chromatograms and it was decided to perform direct sequencing analysis for the patient and control groups on this region. The majority of these variants were identified only in the Black South African population, including $(CGG)_9$, -23A→G, IVS2+21T→C, IVS3+111A→G, L129 and IVS6-2A→T. The intronic variants IVS2+21T→C and IVS3+111A→G occurred together in a Black patient and were absent in the control population. The base changes are not expected to alter gene splicing and were not investigated further. The potential disease-causing splice variant IVS6-2A→T identified in only one patient was absent in the control population.

Statistical analysis was performed for the various polymorphisms identified and the only significant association achieved in the Caucasian population was for the polymorphism within exon 6 (V221) (allele frequency: $p<0.03$, χ^2 with Yates' correction = 3.14; genotype frequency: $p<0.02$, χ^2 with Yates' correction = 3.98) when comparing the patient group with the population-matched control group.

Several statistically significant associations were observed for the polymorphisms identified in the Black population when comparing the patient group with the population-matched control group. These associations included the following variants: -23A→G (allele frequency: $p<0.0007$, χ^2 with Yates' correction = 10.53; genotype frequency: $p<0.001$, χ^2 = 12.82), IVS1-24G/C (allele frequency: $p<0.0001$, χ^2 with Yates' correction = 14.71; genotype frequency: $p<0.001$, χ^2 = 13.09), L129 (allele frequency: $p<0.04$, χ^2 with Yates' correction = 2.86; genotype frequency: $p<0.02$, χ^2 = 5.10) and V221 (allele

152

frequency: p<0.04, χ^2 with Yates' correction = 3.15; genotype frequency: p<0.03, χ^2 with Yates' correction = 3.32).

HAMP gene

Mutation analysis of the *HAMP* gene revealed a novel variant identified in the 3'-UTR region of the gene, causing a C to T substitution (3'-UTR+131). This variant was present in an individual who was also heterozygous for both the C282Y and H63D mutations of the *HFE* gene.

DCYTB gene

The previously described polymorphism S266N within exon 4 of *DCYTB* was identified at a high frequency in our study population, at similar frequencies in the respective Caucasian and Black populations (Table 3.5.1). Additionally, we identified a new polymorphism within intron 2, causing a T to C transition (IVS2+8), which was also common in our study population. Three novel mutations were identified only in the patient group: 1) a C to G transversion within intron 1 (IVS1-4), 2) a G to A substitution at amino acid 226 resulting in an arginine to histidine substitution (R226H), and 3) a G to C transversion within intron 3 (IVS3-32). The IVS1-4C→G variant was identified in 4 of 56 (7.1%) and mutation R226H was identified in 1 of 56 Caucasian patient referrals. Variant IVS3-32G→C was identified in 4 of 11 (36.4%) Black HH patients and was absent in the population-matched control group. A poly T region within intron 2 showed high variability in our study population and ranged from 12 to 21 Ts in this region. No statistically significant associations were observed for the polymorphisms identified in

the *DCYTB* gene, although it is highly likely that the mutations only identified in patients and not controls might contribute to the disease phenotype.

SLC11A1 gene

Mutation analysis of the promoter region by HEX-SSCP identified the -237C→T polymorphism (Lewis *et al* 1996) in the heterozygous state in a single UK patient (1 of 81, 1.23%), while detected in 10 of 91 (10.99%) population-matched control individuals. Only one of the control individuals was homozygous (1 of 91, 1.10%) for the polymorphism. RFLP analysis revealed only previously described alleles of the 5'-(GT)n repeat in the UK study cohort, including alleles 1, 2, 3 (Blackwell *et al* 1995) and 5 (Graham *et al* 2000).

The allele and genotype frequencies for *SLC11A1* promoter variants are shown in Tables 3.5.2 and 3.5.3. Statistically significant under representation was observed for both allele ($p<0.002$, χ^2 with Yates' correction = 7.75, df=1) and genotype ($p<0.01$, χ^2 = 8.79, df=2) frequencies for the −237C→T polymorphism in the patient group compared with the controls. This association was particularly evident in males for both allele ($p<0.002$, χ^2 with Yates' correction = 6.98, df=1) and genotype ($p<0.01$, χ^2 = 7.84, df=2) frequencies compared with the respective gender control group. No statistically significant differences were observed for the variants of the 5'-(GT)n repeat polymorphism (data not shown). However, when stratified according to the presence of the −237C→T polymorphism, significance was achieved for allele 3 compared with the controls ($p<0.005$, χ^2 with Yates' correction = 7.67) and especially in males,

154

representing a larger sample size, when compared with the respective gender control group (p<0.01, χ^2 with Yates' correction = 6.65). Re-assessment of the promoter alleles according to expression profiles upon iron loading previously determined for allele 3 (chapter 3.1), demonstrated a significant association between allele 3 and iron overload in patients without two copies of the C282Y mutation (p<0.01, χ^2 with Yates' correction = 6.66, df=1). When investigating the respective gender groups, significant association was observed for males (p<0.03, χ^2 with Yates' correction = 4.84, df=1) compared with the respective gender control group for allele 3 (in the absence of the $-237C \rightarrow T$ polymorphism).

Table 3.5.2. Allele frequencies of the various *SLC11A1* polymorphisms identified in individuals from the United Kingdom

VARIANTS	[a]ALLELE FREQUENCIES			
	Control (2n=182)		Iron overload patients (2n=162)	
Promoter alleles	Male (2n=82)	Female (2n=100)	Male (2n=118)	Female (2n=44)
Allele 1	-	0.020	-	
Allele 2	0.354	0.270	0.246	0.295
Allele 3	0.646	0.700	0.754	0.659
Allele 5	-	0.010	-	0.045
-237C→T	0.085	0.060	0.008	-

[a]Allele frequencies of polymorphic alleles denoted, Control, population-matched control group, Allele 1, t(gt)$_5$ac(gt)$_5$ac(gt)$_{11}$ggcaga(g)$_6$, Allele 2, t(gt)$_5$ac(gt)$_5$ac(gt)$_{10}$ggcaga(g)$_6$, Allele 3, t(gt)$_5$ac(gt)$_5$ac(gt)$_9$ggcaga(g)$_6$, Allele 5, t(gt)$_4$ac(gt)$_5$ac(gt)$_{10}$ggcaga(g)$_6$

Table 3.5.3. Genotype frequencies of the various *SLC11A1* polymorphisms identified in the population from the United Kingdom

Promoter alleles	Control (n=91)		Iron overload patients (n=81)	
	Males (n=41)	Females (n=50)	Males (n=59)	Females (n=22)
Allele 2; Allele 2	0.15 (6)	0.02 (1)	0.03 (2)	0.14 (3)
Allele 2; Allele 3	0.41 (17)	0.50 (25)	0.42 (25)	0.32 (7)
Allele 3; Allele 3	0.44 (18)	0.42 (21)	0.54 (32)	0.45 (10)
Allele 3; Allele 5	-	0.02 (1)	-	0.09 (2)
Allele 3; Allele 1	-	0.04 (2)	-	-
-237CC	0.85 (35)	0.88 (44)	0.98 (58)	1.00 (22)
-237CT	0.12 (5)	0.12 (6)	0.02 (1)	-
-237TT	0.02 (1)		-	-
Allele 2; Allele 3; -237C/T	0.02 (1)	0.02 (1)	-	-
Allele 3; Allele 3; -237C/T	0.10 (4)	0.10 (5)	0.02 (1)	-
Allele 3; Allele 3; -237T/T	0.02 (1)	-	-	-

Control, population-matched control group, Allele 1, $t(gt)_5ac(gt)_5ac(gt)_{11}ggcaga(g)_6$, Allele 2, $t(gt)_5ac(gt)_5ac(gt)_{10}ggcaga(g)_6$, Allele 3, $t(gt)_5ac(gt)_5ac(gt)_9ggcaga(g)_6$, Allele 5, $t(gt)_4ac(gt)_5ac(gt)_{10}ggcaga(g)_6$. The number of individuals for each genotype is indicated in brackets.

3.5.5. DISCUSSION

Polymorphisms are sequence variations that are present in the general population with no obvious deleterious effect on protein function. However, several studies have

156

recently demonstrated the involvement of low-penetrance polymorphisms in the clinical expression of complex diseases (Borrego *et al* 1999, Cambien *et al* 1999, Cargill *et al* 1999, Fitze *et al* 1999, Yang *et al* 2001, Rosenberg *et al* 2002). In this study, statistically significant associations were observed for many of the polymorphisms identified, especially in the Black South African population. Because of small sample size, these observations could be due to chance although the possibility of low penetrance mutations contributing to the disease phenotype cannot be ruled out. The likelihood that the polymorphisms identified may be in linkage disequilibrium with other disease-causing loci should also be considered. Functional studies need to be performed to elucidate the role of these polymorphisms, potentially contributing to the haemochromatosis phenotype in the Black South African population.

A likely candidate for contributing to iron overload, in the Black South African population, is the $(CGG)n$ microsatellite identified in the promoter region of the *SLC11A3* gene. Previously, seven $[(CGG)_7]$ and eight $[(CGG)_8]$ repeats were identified in this region. Here we identified nine $[(CGG)_9]$ repeats in the Black South African population and a marginal significant association was observed for this polymorphism. Of particular interest is the $-23A{\rightarrow}G$ polymorphism, identified in 8 of 11 Black HH patients and 1 of 20 controls in the same population, which appears to be in *cis* with the $(CGG)_7$ allele. It would be interesting to determine if an allelic effect similar to that recently observed for the $-237C/T$ polymorphism in the *SLC11A1* gene applies, where expression of allele 3 of the promoter $(GT)n$ repeat is reverted in the presence of this variant, resulting in expression profiles similar to that of allele 2 (chapter 3.1).

Possible non-*HFE* disease-causing mutations, not detected in the control population, were identified in 11 of 67 (16%) patients investigated using dHPLC analysis. These included 6 of 56 (11%) in the Caucasian population [*HAMP*: 3'-UTR+131C→T (1 of 56, 2%); *DCYTB*: IVS1-4C→G (4 of 56, 7%), R226H (1 of 56, 2%)] and 5 of 11 (45%) in the Black population [*SLC11A3*: IVS6-2A→T (1 of 11, 9%); *DCYTB*: IVS3-32G→C (4 of 11, 36%)].

Intronic mutations identified in the *SLC11A3* (IVS6-2A→T) and *DCYTB* (IVS1-4C→G, IVS3-32G→C) genes may cause splice defects resulting in abolishment or at least a reduction in the amount of mature mRNA generated. It was estimated that 15% of all point mutations in an intron result in a mRNA splicing defect, causing human disease (Krawczak *et al* 1992). Efficient splicing is dependent on accurate cleavage and rejoining which appears to be determined by the GT and AG dinucleotide present at the 5' and 3' exon/intron boundaries, respectively. However, these sequence motifs are not the only determinants of accurate splicing (Breathnach and Chambon 1981, Mount 1982, Padgett *et al* 1986). A branch site 5' from the exon (estimated to be between base pairs 20 and 40 in the intron), with a conserved A residue serving as a branch point, and a polypyrimidine tract preceding the 3' AG dinucleotide have also been identified to affect splicing. The *DCYTB* IVS3-32G→C variant could affect splicing since it is included within the branch site area. The IVS1-4C→G variant identified in *DCYTB* precedes an exon encoding for a region of b561 that is thought to be related to substrate binding or recognition and variant IVS3-32G→C precedes the exon encoding

the sixth transmembrane domain. Splicing defects are a common cause of human disease resulting in an exon not being recognised (exon skipping) and thereby excluded from the mature mRNA transcript. Variant IVS6-2A→T of the *SLC11A3* gene precedes an exon encoding the second intracellular loop between transmembrane domains four and five.

Alternative splicing of many pre-mRNAs is also affected by the intracellular concentrations of antagonistic splicing factors of the SR family and hnRNAPA1 (Zahler *et al* 1993, Caceres *et al* 1994, Muro *et al* 1999). Predictive results for the effect of the R226H variant, located at the C-terminal of the *DCYTB* gene, on these factors were generated by using the ESEfinder (ESE – Exonic Splice Element) program (http://exon.cshl.edu/ESE/). A SF2/ASF and SC35 binding site was abolished in the presence of the mutated allele (A) which might result in generating aberrant mRNAs that are either unstable or code for defective or deleterious protein isoforms. Functional studies needs to be performed to elucidate the role of all the mutations identified and to confirm results obtained using the ESEfinder program. *DCYTB* is responsible for the duodenal ferric reductase activity with protein and mRNA levels increasing in iron deficiency conditions and reduces upon iron loading, therefore regulating iron status (McKie *et al* 2001, 2002). The mutations identified in *DCYTB* could possibly disrupt the gene product, resulting in an inability of the gene to regulate intracellular iron levels.

The patient with the 3'UTR+131C→T variant, a 23 year old Caucasian male, presented with high serum ferritin levels [226.8 ng/ml (reference range 20-100 ng/ml)] and was

found to be heterozygous for both the C282Y and H63D mutations that are rarely (<1% of cases) associated with iron overload. Subsequent analysis showed that his father and younger brother were H63D heterozygotes and his mother and other brother were heterozygous for both the C282Y and H63D mutations. The 3'UTR+131C→T mutation was also present in the proband's father who presented with normal iron status. The other family members' iron status profiles were not available and were not available for testing of *HAMP*. It seems likely that this variant is in linkage disequilibrium with an unknown mutation contributed to the elevated iron levels in the *HFE* compound heterozygote due to digenic inheritance (Merryweather-Clarke *et al* 2003).

Several iron response elements (IREs) have been shown to be located in the 3'UTR region of genes, regulating iron metabolism. We speculate that the mutation identified in *HAMP* in the 3'-UTR of the gene (3'UTR+131C→T), could possibly disrupt an unidentified iron response element resulting in mRNA degradation. The patient with this mutation, a 23 year old Caucasian male, presented with high serum ferritin levels [226.8 ng/l (reference range 20-100 ng/l)] and was found to be heterozygous for both the C282Y and H63D mutations that are rarely (<1% of cases) associated with iron overload. Subsequent analysis showed that his father and younger brother were H63D heterozygotes and his mother and other brother was heterozygous for both the C282Y and H63D mutations. The 3'UTR+131C→T mutation was also present in the proband's father who presented with normal iron status. The other family members' iron status profiles were not available and were not available for testing of *HAMP*. It seems likely

that this mutation contributed to the elevated iron levels in the *HFE* compound heterozygote due to digenic inheritance (A Merryweather-Clarke, in press).

The *SLC11A1* gene is implicated in iron transport and influences intracellular iron levels (Barton *et al* 1999b, Biggs *et al* 2001). High expression of the *SLC11A1* gene in macrophages is associated with an increase in iron efflux and results in low intracellular iron levels. Similarly, elevated intracellular iron levels were observed with low *SLC11A1* activity. The *SLC11A1* gene was investigated by Lee *et al* (2002) in haemochromatosis patients homozygous for the *HFE* C282Y mutation. Their results concerning the 5'-(GT)n repeat polymophism were inconclusive and no other associations were observed. The likelihood that *SLC11A1* could contribute, as a possible modifier locus, to the clinical variability observed in HH penetrance therefore needs to be investigated further taking our *in vitro* observations (chapter 3.1) into consideration. Our data in patients with primary iron overload in the absence of C282Y homozygosity, indicated statistically significant differences compared with population-matched controls for the 5'(GT)n promoter repeat polymorphism. Association of allele 3 of the *SLC11A1* gene with iron overload can be explained by the fact that allele 3, driving high expression of the gene, is associated with low intracellular iron levels and higher levels of iron in macrophages, limited by the host iron-witholding mechanism. Barton *et al* (1999b) showed that higher influx of iron into the cytosol was obtained in unstimulated *Slc11a1* expressing cells compared to non-expressing cells, therefore contributing to regulation of iron balance or metabolism. A feature of haemochromatosis is that macrophages fail to regulate iron release (Brock 2000) and this could possibly be a direct result of *SLC11A1* allele 3 gene

161

expression on macrophage function. Similar studies are underway in a South African population to investigate whether the association observed can be replicated in another population.

Follow-up studies need to be performed to elucidate the mode of inheritance, by including family members of these individuals, and correlate iron parameters with mutations identified in affected individuals. It is well known that autosomal recessive inheritance is associated with *HFE* mutations (Feder *et al* 1996) while mutations in the *SLC11A3* gene are associated with an autosomal dominant form of inheritance (Njajou *et al* 2001; Montosi *et al* 2001, Wallace *et al* 2002, Devalia *et al* 2002, Roetto *et al* 2002). *DCYTB* mainly causes microcytic anaemia in mice and no association with mutations in patients presenting with iron overload has previously been reported in relation to this gene. Identification of two novel variants in this gene, R226H and IVS1-4C→G, therefore represents an important finding in the Caucasian population. These two mutations are likely to explain the non-*HFE* iron overload in up to 11% of South African Caucasians due to their potential role in exon/intron splicing.

Iron overload in African patients could not be ascribed to the common *HFE* mutations, but possible mutations in the *HFE* promoter region, affecting regulation of the gene, and the possible involvement of *SLC11A3* and *DCYTB* warrants further investigation in this context. The data presented are in accordance with various other studies, demonstrating the significance of multiple genes in iron regulation.

3.5.6. ACKNOWLEDGEMENTS

CHAPTER 4

CONCLUSIONS AND FUTURE PROSPECTS

4.1. IN CONCLUSION

Our increasing understanding of the molecular basis of genetic disorders of iron metabolism in mammalian species is providing new insight into genes involved in mediating intestinal iron absorption. All organisms are dependant on iron for survival and it is essential for functions such as the production of cellular iron, oxygen transport, DNA synthesis and various enzymatic reactions. It has become clear that the control of iron metabolism is interconnected with other cellular pathways and further analysis of these aspects are likely to provide a better understanding of pathological processes in which iron may participate. This study focused on mutational analysis of the iron transporter solute carrier family 11 member 1 gene (*SLC11A1*) known to have various pleiotropic effects on macrophage function, within the context of various disease phenotypes where iron may be involved.

4.1.1. *In vitro* expression of the *SLC11A1* gene 5'-(GT)n promoter alleles

In vitro studies, using luciferase reporter gene constructs, showed that the naturally occurring alleles of the *SLC11A1* gene differed in their ability to drive gene expression upon iron loading. This was the first study to investigate the effect of iron loading on expression of the naturally occurring alleles of the *SLC11A1* 5'-(GT)n promoter repeat polymorphism. High gene expression was observed for allele 3, which is in accordance with its role in autoimmune disease susceptibility and low gene expression, associated with infectious disease susceptibility, was observed for allele 2. The most striking finding was the opposite effect observed for allele 3 in the presence of the $-237C \rightarrow T$

165

polymorphism. An inverse in allele 3 gene expression was observed in the presence of the −237C→T polymorphism, which is similar to that observed for allele 2 known to be associated with infectious disease susceptibility. Even lower expression was observed for allele 3 in association with the −237C→T polymorphism compared with allele 2. These observations also demonstrated that the −237C→T polymorphism is functional and should be included in future association studies for assessment of the *SLC11A1* gene 5'-(GT)n alleles in relation with autoimmune and infectious disease susceptibility. The effect observed for allele 3 in association with the −237C→T polymorphism on gene expression could explain the discrepancies observed for various autoimmune and infectious disease association studies. Several studies have not shown association with the 5'-(GT)n repeat (Cervino *et al* 2000, Singal *et al* 2000) and re-investigation of these alleles in context of expression profiles observed with the −237C→T polymorphism may be very informative. Also, the opposite effect mediated by the −237C→T polymorphism in association of allele 3, suggests that the −237C→T polymorphism could possibly affect an iron response element (IRE). No conventional IREs have thus far been identified in the promoter region of *SLC11A1*, but this observation raise the possibility of unidentified IREs in this region. Another possibility could be that novel/unknown consensus sequences for IREs are present in the promoter region of *SLC11A1* and is in close proximity of the −237C→T polymorphism.

166

4.1.2. The *SLC11A1* gene and the diseases investigated

The use of genetic association studies are well recognised for identifying genetic variants that alter susceptibility to complex diseases (Risch and Merikangas 1996, Campbell and Rudan 2002). Several studies have also demonstrated the involvement of low-penetrance polymorphisms in the clinical expression of complex diseases (Borrego *et al* 1999, Cambien *et al* 1999, Cargill *et al* 1999, Fitze *et al* 1999, Yang *et al* 2001, Rosenberg *et al* 2002).

The possible involvement of the *SLC11A1* gene as a modifier locus was investigated in diseases including oesophageal cancer (OC), inflammatory bowel disease (IBD) and hereditary haemochromatosis (HH), due to its role in inflammation and iron transport. The $-237C{\rightarrow}T$ polymorphism was under represented in the respective patient groups studied compared with population-matched controls. A significant association was achieved for allele 3 of the *SLC11A1* gene 5'-(GT)n polymorphism only when stratified according to the presence of the $-237C{\rightarrow}T$ polymorphism for all diseases investigated. Re-assessment of the promoter alleles according to expression profiles previously demonstrated, showed a statistically significant association for allele 3 with OC and non-HFE iron overload compared with the respective population-matched controls. The 5'-(GT)n alleles were however not associated with IBD when allele 3 occurring together with the -237 T-allele, was combined with allele 2 based on similar expression profiles, but this may be due to the small sample size.

167

i) Oesophageal cancer (OC)

An increasing number of individuals in the Coloured South African population are currently developing OC. Patients with OC present with symptoms such as difficulty to swallow and loss of weight and, at the time of diagnosis, have deficient diets. It is therefore difficult to assess iron parameters accurately in these patients because iron deficiencies observed could be due to dietary restrictions. Two polymorphic alleles of the SLC11A1 gene (IVS1-28C→T and allele 3) were shown to be associated with OC in our study population. Association with allele 3, driving high SLC11A1 gene expression and being associated with autoimmune diseases, suggests the involvement of inflammation and generation of reactive oxygen species in the development of OC in our study population. Association with the variant in intron 1, IVS1-28C→T, found to be in linkage disequilibrium with allele 3, could also be associated with an unidentified functional mutation within another gene in close proximity of the SLC11A1 gene. However, after extensive bioinformatic analysis (results not shown), no obvious candidate genes were found in close proximity of the SLC11A1 gene. Several genes with unknown function were detected and therefore association of these genes with OC susceptibility cannot be excluded.

Exposure to cytotoxic agents released during inflammation, such as tumour necrosis factor-α (TNF-α), can lead to the development of cancer (Ruuls and Sedgwick 1999). In this respect, the involvement of the SLC11A1 gene in OC is further highlighted because individuals homozygous for allele 3 produce higher levels of TNF-α, a proinflammatory cytokine (Blackwell et al 2003). It therefore seems possible that the predominance of

allele 3 observed in OC patients may be related to cancer development via a mechanism involving TNF-α inflammatory response.

ii) Inflammatory bowel disease (IBD)

Frequent co-existence of IBD with other autoimmune conditions such as multiple sclerosis (Minuk and Lewkonia 1986, Agranof and Schon 1995) where iron dysregulation appears to be involved (Kotze et al 2001, Kotze et al 2003), raises the possibility that common features shared by these two conditions (Lossos et al 1995) could be explained by a related autoimmune process mediated, at least in part, by the SLC11A1 gene. The role of this gene in many clinically distinct autoimmune and infectious diseases may suggest a mechanistic continuum of the same disease process involving iron dysregulation.

Linkage demonstrated between the SLC11A1 gene and IBD susceptibility (Hofmeister et al 1997, Kojima et al 2001) is in general agreement with our findings, demonstrating the significance of the -237C$\rightarrow$T polymorphism as a modifying factor, and to a lesser extent the possible involvement of allele 3 of the 5'-(GT)n polymorphism. Although this may be related to the many pleiotropic effects of this gene on macrophage function (Searle and Blackwell 1999, Marquet et al 1999), the in vitro studies performed in this analysis further highlight the crucial role iron plays in the disease process underlying clinical expression of CD and possibly also UC. Unfortunately iron status was not available for the patients included in this study.

Again, association with allele 3 of *SLC11A1* in this group was only demonstrated when stratified according to the presence of the $-237C{\rightarrow}T$ polymorphism. This is in accordance with its role in autoimmune disease susceptibility. Several other genes have also been implicated in CD pathogenesis, including the *CARD15* gene. Wild-type *CARD15* enhances NF-κB expression in macrophages, which initiates immune-cell apoptosis (Hugot *et al* 2001, Ogura *et al* 2001b). Low frequencies of the common *CARD15* mutations were identified in our study population and an unexpected association of a previously described variant, A725G, was demonstrated with UC. The common *CARD15* mutations probably play a less important role in the development of IBD in non-Caucasian South African populations and other genes implicated in IBD susceptibility should be considered. However, the possibility of other variants in the *CARD15* gene contributing to IBD in the South African population cannot be excluded and should also be investigated.

Involvement of the *TNF-α* gene in IBD susceptibility is also well documented (Ruuls and Sedgwick 1999, Koss *et al* 2000, Sashio *et al* 2002) and it is clear that interaction between various genes and environmental factors contribute to this disease phenotype. An association with *TNF-α*, using linkage studies in an Australian population, has recently been observed in families with both CD and UC (O'Callaghan *et al* 2003). These authors have shown that association of a *TNF-α* promoter variant ($-857C{\rightarrow}T$) with IBD was more prevalent in families with *CARD15* risk alleles ($p<0.0001$) segregating with the disease, suggesting that both *TNF-α* and *CARD15* interact with NF-κB in macrophages. As previously mentioned, individuals homozygous for allele 3 of

SLC11A1 express higher levels of TNF-α, possibly due to the pleiotropic effects of *SLC11A1* on *TNF-α* as a contributing factor of the disease phenotype. We have shown for the first time that presence of the -237C→T polymorphism is likely to protect against IBD and should be considered as a potential modifying factor in disease expression.

Signal transduction pathways such as NF-κB shows rapid and sustained translocation to the nucleus in *Slc11a1* wildtype compared to mutant macrophages after stimulation of the macrophage with LPS, mediated by oxygen free radicals. Results from a previous study (Searle and Blackwell 1999) and our current results of the luciferase reporter gene constructs (Chapter 3.1), demonstrated differential expression of the *SLC11A1* gene upon stimulation of macrophages with LPS. These results suggested differential expression of LPS-related response elements (NF-κB, AP-1 and NF-IL6) by the various alleles of the 5'-(GT)n repeat promoter polymorphism of *SLC11A1*. The gene could therefore act as a modifier locus together with *CARD15* and *TNF-α* expression all being expressed in macrophages, thereby interacting through a common pathway to cause IBD. In turn, the effects mediated by *SLC11A1* on *TNF-α* could possibly directly alter *CARD15* gene expression or *vise versa*.

iii) Hereditary haemochromatosis (HH)

We hypothesised that the involvement of the *SLC11A1* gene in iron transport could contribute to the variability observed in HH phenotypes. By investigating individuals from the United Kingdom presenting with primary iron overload in the absence of homozygosity for the *HFE* C282Y mutation, under representation of the −237C→T

variant in the promoter region of the *SLC11A1* gene was demonstrated. Iron-overload observed in patients with haemochromatosis can affect the immune system by exposing the system to higher levels of iron, leading to altered function of the immune system or by facilitating the availability of iron to micro-organisms (Brock 2000). A feature of haemochromatosis is that macrophages fail to regulate iron release and this could possibly be a direct result of *SLC11A1* allele 3 gene expression on macrophage function.

Upon investigating other genes involved in regulation and transport of iron, including *HFE*, *HAMP*, *SLC11A3* and *DCYTB*, potential disease-causing mutations were identified in 16% of our study population. These mutations could possibly explain the iron overload in 11% White and 45% Black South African patients studied. The common *HFE* C282Y mutation is not associated with African iron overload and we demonstrated possible involvement for the *SLC11A3* and *DCYTB* genes in Black patients with primary iron overload for the first time. The possible involvement of the *DCYTB* gene in iron overload is also documented for the first time and further highlights the heterogeneity of HH.

Delineation of effects of the mutations identified on the protein structure could broaden our understanding of the biological role of these genes in iron homeostasis. The identification of mutations in patients with primary iron overload in the genes investigated would in future lead to improved and accurate counselling of patients and their families, and ultimately contributing to optimal therapy targeted at the cause of the disease. Following further studies, these results could lead to formulating a

comprehensive DNA diagnostic test for individuals presenting with primary iron overload in the South African population.

4.1.3. The *SLC11A1* gene

Although association of *SLC11A1* with susceptibility to various autoimmune and infectious diseases may be related to the many pleiotropic effects of the *SLC11A1* gene on macrophage function, it is highly likely that its involvement in iron transport may contribute directly to most of these disease phenotypes. This is based on previous findings (Atkinson *et al* 1997, Atkinson and Barton 1998, 1999, Barton *et al* 1999b, Goswami *et al* 2001, Mulero *et al* 2002) and the results obtained in our study demonstrating differential expression of *SLC11A1* alleles in the presence of ferric ammonium citrate. The functional significance of the *SLC11A1* promoter repeat polymorphism in infectious and autoimmune disease susceptibility therefore seems to be directly linked to iron regulation. A significant inflammatory process also interferes in iron delivery for heme synthesis, resulting in altered iron levels. Haemoproteins are involved in various biological functions including oxygen binding (haemoglobins), oxygen metabolism (oxidases, peroxidases, catalases and hydroxylases) and electron transfer (cytochromes).

The primary function of the *SLC11A1* gene is to transport divalent cations (such as Fe^{2+}, Zn^{2+} and Mn^{2+}). The secondary function of these metal ions should not be excluded, since these may contribute to the pleiotropic effects observed. The *Slc11a1* gene also functions as a pH-dependent manganese transporter at the phagsomal membrane and Jabado *et al* (2000) suggested that by transporting Mn^{2+} from the phagosomal lumen to

the cytoplasm deprives ingested pathogens of Mn^{2+}, an essential co-factor of superoxide dismutase. Similar arguments can be proposed for the transport of Fe^{2+} or Zn^{2+} across the phagosomal membrane which together may be essential for pathogen survival or expression of virulence determinants responsible for inhibition of phagolysosome maturation (Clemens and Horwitz 1995, Strugill-Koszycki *et al* 1996, Jabado *et al* 2000).

Low expression of the *Slc11a1* gene in mice, similar to allele 2 of *SLC11A1* in humans, is associated with iron influx, resulting in high intracellular iron levels and high transferrin saturation. Higher expression of the gene in mice, similar to allele 3 of *SLC11A1* in humans, is associated with efflux of iron and therefore low intracellular iron levels. Based on these results, we could speculate that individuals heterozygous for alleles 2 and 3 of human *SLC11A1* can regulate iron influx/efflux depending on the need for iron in either direction. A built-up of iron in and around macrophages is also associated with inflammation and can lead to anaemia of chronic disease. This could be as a consequence of invading pathogens resulting in a blockage of iron release (iron-withholding mechanism). Barton *et al* (1999b) have shown, by using the fluorescent iron sensor calcein (Breuer *et al* 1995), that higher influx of iron into the cytosol was obtained in unstimulated *Slc11a1* expressing cells compared to non-expressing cells. In this respect, iron can be stored in ferritin and exported to the circulation via *SLC11A3*, if required. Higher levels of iron in the cytosol can also lead to disruption of the heme biosynthesis pathway and cause the various pleiotropic effects on macrophage activation. Elucidation of the influence of genes affecting iron stores may therefore provide a route to understanding genetic differences in susceptibility to a range of common diseases.

174

The importance of the various pleiotropic effects of *SLC11A1* in disease susceptibility has not yet been fully elucidated. In mice, these pleiotropic effects mediated by the *Slc11a1* gene, includes up-regulation of class II I-A antigen expression, nitric oxide, protein kinase C and antigen presentation (Blackwell *et al* 1991, Buschman *et al* 1995, Zwilling and Hilburger 1994). Several of these pleiotropic effects are speculated to be the result of mRNA stability, where high cellular iron levels cause low mRNA stability and *vise versa* (Blackwell and Searle 1999). In the presence of an inflammatory stimulus, an increase in *Slc11a1* expression has been observed (Govoni *et al* 1997), suggesting that *Slc11a1* expression is modified in the presence or absence of an inflammatory signal. It can also be suggested that the source/origin of the pathogen mediating an inflammatory response can differentially affect expression of the *SLC11A1* gene, thereby contributing to the many pleiotropic effects observed.

Although it seems likely that iron transport regulated by *SLC11A1* causes the pleiotropic effects associated with macrophage activation, the direct influence of these pleiotropic effects should not be excluded in contributing to susceptibility of inflammatory and infectious diseases. It is clear that the complexity of several genes involved in regulating iron homeostasis complicates disease association studies. A better understanding of these aspects is likely to be obtained by using microarray analysis. Microarrays offer the advantage to simultaneously investigate the pattern of expression of thousands of genes. This pattern of expression is biologically informative and provides clues to the function of the gene. Proteins interact directly with each other, with more than one

member of a subunit complex required for disease expression, and may operate in distinct but complementary pathways. A dysfunction in one pathway may be rescued by an alternative route, which also could be compromised by the occurrence of disruptive mutations.

Based on the data obtained in this study we conclude that the naturally occurring alleles of the repeat promoter polymorphism of the *SLC11A1* gene most likely acts as a modifying factor in diseases involving iron homeostasis, especially in genetically predisposed individuals. It is clear that both environmental factors and the influence of other genes affect expression of the *SLC11A1* gene. In order to determine how susceptibility alleles contributing to the various diseases implicated may interact and give rise to a multifactorial phenotype, it is important to understand the mode of action of the *SLC11A1* gene as a potential modifier gene. Our results have contributed to unraveling the mechanisms underlying *SLC11A1* gene regulation. This study has led to a better understanding of the molecular basis of autoimmune and infectious disease susceptibility and may eventually contribute to the elucidation of disease mechanisms related to defective macrophage function.

4.2. FUTURE PROSPECTS

4.2.1. Oesophageal cancer (OC)

To further investigate the involvement of OC in the genetically distinct populations of South Africa, our study population needs to be extended to include individuals of the White and Black South African populations. A previous study including the Black South African population confirmed association of variants of the *SLC11A1* gene with OC susceptibility (du Plessis 2000). The possible effect of the IVS1-28C→T polymorphism, as a branch point mutation that may affect splicing, will be investigated by using the minigene system (Cartegni *et al* 2002).

Human papillomavirus infections have been implicated as a risk factor for the development of squamous cell carcinomas that may possibly increase the risk of OC development (Chang *et al* 1993, Cooper 1995). An intriguing possibility that requires further investigation is whether allele 2 of the *SLC11A1* 5'-(GT) repeat polymorphism shown to be associated with infectious diseases, would predispose individuals to human papillomavirus infection.

4.2.2. Inflammatory bowel disease (IBD)

The IBD patient groups will be extended by including more patients from the different ethnic groups in the South African population, to further investigate the involvement of the *SLC11A1* gene. Mutation analysis of the *CARD15* and *TNF-α* genes will be performed to identify possible mutations contributing to IBD in the South African population and gene-gene interaction will be investigated for these genes and *SLC11A1*.

177

The clinical data obtained for our study population was very limited and a challenge for the future would be to obtain well-characterised patient and control groups. As shown in a previous study, disease type/behaviour (including fistulising, stricturing and inflammation) and location of the disease (including colonic, enterocolonic and small bowel) may be preferentially associated with one of the three common *CARD15* mutations (Sambuelli *et al* 2003). A well-characterised patient group could shed more light on the low frequency obtained in our study population for the three common *CARD15* mutations.

4.2.3. Hereditary haemochromatosis (HH)

Our study involving the *SLC11A1* gene analysed in patients with apparent non-HFE iron overload (although some individuals included for analysis were C282Y heterozygotes) will be extended by screening individuals from the South African population with primary iron overload. The South African study population to be recruited for further study would also include C282Y homozygotes with variable clinical expression, ranging from organ damage at a relatively young age to non-expression in advanced age. Association of allele 3 in these groups could establish an association of *SLC11A1* with primary iron overload. Functional studies need to be performed to elucidate the role of potential disease-causing mutations identified in the *SLC11A3*, *HAMP* and *DCYTB* genes. Promoter variants identified in the *SLC11A3* gene will be investigated for functionality by using luciferase reporter gene constructs. In addition, the results obtained using the ESE finder program to identify potential splice variants will be investigated by using the minigene system. Further investigation of a well-characterised study population could shed more light on the function of these genes. Extensive screening of these genes in

178

individuals with iron overload in especially the Black South African population will be performed in future to elucidate the involvement of these genes in this population group. We neglected to investigate the promoter regions of *HFE*, *HAMP* and *DCYTB* but the intension is to include these gene regions for mutational analysis in future studies.

Despite the many limitations of the study, such as small sample size, lack of extensive clinical information on age-related disease expression, and the fact that iron status was not determined, this investigation represents a significant contribution to current knowledge on the role of iron or iron-related genes in health and disease.

CHAPTER 5

REFERENCES

Aballay A, Arenas NG, Quest AF, Mayorga LS. A factor with a zinc- and phorbol ester-binding domain is necessary for endosome fusion. Exp Cell Res 235: 28-34, 1997.

Aballay A, Sarrouf MN, Colombo MI, Stahl PD, Mayorga LS. Zn^{2+} depletion blocks endosome fusion. Biochem J 312: 919-923, 1995.

Abboud S, Haile DJ. A novel mammalian iron-regulated protein involved in intracellular iron metabolism. J Biol Chem 275: 19906-19912, 2000.

Abel L, Sanchez FO, Oberti J, Thuc NV, Hoa LV, Lap VD, Skamene E, Lagrange PH, Schurr E. Susceptibility to leprosy is linked to the human *NRAMP1* gene. J Infect Dis 177: 133-145, 1998.

Adams P, Brissot P, Powell W. EASL International Consensus Conference on Haemochromatosis: Part II Expert document. J Hepatol 33: 485-504, 2000.

Adams PC, Campion ML, Gandon G, Le Gall J-Y, David V, Jouanolle A-M. Clinical and family studies in genetic hemochromatosis: microsatellite and *HFE* studies in five atypical families. Hepatol 26: 986-990, 1997.

Adelaide J, Monges G, Derderian C, Seitz JF, Birnbaum D. Oesophageal cancer and amplification of the human cyclin D gene *CCND1/PRAD1*. Br J Cancer 71: 64-68, 1995.

Agranof D, Schon F. Are focal white matter lesions in patients with inflammatory bowel disease linked to multiple sclerosis? Lancet 346: 190-191, 1995.

Alcais A, Sanchez FO, Thuc NV, Lap VD, Oberti J, Lagrange PH, Schurr E, Abel L. Granulomatous reaction to intradermal injection of lepromin (Mitsuda reaction) is linked to the human *NRAMP1* gene in Vietnamese leprosy sibships. J Infect Dis 181: 302-308, 2000.

Alford CE, King TE, Campbell PA. Role of transferrin, transferrin receptors, and iron in macrophage listericidal activity. J Exp Med 174: 459-466, 1991.

Amer MH, El-Yazigi A, Hannan MA, Mohamed ME. Water contamination and esophageal cancer at Gassim Region, Saudi Arabia. Gastroenterology 98: 1141-1147, 1990.

Annese V, Latiano A, Bovio P, Forabosco P, Piepoli A, Lombardi G, Andreoli A, Astegiano M, Gionchetti P, Riegler G, Sturniolo GC, Clementi M, Rappaport E, Fortina P, Devoto M, Gasparini P, Andriulli A. Genetic analysis in Italian families with inflammatory bowel disease supports linkage to the IBD1 locus- a GISC study. Eur J Hum Genet 7: 567-573, 1999.

Arias M, Rojas M, Zabaleta J, Rodriguez JI, Paris SC, Barrera LF, Garcia LF. Inhibition of virulent *Mycobacterium tuberculosis* by Bcg(r) and Bcg(s) macrophages correlates with nitric oxide production. J Infect Dis 176: 1552-1558, 1997.

Atkinson PGP, Barton CH. Ectopic expression of *Nramp1* in COS-1 cells modulates iron accumulation. FEBS Lett 425: 239-242, 1998.

Atkinson PG, Barton CH. High level expression of Nramp1G169 in RAW264.7 cell transfectants: analysis of intracellular iron transport. Immunology 96: 656-662, 1999.

Atkinson PGP, Blackwell JM, Barton CH. *Nramp1* locus encodes a 65 kDa interferon-γ inducible protein in murine macrophages. Biochem J 325: 779-786, 1997.

Bacon BR, Sadiq SA. Hereditary hemochromatosis: presentation and diagnosis in the 1990s. Am J Gastroenterol 92: 784-789, 1997.

Bahram S, Gilfillan S, Kuhn LC, Moret R, Schulze JB, Lebeau A, Schumann K. Experimental hemochromatosis due to MHC class I HFE deficiency: immune status and iron metabolism. Proc Natl Acad Sci USA 96: 13312-13317, 1999.

Barrett MT, Galipeau PC, Sanchez CA, Emond MJ, Reid BJ. Determination of the frequency of loss of heterozygosity in esophageal adenocarcinoma by cell sorting, whole genome amplification and microsatellite polymorphisms. Oncogene 12: 1873-1878, 1996.

Barton CH, Biggs TE, Baker ST, Bowen H, Atkinson PG. *Nramp1*: a link between intracellular iron transport and innate resistance to intracellular pathogens. J Leukoc Biol 66: 757-762, 1999b.

Barton JC, Sawada-Hirai R, Rothenberg BE, Acton R. Two novel missense mutations of the *HFE* gene (I105C and G93R) and identification of the S65C mutation in Alabama hemochromatosis probands. Blood Cells Mol Dis 25: 146-154, 1999a.

Barton CH, White JK, Roach TI, Blackwell JM. NH2-terminal sequence of macrophage-expressed natural resistance-associated macrophage protein (Nramp) encodes a proline/serine-rich putative Src homology 3-binding domain. J Exp Med 179: 1683-1687, 1994.

Barton CH, Whitehead SH, Blackwell JM. Nramp transfection transfers *Ity/Lsh/Bcg*-related pleiotropic effects on macrophage activation: Influence on oxidative burst and Nitric oxide pathway. Mol Med 1: 267-279, 1995.

Bassuny WM, Ihara K, Matsuura N, Ahmed S, Kohno H, Kuromaru R, Miyako K, Hara T. Association study of the *NRAMP1* gene promoter polymorphism and early-onset type 1 diabetes. Immunogenetics 54: 282-285, 2002.

Bellamy R. The natural resistance-associated macrophage protein and susceptibility to intracellular pathogens. Microbes Infect 1: 23-27, 1999.

Bellamy R, Ruwende C, Corrah T, McAdam KP, Whittle HC, Hill AV. Variations in the *NRAMP1* gene and susceptibility to tuberculosis in West Africans. N Engl J Med 338: 640-644, 1998.

Beutler E. The HFE Cys282Tyr mutation as a necessary but not sufficient cause of clinical hereditary haemochromatosis. Blood 101: 3347-3350, 2003.

Beutler E, West C. New diallelic markers in the HLA region of chromosome 6. Blood Cells Mol Dis 23: 219-229, 1997.

Biggs TE, Baker ST, Botham MS, Dhital A, Barton CH, Perry VH. *Nramp1* modulates iron homoeostasis *in vivo* and *in vitro*: evidence for a role in cellular iron release involving de-acidification of intracellular vesicles. Eur J Immunol 31: 2060-2070, 2001.

Binder V. Genetic epidemiology in inflammatory bowel disease. Dig Dis 16: 351-355, 1998.

Blackwell JM. The macrophage resistance gene *Lsh/Ity/Bcg*. Res Immunol 140: 767-769, 1989.

Blackwell JM. Structure and function of the natural-resistance-associated macrophage protein (Nramp1), a candidate protein for infectious and autoimmune disease susceptibility. Mol Med Today 2: 205-211, 1996.

Blackwell JM. Genetics and genomics in infectious disease susceptibility. Trends Mol Med 7: 521-526, 2001.

Blackwell JM, Barton CH, White JK, Searle S, Baker A-M, Williams H, Shaw M-A. Genomic organisation and sequence of the human *NRAMP* gene: identification and mapping of a promoter region polymorphism. Molec Med 1: 194-205, 1995.

Blackwell JM, Goswami T, Evans CAW, Sibthorpe D, Papo N, White JK, Searle S, Miller EN, Peacock CS, Mohammed H, Ibrahim M. *SLC11A1* (formerly *NRAMP1*) and disease resistance. Cellular Microbiology 3: 773-784, 2001.

Blackwell JM, Roach TI, Atkinson SE, Ajioka JW, Barton CH, Shaw MA. Genetic regulation of macrophage priming/activation: the *Lsh* gene story. Immunol Lett 30: 241-248, 1991.

Blackwell JM, Roberts CW, Roach TIA, Alexander J. Influence of macrophage resistance gene *Lsh/Ity/Bcg* (candidate *Nramp*) on *Toxoplasma gondii* infection in mice. Clin Exp Immunol 97: 107-112, 1994.

Blackwell JM, Searle S. Genetic regulation of macrophage activation: understanding the function of *Nramp1* (=*Ity/Lsh/Bcg*). Immunol Lett 65: 73-80, 1999.

Blackwell JM, Searle S, Goswami T, Miller EN. Understanding the multiple functions of *Nramp1*. Microbes and Infection 2: 317-321, 2000.

Blackwell JM, Searle S, Mohamed H, White JK. Divalent cation transport and susceptibility to infectious and autoimmune disease: continuation of the *Ity/Lsh/Bcg/Nramp1/Slc11a1* gene story. Immunol Lett 85: 197-203, 2003.

Blackwell JM, Toole S, King M, Dawda P, Roach TI, Cooper A. Analysis of *Lsh* gene expression in congenic B10.L-Lshr mice. Curr Top Microbiol Immunol 137: 301-309, 1988.

Blot WJ. Alcohol and cancer. Cancer Res 52: 2119s-2123s, 1992.

Blot WJ. Esophageal cancer trends and risk factors. Semin Oncol 21: 403-410, 1994.

Blot WJ, Fraumeni JF Jr. Trends in esophageal cancer mortality among US blacks and whites. Am J Public Health 77: 296-298, 1987.

Booker GW, Gout I, Downing AK, Driscoll PC, Boyd J, Waterfield MD, Campbell ID. Solution structure and ligand-binding site of the SH3 domain of the p85 alpha subunit of phosphatidylinositol 3-kinase. Cell 73: 813-822, 1993.

Borrego S, Saez ME, Ruiz A, Gimm O, Lopez-Alonso M, Antinolo G, Eng C. Specific polymorphisms in the *RET* proto-oncogene are over-represented in patients with Hirschsprung disease and may represent loci modifying phenotypic expression. J Med Genet 36: 771-774, 1999.

Bothwell TH, Charlton RW, Motulsky AG. Hemochromatosis. In: The Metabolic and Molecular Basis of Inherited Disease. Ed. By CR Scriver, AL Beaudet, WS Sly and D Valle, 2237-2269. McGraw-Hill, New York, 1995.

Bothwell TH, MacPhail AP. Hereditary hemochromatosis: etiologic, pathologic, and clinical aspects. Semin Hematol. 35: 55-71, 1998.

Bradley DJ. Genetic control of natural resistance to *Leishmania donovani*. Nature 250: 353-354, 1974.

Brant SR, Fu Y, Fields CT, Baltazar R, Ravenhill G, Pickles MR, Rohal PM, Mann J, Kirschner BS, Jabs EW, Bayless TM, Hanauer SB, Cho JH. American families with Crohn's disease have strong evidence for linkage to chromosome 16 but not chromosome 12. Gastroenterology 115: 1056-1061, 1998.

Breathnach R, Chambon P. Organisation and expression of eukaryotic split genes coding for proteins. Annu Rev Biochem 50: 349-383, 1981.

Breuer W, Epsztejn S, Cabantchik ZI. Iron acquired from transferrin by K562 cells is delivered into a cytoplasmic pool of chelatable iron(II). J Biol Chem 270: 24209-24215, 1995.

Brock JH. Role of iron in infections and immunity. In: Hemochromatosis genetics, pathophysiology, diagnosis, and treatment. Ed. By J.C. Barton and C.Q. Edwards,371-380. Cambridge University Press, 2000.

Brown IN, Glynn AA, Plant J. Inbred mouse strain resistance to *Mycobacterium lepraemurium* follows the *Ity/Lsh* pattern. Immunology 47: 149-156, 1982.

Brown DH, LaFuse W, Zwilling BS. Cytokine-mediated activation of macrophages from *Mycobacterium bovis* BCG –resistant and –susceptible mice: differential effects of corticosterone on antimycobacterial activity and expression of the *Bcg* gene (candidate *Nramp*). Infect Immun 63: 2983-2988, 1995.

Brown DH, Lafuse WP, Zwilling BS. Stabilized expression of mRNA is associated with mycobacterial resistance controlled by *Nramp1*. Infect Immun 65: 597-603, 1997.

Buschman E, Skamene E. From *Bcg/Lsh/Ity* to *Nramp1*: three decades of search and research. Drug Metab Dispos 29: 471-473, 2001.

Buschman E, Schurr E, Skamene E. Constitutional resistance to mycobacterial infections: Role of genetic factors in resistance to infection. In: Constitutional Resistance to infection. Ed. By Verdiun CM, Watson DA, van Dijk H, Verhoef J, 55-81, R.G. Landes Company, Austin, TX, 1995.

Butt C, Zheng H, Randell E, Robb D, Parfrey P, Xie Y. Combined carrier status of prothrombin 20210A and factor XIII-A Leu34 alleles as a strong risk factor for myocardial infarction: evidence of a gene-gene interaction. Hemostasis Thrombosis and vascular biology 101: 3037-3041, 2003.

Buu A, Garreau H, Jacquet M. The glyceraldehyde-3-phosphate dehydrogenase binds *in vitro* to the SH3 domain of *Saccharomyces cerevisiae* Cdc25p. C R Acad Sci III 318: 665-669, 1995.

Caceres JF, Stamm S, Helfman DM, Krainer AR. Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. Science 265: 1706-1709, 1994.

Cairo G, Conte D, Bianchi L, Fraquelli M, Recalcati S. Reduced serum ceruloplasmin levels in hereditary haemochromatosis. Br J Haematol 114: 226-229, 2001.

Calver GA, Kenny CP, Kushner DJ. Inhibition of the growth of *Neisseria meningitides* by reduced ferritin and other iron-binding agents. Infect Immunol 25: 880-890, 1979.

Camaschella C, Roetto A, Cali A, De Gobbi M, Garozzo G, Carella M, Majorano N, Totaro A, Gasparini P. The gene *TFR2* is mutated in a new type of haemochromatosis mapping to 7q22. Nat Genet 25: 14-15, 2000.

Camaschella C, Roetto A, Cicilano M, Pasquero P, Bosio S, Gubetta L, Di Vito F, Girelli D, Totaro A, Carella M, Grifa A, Gasparini P. Juvenile and adult hemochromatosis are distinct genetic disorders. Eur J Hum Genet 5: 371-375, 1997.

Cambien F, Poirier O, Nicaud V, Herrmann SM, Mallet C, Ricard S, Behague I, Hallet V, Blanc H, Loukaci V, Thillet J, Evans A, Ruidavets JB, Arveiler D, Luc G, Tiret L. Sequence diversity in 36 candidate genes for cardiovascular disorders. Am J Hum Genet 65: 183-191, 1999.

Campbell H, Rudan I. Interpretation of genetic association studies in complex disease. Pharmacogenomics J 2: 349-360, 2002.

Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, Patil N, Shaw N, Lane CR, Lim EP, Kalyanaraman N, Nemesh J, Ziaugra L, Friedland L, Rolfe A, Warrington J, Lipshutz R, Daley GQ, Lander ES. Characterisation of single-nucleotide polymorphisms in coding regions of human genes. Nat Genet 22: 231-238, 1999.

Cartegni L, Chew SL, Krainer AR. Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat Rev Genet 3: 285-298, 2002.

Cavanaugh JA, Callen DF, Wilson SR, Stanford PM, Sraml ME, Gorska M, Crawford J, Whitmore SA, Shlegel C, Foote S, Kohonen-Corish M, Pavli P. Analysis of Australian Crohn's disease pedigrees refines the localization for susceptibility to inflammatory bowel disease on chromosome 16. Ann Hum Genet 62: 291-298, 1998.

Cavanaugh J, IBD International Genetics Consortium. International collaboration provides convincing linkage replication in complex disease through analysis of a large pooled data set: Crohn disease and chromosome 16. Am J Hum Genet 68: 1165-1171, 2001.

Cazzola M, Ponchio L, de Benedetti F, Ravelli A, Rosti V, Beguin Y, Invernizzi R, Barosi G, Martini A. Defective iron supply for erythropoiesis and adequate endogenous erythropoietin production in the anemia associated with systemic-onset juvenile chronic arthritis. Blood 87: 4824-4830, 1996.

Cellier M, Govoni G, Vidal S, Kwan T, Groulx N, Liu J, Sanchez F, Skamene E, Schurr E, Gros P. Human natural resistance-associated macrophage protein: cDNA cloning, chromosomal mapping, genomic organisation, and tissue-specific expression. J Exp Med 180: 1741-1752, 1994.

Cellier M, Prive G, Belouchi A, Kwan T, Rodrigues V, Chia W, Gros P. Nramp defines a family of membrane-proteins. Proc Natl Acad Sci USA 92: 10089-10093, 1995.

Cellier M, Shustik C, Dalton W, Rich E, Hu J, Malo D, Schurr E, Gros P. Expression of the human NRAMP1 gene in professional primary phagocytes: studies in blood cells and in HL-60 promyelocytic leukaemia. J Leukoc Biol 61: 96-105, 1997.

Cervino AC, Lakiss S, Sow O, Hill AV. Allelic association between the NRAMP1 gene and susceptibility to tuberculosis in Guinea-Conakry. Ann Hum Genet 64: 507-512, 2000.

Chang F, Syrjänen S, Shen Q, Wang L, Syrjänen K. Screening for human papillomavirus infections in esophageal squamous cell carcinomas by in situ hybridisation. Cancer 72: 2525-2530, 1993.

Chen X, Ding YW, Yang G, Bondoc F, Lee MJ, Yang CS. Oxidative damage in an esophageal adenocarcinoma model with rats. Carcinogenesis 21: 257-263, 2000.

Chen X, Yang GY, Ding WY, Bondoc F, Curtis SK, Yang CS. An esophagogastroduodenal anastomosis model for esophageal adenocarcinogenesis in rats and enhancement by iron overload. Carcinogenesis 20: 1801-1808, 1999.

Cho J, Nicolae D, Gold L, Fields C, LaBuda M, Rohal P, Pickles M, Qin L, Fu Y, Mann J, Kirschner B, Jabs E, Weber J, Hanauer S, Bayless T, Brant S. Identification of novel susceptibility loci for inflammatory bowel disease on chromosomes 1p, 3q, and 4q: evidence for epistasis between 1p and IBD1. Proc Nat Acad Sci USA 95: 7502-7507, 1998.

Cho J, Nicolae D, Ramos R, Fields C, Rabenau K, Corradino S, Brant S, Espinosa R, LeBeau M, Hanauer S, Bodzin J, Bonen D. Linkage and linkage disequilibrium in chromosome band 1p36 in American Chaldeans with inflammatory bowel disease. Hum. Molec Genet 9: 1425-1432, 2000.

185

Clemens DL, Horwitz MA. Characterization of the Mycobacterium tuberculosis phagosome and evidence that phagosomal maturation is inhibited. J Exp Med 181: 257-270, 1995.

Cohen G. The Fenton reaction. In: Handbook of Methods for Oxygen Radical Research. Ed. By Greenwald RA, 55-64. CRC Press, Boca Raton, 1985.

Cooper K. p53 mutations in human papillomavirus-associated oesophageal squamous cell carcinoma. Br J Can 72: 1337, 1995.

Curran ME, Lau KF, Hampe J, Schreiber S, Bridger S, Macpherson AJ, Cardon LR, Sakul H, Harris TJ, Stokkers P, Van Deventer SJ, Mirza M, Raedler A, Kruis W, Meckler U, Theuer D, Herrmann T, Gionchetti P, Lee J, Mathew C, Lennard-Jones J. Genetic analysis of inflammatory bowel disease in a large European cohort supports linkage to chromosomes 12 and 16. Gastroenterology 115: 1066-1071, 1998.

Derman DP, Bothwell TH, Torrance JD, Bezwoda WR, MacPhail AP, Kew MC, Sayers MH, Disler PB, Charlton RW. Iron absorption from maize (Zea mays) and sorghum (Sorghum vulgare) beer. Br J Nutr 43: 271-279, 1980.

Devalia V, Carter K, Walker AP, Perkins SJ, Worwood M, May A, Dooley JS. Autosomal dominant reticuloendothelial iron overload associated with a 3-base pair deletion in the ferroportin 1 gene (*SLC11A3*). Blood 100: 695-697, 2002.

De Villiers JN, Hillermann R, Loubser L, Kotze MJ. Spectrum of mutations in the *HFE* gene implicated in haemochromatosis and porphyria. Hum Mol Genet 8: 1517-1522, 1999.

Donovan A, Brownlie A, Zhou Y, Shepard J, Pratt SJ, Moynihan J, Paw BH, Drejer A, Barut B, Zapata A, Law TC, Brugnara C, Lux SE, Pinkus GS, Pinkus JL, Kingsley PD, Palis J, Fleming MD, Andrews NC, Zon LI. Positional cloning of zebrafish ferroportin1 identifies a conserved vertebrate iron exporter. Nature 403: 776-781, 2000.

Douabin V, Moirand R, Jouanolle A, Brissot P, Le Gall J, Deugnier Y, David V. Polymorphisms in the *HFE* gene. Hum Hered 49: 21-26, 1999.

Duchmann R, Kaiser I, Hermann E, Mayet W, Ewe K, Meyer zum Buschenfelde K. Tolerance exists towards resident intestinal flora but is broken in active inflammatory bowel disease (IBD). Clin Exp Immunol 102: 448-455, 1995.

Duerr RH, Barmada MM, Zhang L, Davis S, Preston RA, Chensny LJ, Brown JL, Ehrlich GD, Weeks DE, Aston CE. Linkage and association between inflammatory bowel disease and a locus on chromosome 12. Am J Hum Genet 63: 95-100, 1998.

Du Plessis L. Molecular investigation of oesophageal cancer in South Africa: A search for new candidate genes. University of Stellenbosch, December 2000.

Du Plessis L, Dietzsch E, Van Gele M, Van Roy N, Van Helden P, Parker MI, Mugwanya DK, De Groot M, Marx MP, Kotze MJ, Speleman F. Mapping of novel regions of DNA gain and loss by comparative genomic hybridisation in esophageal carcinoma in the Black and Colored populations of South Africa. Cancer Res 59: 1877-1883, 1999.

Edwards CQ, Griffen LM, Goldgar D, Drummond C, Skolnick MH, Kushner JP. Prevalence of hemochromatosis among 11 065 presumably healthy blood donors. N Engl J Med 318: 1355-1362, 1988.

Esposito L, Hill NJ, Pritchard LE, Cucca F, Muxworthy C, Merriman ME, Wilson A, Julier C, Delepine M, Tuomilehto J, Tuomilehto-Wolf E, Ionesco-Tirgoviste C, Nistico' L, Buzzetti R, Pozzilli P, Ferrari M, Bosi E, Pociot F, Nerup J, Bain SC, Todd JA. Genetic analysis of chromosome 2 in type 1 diabetes: analysis of putative loci IDDM7, IDDM12, and IDDM13 and candidate genes *NRAMP1* and *IA-2* and the *interleukin-1* gene cluster. IMDIAB Group Diabetes 47: 1797-1799, 1998.

Evans CAW, Harbuz MS, Ostenfeld T, Norrish A, Blackwell JM. *Nramp1* is expressed in neurons and is associated with behavioural and immune responses to stress. Neurogenetics 3: 69-78, 2001.

Farmer R, Michener W, Mortimer E. Studies of family history among patients with inflammatory bowel disease. Clin Gastroent 9: 271-278, 1980.

Feder JN, Gnirke A, Thomas W, Tsuchihashi Z, Ruddy DA, Basava A, Dormishian F, Domingo R Jr, Ellis MC, Fullan A, Hinton LM, Jones NL, Kimmel BE, Kronmal GS, Lauer P, Lee VK, Loeb DB, Mapa FA, McClelland E, Meyer NC, Mintier GA, Moeller N, Moore T, Morikang E, Wolff RK. A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. Nat Genet 13: 399-408, 1996.

Feder JN, Penny DM, Irrinki A, Lee VK, Lebron JA, Watson N, Tsuchihashi Z, Sigal E, Bjorkman PJ, Schatzman RC. The hemochromatosis gene product complexes with the transferrin receptor and lowers its affinity for ligand binding. Proc Natl Acad Sci USA 95: 1472-1477, 1998.

Fiocchi C. Inflammatory bowel disease: etiology and pathogenesis. Gastroenterology 115: 182-205, 1998.

Fitze G, Schreiber M, Kuhlisch E, Schackert HK, Roesner D. Association of *RET* protooncogene codon 45 polymorphism with Hirschsprung disease. Am J Hum Genet 65: 1469-1473, 1999.

Fleming RE, Holden CC, Tomatsu S, Waheed A, Brunt EM, Britton RS, Bacon BR, Roopenian DC, Sly WS. Mouse strain differences determine severity of iron accumulation in *Hfe* knockout model of hereditary hemochromatosis. Proc Natl Acad Sci USA 98: 2707-2711, 2001.

Gao PS, Fujishima S, Mao XQ, Remus N, Kanda M, Enomoto T, Dake Y, Bottini N, Tabuchi M, Hasegawa N, Yamaguchi K, Tiemessen C, Hopkin JM, Shirakawa T, Kishi F. Genetic variants of *NRAMP1* and active tuberculosis in Japanese populations. International Tuberculosis Genetics Team. Clin Genet 58: 74-76, 2000.

Gilberts EC, Greenstein AJ, Katsel P, Harpaz N, Greenstein RJ. Molecular evidence for two forms of Crohn disease. Proc Natl Acad Sci USA 91: 12721-12724, 1994.

Goldstein SR, Yang GY, Chen X, Curtis SK, Yang CS. Studies of iron deposits, inducible nitric oxide synthase and nitrotyrosine in a rat model for esophageal adenocarcinoma. Carcinogenesis 19: 1445-1449, 1998.

Goswami T, Bhattacharjee A, Babal P, Searle S, Moore E, Li M, Blackwell JM. Natural-resistance-associated macrophage protein 1 is an H+/bivalent cation antiporter. Biochem J 354: 511-519, 2001.

Goto Y, Buschman E, Skamene E. Regulation of host resistance to Mycobacterium intracellulare in vivo and in vitro by the Bcg gene. Immunogenetics 30: 218-221, 1989.

Govoni G, Canonne-Hergaux F, Pfeifer CG, Marcus SL, Mills SD, Hackam DJ, Grinstein S, Malo D, Finlay B, Gros P. Functional expression of *Nramp1 in vitro* in the murine macrophage line RAW264.7. Infect Immun 67: 2225-2232, 1999.

Govoni G, Gauthier S, Billia F, Iscove NN, Gros P. Cell-specific and inducible *Nramp1* gene expression in mouse macrophages *in vitro* and *in vivo*. J Leukoc Biol 62: 277-286, 1997.

187

Govoni G, Vidal S, Gauthier S, Skameme E, Malo D, Gros P. The *Bcg/Ity/Lsh* locus: Genetic transfer of resistance to infections in C57BL/6J mice transgenic for the *Nramp1* Gly169 allele. Infect and Immun 64: 2923-2929, 1996.

Graham AM, Dollinger MM, Howie SEM, Harrison DJ. Identification of novel alleles at a polymorphic microsatellite repeat region in the human *NRAMP1* gene promoter: analysis of allele frequencies in primary biliary cirrhosis. J Med Genet 37: 150-152, 2000.

Green R, Esparza I, Schreiber R. Iron inhibits the non-specific tumoricidal activity of macrophages. A possible contributory mechanism for neoplasia in haemochromatosis. Ann N Y Acad Sci 526: 301-309, 1988.

Greenstein A, Lachman P, Sachar D, Springhorn J, Heimann T, Janowitz H, Aufses A. Perforating and non-perforating indications for repeated operations in Crohn's disease: evidence for two clinical forms. Gut 29: 588-592, 1988.

Greenwood CM, Fujiwara TM, Boothroyd LJ, Miller MA, Frappier D, Fanning EA, Schurr E, Morgan K. Linkage of tuberculosis to chromosome 2q35 loci, including *NRAMP1*, in a large aboriginal Canadian family. Am J Hum Genet 67: 405-416, 2000.

Griffiths A. Inflammatory Bowel Disease. Adolesc Med 6: 351-368, 1995.

Gross CN, Irrinki A, Feder JN, Enns CA. Co-trafficking of HFE, a nonclassical major histocompatibility complex class I protein, with the transferrin receptor implies a role in intracellular iron regulation. J Biol Chem 273: 22068-22074, 1998.

Gros P, Skamene E, Forget AI. Genetic control of natural resistance to *Mycobacterium bovis* (*BCG*) in mice. J Immunol 127: 2417-2421, 1981.

Gruenheid S, Cellier M, Vidal S, Gros P. Identification and characterization of a second mouse *Nramp* gene. Genomics 25: 514-525, 1995.

Gruenheid S, Pinner E, Desjardins M, Gros P. Natural resistance to infections with intracellular pathogens: the Nramp1 protein is recruited to the membrane of the phagosome. J Exp Med 185: 717-730, 1997.

Gunshin H, Mackenzie B, Berger UV, Gunshin Y, Romero MF, Boron WF, Nussberger S, Gollan JL, Hediger MA. Cloning and characterisation of a mammalian proton-coupled metal-ion transporter. Nature 388: 482-488, 1997.

Gustavisson P, Hogstedt C, Evanoff B. Increased risk of esophageal cancer among workers exposed to combustion products. Arch Environ Health 48: 243-244, 1993.

Haile DJ. Assignment of *Slc11a3* to mouse chromosome 1 band 1B and *SLC11A3* to human chromosome 2q32 by in situ hybridization. Cytogenet Cell Genet 88: 328-329, 2000.

Hampe J, Cuthbert C, Croucher P, Mirza M, Mascheretti S, Fisher S, Frenzel H, King K, Hasselmeyer A, MacPherson A, Bridger S, Van Deventer S, Forbes A, Nikolaus S, Lennard-Jones J, Foelsch U, Krawczak M, Lewis C, Schreiber S, Mathew C. Association between insertion mutation in *NOD2* gene and Crohn's disease in German and British populations. Lancet 357: 1925-1928, 2001.

Hampe J, Frenzel H, Mirza M, Croucher P, Cuthbert A, Mascheretti S, Huse K, Platzer M, Bridger S, Meyer B, Nurnberg P, Stokkers P, Krawczak M, Mathew C, Curran M, Schreiber S. Evidence for a

NOD2-independent susceptibility locus for inflammatory bowel disease on chromosome 16p. Proc Natl Acad Sci USA 99: 321-326, 2002.

Harries A, Baird A, Rhodes J. Non-smoking: a feature of ulcerative colitis. Br Med J 284: 706, 1982a.

Harries A, Jones L, Heatley R, Rhodes J. Smoking habits and inflammatory bowel disease: effect on nutrition. Br Med J 284: 1161, 1982b.

Heikenen J, Werlin S, Brown C, Balint J. Presenting symptoms and diagnostic lag in children with inflammatory bowel disease. Inflamm Bowel Dis 5: 158-160, 1999.

Henz S, Reichen J, Liechti-Gallati S. HLA-H gene mutations and haemochromatosis: the likely association of H63D with mild phenotype and the detection of S65C, a novel variant in exon 2. J Hepatol 26 (Suppl 1): 57, 1997.

Hermanek P, Sobin LH. TNM classification of malignant tumors (eds), New York: Springer-Verlag, Ed 4, 40-42, 1987.

Hille JJ, Markowitz S, Margolius KA, Isaacson C. Human papillomavirus and carcinoma of the esophagus. N Engl J Med 312: 1707, 1985.

Hofmeister A, Neibergs H, Pokorny R, Galandiuk S. The natural resistance-associated macrophage gene is associated with Crohn's disease. Surgery 122: 173-179, 1997.

Höhler T, Leininger S, Schneider PM. A new polymorphism in the human *HFE* gene. Immunogenetics 49: 823-824, 1999.

Hsing AW, McLaughlin JK, Olsen JH, Mellemkjar L, Wacholder S, Fraumeni JF Jr. Cancer risk following primary hemochromatosis: a population-based cohort study in Denmark. Int J Cancer 60: 160-162, 1995.

Hugot J, Chamaillard M, Zouali H, Lesage S, Cezard J, Belaiche J, Almer S, Tysk C, O'Morain C, Gassull M, Binder V, Finkel Y, Cortot A, Modigliani R, Laurent-Puig P, Gower-Rousseau C, Macry J, Colombel J-F, Sahbatou M, Thomas G. Association of *NOD2* leucine –rich repeat variants with susceptibility to Crohn's disease. Nature 411: 599-603, 2001.

Hugot J, Laurent-Puig P, Gower-Rousseau C, Olson JM, Lee J, Beaugerie L, Naom I, Dupas J, Van Gossum A, Orholm M, Bonaiti-Pellie C, Weissenbach J, Mathew C, Lennard-Jones J, Cortot A, Colombel J, Thomas G. Mapping of a susceptibility locus for Crohn's disease on chromosome 16. Nature 379: 821-823, 1996.

Huot AE, Gundel RM, Hacker MP. Effect of erythrocytes on alveolar macrophage cytostatic activity induced by bleomycin lung damage in rats. Cancer Res 50: 2351-2355, 1990.

IBD International Genetics Consortium. International collaboration provides convincing linkage replication in complex disease through analysis of a large pooled data set: Crohn's disease and chromosome 16. Am J Hum Genet 68: 1165-1171, 2001.

Ikuta K, Fujimoto Y, Suzuki Y, Tanaka K, Saito H, Ohhira M, Sasaki K, Kohgo Y. Overexpression of hemochromatosis protein, HFE, alters transferrin recycling process in human hepatoma cells. Biochim Biophys Acta 1496: 221-231, 2000.

International Agency for Research an Cancer (IARC): Tobacco smoking, IARC Monogr on evaluation of carcinogenic risks to humans Vol 38, Lyon, France, 1986.

189

International Agency for Research an Cancer (IARC): Alcohol drinking, IARC Monogr an evaluation of carcinogenic risks to humans Vol 44, Lyon, France, 1988.

Isaacson C, Bothwell TH, MacPhail AP, Simon M. The iron status of urban black subjects with carcinoma of the oesophagus. S Afr Med J 67: 591-593, 1985.

Isshiki H, Akira S, Tanabe O, Nakaji T, Shimamoto T, Hirano T, Kishimoto T. Constitutive and interleukin-1 (IL-1)-inducible factors interact with the IL-1-response element in the *IL-6* gene. Molec Cell Biol 10: 2757-2764, 1990.

Jabado N, Jankowski A, Dougaparsad S, Picard V, Grinstein S, Gros P. Natural resistance to intracellular infections: natural Resistance-associate Macrophage protein 1 (*NRAMP1*) functions as a pH dependent manganese transporter at the phagosomal membrane. J Exp Med 192: 1237-1247, 2000.

Kaplan J, Kushner JP. Mining the genome for iron. Nature 403: 711-713, 2000.

Kehrer JP. The Haber-Weiss reaction and mechanisms of toxicity. Toxicology 149: 43-50, 2000.

Kerppola RE, Ames GF-L. Topology of the hydrophobic membrane-bound components of the histidine periplasmic permease: comparison with other members of the family. J Biol Chem 267: 2329-2336, 1992.

Kishi F. Isolation and characterization of human Nramp cDNA. Biochem Biophys Res Commun 204: 1074-1080, 1994.

Kishi F, Tanizawa Y, Nobumoto M. Structural analysis of human natural resistance-associated macrophage protein 1 promoter. Mol Immunol 33: 265-268, 1996.

Klemsz MJ, McKercher SR, Celada A, Van Beveren C, Maki RA. The macrophage and B cell-specific transcription factor PU.1 is related to the ets oncogene. Cell 61: 113-124, 1990.

Klimstra DS. Pathologic prognostic factors in esophageal carcinoma. Semin Oncol 21: 425-430, 1994.

Kojima Y, Kinouchi Y, Takahashi S, Negoro K, Hiwatashi N, Shimosegawa T. Inflammatory bowel disease is associated with a novel promoter polymorphism of natural resistance-associated macrophage protein 1 (*NRAMP1*) gene. Tissue Antigens 58: 379-384, 2001.

Koss K, Satsangi J, Fanning GC, Welsh KI, Jewell DP. Cytokine (TNF-alpha, LT-alpha, and IL-10) polymorphisms in inflammatory bowel diseases and normal controls: differential effects on production and allele frequencies. Genes Immun 1: 185-190, 2000.

Kotze MJ, de Villiers JNP, Rooney RN, Grobbelaar JJ, Mansvelt EPG, Bouwens CS, Carr J, Stander I, du Plessis L. Analysis of the *NRAMP1* gene implicated in iron transport: Association with multiple sclerosis and age effects. Blood Cells Mol Dis 27: 44-53, 2001.

Kotze MJ, de Villiers JNP, Zaahl MG, Robson KJH. The role of iron metabolism in multiple sclerosis. In: Metal Ions and Neurodegenerative Disorders. Ed. By Zatta P, World Sci. Singapore, London, 2003.

Kotze MJ, Theart L, Callis M, Peeters A, Thiart R, Langenhoven E. Nonradioactive multiplex PCR screening strategy for the simultaneous detection of multiple low-density lipoprotein receptor gene mutations. PCR Methods Applic 4: 352-356, 1995.

Krause A, Neitz S, Magert HJ, Schulz A, Forssmann WG, Schulz-Knappe P, Adermann K. LEAP-1, a novel highly disulfide-bonded human peptide, exhibits antimicrobial activity. FEBS Lett 480: 147-150, 2000.

Krawczak M, Reiss J, Cooper DN. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. Hum Genet 90: 41-54, 1992.

Kyle J. Crohn's disease in the northeastern and northern Isles of Scotland: an epidemiological review. Gastroenterol 103: 392-399, 1992.

Lang T, Prina E, Sibthorpe D, Blackwell JM. Nramp1 transfection transfers *Ity/Lsh/Bcg*-related pleiotropic effects on macrophage activation: Influence on antigen processing and presentation. Infect Immun 65: 380-386, 1997.

La Vecchia C, Negri E. The role of alcohol in oesophageal cancer in non-smokers, and of tobacco in non-drinkers. Int J Cancer 43: 784-785, 1989.

Leclercq V, Lebastard M, Belkaid Y, Louis J, Milon G. The outcome of the parasitic process initiated by Leishmania infantum in laboratory mice: a tissue-dependent pattern controlled by the Lsh and MHC loci. J Immunol 157: 4537-4545, 1996.

Lee P, Gelbart T, West C, Halloran C, Beutler E. Seeking candidate mutations that affect iron homeostasis. Blood Cells Mol Dis 29: 471-487, 2002.

Lernmark A, Ott J. Sometimes it's hot, sometimes it's not. Nat Genet 19: 213-214, 1998.

Lesage S, Zouali H, Cezard JP, Colombel JF, Belaiche J, Almer S, Tysk C, O'Morain C, Gassull M, Binder V, Finkel Y, Modigliani R, Gower-Rousseau C, Macry J, Merlin F, Chamaillard M, Jannot AS, Thomas G, Hugot JP; EPWG-IBD Group, EPIMAD Group, GETAID Group. *CARD15/NOD2* mutational analysis and genotype-phenotype correlation in 612 patients with inflammatory bowel disease. Am J Hum Genet 70: 845-857, 2002.

Levy JE, Montross LK, Cohen DE, Fleming MD, Andrews NC. The C282Y mutation causing hereditary hemochromatosis does not produce a null allele. Blood 94: 9-11, 1999.

Lewis LA, Victor TC, Helden EG, Blackwell JM, da Silva-Tatley F, Tullett S, Ehlers M, Beyers N, Donald PR, van Helden PD. Identification of C to T mutation at position -236 bp in the human *NRAMP1* gene promoter. Immunogenetics 44: 309-311, 1996.

Lewkonia R, McConnell R. Familial inflammatory bowel disease-heredity or environment? Gut 17: 237-243, 1976.

Liang YY and Lu SH. Expression of *C-myc* and *HER-1* genes in the development of human esophageal cancer. Chin J Oncol 13: 168-170, 1991.

Liaw YS, Tsai-Wu JJ, Wu CH, Hung CC, Lee CN, Yang PC, Luh KT, Kuo SH. Variations in the *NRAMP1* gene and susceptibility of tuberculosis in Taiwanese. Int J Tuberc Lung Dis 6: 454-460, 2002.

Liu J, Fujiwara TM, Buu NT, Sanchez FO, Cellier M, Paradis AJ, Frappier D, Skamene E, Gros P, Morgan K, Schurr E. Identification of polymorphisms and sequence variants in the human homologue of the mouse natural resistance-associated macrophage protein gene. Am J Hum Genet 56: 845-853, 1995.

Loftus E Jr, Silverstein M, Sandborn W, Tremaine W, Harmsen W, Zinsmeister A. Crohn's disease in Olmsted County, Minnesota, 1940-1993: incidence, prevalence, and survival. Gastroenterol 114: 1161-1168, 1998.

Lossos A, River Y, EliakimA, Steiner I. Neurological aspects of inflammatory bowel disease. Neurology 45: 416-421, 1995.

Loubser O, Marais AD, Kotze MJ, Godenir N, Thiart R, Scholtz CL, de Villiers JN, Hillermann R, Firth JC, Weich HF, Maritz F, Jones S, van der Westhuyzen DR. Founder mutations in the LDL receptor gene contribute significantly to the familial hypercholesterolemia phenotype in the indigenous South African population of mixed ancestry. Clinl Genet 55: 340-345, 1999.

Lu SH. Alterations of oncogenes and tumor suppressor genes in esophageal cancer in China. Mutat Res 462: 343-353, 2000.

Lu SH, Hseih LL, Luo FC, Weinstein IB. Amplification of the EGFreceptor and C-myc genes in human esophageal cancer. Int J Cancer 42: 502-505, 1988.

MacPhail AP, Simon MO, Torrance JD, Charlton RW, Bothwell TH, Isaacson C. Changing patterns of dietary iron overload in Black South Africans. Am J Clin Nutr 32: 1272-1278, 1979.

Maliarik MJ, Chen KM, Sheffer RG, Rybicki BA, Major ML, Popovich J Jr, Iannuzzi MC. The natural resistance-associated macrophage protein gene in African Americans with sarcoidosis. Am J Respir Cell Mol Biol 22: 672-675, 2000.

Malo D, Vogan K, Vidal S, Hu J, Cellier M, Schurr E, Fuks A, Bumstead N, Morgan K, Gros P. Haplotype mapping and sequence analysis of the mouse Nramp gene predict susceptibility to infection with intracellular parasites. Genomics 23: 51-61, 1994.

Mandard AM, Hainaut P, Hollstein M. Genetic steps in the development of squamous cell carcinoma of the esophagus. Mutat Res 462: 335-342, 2000.

Mandishona E, MacPhail AP, Gordeuk VR, Kedda MA, Paterson AC, Rouault TA, Kew MC. Dietary iron overload as a risk factor for hepatocellular carcinoma in Black Africans. Hepatology 27: 1563-1566, 1998.

Marquet S, Lepage P, Hudson TJ, Musser JM, Schurr E. Complete nucleotide sequence and genomic structure of the human NRAMP1 gene region on chromosome region 2q35. Mamm Genome 11: 755-762, 2000.

Marquet S, Sanchez FO, Arias M, Rodriguez J, Paris SC, Skamene E, Schurr E, Garcia LF. Variants of the human NRAMP1 gene and altered human immunodeficiency virus infection susceptibility. J Infect Dis 180: 1521-1525, 1999.

Martinez JL, Delgado-Iribarren A, Baquero F. Mechanisms of iron acquisition and bacterial virulence. FEMS Microbiol Rev 6: 45-56, 1990.

Mayeda A, Krainer AR. Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. Cell 68: 365-375, 1992.

McKie AT, Barrow D, Latunde-Dada GO, Rolfs A, Sager G, Mudaly E, Mudaly M, Richardson C, Barlow D, Bomford A, Peters TJ, Raja KB, Shirali S, Hediger MA, Farzaneh F, Simpson RJ. An iron-regulated ferric reductase associated with the absorption of dietary iron. Science 291: 1755-1759, 2001.

McKie AT, Latunde-Dada GO, Miret S, McGregor JA, Anderson GJ, Vulpe CD, Wrigglesworth JM, Simpson RJ. Molecular evidence for the role of a ferric reductase in iron transport. Biochem Soc Trans 30: 722-724, 2002.

McKie AT, Marciani P, Rolfs A, Brennan K, Wehr K, Barrow D, Miret S, Bomford A, Peters TJ, Farzaneh F, Hediger MA, Hentze MW, Simpson RJ. A novel duodenal iron-regulated transporter, IREG1, implicated in the basolateral transfer of iron to the circulation. Mol Cell 5: 299-309, 2000.

192

Meisner SJ, Mucklow S, Warner G, Sow SO, Lienhardt C, Hill AV. Association of NRAMP1 polymorphism with leprosy type but not susceptibility to leprosy per se in west Africans. Am J Trop Med Hyg 65: 733-735, 2001.

Mercier G, Bathelier C, Lucotte G. Frequency of the C282Y mutation of hemochromatosis in five French populations. Blood Cells Mol Dis 24: 165-166, 1998.

Merryweather-Clarke AT, Cadet E, Bomford A, Capron D, Viprakasit V, Miller A, McHugh PJ, Chapman RW, Pointon JJ, Wimhurst VLC, Livesey KJ, Tanphaichitr V, Rochette J, Robson KJH. Digenic inheritance of mutations in *HAMP* and *HFE* results in different types of haemochromatosis. Hum Mol Genet, in press.

Merryweather-Clarke AT, Pointon JJ, Jouanolle AM, Rochette J, Robson KJ. Geography of *HFE* C282Y and H63D mutations. Genet Test 4: 183-198, 2000.

Milani MY, Kotze MJ. Molecular diagnosis of hereditary haemochromatosis--identify an affected person and save a family. S Afr Med J 89: 263-264, 1999.

Miller S, Dykes D, Polesky H. A simple salting out procedure for extracting DNA from human nucleated cells. Nucl Acids Res 16: 1215, 1988.

Minuk GY, Lewkonia RM. Possible familial association of multiple sclerosis and inflammatory bowel disease. N Engl J Med 314: 586, 1986.

Montosi G, Donovan A, Totaro A, Garuti C, Pignatti E, Cassanelli S, Trenor CC, Gasparini P, Andrews NC, Pietrangelo A. Autosomal-dominant hemochromatosis is associated with a mutation in the ferroportin (*SLC11A3*) gene. J Clin Invest 108: 619-623, 2001.

Moskaluk CA, Hu J, Perlman EJ. Comparative genomic hybridization of esophageal and gastroesophageal adenocarcinomas shows consensus areas of DNA gain and loss. Genes Chromosomes Cancer 22: 305-311, 1998.

Mount S. A catalogue of splice junction sequences. Nucleic Acids Res 10: 459-472, 1982.

Muir WA, McLaren GD, Braun W, Askari A. Evidence for heterogeneity in hereditary hemochromatosis. Evaluation of 174 persons in nine families. Am J Med 76: 806-814, 1984.

Mulero V, Searle S, Blackwell JM, Brock JH. Solute carrier 11a1 (*Slc11a1*; formerly *Nramp1*) regulates metabolism and release of iron acquired by phagocytic, but not transferrin-receptor-mediated, iron uptake. Biochem J 363: 89-94, 2002.

Muro AF, Caputi M, Pariyarath R, Pagani F, Buratti E, Baralle FE. Regulation of fibronectin EDA exon alternative splicing: possible role of RNA secondary structure for enhancer display. Mol Cell Biol 19: 2657-2671, 1999.

Murray MJ, Murray AB. Starvation suppression and refeeding activation of infection; an ecological necessity? Lancet 1: 123-125, 1977.

Murray MJ, Murray A.B, Murray CJ. An ecological interdependence of diet and disease? A study of infection in one tribe consuming two different diets. Am J Clin Nutr 33: 697-701, 1980.

Nicolas G, Bennoun M, Devaux I, Beaumont C, Grandchamp B, Kahn A, Vaulont S. Lack of hepcidin gene expression and severe tissue iron overload in upstream stimulatory factor 2 (*USF2*) knockout mice. Proc Natl Acad Sci USA 98: 8780-8785, 2001.

Nielsen OJ, Andersen LS, Hansen NE and Hansen TM. Serum transferrin receptor levels in anemic patients with rheumatoid arthritis. Scand J Clin Lab Invest 54: 75-82, 1994.

Njajou OT, Vaessen N, Joosse M, Berghuis B, van Dongen JW, Breuning MH, Snijders PJ, Rutten WP, Sandkuijl LA, Oostra BA, van Duijn CM, Heutink P. A mutation in *SLC11A3* is associated with autosomal dominant hemochromatosis. Nat Genet 28: 213-214, 2001.

Norell S, Ahlbom A, Lipping H, Osterblom L. Oesophageal cancer and vulcanisation work. Lancet 1: 462-463, 1983.

O'Callaghan NJ, Adams KE, Walker EJ, van Heel DA, Cavanaugh JA. Association of TNF-α -857 with IBD in the Australian population. Digestive Disease Week M1534, 524, 2003.

Ogura Y, Bonen D, Inohara N, Nicolae D, Chen F, Ramos R, Britton H, Moran T, Karaliuskas R, Duerr R, Achkar J, Brant S, Bayless T, Kirschner B, Hanauer S, Nunez G, Cho J. A frameshift mutation in *NOD2* associated with susceptibility to Crohn's disease. Nature 411: 603-606, 2001a.

Ogura Y, Inohara N, Benito A, Chen FF, Yamaoka S, Nunez G. *Nod2*, a Nod1/Apaf-1 family member that is restricted to monocytes and activates NF-κB. J Biol Chem 276: 4812-4818, 2001b.

Oldenburg B, Koningsberger JC, van Berge Henegouwen GP, van Asbeck BS, Marx JJM. Iron and inflammatory bowel disease. Aliment Pharmacol Ther 15: 429-438, 2001.

Oppenheimer SJ, Mac Farlane SB, Moody JB, Bunari O, Hendrickse RG. Effects of iron prophylaxis on morbidity due to infectious disease: report on clinical studies in Papua New Guinea. Trans R Soc Trop Med Hyg 80: 596-602, 1986.

Orholm M, Munkholm P, Langholz E, Nielsen O, Sorensen T, Binder V. Familial occurrence of inflammatory bowel disease. New Eng J Med 324: 84-88, 1991.

Orlando RC. Mechanisms of epithelial injury and inflammation in gastrointestinal diseases. Rev Gastroenterol Disord 2: S2-8, 2002.

Paavola P, Helio T, Kiuru M, Halme L, Turunen U, Terwilliger J, Karvonen AL, Julkunen R, Niemela S, Nurmi H, Farkkila M, Kontula K. Genetic analysis in Finnish families with inflammatory bowel disease supports linkage to chromosome 3p21. Eur J Hum Genet 9: 328-334, 2001.

Padgett R, Grabowski P, Konarska M, Seiler S, Sharp P. Splicing of messenger RNA precursors. Annu Rev Biochem 55: 1119-1150, 1986.

Pagani F, Stuani C, Tzetis M, Kanavakis E, Efthymiadou A, Doudounakis S, Casals T, Baralle FE. New type of disease causing mutations: the example of the composite exonic regulatory elements of splicing in CFTR exon 12. Hum Mol Genet 12: 1111-1120, 2003.

Park CH, Valore EV, Waring AJ, Ganz T. Hepcidin, a urinary antimicrobial peptide synthesized in the liver. J Biol Chem 276: 7806-7810, 2001.

Parkes M, Barmada M, Satsangi J, Weeks D, Jewell D, Duerr R. The IBD2 locus shows linkage heterogeneity between ulcerative colitis and Crohn disease. Am J Hum Genet 67: 1605-1610, 2000.

Parkin DM, Laara E, Muir CS. Estimates of the worldwide frequency of sixteen major cancers in 1980. Int J Cancer 41: 184-197, 1988.

Parkkila S, Waheed A, Britton RS, Bacon BR, Zhou XY, Tomatsu S, Fleming RE, Sly WS. Association of the transferrin receptor in human placenta with HFE, the protein defective in hereditary hemochromatosis. Proc Natl Acad Sci USA 94: 13198-13202, 1997.

Pigeon C, Ilyin G, Courselaud B, Leroyer P, Turlin B, Brissot P, Loreal O. A new mouse liver-specific gene, encoding a protein homologous to human antimicrobial peptide hepcidin, is overexpressed during iron overload. J Biol Chem 276: 7811-7819, 2001.

Piperno A, Arosio C, Fossati L, Vigano M, Trombini P, Vergani A, Mancia G. Two novel nonsense mutations of HFE gene in five unrelated Italian patients with hemochromatosis. Gastroenterol 119: 441-445, 2000.

Piperno A, Sampietro M, Pietrangelo A, Arosio C, Lupica L, Montosi G, Vergani A, Fraquelli M, Girelli D, Pasquero P, Roetto A, Gasparini P, Fargion S, Conte D, Camaschella C. Heterogeneity of hemochromatosis in Italy. Gastroenterol 114: 996-1002, 1998.

Plant J, Glynn AA. Natural resistance to Salmonella infection, delayed hypersensitivity and Ir genes in different strains of mice. Nature 248: 346-347, 1974.

Podolsky DK. The current future understanding of inflammatory bowel disease. Best Pract Res Clin Gastroenterol 16: 933-943, 2002.

Pointon JJ. The Genetics of Haemochromatosis. University of Oxford, Michaelmas 2001.

Pointon JJ, Wallace D, Merryweather-Clarke AT, Robson KJH. Uncommon mutations in the hemochromatosis gene. Genet Test 4: 151-161, 2000.

Powell LW, Jazwinska EC, Halliday JW. Primary iron overload. In: Iron Metabolism in Health and Disease. Ed. By JH Brock, JW Halliday, MJ Pippard, LW Powell, 228-270, London, Saunders, 1994.

Puliti M, Radzioch D, Mazzolla R, Barluzzi R, Bistoni F, Blasi E. Influence of the Bcg locus on macrophage response to the dimorphic fungus Candida albicans. Infect Immun 63: 4170-4173, 1995.

Rao PH, Mathew S, Kelsen DP, Chaganti RS. Cytogenetics of gastric and esophageal adenocarcinomas 3q deletion as a possible primary chromosomal change. Cancer Genet Cytogenet 81: 139-143, 1995.

Reif S, Klein I, Arber N, Gilat T. Lack of association between smoking and inflammatory bowel disease in Jewish patients in Israel. Gastroenterol 108: 1683-1687, 1995.

Rhodes D, Raha-Chowdhury R, Cox TM, Trowsdale J. Homozygosity for the predominant Cys282Tyr mutation and absence of disease expression in hereditary haemochromatosis. J Med Genet 34: 761-764, 1997.

Rich A, Nordheim A, Wang AH. The chemistry and biology of left-handed Z-DNA. Annu Rev Biochem 53: 791-846, 1984.

Rioux JD, Daly MJ, Green T, Stone V, Lander ES, Hudson TJ, Steinhart AH, Bull S, Cohen Z, Greenberg G, Griffiths A, McLeod R, Silverberg M, Williams CN, Siminovitch KA. Absence of linkage between inflammatory bowel disease and selected loci on chromosomes 3, 7, 12, and 16. Gastroenterol 115: 1062-1065, 1998.

Rioux J, Daly M, Silverberg M, Lindblad K, Steinhart H, Cohen Z, Delmonte T, Kocher K, Miller K, Guschwan S, Kulbokas E, O'Leary S. Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. Nature Genet 29: 223-228, 2001.

Rioux J, Silverberg M, Daly M, Steinhart A, McLeod R, Griffiths A, Green T, Brettin T, Stone V, Bull S, Bitton A, Williams C, Greenberg G, Cohen Z, Lander E, Hudson T, Siminovitch K. Genomewide search in Canadian families with inflammatory bowel disease reveals two novel susceptibility loci. Am J Hum Genet 66: 1863-1870, 2000.

Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science 273: 1516-1517, 1996.

Roach TI, Kiderlen AF, Blackwell JM. Role of inorganic nitrogen oxides and tumor necrosis factor alpha in killing *Leishmania donovani* amastigotes in gamma interferon-lipopolysaccharide-activated macrophages from *Lshs* and *Lshr* congenic mouse strains. Infect Immun 59: 3935-3944, 1991.

Rochette J, Pointon JJ, Fisher CA, Perera G, Arambepola M, Arichchi DS, De Silva S, Vandwalle JL, Monti JP, Old JM, Merryweather-Clarke AT, Weatherall DJ, Robson KJ. Multicentric origin of hemochromatosis gene (*HFE*) mutations. Am J Hum Genet 64: 1056-1062, 1999.

Rodriguez MR, Gonzalez-Escribano MF, Aguilar F, Valenzuela A, Garcia A, Nunez-Roldan A. Association of *NRAMP1* promoter gene polymorphism with the susceptibility and radiological severity of rheumatoid arthritis. Tissue Antigens 59: 311-315, 2002.

Roetto A, Merryweather-Clarke AT, Daraio F, Livesey K, Pointon JJ, Barbabietola G, Piga A, Mackie PH, Robson KJ, Camaschella C. A valine deletion of ferroportin 1: a common mutation in hemochromastosis type 4. Blood 100: 733-734, 2002.

Roetto A, Papanikolaou G, Politou M, Alberti F, Girelli D, Christakis J, Loukopoulos D, Camaschella C. Mutant antimicrobial peptide hepcidin is associated with severe juvenile hemochromatosis. Nat Genet 33: 21-22, 2003.

Roetto A, Totaro A, Piperno A, Piga A, Longo F, Garozzo G, Cali A, De Gobbi M, Gasparini P, Camaschella C. New mutations inactivating transferrin receptor 2 in hemochromatosis type 3. Blood 97: 2555-2560, 2001.

Rogers AE, Zeisel SH, Groopman J. Diet and carcinogenesis. Carcinogenesis 14: 2205-2217, 1993.

Rosen N. The molecular basis for cellular transformation: implications for esophageal carcinogenesis. Semin Oncol 21: 416-424, 1994.

Rosenberg N, Murata M, Ikeda Y, Opare-Sem O, Zivelin A, Geffen E, Seligsohn U. The frequent 5,10-methylenetetrahydrofolate reductase C677T polymorphism is associated with a common haplotype in whites, Japanese, and Africans. Am J Hum Genet 70: 758-762, 2002.

Rosenblum-Vos LS, Meltzer SJ, Leana-Cox J, Schwartz S. Cytogenetic studies of primary cultures of esophageal squamous cell carcinoma. Cancer Genet Cytogenet 70: 127-131, 1993.

Rosmorduc O, Poupon R, Nion I, Wendum D, Feder J, Bereziat G, Hermelin B. Differential *HFE* allele expression in hemochromatosis heterozygotes. Gastroenterol 119: 1075-1086, 2000.

Ruuls SR, Sedgwick JD. Unlinking tumor necrosis factor biology from the major histocompatibility complex: lessons from human genetics and animal models. Am J Hum Genet 65: 294-301, 1999.

Ryu S, Park YK, Bai GH, Kim SJ, Park SN, Kang S. 3'UTR polymorphisms in the *NRAMP1* gene are associated with susceptibility to tuberculosis in Koreans. Int J Tuberc Lung Dis 4: 577-580, 2000.

Sachot S, Moirand R, Jouanolle AM, Mosser J, Fergelot P, Deugnier Y, Brissot P, le Gall JY, David V. Low penetrant hemochromatosis phenotype in eight families: no evidence of modifiers in the MHC region. Blood Cells Mol Dis 27: 518-529, 2001.

Sambuelli AM, Murillo L, Negreira S, Gil A, Huernos S, Goncalves S, Vazquez H, Maurino E, Bai JC, Pena AS. Correlation between mutations of *NOD2/CARD15* gene and clinical patterns of Crohn's disease (CD). Digestive Disease Week, Orlando, Florida: M1568, 528, 2003.

Sanderson I. Chronic inflammatory bowel disease. Clin Gastroenterol 15: 71-87, 1986.

Sanjeevi CB, Miller EN, Dabadghao P, Rumba I, Shtauvere A, Denisova A, Clayton D, Blackwell JM. Polymorphism at *NRAMP1* and D2S1471 loci associated with juvenile rheumatoid arthritis. Arthritis Rheum 43: 1397-1404, 2000.

Sarlomo-Rikala M, El-Rifai W, Lahtinen T, Andersson LC, Miettinen M, Knuutila S. Different patterns of DNA copy number changes in gastrointestinal stromal tumors, leiomyomas, and schwannomas. Hum Pathol 29: 476-481, 1998.

Sartor R. Current concepts of the etiology and pathogenesis of ulcerative colitis and Crohn's disease. Gastroenterol Clin North Am 24: 475-507, 1995.

Sashio H, Tamura K, Ito R, Yamamoto Y, Bamba H, Kosaka T, Fukui S, Sawada K, Fukuda Y, Tamura K, Satomi M, Shimoyama T, Furuyama J. Polymorphisms of the *TNF* gene and the TNF receptor superfamily member 1B gene are associated with susceptibility to ulcerative colitis and Crohn's disease, respectively. Immunogenetics 53: 1020-1027, 2002.

Satsangi J, Grootscholten C, Holt H, Jewell D. Clinical patterns of familial inflammatory bowel disease. Gut 38: 738-741, 1996b.

Satsangi J, Jewell D, Bell J. The genetics of inflammatory bowel disease. Gut 40: 572-574, 1997.

Satsangi J, Jewell D, Rosenberg W, Bell J. Genetics of inflammatory bowel disease. Gut 35: 696-700, 1994.

Satsangi J, Parkes M, Louis E, Hashimoto L, Kato N, Welsh K, Terwilliger J, Lathrop G, Bell J, Jewell D. Two stage genome-wide search in inflammatory bowel disease provides evidence for susceptibility loci on chromosomes 3, 7 and 12. Nature Genet 14: 199-202, 1996a.

Schurr E, Malo D, Radzioch D, Buschman E, Morgan K, Gros P, Skamene E. Genetic control of innate resistance to mycobacterial infections. Immunol Today 12: A42-45, 1991.

Searle S, Blackwell JM. Evidence for a functional repeat polymorphism in the promoter of the human *NRAMP1* gene that correlates with autoimmune versus infectious disease susceptibility. J Med Genet 36: 295-299, 1999.

Searle S, Bright NA, Roach TI, Atkinson PG, Barton CH, Meloen RH, Blackwell JM. Localisation of Nramp1 in macrophages: modulation with activation and infection. J Cell Sci 111: 2855-2866, 1998.

Selikoff IJ, Hammond EC, Seidman H. Mortality experience of insulation workers in the United States and Canada, 1943-1976. Ann N Y Acad Sci 330: 91-116, 1979.

Shanahan F. Progress in the pathogenesis of ulcerative colitis. In: Inflammatory bowel disease. Ed. By Tytgat G, Bartelsman J, Van Deventer S, 32-39, Kluwer, Dordrecht, 1995.

Shanahan F, Randolph L, King R, Oseas R, Brogan M, Witkop C, Rotter J, Targan S. Hermansky-Pudlak syndrome: an immunologic assessment of 15 cases. Am J Med 85: 823, 1988.

Shatwell KP, Dancis A, Cross AR, Klausner RD, Segal AW. The FRE1 ferric reductase of *Saccharomyces cerevisiae* is a cytochrome b similar to that of NADPH oxidase. J Biol Chem 271: 14240-14244, 1996.

Shaw MA, Clayton D, Atkinson SE, Williams H, Miller N, Sibthorpe D, Blackwell JM. Linkage of rheumatoid arthritis to the candidate gene *NRAMP1* on 2q35. J Med Genet 33: 672-677, 1996.

197

Shaw MA, Clayton D, Blackwell JM. Analysis of the candidate gene *NRAMP1* in the first 61 ARC National Repository families for rheumatoid arthritis. J Rheumatol 24: 212-214, 1997.

Shimada M, Yanagisawa A, Kato Y, Inoue M, Shiozaki H, Monden M, Nakamura Y. Genetic mechanisms in esophageal carcinogenesis: frequent deletion of 3p and 17p in premalignant lesions. Genes Chromosomes Cancer 15: 165-169, 1996.

Simon M, Pawlotsky Y, Bourel M, Fauchet R, Genetet B. Letter: Idiopathic hemochromatosis associated with HL-A 3 tissular antigen. Nouv Presse Med 4: 1432, 1975.

Singal DP, Li J, Zhu Y, Zhang G. *NRAMP1* gene polymorphisms in patients with rheumatoid arthritis. Tissue Antigens 55: 44-47, 2000.

Skamene E, Gros P, Forget A, Kongshavn PA, St Charles C, Taylor BA. Genetic regulation of resistance to intracellular pathogens. Nature 297: 506-509, 1982.

Skamene E, Gros P, Forget A, Patel PJ, Nesbitt MN. Regulation of resistance to leprosy by chromosome 1 locus in the mouse. Immunogenetics 19: 117-124, 1984.

Smith PG. Late effects of x-ray treatment of anklyosing spondylitis. In: Radiation Carcinogenesis. Ed. By Boice J, Fraumeni J,107-118, Raven, New York, 1984.

Snook J, de Silva H, Jewell D. The association of autoimmune disorders with inflammatory bowel disease. Q J Med 72: 835-840, 1989.

Somerville K, Logan R, Edmond M, Langman M. Smoking and Crohn's disease. Br Med J 289: 954-956, 1994.

Steinmetz KA, Potter JD. Vegetables, fruit, and cancer. I. Epidemiology. Cancer Causes Control 2: 325-357, 1991.

Stokkers P, Curvers W, Vogels E, Hommes D, van Deventer S. Low prevalence of NOD2 variants in Dutch cohort of sporadic Crohn's disease. Digestive Disease Week, Orlando, Florida: M1554, 526, 2003.

Stokkers P, de Heer K, Leegwater A, Reitsma P, Tytgat G, van Deventer S. Inflammatory bowel disease and the genes for the natural resistance-associated macrophage protein 1 and the interferon-γ receptor 1. Int J Colorect Dis 14: 13-17, 1999.

Strugill-Koszycki S, Schaible U, Russell DG. Mycobacterium-containing phagosomes are accessible to early endosomes and reflect a transitional state in normal phagosome biogenesis. EMBO J 24: 6960-6968, 1996.

Supek F, Supekova L, Nelson H, Nelson N. A yeast manganese transporter related to the macrophage protein involved in conferring resistance to mycobacteria. Proc Natl Acad Sci USA 93: 5105-5110, 1996.

Targan S. Definition of the pathogenesis of Crohn's disease: evolution of a hypothetical model. In: Inflammatory bowel disease. Ed. By Tytgat G, Bartelsman J, Van Deventer S, 40-47, Kluwer, Dordrecht, 1995.

Targan S, Murphy L. Clarifying the causes of Crohn's. Nat Med 1: 1241-1243, 1995.

Tarmin L, Yin J, Zhou X, Suzuki H, Jiang HY, Rhyu MG, Abraham JM, Krasna MJ, Cottrell J, Meltzer SJ. Frequent loss of heterozygosity on chromosome 9 in adenocarcinoma and squamous cell carcinoma of the esophagus. Cancer Res 54: 6094-6096, 1994.

The UK Haemochromatosis Consortium. A simple genetic test identifies 90% of UK patients with haemochromatosis. Gut 41: 841-844, 1997.

Togokuni S. Iron induced carcinogenesis: the role of redox regulation. Free Radic Biol Med 20: 553-566, 1996.

Tysk C, Lindberg E, Jarnerot G, Floderus-Myrhed B. Ulcerative colitis and Crohn's disease in unselected population of monozygotic and dizygotic twins: a study of heritability and the influence of smoking. Gut 29: 990-996, 1988.

Van Asbeck BS, van Verbrugh HA, Oost B.A, Marx JJM, Imhof HW and Verhoef J. Listeria monocytogenes meningitis and decreased phagocytosis associated with iron overload. Br Med J 284: 542-544, 1982.

Van Heel DA, McGovern DP, Cardon LR, Dechairo BM, Lench NJ, Carey AH, Jewell DP. Fine mapping of the IBD1 locus did not identify Crohn disease-associated NOD2 variants: implications for complex disease genetics. Am J Med Genet 111: 253-259, 2002.

Van Heel DA, McGovern DP, Jewell DP. Crohn's disease: genetic susceptibility, bacteria and innate immunity. Lancet 357: 1902-1904, 2001.

Van Montfrans C, Peppelenbosch M, te Velde AA, van Deventer S. Inflammatory signal transduction in Crohn's disease and novel therapeutic approaches. Biochem Pharmacol 64: 789-795, 2002.

Van Rensburg SJ. Epidemiologic and dietary evidence for a specific nutritional predisposition to esophageal cancer. J Natl Cancer Inst 67: 243-251, 1981.

Vermeire S, Peeters M, Vlietinck R, Parkes M, Satsangi J, Jewell D, Rutgeerts P. Exclusion of linkage of Crohn's disease to previously reported regions on chromosomes 12, 7, and 3 in the Belgian population indicates genetic heterogeneity. Inflamm Bowel Dis 6: 165-170, 2000.

Vidal SM, Belouchi A, Cellier M, Beatty B, Gros P. Cloning and characterisation of a second human NRAMP gene on chromosome 12q13. Mamm Genome 6: 224-230, 1995.

Vidal SM, Malo D, Vogan K, Skamene E, Gros P. Natural resistance to infection with intracellular parasites: isolation of a candidate for Bcg. Cell 73: 469-485, 1993.

Vulpe CD, Kuo YM, Murphy TL, Cowley L, Askwith C, Libina N, Gitschier J, Anderson GJ. Hephaestin, a ceruloplasmin homologue implicated in intestinal iron transport, is defective in the sla mouse. Nat Genet 21: 195-199, 1999.

Wallace DF, Dooley JS, Walker AP. A novel mutation of HFE explains the classical phenotype of genetic hemochromatosis in a C282Y heterozygote. Gastroenterology 116: 1409-1412, 1999.

Wallace DF, Pedersen P, Dixon JL, Stephenson P, Searle JW, Powell LW, Subramaniam VN. Novel mutation in ferroportin1 is associated with autosomal dominant hemochromatosis. Blood 100: 692-694, 2002.

Weber J, Werre JM, Julius HW, Marx JJ. Decreased iron absorption in patients with active rheumatoid arthritis, with and without iron deficiency. Ann Rheum Dis 47: 404-409, 1988.

Weiss G. Iron and immunity: a double-edged sword. Eur J Clin Invest 32 (suppl 1): 70-78, 2002.

Weiss G, Fuchs D, Hausen A, Reibnegger G, Werner ER, Werner-Felmayer G, Wachter H. Iron modulates interferon-gamma effects in the human myelomonocytic cell line THP-1. Exp Hematol 20: 605-610, 1992.

Weitzman J. Bowel disease gets the Nod. TRENDS in Mol Med 7: 335-336, 2001.

Whang-Peng J, Banks-Schlegel SP, Lee EC. Cytogenetic studies of esophageal carcinoma cell lines. Cancer Genet Cytogenet 45: 101-120, 1990.

White JK, Shaw MA, Barton CH, Cerretti DP, Williams H, Mock BA, Carter NP, Peacock CS, Blackwell JM. Genetic and physical mapping of 2q35 in the region of the *NRAMP* and *IL8R* genes: identification of a polymorphic repeat in exon 2 of *NRAMP*. Genomics 24: 295-302, 1994.

Worwood M, Jackson HA, Feeney GP, Edwards C, Bowen DJ. A single tube heteroduplex PCR for the common HFE genotypes. Blood 94 (S1): 405a, 1999.

Worwood M, Shearman JD, Wallace DF, Dooley JS, Merryweather-Clarke AT, Pointon JJ, Rosenberg WMC, Bowen DJ, Burnett AK, Jackson HA, Lawless S, Raha-Chowdhury R, Partridge J, Williams R, Bomford A, Walker AP, Robson KJH. A simple genetic test identifies 90% of UK patients with haemochromatosis. Gut 41: 841-844, 1997.

Wright AC, Simpson LM, Oliver JD. Role of iron in the pathogenesis of the Vibrio vulnificus infections. Infect Immunol 34: 503-507, 1981.

Yamazaki K, Takazoe M, Tanaka T, Kazumori T, Nakamura Y. Absence of mutation in the *NOD2/CARD15* gene among 483 Japanese patients with Crohn's disease. J Hum Genet 47: 469-472, 2002.

Yang YS, Kim SJ, Kim JW, Koh EM. *NRAMP1* gene polymorphisms in patients with rheumatoid arthritis in Koreans. J Korean Med Sci 15: 83-87, 2000.

Yang H, McElree C, Roth M, Shanahan F, Targan S, Rotter J. Familial empirical risks for inflammatory bowel disease: differences between Jews and non-Jews. Gut 34: 517-524, 1993.

Yang H, Rotter J. Genetics of inflammatory bowel disease. In: Inflammatory Bowel Disease: From Bench to Bedside. Ed. By Targan S, Shanahan F, 32-64, Williams and Wilkins, Baltimore, 1993.

Yang Z, Sugawara M, Ponath PD, Wessendorf L, Banerji J, Li Y, Strominger JL. Interferon gamma response region in the promoter of the human *DPA* gene. Proc Natl Acad Sci USA 87: 9226-9230, 1990.

Yang CS, Wang ZY. Tea and cancer. J Natl Cancer Inst 85: 1038-1049, 1993.

Yang Z, Zhu T, Ma G, Yin H, Qian W, Zhang F, Cao K, Ma W. Apolipoprotein E polymorphism in the early onset of coronary heart disease. Chin Med J (Engl) 114: 983-985, 2001.

Zahler AM, Neugebauer KM, Lane WS, Roth MB. Distinct functions of SR proteins in alternative pre-mRNA splicing. Science 260: 219-222, 1993.

Zhou XY, Tomatsu S, Fleming RE, Parkkila S, Waheed A, Jiang J, Fei Y, Brunt EM, Ruddy DA, Prass CE, Schatzman RC, O'Neill R, Britton RS, Bacon BR, Sly WS. *HFE* gene knockout produces mouse model of hereditary hemochromatosis. Proc Natl Acad Sci USA 95: 2492-2497, 1998.

Zouali H, Chamaillard M, Lesage S, Cezard JP, Colombel JF, Belaiche J, Almer S, Tysk C, Montague S, Gassull M, Christensen S, Finkel Y, Gower-Rousseau C, Modigliani R, Macry J, Selinger-Leneman H, Thomas G, Hugot JP; GETAID. Genetic refinement and physical mapping of a chromosome 16q candidate region for inflammatory bowel disease. Eur J Hum Genet 9: 731-742, 2001.

Zwilling BS, Hilburger ME. Macrophage resistance genes: *Bcg/Ity/Lsh*. Immunol Ser 60: 233-245, 1994.

Zwilling BS, Kuhn DE, Wikoff L, Brown D, Lafuse W. Role of iron in *Nramp1*-mediated inhibition of mycobacterial growth. Infect Immun 67: 1386-1392, 1999.

APPENDIX A

ABSTRACTS FOR CONFERENCE CONTRIBUTIONS DURING

PH D STUDY

ABSTRACT 1

Poster presentations at:

1. The American Society of Human Genetics 52nd Annual Meeting, Baltimore, Maryland, USA, 15-19 October 2002.

2. British Human Genetics Conference, York, United Kingdom, 23-25 September 2002.

Published abstracts:

Zaahl MG, Peeters AV, Warnich L, Kotze MJ, Robson KJH. Transcriptional activity of variants of the Z-DNA forming *NRAMP1* promoter repeat polymorphism. American Journal of Human Genetics Suppl to vol 71 (4): abstract 924, pg 330, 2002.

Zaahl MG, Peeters AV, Warnich L, Kotze MJ, Robson KJH. Transcriptional activity of variants of the Z-DNA forming *NRAMP1* promoter repeat polymorphism. Journal of Medical Genetics Suppl 1: abstract 3.50, pg S74, 2002.

ABSTRACT 2

Poster presentation at:

1. Digestive Disease Week, Orlando, Florida, USA, 17-22 May 2003.

Oral presentation at:

1. 10th Biennial congress of the Southern African Society of Human Genetics (SASHG), Durban, South Africa, 11-14 May 2003.

ABSTRACT 3

Poster presentations at:

1. Joint Neurology Association Congress, Kramer Building, Middle Campus, University of Cape Town, South Africa, 29 January to 1 February 2003.

2. 10th Biennial congress of the Southern African Society of Human Genetics (SASHG), Durban, South Africa, 11-14 May 2003.

ABSTRACT 4

Poster presentations at:

1. Bioiron 2003, Washington DC, USA, 4-9 May 2003.

2. 10th Biennial congress of the Southern African Society of Human Genetics (SASHG), Durban, South Africa, 11-14 May 2003.

ABSTRACT 5

Poster presentations at:

1. Bioiron 2003, Washington DC, USA, 4-9 May 2003.

The name of the presenting author is underlined.

ABSTRACT 1

Transcriptional activity of variants of the Z-DNA forming *NRAMP1* promoter repeat polymorphism

Zaahl MG[1,2], Warnich L[2], Kotze MJ[3], Robson KJH[4]

[1]Division of Human Genetics, Faculty of Health Sciences, University of Stellenbosch, Tygerberg, [2]Department of Genetics, University of Stellenbosch, Stellenbosch, [3]Genecare Molecular Genetics, Christiaan Barnard Annexe, Cape Town, SOUTH AFRICA and [4]MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Oxford, UNITED KINGDOM.

Background: *NRAMP1* is thought to be involved in iron homeostasis. Previous studies indicated that allele 2 $[t(gt)_5(gt)_5(gt)_{10}]$ of a Z-DNA forming repeat polymorphism in the promoter region is associated with infectious disease susceptibility and allele 3 $[t(gt)_5(gt)_5(gt)_9]$ with autoimmune disease susceptibility

Aim: To determine the effect of the various repeat polymorphism alleles on *NRAMP1* expression.

Materials and methods: The variants were cloned in a luciferase-reporter vector (pGL2) and transfected into U937 and THP-1 cell lines using FUGENE 6. Cells were incubated in the absence and presence of exogenous stimuli, including interferon-γ, bacterial lipopolysaccharide, bacterial lipopolysaccharide plus interferon-γ, and ferric ammonium citrate. Luciferase activity in harvested cells was determined luminometrically.

Results and Discussion: Six promoter variants were studied, including four previously described variants [alleles 2 and 3, allele 4 $(t(gt)_5(gt)_5(gt)_4)$ and allele 5 $(t(gt)_4(gt)_5(gt)_{10})$]

and two novel variants [allele 6 $(t(gt)_5(gt)_5(gt)_6)$ and allele 7 $(t(gt)_5(gt)_5(gt)_8)$], and one previously described base pair substitution (-237C/T). Ferric ammonium citrate caused a five-fold enhancement of luciferase reporter gene expression for allele 3 with U937 cells. Interferon-γ and bacterial lipopolysaccharide caused a ten-fold enhancement of gene expression for allele 5 and a five fold-enhancement for the addition of interferon-γ for the –237C/T polymorphism in THP-1 cells.

Conclusion

The alleles differ in their ability to drive gene expression between constructs and cell lines. These results support the hypothesis that this functional repeat polymorphism could contribute directly to the association of variants with infectious or autoimmune disease.

ABSTRACT 2

Analysis of the *SLC11A1* gene in South African patients with inflammatory bowel disease

M. G. Zaahl[1], **A.V. Peeters**[1], **L. Warnich**[1], **T. Winter**[2], **K.J.H. Robson**[3], **M. J. Kotze**[4]

[1]Department of Genetics, University of Stellenbosch, Stellenbosch, [2]Gastrointestinal Clinic, Groote Schuur Hospital, Cape Town, South Africa, [3]MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Oxford, United Kingdom and [4]Genecare Molecular Genetics, Christiaan Barnard Memorial Hospital, Cape Town, South Africa.

Background: Inflammatory bowel disease (IBD) is subdivided into Crohn disease (CD) and Ulcerative colitis (UC) and occurs due to chronic inflammation of the intestines. The aetiology of IBD remains unknown but it has been suggested that genetic, infectious and environmental factors contribute to the disease phenotype. The *solute carrier family 11 member 1 (SLC11A1)* gene, previously known as the natural resistance associated macrophage protein 1 (*NRAMP1*) gene, has previously been implicated in Crohn disease, but this association could not be confirmed in all the populations studied.

Aim: To investigate the likelihood that variation in the *SLC11A1* gene may underlie susceptibility to IBD in South African patients and to determine the effect of the various repeat polymorphism alleles on gene expression.

Materials and Methods: DNA extraction was performed on blood samples of 77 IBD patients, including 42 CD and 35 UC patients, and 110 control individuals. PCR amplification was performed on the promoter region and exon 2 of *SLC11A1* followed by heteroduplex single-strand conformation polymorphism (HEX-SSCP) and/or restriction fragment length polymorphism (RFLP) analysis using *RsaI. SLC11A1* promoter region

variants were cloned in a luciferase-reporter vector (pGL2) and transfected into U937 and THP-1 cell lines using FUGENE 6. Cells were incubated in the absence and presence of exogenous stimuli, including interferon-γ (IFN), bacterial lipopolysaccharide (LPS) and ferric ammonium citrate (FAC). Luciferase activity in harvested cells was determined luminometrically.

Results and Discussion: HEX-SSCP and RFLP analysis revealed several novel variants in the promoter region (-8G/A; t(gt)$_5$ac(gt)$_5$ac(gt)$_6$ggcaga(g)$_6$; t(gt)$_5$ac(gt)$_5$ac(gt)$_8$ggcaga(g)$_6$) and exon 2 (112G/A; IVS1-28C/T; 147delGACCAGCCC; 156insGACCAGCCCAG). Statistically significant differences were obtained for the total IBD patient group ($p < 0.0007$) and CD ($p < 0.002$) and UC ($p < 0.05$) patients groups, respectively, compared with the control group for the –237C/T promoter polymorphism. Six promoter variants were cloned. FAC caused a five-fold enhancement of luciferase reporter gene expression for allele 3 compared to the various constructs. This was statistically significant for all constructs ($p < 0.05$).

Conclusion: The alleles differ in their ability to drive gene expression upon iron loading. This finding provides direct support for the hypothesis that iron dysregulation may underlie the allelic association of the functional *SLC11A1* promoter polymorphism with infectious and autoimmune diseases.

ABSTRACT 3

Microarray analysis and microsatellite genotyping in South African patients with multiple sclerosis: Identification of a significant association with the gene encoding monocyte chemotactic protein-3

<u>Zaahl MG</u>[1,2], Merryweather-Clarke AT[3], de Villiers JNP[1,4], Pointon JJ[3], Smith L[5], Freeman T[6], Williams D[6], Pritchard C[5], Warnich L[2], Robson KJH[3], Kotze MJ[4]

[1]Division of Human Genetics, Faculty of Health Sciences, University of Stellenbosch, Tygerberg, South Africa, [2]Department of Genetics, University of Stellenbosch, Stellenbosch, South Africa, [3]MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Oxford, UK, [4]Genecare Molecular Genetics, Christiaan Barnard Memorial Hospital, Cape Town, South Africa, [5]MRC Mammalian Genetics Unit, Harwell, Didcot, UK, and [6]UK MRC HGMP Resource Centre, Hinxton, Cambridge, UK.

Multiple genes and environmental factors are believed to contribute to the aetiology of multiple sclerosis (MS), a chronic inflammatory disease of the central nervous system. In this study microarray technology, which makes it possible to simultaneously study the expression of thousands of genes in a single experiment, has been applied to identify genes underlying susceptibility to MS in the genetically homogeneous Afrikaner population of South Africa. RNA was extracted from cultured human fibroblasts, and fluorescently labelled single stranded cDNA was generated using either reverse transcriptase incorporation of Cy-3 or Cy-5 dCTP into first strand cDNA, or Klenow labelling following a cDNA amplification step. The fluorescently labelled cDNAs of the patient and control samples were hybridised to the immobilised DNAs on the array using an automated slide processor. Following hybridisation, the DNA microarrays were

scanned to monitor the fluorescence of each target that was successfully hybridised to the immobilised probe. Data were interpreted using the Applied Biosystems Quantarray package, and further data analysis was performed in Genespring (Silicon Genetics) and Microsoft Excel. The resulting gene expression ratios indicated significant under- and over-expression of various genes involved in regulating macrophage function in the MS compared with the control samples. Over-expression of monocyte chemotactic protein-3 (MCP-3), known to attract immune cells to the site of inflammation, represented the most significant finding (p<0.002). The results were confirmed with quantitative methods, including real-time polymerase chain reaction (PCR) and northern blot analysis. Subsequent analysis of the CA/GA microsatellite polymorphism in the promoter-enhancer region of the *MCP-3* gene revealed a significant difference in allelic distribution between 117 South African MS patients and 73 population-matched controls ($p<0.05$, 3df, $\chi^2=9.14$). Sequencing of the coding region of three MS patients showing over-expression of MCP-3 did not reveal any sequence changes, which is in accordance with involvement of the promoter region in MCP-3 over-expression. Since the data presented in this study confirm previous findings of a significant association between MS and the *MCP-3* gene, a comprehensive genetic test based on these and related findings will be developed for diagnostic purposes.

210

ABSTRACT 4

HFE, *DCYTB*, *IREG1* and *HAMP* mutations in primary iron overload patients in the South African population

Zaahl MG[1,2], Merryweather-Clarke AT[3], Pointon JJ[3], van der Merwe S[4], Warnich L[2], Kotze MJ[5], Robson KJH[3]

[1]Division of Human Genetics, Faculty of Health Sciences, University of Stellenbosch, Tygerberg, [2]Department of Genetics, University of Stellenbosch, Stellenbosch, SOUTH AFRICA, [3]MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Oxford, UNITED KINGDOM, [4]Department of Internal Medicine, University of Pretoria, Pretoria and [5]Genecare Molecular Genetics (Pty) Ltd, Christiaan Barnard Memorial Hospital, Cape Town, SOUTH AFRICA

Background: Our increasing understanding of the molecular basis of genetic disorders of iron metabolism in mammalian species is providing new insight into genes involved in mediating intestinal iron absorption. It has become clear that the control of iron metabolism is interconnected with other cellular pathways and further analysis of these aspects are likely to provide a better understanding of pathological processes in which iron may participate.

Aim: To identify the possible involvement of genes involved in primary iron overload including *DCYTB*, *HAMP*, *IREG1* and *HFE* in patients referred for a diagnosis of hemochromatosis who tested negative for the most frequent *HFE* mutation C282Y.

Materials and methods: The study population consists of 72 patients with primary iron overload and 50 population-matched control individuals. Polymerase chain reaction (PCR) amplification of genomic DNA was performed on the coding and promoter regions of *DCYTB*, *HAMP*, *IREG1* and *HFE* using intronic primers. The PCR products were

subjected to denaturing high performance liquid chromatography (dHPLC) analysis and those with variation in wave patterns were subjected to sequencing analysis. Statistically significant differences were determined by the Fisher exact test and/or chi-squared (χ^2) analysis with Yates' correction. A P value smaller than 0.05 was regarded as statistically significant.

Results and Discussion: Mutation analysis of the various genes revealed several variations with 6 novel mutations in *IREG1*, 2 in *HAMP*, 2 in *DCYTB* and 2 in *HFE* as well as novel polymorphisms in all the genes investigated. Several previously described polymorphisms were also detected, including 7 *HFE* and 4 *IREG1* polymorphisms. Statistically significant differences were observed when comparing the patient and population-matched control groups for a polymorphism identified in the promoter region of *IREG1* (p<0.003, χ^2 = 8.51, 1df).

Conclusion: The identification of several mutations in the investigated genes in 17% of the patient group indicates the involvement of these genes in iron homeostasis.

ABSTRACT 5

Heterozygosity for novel hepcidin mutations may modify the phenotype of *HFE* C282Y heterozygotes

Alison T Merryweather-Clarke[1], Monique G Zaahl[2,3], Estelle Cadet[4], Vip Viprakasit[1,5], Schalk van der Merwe[6], Dominique Capron[7], Adrian Bomford[8], Jennifer J Pointon[1], Victoria LC Wimhurst[1], Louise Warnich[3], Voravarn Tanphaichitr[5], Maritha J Kotze[9], Jacques Rochette[4], <u>Kathryn JH Robson</u>[1].

[1]MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, Headley Way, Oxford OX3 9DS, UK; [2]Division of Human Genetics, Faculty of Health Sciences, University of Stellenbosch, Tygerberg, South Africa (SA); [3]Department of Genetics, University of Stellenbosch, Stellenbosch, SA; [4]Génétique Médicale, Faculté de Médicine, Université Jules Verne de Picardie, Amiens, France; [5]Department of Paediatrics, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand; [6]Department of Internal Medicine, University of Pretoria, SA; [7]Hépato-gastro-entérologie, Faculté de Médicine, Université Jules Verne de Picardie, Amiens, France; [8]Institute of Liver Studies, Kings College Hospital, London, UK; [9]Genecare Molecular Genetics, Christiaan Barnard Memorial Hospital, Cape Town, SA.

Hepcidin antimicrobial peptide (HAMP) is secreted by the liver in response to increased iron stores, resulting in decreased absorption of dietary iron. Juvenile haemochromatosis patients have been described who are homozygous for mutations in the hepcidin gene (Roetto et al, 2003). We report here a spectrum of hepcidin mutations found in patients with varying severity of iron loading, some of whom are heterozygous for the *HFE* C282Y mutation, and suggest that the phenotype of *HFE* heterozygotes may be modulated by *HAMP* mutations.

The first patient, BS, is a 65 year old British male who had a triple angioplasty at age 55. He had a serum ferritin (sF) of 1000µg/l, and is heterozygous for *HFE* C282Y. DHPLC analysis of *HAMP* exons 2 & 3 revealed a novel heteroduplex pattern. Sequencing revealed that BS is heterozygous for a missense mutation in exon 3 of *HAMP*. BS also has two children who are C282Y homozygotes, both undergoing venesection, a 35 year

old son whose serum ferritin on diagnosis was 1000µg/l and who lacks the *HAMP* mutation, and a 39 year old daughter with rheumatoid arthritis who is heterozygous for the *HAMP* mutation.

The second patient is from Picardy, MP, aged 46 years, who is also a *HFE* C282Y heterozygote. At age 27 he was found to have congestive heart failure, with general fatigue and dyspnoea as the main symptoms. He also presented with hypogonadotrophic hypogonadism, type I diabetes, skin pigmentation and hepatomegaly. Transferrin saturation (TSat) was 92%, sF 1645µg/l, serum glucose 10µM. Fibrosis was detected on liver biopsy, when hepatic iron concentration was 481 µM/g dry liver and HII=17.8. Removal of 27g of iron over 4 years restored his TSat to 30% and sF to 40µg/l. A total of about 80g iron has been removed over during 19 years of inconsistent attendance for phlebotomy. Unless at least 0.5g of iron is removed per month, his TSat goes up to >80% in about 3 months. DHPLC and sequencing analysis of *HAMP* revealed heterozygosity for a deletion in exon 2 which also involves a splice site. The resulting transcript is likely to be extended to 179 codons due to a frameshift, completely disrupting the active peptide. MP's mother is aged 86 years with no major clinical problems. She lacks *HFE* C282Y but is heterozygous for *HFE* H63D, and also heterozygous for the *HAMP* mutation. Her transferrin saturation is 35%, ferritin 52µg/l, normal transaminases, serum glucose 6µM, red cells 4.25×10^{12}/l, Hb13g/dl, MCV 82fl. Two further pairs of related patients with iron overload from South Africa, three of whom are *HFE* C282Y and H63D compound heterozygotes and the other a *HFE* H63D

heterozygote, were found to be heterozygotes for two different pairs of *HAMP* mutations and polymorphisms following DHPLC and sequencing.

Finally, a young Thai haemoglobin E patient who has never been transfused, with an sF of 461μg/l, was found to be heterozygous for a polymorphism in *HAMP* intron 1.

We propose that the phenotype of C282Y heterozygotes may be modified by heterozygosity for mutations which disrupt the function of hepcidin in iron homeostasis, with the severity of iron overload corresponding to the severity of the *HAMP* mutation.