

Mutation analysis of four genes implicated in iron homeostasis in porphyria cutanea tarda (PCT) patients

Nicola Panton

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science (MSc)
in Genetics at the University of Stellenbosch.



Supervisor: Dr MG Zaahl

Co-supervisors: Prof L Warnich and Prof RJ Hift

March 2008

DECLARATION

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

Signature:

Date:

Summary

The porphyrias are a group of genetic diseases resulting from the accumulation of haem precursors due to defective enzyme activity in either one of the last seven enzymes in the haem biosynthesis pathway. One of the common hepatic porphyrias, porphyria cutanea tarda (PCT), arises from the inhibition of uroporphyrinogen decarboxylase (UROD) activity. It is characterised by excessive urinary and hepatic excretion of uroporphyrinogens and manifests cutaneously in the form of dermatitis. Two main forms of PCT have been described, namely familial PCT (fPCT) and sporadic PCT (sPCT). PCT is a complex disease and a few genetic (including modifier loci) and environmental precipitating factors have been implicated in the aetiology of PCT. An important exacerbating factor, iron overload, is observed in the majority of PCT patients.

The aim of this study was to determine whether DNA sequence variation in the 5' untranslated regulatory region of four genes involved in iron metabolism i.e. *CP*, *CYBRD1*, *HAMP* and *SLC40A1* may in any way be associated with PCT. The study cohort consisted of 74 patients from three diverse South African populations including 15 Black (eight males and seven females), 30 Caucasian (13 male and 17 females) and 29 Coloured (18 males and 11 females) individuals as well as 132 population-matched controls. The promoter region of the selected genes were screened for variation utilising the techniques of polymerase chain reaction (PCR) amplification, heteroduplex single-stranded conformational polymorphism (HEX-SCCP) analysis, restriction fragment length polymorphism (RFLP) analysis and bi-directional semi-automated DNA sequencing.

Twenty three previously described and eleven novel variants were identified. The novel variants comprised *CYBRD1*: -1540G/A, -1474G/A, -1452T/C, -1346T/C, -1272T/C, -645T/C; G(T)₈G(T)_nG(T)_nG(T)₉; *HAMP*: -429G/T and *SLC40A1*: -1461T/C, -1399G/A, -524C/T. Statistically significant associations were observed at a number of loci. *In silico* analysis revealed several putative transcription factor binding sites (TFBSs) spanning the regions of variation. The disruption of an existing (or creation of a novel) TFBS is thought to occur in the presence of a

variant in a number of instances. This may lead to the manipulation of transcription rates, thereby depicting a possible mechanism for gene dysregulation.

The study presented here was undertaken as a preliminary investigation to determine the contribution (if any) of variants in the regulatory regions of candidate genes in iron metabolism in South African PCT patients. Considering the increasing incidence of PCT, in particular the Black South African population, it is necessary to elucidate the underlying mechanisms of iron overload in PCT patients. The propitious findings signified in the study, in conjunction with phenotype-genotype correlations, will assist in clarifying the association between iron overload and PCT.

Opsomming

Die porfirieë is 'n groep genetiese siektes wat ontstaan as gevolg van defektiewe ensiemaktiwiteit in een van die finale sewe ensieme in die heembiosintese padweg, wat dan 'n opeenhoping van heemvoorlopers veroorsaak. Een van die algemene hepatiese porfirieë is porfirie cutanea tarda (PCT), wat ontstaan as gevolg van die inhibering van uroporfirinogeen dekarboksilase (UROD) aktiwiteit. PCT is gekenmerk deur oormatige urinêre en hepatiese uitskeiding van uroporfirinogene en manifesteer kutaan as dermatitis. Twee hoof vorme van PCT is beskryf, naamlik oorerflike PCT(oPCT) en sporadiese PCT (sPCT). PCT is 'n komplekse siekte en die ontwikkeling van PCT is al toegeskryf aan beide genetiese en omgewingsfaktore. 'n Belangrike verslegtende faktor, ysteroorlading, is waargeneem in die meerderheid van PCT pasiënte.

Die doel van die huidige studie was om te bepaal of DNS volgorde variasie in die 5' onvertaalde regulatoriese area van vier gene betrokke by ystermetabolisme (*CP*, *CYBRDI*, *HAMP* en *SLC40A1*) moontlik met PCT assosieer kan word. Die studiegroep het bestaan uit 74 pasiënte vanuit drie verskeie Suid-Afrikaanse populasies, insluitend 15 Swart (8 manlike en 7 vroulike), 30 Kaukasiër (13 manlike en 17 vroulike) en 29 Kleurling (18 manlike en 11 vroulike) individue, asook 132 onverwante populasie-gepaste kontroles.

Drie-en-twintig reeds-beskryfde en elf nuwe variante is geïdentifiseer. Die nuwe variante sluit in ***CYBRDI***: -1540G/A, -1477G/A, -1452T/C -1346T/C, -645T/C; G(T)₈G(T)_nG(T)_nG(T)₉; ***HAMP***: -429 G/T; ***SLC40A1***: -1461 T/C, -1399G/A, -750G/A, -524C/T. Statisties betekenisvolle assosiasie is waargeneem by 'n aantal loci. *In silico* analise het potensiële transkripsiefaktor bindingsetels (TFBSs) wat strek oor die areas van variasie, uitgelig. Die onderbreking van 'n bestaande (of ontwikkeling van 'n nuwe) TFBS vind vermoedelik plaas in die aanwesigheid van 'n variant in 'n aantal gevalle. Dit kan moontlik lei tot die manipulasie van transkripsie tempo's, wat 'n moontlike meganisme vir geen disregulasie uitbeeld.

Die studie hier voorgelou is onderneem as 'n voorlopige ondersoek, om te bepaal of variante in die regulatoriese areas van kandidaat gene in ystermetabolisme 'n bydrae (indien enige) lewer tot PCT in Suid-Afrikaanse pasiënte. Gesien in die lig van verhoogde voorkoms van PCT, veral in die Swart Suid-Afrikaanse populasie, is dit nodig om die onderliggende meganisme van ysteroorlading in PCT pasiënte te verduidelik. Die belowende bevindings van hierdie studie, saam met fenotipe-genotipe korrelasie, sal help om die verhouding tussen ysteroorlading en PCT te verklaar.

Dedicated to the late Frances Agnes Panton

Each of us has a personal calling that is as unique as a fingerprint. The best way to succeed is to discover what you love and then find a way to offer it to others and also allowing the energy of the universe to lead you
-Oprah Winfrey-

ACKNOWLEDGEMENTS

I would like to express appreciation towards the following individuals and institutions:

The National Research Foundation (NRF) (Thuthuka), Harry Crossley Foundation, Ernst and Ethel Erikson Trust and the University of Stellenbosch for funding this study.

The University of Stellenbosch and the Department of Genetics for providing the infrastructure to complete this study.

My supervisor, Dr MG Zaahl, for creating the opportunity for me to conduct the research in her laboratory and for proofreading this thesis.

My co-supervisor, Prof LW Warnich, for proofreading this thesis and for challenging discussions.

My co-supervisor, Prof RJ Hift, for providing the clinical samples, for proofreading this thesis, and for interesting discussions.

Mrs E Dietzsch for proofreading this thesis, her advice and continual encouragement.

Mr W Botes for his guidance on statistical analysis.

Ms V Human, my lab colleague, for her guidance on trouble shooting.

Dr M Venter for his help with bioinformatic analysis.

My fellow colleagues in lab 242, for providing a cheerful and supportive working environment.

My friends, especially Nicóla and Katie and my only brother, Mark, for their emotional support and encouragement.

My furry friends, Zorro, Mitz, Gingy and Zelda for keeping me company whilst writing my thesis.

My mother and father for providing financial assistance, unwavering support, persistent encouragement and for showing a keen interest in my passion.

TABLE OF CONTENTS

List of Abbreviations and Symbols	i
List of Figures	x
List of Tables	xii

CHAPTER 1: LITERATURE REVIEW

1.1 INTRODUCTION	1
1.2 PORPHYRIAS	3
1.2.1 PORPHYRIA CUTANEA TARDA (PCT)	5
1. SPORADIC PCT (sPCT)	6
2. FAMILIAL PCT (fPCT)	7
3. HEPATOERYTHROPOIETIC PORPHYRIA (HEP)	7
1.2.5 SYMPTOMS AND DIAGNOSIS	7
1.2.6 PRECIPITATING FACTORS ASSOCIATED WITH PCT	10
1.2.6.1 Environmental factors	10
i) Alcohol misuse	10
ii) Toxic chemicals	10
iii) Oestrogen therapy	11
iv) Hepatitis C virus (HCV)	11
v) Human immunodeficiency virus (HIV)	11
vi) Liver disease	12
vii) Systemic lupus erythematosus (SLE) and Lymphoma	12
viii) Smoking	12
iv) Low vitamin C and Caretonoid status	13
1.2.6.2 Genetic factors	13
i) Uroporphyrinogen decarboxylase (<i>UROD</i>) mutations	13
ii) High iron (<i>HFE</i>) mutations	14
1.2.7 TREATMENT AND CLINICAL MANAGEMENT	14
1.3 IRON METABOLISM	15
1.3.1 IRON ABSORPTION	17
1.3.1.1 Dietary iron uptake	19
i) Haem iron uptake	19
ii) Non-haem iron uptake	20
iii) Paraferitin-mediated iron uptake	20

1.3.1.2 Iron storage and translocation in enterocytic cells	21
1.3.1.3 Iron export	21
1.3.2 TRANSFERRIN RECEPTOR MEDIATED UPTAKE	22
1.3.3 IRON STORAGE	25
1.3.4 IRON RECYCLING	25
1.3.5 IRON EXCRETION	26
1.3.6 IRON HOMEOSTASIS	27
1.3.6.1 Iron regulation at the cellular level	27
1.3.6.2 Iron regulation at the systemic level	28
1.4 IRON AND PCT	32
1.5 GENES INVOLVED IN IRON HOMEOSTASIS	34
1.5.1 CERULOPLASMIN (<i>CP</i>)	35
1.5.2 CYTOCHROME B REDUCTASE 1 (<i>CYBRD1</i>)	36
1.5.3 HEPCIDIN ANTIMICROBIAL PEPTIDE (<i>HAMP</i>)	37
1.5.4 SOLUTE CARRIER FAMILY 40 MEMBER1 (<i>SCL40A1</i>)	38
1.6 REGULATION OF GENE EXPRESSION	39
1.6.1 TRANSCRIPTION FACTORS	40
1.6.1.1 Transcription factors and iron metabolism	41
1.7 GENETIC VARIATION	43
1.7.1 GENETIC VARIATION IN THE PROMOTER REGION	44
1.7.1.1 Regulatory Single nucleotide polymorphisms (RSNPs)	44
1.8 HYPOTHESIS AND OBJECTIVES	44
CHAPTER 2: MATERIALS AND METHODS	
2.1 STUDY COHORT	46
2.2 DNA ISOLATION AND PURIFICATION	47
2.2.1 TOTAL GENOMIC DNA (gDNA) ISOLATION FROM WHOLE BLOOD	47
2.2.2 gDNA PURIFICATION	48
2.3 POLYMERASE CHAIN REACTION (PCR) AMPLIFICATION	49
2.3.1 OLIGONUCLEOTIDE PRIMERS	49
2.3.2 PCR REACTION CONDITIONS	54
2.4 AGAROSE GEL ELECTROPHORESIS	55
2.5 HETERODUPLEX SINGLE-STRANDED CONFORMATION POLYMORPHISM (HEX-SSCP) ANALYSIS	55
2.6 SEMI-AUTOMATED DNA SEQUENCING	56
2.6.1 PURIFICATION OF PCR PRODUCTS	56

2.6.2 CYCLE SEQUENCING	57
2.7 RESTRICTION LENGTH POLYMORPHISM (RFLP) ANALYSIS	57
2.8 STATISTICAL ANALYSIS	59
2.9 BIOINFORMATIC ANALYSIS	60
CHAPTER 3: RESULTS AND DISCUSSION	61
MUTATION ANALYSIS OF THE PROMOTER REGION OF FOUR GENES IMPLICATED IN IRON HOMEOSTASIS (<i>CP</i> , <i>CYBRD1</i> , <i>HAMP</i> AND <i>SLC40A1</i>) IN PORPHYRIA CUTANEA TARDA (PCT) PATIENTS.	
CHAPTER 4: CONCLUSIONS AND FUTURE PROSPECTS	109
CHAPTER 5: REFERENCES	116
5.1 GENERAL REFERENCES	116
5.2 ELECTRONIC REFERENCES	147
APPENDICES	
APPENDIX 1: LIST OF REAGENTS AND CHEMICALS USED IN THIS STUDY	148
APPENDIX 2: PROMOTER SEQUENCES OF <i>CP</i>, <i>CYBRD1</i>, <i>HAMP</i> and <i>SLC40A1</i> INDICATING THE POSITION OF DESIGNED PRIMER SETS AND VARIANTS IDENTIFIED IN THE CURRENT STUDY.	150
APPENDIX 3: HEX-SCCP GEL IMAGES AND SEQUENCING CHROMATOGRAMS OF VARIANTS IDENTIFIED IN THIS STUDY	155
APPENDIX 4: ABSRACT OF ORAL PRESENTATION AT 11th SOUTH AFRICAN SOCIETY OF HUMAN GENETICS (SASHG) CONGRESS, GOLDEN GATE, SOUTH AFRICA, 1-3 MARCH 2007	165

LIST OF SYMBOLS AND ABBREVIATIONS

5'	5-prime end
3'	3-prime end
χ^2	Chi-square
%	Percentage
=	Equal to
>	Greater than
<	Less than
\pm	Plus-minus
α	Alpha
β	Beta
ρ	Phi
δ	Delta
μ	Micro(10^{-6})
©	Copyright
°C	Degrees celcius
μ l	Microlitre
μ l/ml	Microlitre per millilitre
μ g/ml	Microgram per millilitre
μ M	Micromolar
®	Registered trademark
™	Trademark
A	
A	Adenine
AA	Acrylamide
AFLP	Amplified fragment length polymorphism
AHR	Aryl-hydrocarbon receptor
AgNO ₃	Silver nitrate
ALA	Delta-aminolevulinic acid
ALAS	Delta-aminolevulinic acid synthase
ALP	Alkaline phosphatase
ALS	Aspartate aminotransferase

ALT	Alanine aminotransferase
AP1	Activator protein 1
APOE	Apolipoprotein E
APS	Ammonium persulphate
ASO	Allele specific oligonucleotide
ATF	Activating transcription factor
ATP	Adenosine 5'-triphosphate
B	
BA	Boric Acid
BAA	N, N' methylenebisacrylamide
<i>BfaI</i>	<i>Bacteroides fragilis</i> , 1 st enzyme
bHLH-ZIP	Basic helix-loop-helix leucine zipper
BLAST	Basic local alignment search tool
bp	Base pair
<i>BsrDI</i>	<i>Bacillus stearothermophilus</i> D70, 1 st enzyme
BSA	Bovine serum albumin
<i>BstUI</i>	<i>Bacillus stearothermophilus</i> UI, 1 st enzyme
C	
C	Cytosine
C282Y	Cysteine282Tyrosine
C/EBP	CCAAT/ enhancer binding protein
C/EBP α	CCAAT/ enhancer binding protein alpha
CBF β	Core-binding factor beta subunit
C ₁₉ H ₁₀ Br ₄ O ₅ S	Bromophenol Blue
CH ₃ (CH ₂) ₁₁ OSO ₃ Na	Sodium dodecyl sulphate
C ₃₁ H ₂₈ N ₂ Na ₄ O ₁₃ S	Xylene cyanol
(CH ₂ OH) ₃ CNH ₂ -Cl	Tris(hydroxymethyl)aminomethane
C ₂ H ₁₈ O ₅ S	Cresol Red
CH ₃ NO	Formamide
C ₇ H ₁₀ O ₂ N ₂	N, N' methylenebisacrylamide
C ₁₀ H ₁₆ N ₂ O ₈	Ethylene diamine tetra-acetic acid
cm	Centimetre

c-Max	Myc associated factor X
c-Myc	Myelocytomatosis
<i>CP</i>	Ceruloplasmin gene
CP	Ceruloplasmin protein
CREB	cAMP response element-binding protein
Cu ²⁺	Copper cupric
<i>CYBRD1</i>	Cytochrome b reductase 1 gene
CYBRD1	Cytochrome b reductase 1 protein
D	
D'	D prime value
dATP	2'- deoxy-adenosine-5'- triphosphate
<i>DCT1</i>	Divalent cation transporter 1 gene
dCTP	2'- deoxy-cytidine-5'- triphosphate
<i>DCYTB</i>	Duodenal cytochrome B gene
dH ₂ O	Distilled water
ddH ₂ O	Double distilled water
del	Deletion
dGTP	2'- deoxy-guanosine-5'- triphosphate
DHPLC	Denaturing high performance liquid chromatography
<i>DMT1</i>	Divalent metal transporter 1 gene
DNA	Deoxyribonucleic acid
dNTP	2'- deoxy-nucleotide-5'- triphosphate
dTTP	2'- deoxy-thymidine-5'- triphosphate
E	
EDTA	Ethylene diamine tetra-acetic acid
ER	Estrogen receptor
ERE	Estrogen responsive element
ESRD	End-stage kidney disease
<i>et al.</i>	<i>Et alia</i> (and others)
EtBr	Ethidium Bromide
EtOH	Ethanol
<i>ETS</i>	<i>ETS</i> oncogene

F

F	Forward primer
Fe	Iron
Fe ²⁺	Ferrous iron
Fe ³⁺	Ferric iron
<i>FECH</i>	Ferrochelatase gene
FISH	Fluorescence <i>in situ</i> hybridisation
FOS	v-FOS FBJ murine osteosarcoma viral oncogene homolog
FOXA1	Forkhead box A1
fPCT	Familial porphyria cutanea tarda
<i>FPN1</i>	Ferroportin 1 gene
Ft	Ferritin
<i>FTL</i>	Ferritin light chain gene
<i>FTH</i>	Ferritin heavy chain gene

G

g	Gram
G	Guanine
<i>GAA</i>	Acid alpha-glucosidase gene
GATA	GATA binding protein
GATA 1	GATA binding protein 1
GATA 2	GATA binding protein 2
GATA 3	GATA binding protein 3
GC	Guanine Cytosine
Gfi1	Growth factor independent 1
GIT	Gastrointestinal tract
GTF	General transcription factor

H

H63D	Histidine63Aspartic acid
<i>HAMP</i>	Hepcidin antimicrobial peptide gene
HAMP	Hepcidin antimicrobial peptide protein
H ₃ BO ₃	Boric Acid
HCB	Hexachlorobenzene

HCHO	Formaldehyde
HCV	Hepatitis C virus
HEP	Hepatoerythropoietic porphyria
<i>HEPC</i>	Hepcidin gene
HEPH	Hephaestin
HEX-SSCP	Heteroduplex single-stranded conformation polymorphism
<i>HFE</i>	High iron gene
HFE	High iron protein
HGMD	Human genome mutation database
HH	Hereditary Haemochromatosis
HIV	Human immunodeficiency virus
<i>HJV</i>	Haemojuvelin gene
HJV	Haemojuvelin protein
HLA	Human leukocyte antigen
HLF	Hepatic leukemia factor
<i>HMOX1</i>	Haem oxygenase 1 gene
HMOX1	Haem oxygenase 1 protein
HNF	Hepatic nuclear factor
HNF1 α	Hepatic nuclear factor 1 alpha
HNF3 β	Hepatic nuclear factor 3 beta
HNF4 α	Hepatic nuclear factor 4 alpha
H ₂ O ₂	Hydrogen peroxide
<i>hpx</i> mice	Hypotransferrinaemic mice
Hr(s)	Hour(s)
HWE	Hardy-Weinberg equilibrium

I

IDT	Intergrated DNA Technologies
<i>i.e</i>	<i>Id est</i> (that is)
IL-10	Interleukin-10
Inr	Initiator
IRE(s)	Iron response element(s)
IRP1	Iron regulatory protein 1
IRP2	Iron regulatory protein 2

J

JUN v-jun avian sarcoma virus 17 oncogene homolog

K

kb Kilobase

KCl Potassium chloride

kDa Kilo Dalton

KHCO₃ Potassium hydrogen carbonate

KH₂PO₄ Potassium di-hydrogen orthophosphate

L

l Litre

LD Linkage disequilibrium

LEAP1 Liver-expressed antimicrobial peptide 1 gene

LOD Logarithm of the Odds

LTF Lactotransferrin (Lactoferrin)

M

m Milli (10⁻³)

M Molar

MAF v-MAF avian musculoaponeurotic fibrosarcoma oncogene homolog

mg Milligram

mg/ml Milligram per millilitre

MgCl₂ Magnesium chloride

min Minutes

ml Millilitre

mM Millimole

Mn²⁺ Manganese

MPO Myeloperoxidase gene

mRNA Messenger ribonucleic acid

MTP1 Metal transporter 1 gene

N

n Nano (10⁻⁹)

n	Number of
N	Adenine/Cytosine/Guanine/Thymine
Na ₂ HPO ₄	Di-sodium hydrogen orthophosphate anhydrous
NaCl	Sodium chloride
NaClO ₄	Sodium acetate
NADPH	Nicotinamide adenosine dinucleotide phosphate
NaOH	Sodium hydroxide
NF-E2	Nuclear factor erythroid 2
NFκB	Nuclear factor kappa- B
NF-Y	Nuclear factor-Y
ng	Nanogram
ng/μl	Nanogram per microlitre
NH ₄ Cl	Ammonium chloride
(NH ₂) ₂ CO	Urea
(NH ₄) ₂ S ₂ O ₈	Ammonium persulphate
Ni ²⁺	Nickel
NOD2	Nucleotide-binding oligomerization domain protein 2
NRAMP1	Natural resistance-associated macrophage protein 1
NRAMP2	Natural resistance-associated macrophage protein 2
NTBI	Non-transferrin-bound iron
O	
O ²⁻	Superoxide
OH	Hydroxyl
OMIM TM	Online mendelian inheritance in man TM
P	
p	Short arm of chromosome
<i>P</i>	Probability value
PAA	Polyacrylamide
Pb ²⁺	Lead plumbous
PBG	Porphobilinogen
<i>PBGD</i>	Porphobilinogen deaminase gene
PBS	Phosphate buffered saline

PCR	Polymerase chain reaction
PCT	Porphyria cutanea tarda
<i>per se</i>	By itself
pg	Page
pH	Potential of hydrogen
PIC	Preinitiation complex
pmol	Pico mole
<i>PTPN22</i>	Protein tyrosine phosphatase nonreceptor type-22 gene
PU.1	Spleen focus forming virus proviral integrating oncogene 1
PWM	Positional weight matrix
Q	
q	Long arm of chromosome
R	
r^2	Correlation coefficient
R	Reverse primer
RACE	Rapid amplification of cDNA ends
RES	Reticuloendothelial system
RET	Rearranged during transfection protooncogene
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
ROS	Reactive oxygen species
rpm	Revolutions per minute
<i>RsaI</i>	<i>Rhodopseudomonas sphaeroides</i> I, 1 st enzyme
rSNP	Regulatory single nucleotide polymorphism
RUNX1	Runt-related transcription factor 1
S	
SA	South African
SSCP	Single-stranded conformational polymorphism
SDS	Sodium dodecyl sulphate
sec	Seconds
<i>SLC40A1</i>	Solute carrier family 40 (Iron regulated transporter) member A1 gene

SLC40A1	Solute carrier family 40 (Iron regulated transporter) member A1 protein
SLE	Systemic lupus erythematosus
SNP	Single nucleotide polymorphism
SP1	Specificity protein 1
sPCT	Sporadic porphyria cutanea tarda
SS African	Sub-Saharan African
STAT1	Signal transducer and activator of transcription 1

T

T	Thymine
T _{A1}	Annealing temperature 1
T _{A2}	Annealing temperature 2
<i>Taq</i>	<i>Thermus aquaticus</i> DNA polymerase
TBE	Tris borate-ethylene diamine tetra-acetic acid buffer
TCF1	Transcription factor 1
TCF14	Transcription factor 14
TEMED	N, N, N', N' Tetramethylethylenediamine
Tf	Transferrin
TF(s)	Transcription factor(s)
TFIIA	Transcription factor IIA
TFIIB	Transcription factor IIB
TFIID	Transcription factor IID
TFIIE	Transcription factor IIE
TFIIF	Transcription factor IIF
TFIIH	Transcription factor IIH
TFBS(s)	Transcription factor binding site(s)
TFPGA	Tools for population genetic association studies
<i>TFR1</i>	Transferrin receptor 1 gene
<i>TFR2</i>	Transferrin receptor 2 gene
T _m	Melting temperature
Tris-HCl	Tris(hydroxymethyl)aminomethane

U

U	Units
---	-------

<i>UROD</i>	Uroporphyrinogen decarboxylase gene
UROD	Uroporphyrinogen decarboxylase protein
USA	United States of America
USF1	Upstream stimulatory factor 1
USF2	Upstream stimulatory factor 2
UTR	Untranslated region
UV	Ultraviolet

V

v	Version
V	Volts
VP	Variegate porphyria
v/v	Volume per volume
<i>vice versa</i>	The other way round

W

w/v	Weight per volume
-----	-------------------

Y

YY1	Ying Yang-1
-----	-------------

Z

Zn ²⁺	Zinc
------------------	------

LIST OF FIGURES**CHAPTER 1: LITERATURE REVIEW**

Figure 1.1	An illustration of the haem biosynthesis pathway and the classification of the human porphyrias	4
Figure 1.2	Images showing cutaneous symptoms of PCT	8
Figure 1.3	A schematic illustration of iron status in healthy individuals	16
Figure 1.4	A schematic illustration of iron distribution and turnover in the human body	16
Figure 1.5	The Fenton Reaction	17
Figure 1.6	A schematic representation of iron absorption in the duodenal enterocyte	18
Figure 1.7	A schematic representation of iron uptake and storage	24
Figure 1.8	A schematic representation of iron recycling	27

LIST OF TABLES

CHAPTER 2: MATERIALS AND METHODS

Table 2.1	Oligonucleotide primer sequences, amplicon sizes and PCR conditions for the amplification of the promoter region of the <i>CP</i> , <i>CYBRD1</i> , <i>HAMP</i> and <i>SLC40A1</i> genes	50
Table 2.2	Four RFLPs detected in the promoter region of the <i>CYBRD1</i> , <i>HAMP</i> and <i>SLC40A1</i> genes	58

CHAPTER 3: RESULTS AND DISCUSSION

Table 3.1	Genotype and polymorphic allele frequencies of the <i>CP</i> , <i>CYBRD1</i> , <i>HAMP</i> and <i>SLC40A1</i> variants identified in patients and control individuals in South African Blacks, Caucasians and Coloureds	66
Table 3.2	The probability (<i>P</i>) values at variant loci in the <i>CP</i> , <i>CYBRD1</i> , <i>HAMP</i> and <i>SLC40A1</i> genes tested for departure from HWE	72
Table 3.3	Allele frequency probability values (<i>P</i>) of variants demonstrating gender specificity in the current study	75
Table 3.4	Comparison of variant (polymorphic) allele frequencies	76
Table 3.5	Predicted TFBSs in the promoter region of the <i>CP</i> , <i>CYBRD1</i> , <i>HAMP</i> , and <i>SLC40A1</i> genes, obtained from data generated by <i>in silico</i> analysis	79

CHAPTER 4: CONCLUSIONS AND FUTURE PROSPECTS

Table 4.1	Variants identified in the current study that warrant further investigation	114
------------------	---	-----

CHAPTER 1: LITERATURE REVIEW



1.1 INTRODUCTION

The concept of a gene as a unit of heredity was first proposed by Johann 'Gregor' Mendel in the 1850's (Weiling 1991). This remarkable discovery provided a crucial foundation for insight into the genetic basis of disease. The “one gene : one disease” hypothesis appropriately explains monogenic diseases. Today, it is acknowledged that the majority of genetic diseases affecting man are not monogenic, but rather polygenic or complex in nature. Complex diseases are caused by the interaction of numerous genes or modifier loci, as well as environmental factors, including diet, hormones, chemical exposure and rearing experiences (Rutter 2007). Diabetes, hypertension, Alzheimer's disease, Hirschsprung's disease, obesity and breast cancer are examples of such disorders (Strachan and Read 1999). Gene variants giving rise to these types of diseases are often subtle and present in the general population. Furthermore, gene variants (or common polymorphisms) may be inherited as so-called haplotypes, which comprise variable combinations of specific closely linked alleles (Brookes 1999). Thus it must be borne in mind that an individual's genetic complement may not necessarily cause disease but rather contributes to increased genomic susceptibility for the disease. In his book “Nature via Nurture”, the author Matt Ridley aptly suggests that “genes are the mechanisms of experience”, i.e. genes (nature) respond to environmental stimuli (nurture) (Ridley 2004).

Since complex genetic disorders are controlled by more than one susceptibility gene, these diseases often display varying phenotypes. The disease porphyria cutanea tarda (PCT), a common hepatic porphyria, is complex in nature. It displays phenotypes of different extremities and does not manifest clinically in the presence of the formerly presumed single causative uroporphyrinogen decarboxylase (*UROD*) gene mutation (Sampietro *et al.* 1999). Several environmental precipitating factors (especially iron overload) have been implicated in the development of this disease and various susceptibility loci [e.g. high iron (*HFE*) gene (Bulaj *et al.* 2000), non aryl-hydrocarbon

receptor (non-*AHR*) (Robinson *et al.* 2002) and cytochrome *CYP1A2* (Christiansen *et al.* 2000)] have been investigated. Two common *HFE* gene mutations associated with iron overload have an increased incidence in PCT patients (Jackson *et al.* 1997; Roberts *et al.* 1997, Mendez *et al.* 1998, Christiansen *et al.* 1999). Presence of the C282Y and H63D loci were demonstrated to be associated with the increased risk of disease (Roberts *et al.* 1997, Sampietro *et al.* 1998). Collectively, these findings indicate that DNA mutations at several genomic sites may influence phenotype (Andrew 1999).

The medical implications of the interplay between genetics and the environment and the clarification of the genotype-phenotype correlation are still in their infancy. The identification of susceptibility loci implicated in complex genetic diseases, will aid molecular biologists in resolving the molecular aetiology of these diseases enabling the development of custom treatment options, informed awareness campaigns and preventative measures.

1.2 PORPHYRIAS

The porphyrias are a group of disorders resulting in the overproduction of haem precursors due to defective enzyme activity in the haem biosynthesis pathway (Elder *et al.* 1972; Brodie *et al.* 1977). This metabolic pathway involves eight enzymes responsible for the sequential conversion of δ -aminolaevulinic acid (ALA) to protoporphyrin and finally haem (Figure 1.1) (Kappas *et al.* 1985). Haem synthesis occurs in all nucleated cells, especially in the liver and erythropoietic tissues. Synthesis is initiated in the mitochondria before entering the cytosol where three stages occur and finally returning to the mitochondria for the final step of haem formation (Nordmann and Puy 2002). ALA synthase (ALAS) is the rate limiting enzyme in this pathway; a deficiency in the erythroid specific ALAS (ALAS2) causes sideroblastic anaemia and not porphyria. A defect in any of the remaining seven enzymes in the haem biosynthesis pathway will cause a different type of porphyria (Anderson *et al.* 2001). The porphyrias are generally classified according to the particular defective enzyme, main site of defect, and whether or not they result in acute attacks (James and Hift 2000). The site of accumulation of porphyrins and their precursors may occur in the liver (hepatic porphyria), where haem is required for the synthesis of haemoproteins, or in red blood cells (erythropoietic porphyria) where haemoglobin is produced.

All of the hepatic porphyrias, except porphyria cutanea tarda (PCT), are characterised by acute attacks, e.g. severe abdominal pain, neurological problems and neuromuscular weakness (Thadani *et al.* 2000). PCT and the erythropoietic porphyrias have mainly cutaneous manifestations (Murphy 1993). Approximately 1% of acute attacks are fatal (Thadani *et al.* 2000); although this rate seems to be decreasing with improved diagnosis and management of porphyria (www.porphyrria.uct.ac.za). All porphyrias share one common clinical symptom, namely photosensitivity of the skin, due to the photodynamic action of the accumulated porphyrins present in the skin (Holti *et al.* 1958).

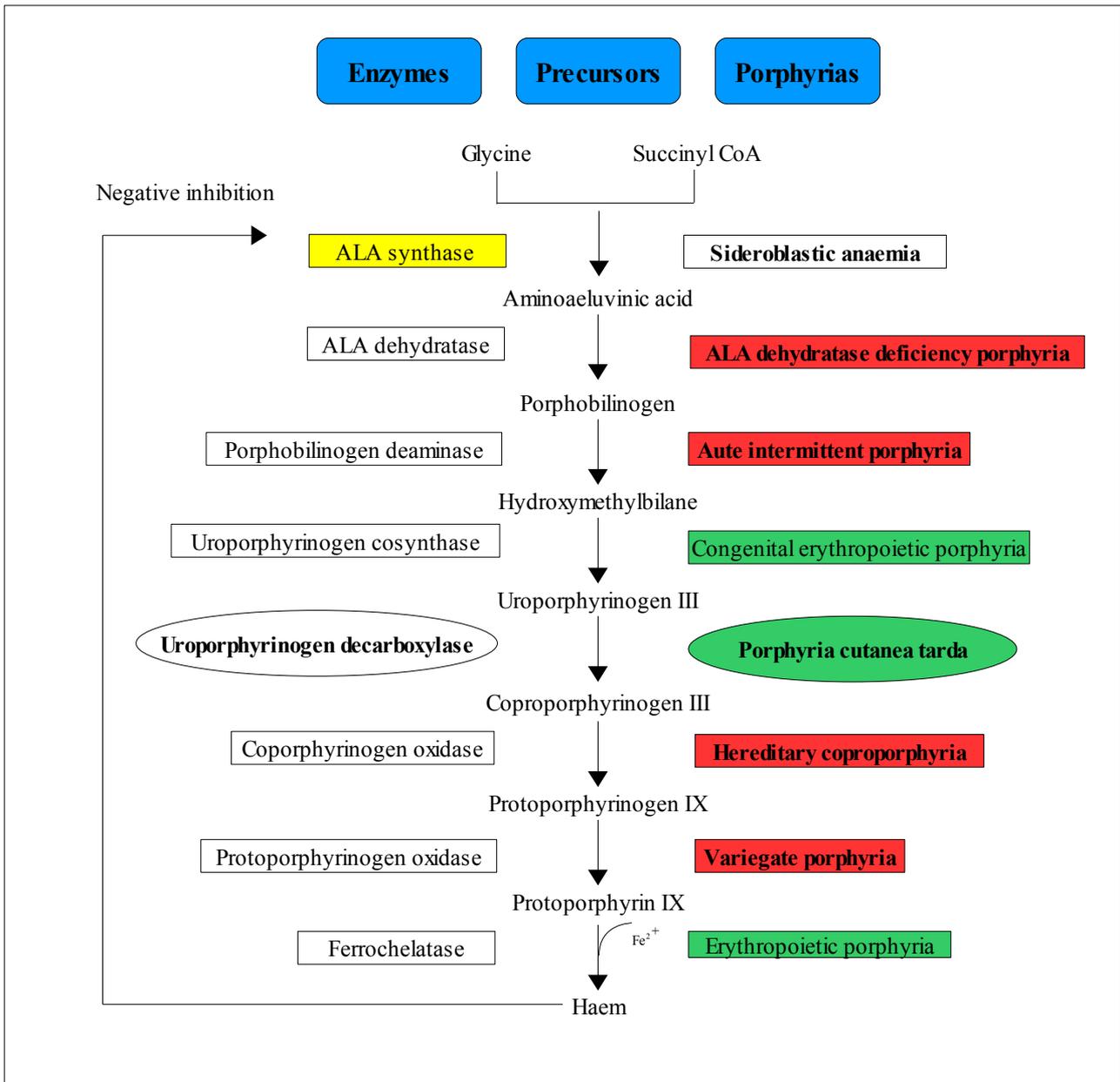


Figure 1.1 An illustration of the haem biosynthesis pathway and the classification of the human porphyrias. A defect in uroporphyrinogen decarboxylase (UROD) enzyme activity causes porphyria cutanea tarda (PCT) as indicated by the ellipse.

blue components of pathway	green non-acute porphyrias	bold text hepatic porphyrias
yellow rate-limiting enzyme	red acute porphyrias	normal text erythropoietic porphyrias

Porphyrins are cyclic structures composed of four pyrrole rings joined by methene bridges (Moss 1987). Porphyrins bind to metals *via* the nitrogen atom in each pyrrole ring to form metalloproteins, which are found in both the animal and plant kingdom (James and Hift 2000). The most important metalloproteins are those that bind to iron to form haem and include haemoglobin, mitochondrial and microsomal cytochromes and myoglobin. Metalloproteins play important roles in electron and oxygen transport, the activation of oxygen and hydrogen peroxidase, hydrogen peroxide degradation, drug metabolism and cell growth (Mauzerall 1998; Tsiftoglou *et al.* 2006). On exposure to light, colourless porphyrin precursors known as porphyrinogens are converted to the red/purple highly fluorescent porphyrins. Since porphyria patients may pass reddish/purple coloured urine (Rich 1999), the disease was aptly named porphyria- a word derived from the Greek word “porphyra” meaning purple.

1.2.1 PORPHYRIA CUTANEA TARDA (PCT)

Waldenstrom (1937) first named this disease when an adult patient presented with dermatitis following exposure to the sun, and with large amounts of uroporphyrin in the urine. Porphyria cutanea tarda (PCT, OMIMTM +176100, 176090) is the most common type of hepatic porphyria, with an estimated prevalence ranging from 1 in 5000 to 1 in 25000 people (Elder 1998; Mendez *et al.* 2005). PCT affects all races, but the disease incidence is highly variable and seems to be population dependent (Mendez *et al.* 2005). PCT has an increased incidence in Europe, North America, the United Kingdom and Argentina (Thadani *et al.* 2000; Mendez *et al.* 2005). In South Africa, PCT has an increased incidence in the Black population (Cripps 1987; Nordmann and Puy 2002; RJ Hift, personal communication). Sporadic cases tend to be more prevalent in men, presumably due to their higher iron stores, whilst familial cases seem to affect males and females equally.

PCT is associated with a deficiency in UROD (EC 4.1.1.37) activity. A reduction in UROD enzyme activity causes an accumulation of uroporphyrinogens and an increased oxidation of uroporphyrinogens to porphyrins (Sampietro *et al.* 1999). This decreased UROD activity *per se* usually does not produce clinical or biochemical symptoms. However, in the presence of additional precipitating factors (discussed in Section 1.2.6), enzyme activity is sufficiently suppressed to cause the onset of PCT. Three forms of PCT have been described, namely sporadic PCT (sPCT; type I), familial PCT (fPCT; type II) and hepatoerythropoietic porphyria (HEP) (de Verneuil *et al.* 1978).

1.2.2 SPORADIC PCT (sPCT)

sPCT is the most frequently observed form of PCT and accounts for between 66% and 90% of cases (Kushner *et al.* 1976; Thadani *et al.* 2000). This type of PCT is associated with a 50% reduction in UROD enzyme activity in the liver only (Elder *et al.* 1978; Felsher *et al.* 1982), but normal UROD protein concentration (Elder *et al.* 1985). Hepatic UROD activity is indirectly inhibited by several environmental and genetic predisposing factors, such as excessive alcohol consumption, oestrogen therapy, selected chemicals [including hexachlorobenzene (HCB)], viral infections [specifically Hepatitis C virus (HCV) and human immunodeficiency virus (HIV)] and mutations in the high iron (*HFE*) gene (Nordmann and Puy 2002). It has been postulated that one or more of the precipitating factors may contribute to the formation of a liver-specific molecule which inhibits or decreases UROD activity (Bulaj *et al.* 2000). In a study of ten unrelated families, performed by Roberts *et al.* (1988) at least some of the sporadic cases were confirmed to be genetically determined. Mutations in the *UROD* gene have not been associated with sporadic PCT (Garey *et al.* 1993), indicating that other genetic factors may be involved in the development of this disease.

1.2.3 FAMILIAL PCT (fPCT)

fPCT accounts for between 10% and 34% of PCT cases (Kushner *et al.* 1976; Thadani *et al.* 2000) and is inherited in an autosomal dominant manner, often with low penetrance (Holti *et al.* 1958; Ziprkowski *et al.* 1966; Benedetto *et al.* 1978). This condition is associated with mutations in the *UROD* gene, resulting in half of the normal *UROD* protein concentration and activity in all tissues. Porphyrin accumulation occurs only in the liver (Bulaj *et al.* 2000). A single mutation in the *UROD* gene alone is not sufficient to promote the onset of PCT. The sporadic and familial forms of PCT tend to manifest following exposure to the same risk factors, although familial onset usually occurs at an earlier age due to the inherited low *UROD* enzyme activity. This observation has led to the assumption that phenotypic expression of PCT requires an as yet unexplained interaction between genetic and environmental factors (Bulaj *et al.* 2000).

1.2.4 HEPATOERYTHROPOIETIC PORPHYRIA (HEP)

A homozygous recessive state for PCT, referred to as hepatoerythropoietic porphyria (HEP) has also been reported (Mendez *et al.* 2005). This form of porphyria is a severe form of cutaneous porphyria that develops in infancy and it is associated with 3-27% *UROD* activity (Elder 1993; Anderson *et al.* 2001). In HEP, the severity of skin disease can lead to photomutilation, resulting in disfigurement of fingers, eyelids, lips, nose and ears. (www.porphyrria.uct.ac.za). In some instances, this may ultimately lead to the complete loss of the appendage.

1.2.5 SYMPTOMS AND DIAGNOSIS

Clinically, PCT is characterised by light-sensitive dermatitis induced by the deposition of uroporphyrins in the skin (Grossman *et al.* 1979). Upon exposure to light (especially ultra violet (UV) light), the porphyrin molecule absorbs the energy of the light ray, causing the molecule's electrons to be more active. With time, the molecule loses the high energy associated with the

electrons and it is this energy in the form of heat and light that causes damage to the skin. Cutaneous symptoms appearing especially on the hands, ears, neck and face include alopecia (loss of hair), skin fragility and hypertrichosis (abnormal hair growth) (Mehanry *et al.* 2003). Exposure to the sun results in the development of white papules (conical elevation of the skin) in areas of bullae (where bones are prominent), especially on the back of the hand, and blisters which ulcerate or form lesions (Figure 1.2) (Nordmann and Puy 2002). The blisters and papules may take many weeks to heal and tend to leave hypo- or hyperpigmented scars (McManus *et al.* 1996; Nordmann and Puy 2002).

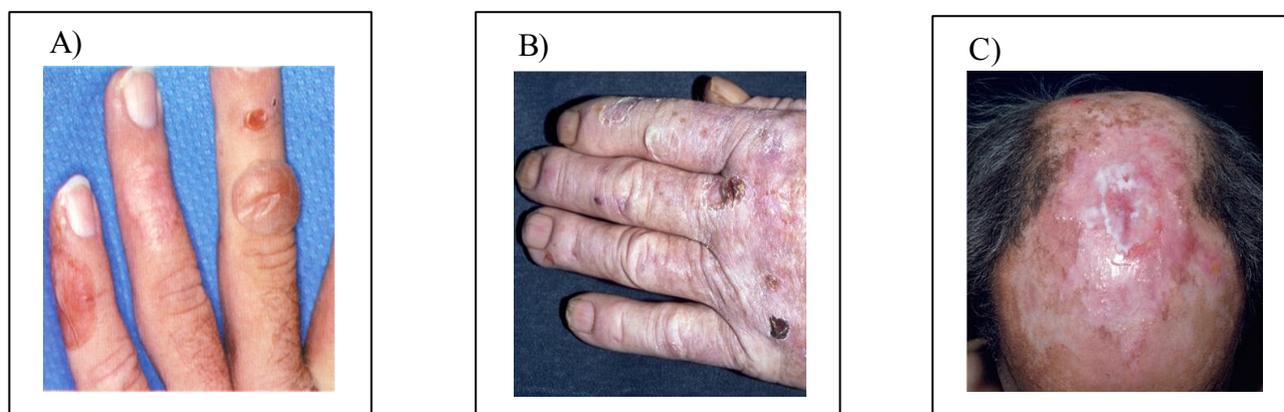


Figure 1.2. Images showing cutaneous symptoms of PCT. A) Blisters and papules are common on the back of the hands. B) Crusty lesions, scars and pigmentation may result from sun exposure on the hands, C) parietal scalp (<http://dermis.net>).

A biochemical basis of PCT may be established by demonstration of a characteristic pattern of accumulation of porphyrins in urine and plasma, particularly uroporphyrin and other water-soluble porphyrins (Kushner *et al.* 1972). The diagnostic routine followed in South Africa requires a blood and urine sample on which specific assays are performed (www.porphyrria.uct.ac.za; Hift 1999). Blood plasma fluorescence scans differentiate between types of porphyria. In PCT patients, a fluorescence scan will display a positive peak at 619 nm. In other countries a slightly wider peak range, 615-620 nm, is used as positive confirmation of PCT (www.porphyrria-europe.com). The

erythrocytes in the blood sample are also tested and a negative erythrocyte fluorescence test eliminates the possibility of an erythropoietic porphyria.

The urine sample is initially screened for porphobilinogen (PBG) [using the Watson-Schwartz reaction (Watson and Schwartz 1941)] and urine porphyrins [using Dean's method (www.porphyrria.uct.ac.za)]. If the test is positive the urine sample is further subjected to chromatographic quantitation of porphyrins. The presence of uroporphyrin, hepatocarbonylic porphyrin, hexacarboxylic porphyrin and pentacarboxylic porphyrin confirms the diagnosis of PCT. Occasionally, a stool sample is also tested on chromatography, as it also displays a characteristic pattern of accumulation. Genetic testing for fPCT is not routinely performed on South African patients.

PCT often manifests in the presence of precipitating factors, it is important to test for the associated conditions. Assessments include HIV and HCV antibody tests to detect the respective viruses; liver biopsy and liver function assays to estimate liver damage and liver enzyme levels (e.g. bilirubin, albumin, aspartate transaminase (AST), alanine transaminase (ALT), and alkaline phosphatase (ALP); mutation screening to determine the presence of hereditary haemochromatosis (HH) associated mutations and serum ferritin and transferrin saturation levels to estimate iron levels (www.porphyrria.uct.ac.za).

1.2.6 PRECIPITATING FACTORS ASSOCIATED WITH PCT

Environmental and genetic factors predispose to both sPCT and fPCT. These factors are considered to contribute in varying degrees and in a cumulative manner, to the pathogenesis of the disease. It thus follows that the presence of a number of risk factors constitutes a greater relative risk, as opposed to the lower risk attained by the presence of a single factor.

1.2.6.1 Environmental factors

i) Alcohol abuse

Alcoholism is frequently associated with liver disease and PCT. It is not a prerequisite for the development of the disease (Elder *et al.* 1972), but is considered to be the most important precipitating factor (Mendez *et al.* 2005). Alcohol may exacerbate PCT in several ways, although the exact mechanisms are not clear. The following hypotheses have been raised and are summarised by Nordmann and Puy (2002): 1) alcohol is thought to reduce haem synthesis by inhibiting one or more of the enzymes (especially UROD and ALA dehydratase) in this pathway. 2) alcohol may cause reduced haem synthesis and/or increased utilisation, and ensuing stimulation of ALA synthase production, resulting in the accumulation of porphyrins and 3) alcohol may stimulate the synthesis of the haemoprotein cytochrome P450 activity leading to an increased haem requirement in the liver. Cessation of alcohol consumption may lead to clinical and biochemical improvement.

ii) Toxic chemicals

Chemicals such as hydroxychlorobenzene, polychlorinated biphenyls (PCBs), polybrominated biphenyls (PBBs), 2,3,7,8-tetrachlorodibenzo-p-dioxin, lindane, and to a lesser extent lead, arsenic and mercury have been shown to be porphyrinogenic in humans (Daniell *et al.* 1997). One of the first chemicals to be identified as an exacerbating factor in PCT was hexachlorobenzene (HCB). A huge outbreak of approximately 4000 cases occurred in Turkey when natives consumed wheat that had been treated with HCB (Nordmann and Puy 2002). The induction of PCT from these chemicals is thought to be mediated through an oxidative reaction catalysed by CYP1A2 (Smith *et al.* 1983).

iii) Oestrogen therapy

Patients receiving oestrogen therapy for prostate and breast cancer or post menopausal symptoms have an increased incidence of PCT (Warin 1963, Grossman *et al.* 1979, Sassa *et al.* 2002).

However, their elevated risk usually occurs in conjunction with one or a number of the other predisposing factors. Consumption of oral contraceptives was previously not considered to be a risk factor (Elder *et al.* 1972), but it has since been proven otherwise (Mor and Capsi 1997).

iv) Hepatitis C virus (HCV)

HCV infection has been proven to be, along with *HFE* mutations, the greatest risk factor for PCT (Egger *et al.* 2002). About 75% of PCT patients from the United States of America and Western European countries have been exposed to the HCV. It has not been ascertained how HCV infection contributes to the development of PCT.

v) Human immunodeficiency virus (HIV)

Between 1987 and 1998 more than 60 cases of PCT associated with HIV infection were reported in European countries (Boisseau *et al.* 1991; Drobacheff *et al.* 1998). Initially it was thought that the presence of PCT in HIV infected patients was coincidental. However, it is now certain that HIV may pathogenically trigger the development of PCT in predisposed individuals, often in conjunction with any of the other precipitating factors. The mechanism by which this occurs is not evident, but it is reasoned that the virus causes damage to the hepatocytes thereby altering porphyrin metabolism and unmasking an existing UROD deficiency.

vi) Liver disease

Liver damage or disease, most commonly cirrhosis, is frequently found in PCT patients (Taddeini and Watson 1968; Cortes *et al.* 1980) and occurs even in the absence of excessive alcohol consumption or chemical poisoning. Patients display impaired liver function (Waldenstrom and Haeger-Aronsen 1960) and hepatomegaly (enlarged liver) is the most frequent histological finding.

vii) Systemic lupus erythematosus (SLE) and Lymphoma

Systemic lupus erythematosus (SLE) is an autoimmune inflammatory disease that affects the connective tissue of joints, tendons, lungs, kidneys, the heart and nervous system, and was first associated with PCT more than fifty years ago (Wolfram 1952). The association between SLE and PCT is rare, even less common than initially thought to be (Sinha *et al.* 1999), and are presumably related due to their common liver-related disease mechanism. Lymphoma comprises several cancers that develop in the lymphatic system. The occurrence of lymphoma in PCT patients is extremely rare and may be coincidental (Lai *et al.* 1984). It is suggested that PCT may manifest paraneoplastically in lymphoma patients .

viii) Smoking

Smoking has been designated as a risk factor in PCT, usually in conjunction with other exacerbating factors (Anderson *et al.* 2001; Egger *et al.* 2002) although the relationship between smoking and PCT is not clearly defined.

ix) Low vitamin C and carotenoid status

Low plasma concentrations of ascorbic acid has been observed in PCT patients and may contribute to the pathogenesis of PCT (Sinclair *et al.* 1997; Gorman *et al.* 2007). Low plasma levels of caretonoid (especially alpha- and beta caretonoid, cryptoxanthin and lycopene), which all possess antioxidative qualities, have also been observed in PCT patients. It is speculated that the oxidative damage caused by excessive porphyrin and iron levels, leads to the depletion of these molecules (Rocchi *et al.* 1995).

1.2.6.2 Genetic factors

i) Uroporphyrinogen decarboxylase (*UROD*) mutations

Uroporphyrinogen decarboxylase is a cytosolic enzyme coded for by the *UROD* gene localised to chromosome 1p34 (de Verneuil *et al.* 1984; Mattei *et al.* 1985; Dubart *et al.* 1986). The coding region of *UROD* consists of ten exons spanning 3.6 kb of DNA (Romana *et al.* 1987). The enzyme catalyses the sequential removal (decarboxylation) of four carboxymethyl side chains of uroporphyrinogen to produce coporphyrinogen in the haem biosynthesis pathway (Mauzerall and Granick 1958).

Phillips *et al.* (2001) noted that *UROD*^{+/-} mice only became porphyric under the pressures of additional precipitating factors. To date, 65 *UROD* mutations have been described in humans (Human Gene Mutation Database (HGMD), <http://www.hgmd.org>), 48 of which have been associated with PCT (Garey *et al.* 1989; Garey *et al.* 1990; McManus *et al.* 1996; Moran-Jimenez *et al.* 1996) and a further ten associated with HEP (de Verneuil *et al.* 1986; Romana *et al.* 1991; de Verneuil *et al.* 1992; McManus *et al.* 1996; Moran-Jimenez *et al.* 1996; Whitby *et al.* 1998; Christiansen *et al.* 1999). These include small and large deletions, insertions, and regulatory, splicing, missense and nonsense mutations. The most frequently observed mutation, associated with PCT, is a splice-site mutation (IVS5+3G→C) which results in the deletion of exon 6 (Garey *et al.* 1990).

ii) High iron gene (*HFE*) mutations

HFE gene mutations associated with HH have been identified in both sporadic and familial PCT and are considered, together with HCV, to be the greatest risk factor for PCT (Bulaj *et al.* 2000). The Cysteine282Tyrosine (C282Y) mutation is the most common *HFE* mutation associated with PCT and observed in 20-44% of patients in northern European countries, the United States of

America (USA) and Australia (Merryweather-Clarke *et al.* 1997; Roberts *et al.* 1997; Sampietro *et al.* 1998; Bulaj *et al.* 2000). However, homozygosity for the C282Y mutation is much higher than the frequency of PCT, thus mutations in the *HFE* gene do not solely account for PCT. The less common Histidine63Aspartic acid (H63D) mutation is prevalent in the Mediterranean communities (Hift *et al.* 2002). The frequencies of this mutation in PCT patients seems to be population dependent. It has also been noticed that patients that have a H63D mutation are likely to have the HCV. Conversely, HCV is less likely to be associated with the C282Y mutation (Toll *et al.* 2006).

1.2.7 TREATMENT AND CLINICAL MANAGEMENT

The principal method of treating PCT is to reduce iron overload by regular phlebotomy (Ippen 1961; Nordmann and Puy 2002). In South African patients, 500 ml of blood is removed fortnightly, usually for a period of three to four months, until serum ferritin levels and transferrin saturation reach the lower end of the normal range (www.porphyrria.uct.co.za). Phlebotomy is contraindicated in cases of anaemia and pulmonary and cardiac disorders; in these instances oral chloroquine is prescribed (Felsher *et al.* 1966; Taljaard *et al.* 1972; Sarkany *et al.* 2001).

Typically in South African patients, both venesection and a 125 mg dose of oral chloroquine three times weekly is prescribed (RJ Hift, personal communication). Chloroquine forms a complex with porphyrin, promoting the release of porphyrin from the liver *via* urinary excretions (Thadani *et al.* 2000). Chloroquine may also act to inhibit the synthesis of uroporphyrin. Management of the cutaneous symptoms include avoidance of/or protection from sunlight, and specialised skin care (Thadani *et al.* 2000). Sunblocks should typically contain high levels of zinc oxide to block out the short UV wavelength A (UVA) and particularly the long UV wavelength B (UVB) (www.porphyrria.uct.ac.za). Care should be taken to protect hands and face from trauma, blistering and scarring. In the event that lesions become infected, a course of antibiotics may be administered.

Improved disease management is achieved by cessation of alcohol and oestrogen intake, and medication aimed at treatment of HCV and/or HIV infections.

1.3 IRON METABOLISM

Iron is a vital mineral required by all living organisms for a number of metabolic and gene regulatory processes. These processes include oxygen and electron transport, the tricarboxylic acid cycle, respiration, DNA, RNA and protein synthesis, and gene expression at the transcriptional and post-transcriptional levels (Roy and Enns 2000; Lieu *et al.* 2001; Li *et al.* 2004; Cairo *et al.* 2006).

Iron is also a fundamental component of various enzymes, including: catalases, cytochromes, oxidases and ribonucleotide reductases (Boldt 1999; Conrad *et al.* 1999; Ponka 1999).

Approximately 12-20 mg dietary iron is consumed daily of which 1-3 mg is absorbed. In healthy adults iron represents between 35 and 45 mg/kg of body weight (Figures 1.3 and 1.4) (Smith 1990; Bothwell *et al.* 1995; Andrews 1999). Haemoglobin, found in erythrocytes, represents between 65% and 70% of total body iron (Anderson *et al.* 2006). Another 10% of iron is present in cytochromes, enzymes and myoglobins, whilst the remaining 20-30% is stored as ferritins in reticuloendothelial macrophages and hepatocytes (Conrad *et al.* 1999). The highest levels of iron are found in the liver, brain, erythrocytes and macrophages (Andrews *et al.* 1999), suggesting that iron has highly specialised and important functions in these tissues.

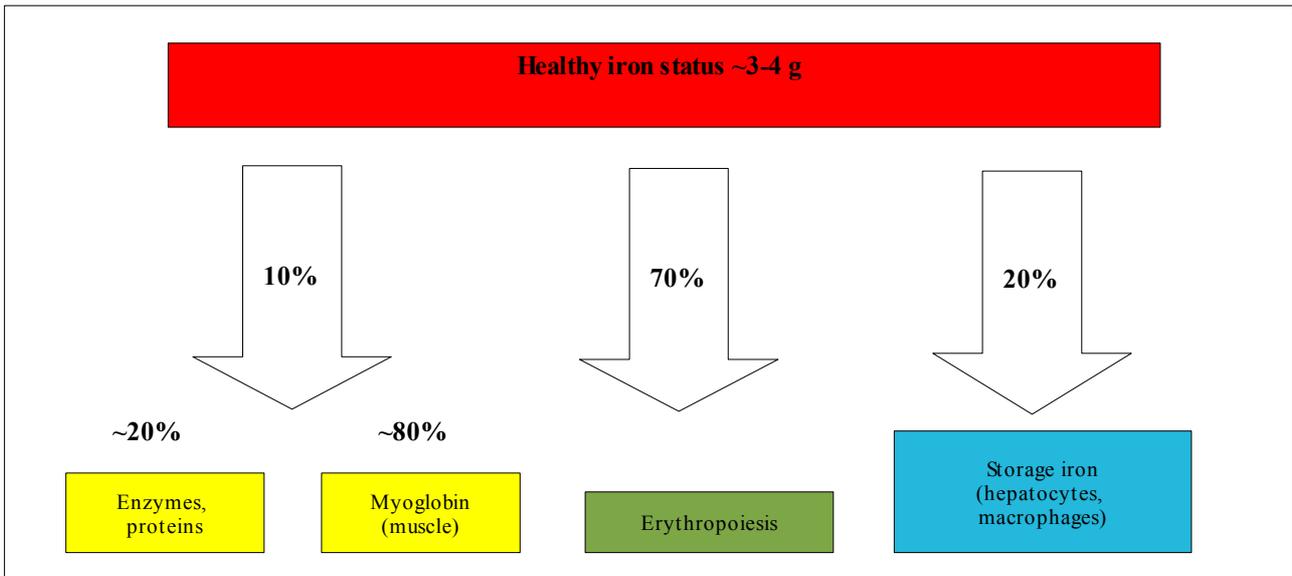


Figure 1.3. A schematic illustration of iron status in healthy individuals. Humans have a total body iron status of 3-4 g. Approximately 10% of this iron is required for enzymes, proteins and myoglobin and 70% is required for the production of new erythrocytes. The remaining 20% is stored in storage cells, predominantly in hepatocytes and macrophages.

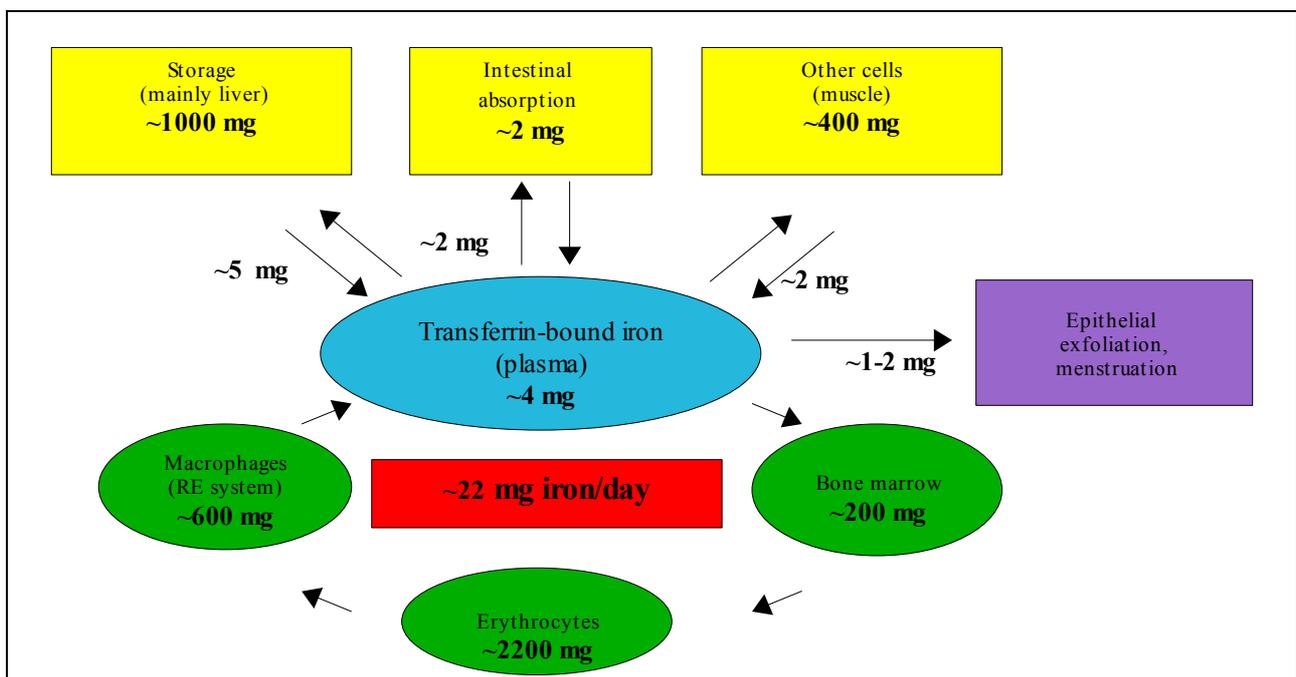


Figure 1.4 A schematic illustration of iron distribution and turnover in the human body. Approximately 4 mg of iron is bound by transferrin. Every day 22 mg of iron is recycled through the reticuloendothelial system. Iron is stored in various organs indicated. Arrows represent the turnover (efflux and/or influx) of iron to and from the organ.

Abbreviations: RE system, reticuloendothelial system.

Iron has the ability to donate and accept electrons in the oxidation-reduction reaction known as the Fenton reaction (Figure 1.5). Herein lies the biochemical importance of iron and its ability to perform the various metabolic processes mentioned (Wessling-Resnick 1999).



Figure 1.5. The Fenton reaction. Iron is capable of accepting and donating electrons to exist in either the oxidised or reduced state.

Abbreviations: Fe^{2+} , ferrous iron; Fe^{3+} , ferric iron; H_2O_2 , hydrogen peroxide; O_2 , oxygen; $\cdot\text{O}_2^-$, superoxide radical; $\cdot\text{OH}$, hydroxyl free radical; OH^- , hydroxyl ion.

Iron metabolism consists of three main processes, i.e. absorption, storage and release (Cairo *et al.* 2006). These processes have to be tightly regulated in order to maintain iron homeostasis and resulting good health. Defects in iron homeostasis lead to either iron deficiency or iron overload disorders. These diseases are categorised in three groups (Lieu *et al.* 2001).

- 1) Diseases associated with a defect occurring in iron absorption;
- 2) Diseases associated with incorrect tissue and cell storage of iron;
- 3) Secondary diseases caused by altered iron levels in tissue and cells.

1.3.1 IRON ABSORPTION

The complete process of iron absorption in the small intestine occurs over three phases. The initial phase refers to the uptake of iron across the brush border (apical membrane) of the duodenal and jejunal enterocytes. Iron is subsequently translocated across the cytosol of the enterocyte where it is exported across the basal membrane into the body's circulation (Figure 1.6) (Lieu *et al.* 2001). Enterocytes situated in the intestinal villi are specialised, polarised cells required for the absorption of iron across the apical membrane (Tapiero *et al.* 2001). For efficient absorption, iron must be in the soluble ferrous (Fe^{2+}) form. Since most dietary iron is in the insoluble inorganic ferric (Fe^{3+})

form, a mucosal feroreductase enzyme, present in the gastrointestinal tract (GIT), is required to reduce ferric iron to ferrous iron (Frazer and Anderson 2003).

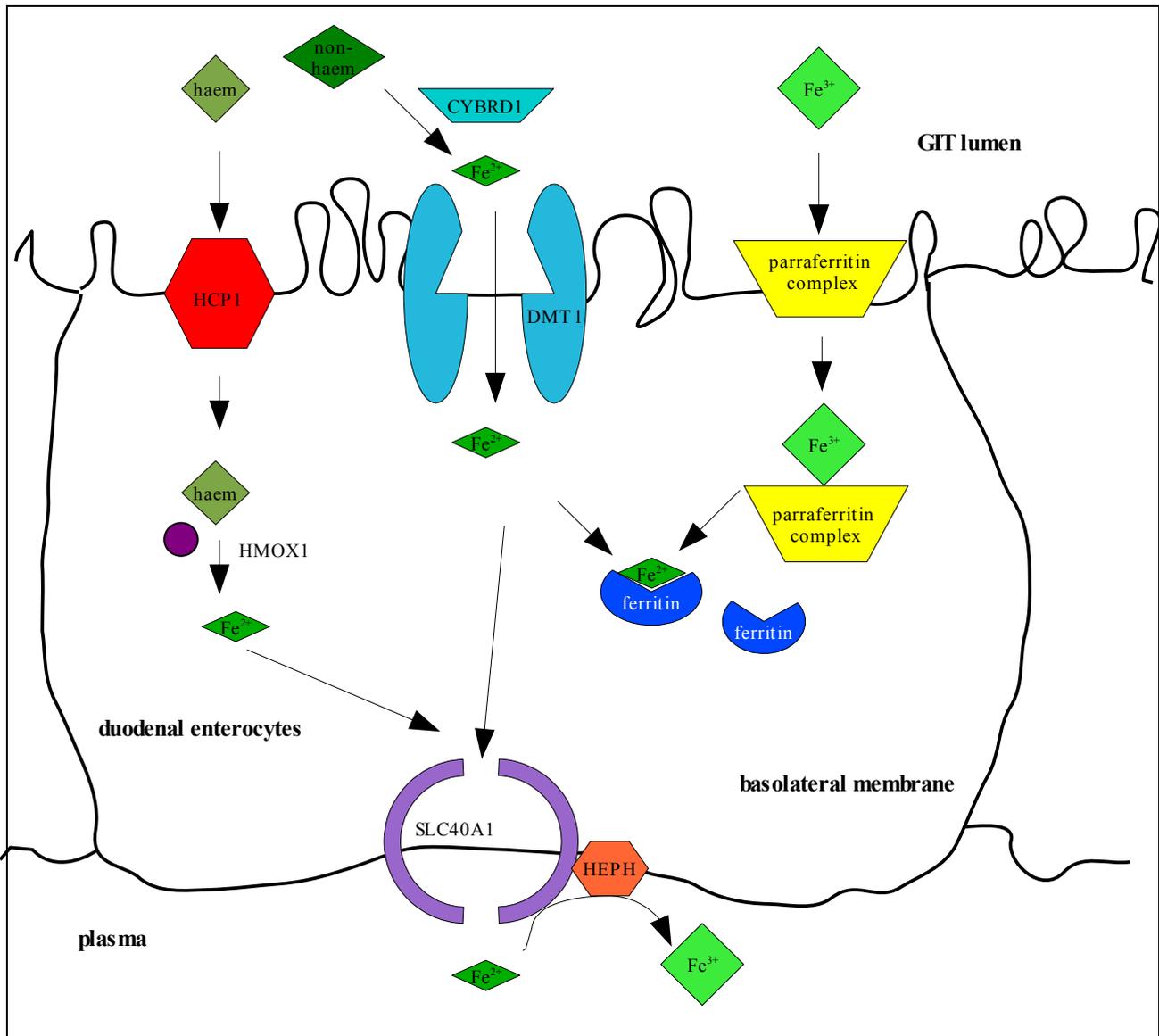


Figure 1.6 A schematic representation of iron absorption in the duodenal enterocyte (Drawing not to scale). Iron absorption refers to the uptake of iron across the apical membrane of the enterocytes present in the duodenum and jejunum, iron translocation across the cytosol and the export of iron across the enterocytic basal membrane into the body's circulation as described.

Abbreviations: CYBRD1, cytochrome b reductase 1 protein; DMT1, divalent metal transporter protein; Fe²⁺, ferrous iron; Fe³⁺, ferric iron; GIT lumen, gastrointestinal tract; HCP1, haem carrier protein 1; HEPH, hephaestin protein; SLC40A1, solute carrier family 40 member 1 protein.

1.3.1.1 Dietary iron uptake

The average daily healthy diet contains approximately 12-20 mg of iron and is present in two forms, haem and non-haem bound irons (Smith 1990). However, only 1-3 mg of iron is absorbed each day. Both these forms of iron are absorbed by the crypt cells (enterocytes) of the duodenum and jejunum (Conrad *et al.* 1987; Wood and Han 1998), but the absorption processes occur by different mechanisms (Lieu *et al.* 2001; Trinder *et al.* 2002).

i) Haem iron uptake

Haem iron is found in myoglobin (muscle) of meat and poultry, and in haemoglobin. This form accounts for 10-20% of iron dietary uptake (Lombard *et al.* 1997). Iron associated with haem is also absorbed more efficiently compared to inorganic (non-haem) iron absorption.

Initially it was thought that haem iron is enzymatically digested in the lumen of the GIT and that haem iron then enters the enterocyte as an intact metalloprotein *via* an unidentified receptor through an internalisation process (Majuri and Grasbeck 1987; Mills and Payne 1995). Once inside the cell, haem oxygenase (HMOX1) would be responsible for the degradation of haem and the release of inorganic iron. The discovery of the novel plasma membrane haem carrier protein 1 (HCP1) has since provided an alternative theory for the way by which haem enters the enterocyte (Latunde-Dada *et al.* 2006). This 54 kDa protein is extremely hydrophobic. It resides on the apical membrane under conditions of iron deficiency and in the cytoplasmic region in states of iron overload. The process of iron uptake across the membrane has yet to be elucidated. However, various findings support the hypothesis that HCP1 plays a significant role in the absorption of haem iron. *HCP1* is highly expressed in the liver, kidney and proximal intestine, in particular the duodenum. A study performed on *Xenopus* oocytes indicates that iron uptake is increased two to three fold when HCP1 is expressed (Latunde-Dada *et al.* 2006). This process was found to be temperature dependent and

saturable, indicative of a carrier mediated process. The mRNA transcripts of *HCPI* are regulated by hypotransferrinaemia and hypoxia, consistent with other genes in the iron metabolism pathway.

ii) Non-haem iron uptake

Non-haem iron is available in cereals, legumes, pulses, fruit and vegetables and constitutes up to 80% of dietary iron (Lombard *et al.* 1997). This form of iron, sometimes referred to as ionic iron, readily forms insoluble complexes at pH 3 in the GIT. Cytochrome b reductase 1 [*CYBRD1*, also referred to as duodenal cytochrome b reductase (*DCYTB*)] expressed in the brush border of the duodenum, possesses ferrireductase activity and is responsible for reducing the insoluble ferric iron into the soluble ferrous iron form *via* the Fenton reaction (McKie *et al.* 2001). The absorptive enterocytic cells do not have transferrin receptors (Pietrangelo *et al.* 1992), indicating that ferrous iron is absorbed by a different mechanism than the conventional transferrin pathway. Ferrous iron is absorbed into enterocytes across the apical membrane by a proton-coupled divalent metal transporter (DMT1) [also referred to as divalent cation transporter 1 (DCT1) or natural resistance-associated macrophage protein 2 (NRAMP2)] (Fleming *et al.* 1997). DMT1 function is optimal at a pH below 6 and this protein is capable of transporting not only ferrous iron, but also a number of divalent cations, such as Mn^{2+} , Cu^{2+} , Zn^{2+} , Ni^{2+} and Pb^{2+} (Gunshin *et al.* 1997).

iii) Paraferitin-mediated iron uptake

Ferric iron can also be transported from the GIT into the enterocytes by the less common and less efficient paraferitin-mediated pathway (Conrad *et al.* 1999). Paraferitin is a 520 kDa membrane complex consisting of flavin mono-oxygenase, mobilferrin and β -integrin. This pathway is not clearly defined but it has been hypothesised that mucin present in the GIT binds and solubilises the ferric iron (Beutler *et al.* 2001) which is then transported to the mobilferrin and β -integrin complexes and internalised. Flavin mono-oxygenase together with NADPH activity is then

responsible for reducing ferric iron to ferrous iron.

1.3.1.2 Iron storage and translocation in enterocytic cells

Once the iron molecule has been transported into the enterocyte *via* one of the three pathways described, it is either translocated to the basolateral membrane for export, or stored in the cell. The mechanism(s) required for this translocation are unknown (Dunn *et al.* 2006). It has been put forward that chaperone proteins may transport the iron to the basolateral membrane (Tapiero *et al.* 2001). Alternatively, protein membrane channels or carriers may assist in the translocation process. Iron is stored in the enterocyte in the form of ferritin. This storage process is discussed in detail in section 1.4.3.

1.3.1.3 Iron export

Iron export or transfer is the final stage of iron absorption and is described as the export of ferrous iron across the basolateral membrane into the plasma. Only one iron exporting protein has been identified to date, ferroportin 1 (Donovan *et al.* 2000, Donovan *et al.* 2006). Ferroportin 1, also known as iron-regulated transporter 1 (*IREG1*), metal transporter protein 1 (*MTP1*) or solute carrier family 40 Member A1 (*SLC40A1*) is responsible for mediating iron transport or efflux from the enterocyte into the plasma (Donovan *et al.* 2000). This 62 kDa protein is present on the basolateral membrane of enterocytes (McKie *et al.* 2000). Ferroportin is responsible for iron export not only from the enterocytes, but also from the Kupffer cells in the liver and from reticuloendothelial macrophages (Abboud and Haile 2000). Once iron has crossed the basolateral membrane of the enterocyte and before it binds to plasma transferrin, it is oxidised by the multi-copper oxidase, hephaestin (*HEPH*) (Vulpe *et al.* 1999). *CP*, a homolog of *HEPH*, is responsible for the oxidation of ferrous to ferric iron after its release from macrophages and Kupffer cells (Harris *et al.* 1998).

1.3.2 TRANSFERRIN RECEPTOR-MEDIATED IRON UPTAKE

Transferrin is a 80 kDa serum protein principally synthesised in the liver (Morgan 1983), and to a lesser extent in the testis, lactating mammary gland, brain and in fetal tissue (Dickson *et al.* 1985; Takeda *et al.* 1998). Transferrin has a high affinity for ferric iron, and in healthy individuals iron circulates in the plasma bound to transferrin between sites of absorption, storage and utilisation (Lieu *et al.* 2001; Hentze *et al.* 2004). By binding iron, transferrin also prevents iron from forming reactive oxygen species (ROS) and subsequent cellular toxicity (Lieu *et al.* 2001). Transferrin consists of two globular domains; each possessing one binding site for one molecule of iron (Yang *et al.* 1984). Most cells present in non-intestinal regions acquire iron from transferrin. Transferrin may exist in three different forms; iron free (apotransferrin), containing one molecule of iron (monoferric) or containing two molecules of iron (diferric). The binding and release of iron from transferrin is pH dependent; at a lowered pH (pH<6.5) iron is released from transferrin. Transferrin enters cells *via* an endocytic pathway involving transferrin receptors to which transferrin binds (Figure 1.7).

There are at least two types of transferrin receptors (Lieu *et al.* 2001) of which transferrin receptors 1 (TfR1) and 2 (TfR2) are the most relevant here. TfR1 is a glycoprotein situated across the cell membrane of most cells and participates in general iron uptake (Hentze *et al.* 2004). The liver, placental tissue, immature erythrocytes and proliferating cells express the highest levels of this receptor (Ponka 1999). This receptor can bind two molecules of transferrin and has the highest affinity for diferric transferrin. This physical interaction is pH dependent, and not temperature or energy dependent (Ciehanover *et al.* 1983). After interaction with various adaptor proteins and *via* coated pits and vesicles, the transferrin receptor-transferrin complex is internalised by an endocytic pathway, a process that is temperature and energy dependent (Thorstensen and Romslo 1990; Ponka 1999). To enable release from the transferrin within the endosome, the pH has to be lowered

significantly. This is achieved by an unknown ATPase proton pump. The transferrin molecule, free of iron, but still attached to its receptors, is returned to the cell surface. Under conditions of a neutral pH, apotransferrin is released from its receptor.

HFE, a regulatory molecule, associates physically with the TfR1 (Parkkila *et al.* 1997; Feder *et al.* 1998), thereby competitively inhibiting the binding of transferrin-bound iron (TBI) to its receptor. (Feder *et al.* 1998; West *et al.* 2001). This association leads to a decreased cycling time of the HFE/TfR1/TBI complex and/or decreases the rate of iron released from transferrin once inside the cell (Siah *et al.* 2006). An association between an intronic *TfR1* mutation (IVS4+198) and sPCT has been described (Lamoril *et al.* 2002).

TfR2 is a TfR1 homologue, and is expressed in the hepatocytes and to a lesser extent other regions of the liver, as well as in the duodenal crypt cells, prostrate, spleen and erythroid cells (Kawabata *et al.* 1999; Kawabata *et al.* 2001). TfR2 does not contain any typical iron response elements (IRE) distinctive of some of the other genes involved in iron metabolism (Calzolari *et al.* 2007). The binding of transferrin to TfR2, like TfR1, is dependent on acidic pH. The remaining binding properties and regulation of expression differ between the two receptors. Notably, TfR2 expression is not regulated by intracellular iron status. Also, TfR2 binds transferrin with a 30-fold lower affinity than TfR1 (Siah *et al.* 2006). TfR2 is associated particularly with iron uptake and storage in hepatocytes and may play a role in liver iron overload (Kawabata *et al.* 1999). However, it seems that TfR polymorphisms are unlikely to be associated with the iron overload observed in PCT (Dereure *et al.* 2001).

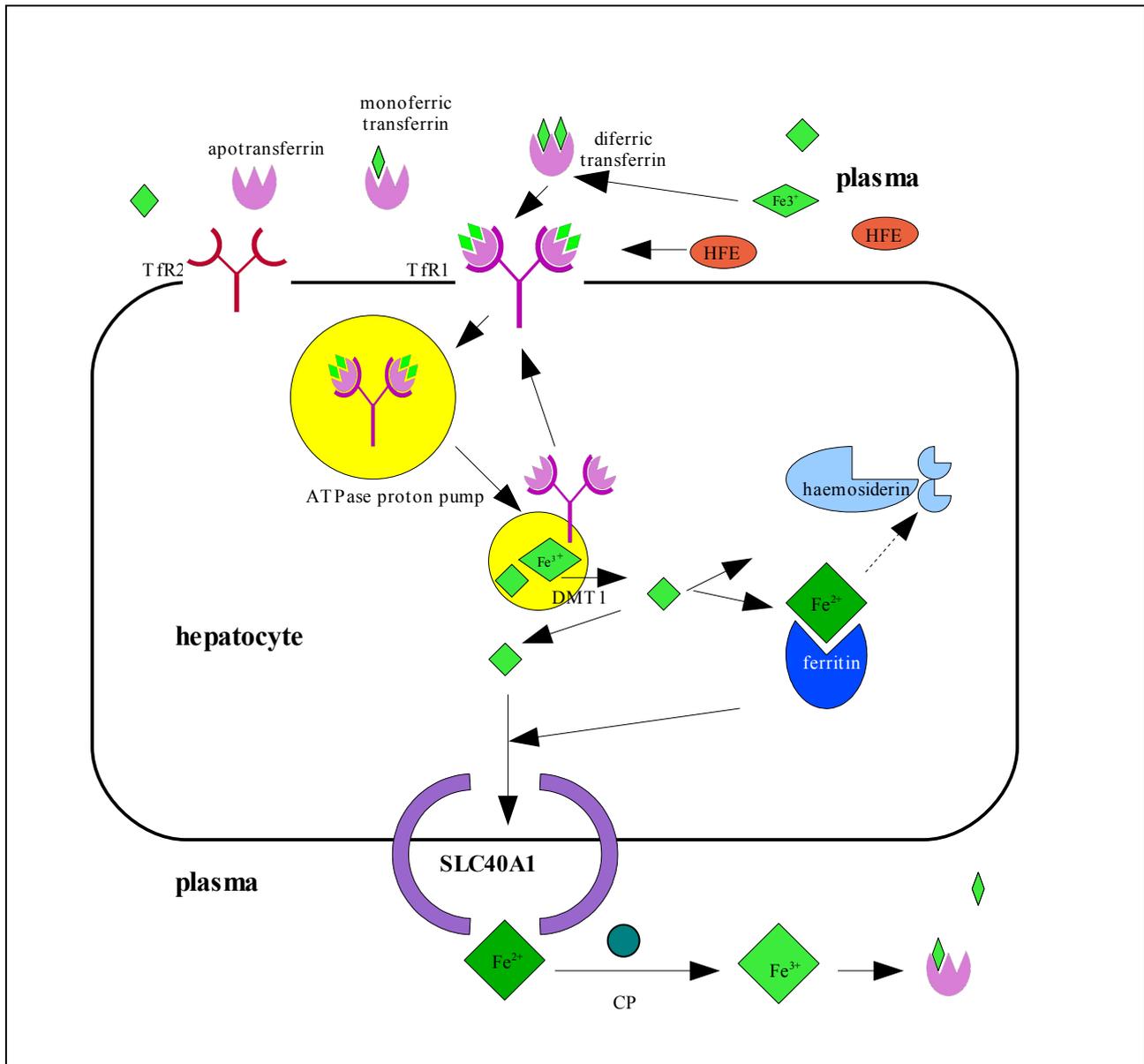


Figure 1.7 A schematic representation of iron uptake and storage (Drawing not to scale). Iron uptake is initiated when iron transported as transferrin interacts directly with the transferrin receptors. It is subsequently internalised into the storage cell as described (Sections 1.4.2 and 1.4.3).

Abbreviations: ATPase, adenosine triphosphatase; CP, ceruloplasmin protein; DMT1, divalent metal transporter 1 protein; Fe^{2+} , ferrous iron; Fe^{3+} , ferric iron; HFE, high iron protein; SLC40A1, solute carrier family 40 member 1 protein; Tfr1, transferrin receptor 1; Tfr2, transferrin receptor 2.

1.3.3 IRON STORAGE

The hepatic parenchyma and reticuloendothelial macrophages of the bone marrow, spleen and liver are the principal sites of iron storage (Knutson and Wessling-Resnick 2003; Cairo *et al.* 2006). The liver and bone marrow of healthy individuals each contains 100-300 mg of iron (Gale *et al.* 1963; Bothwell *et al.* 1979). Once iron has entered the cells *via* the transferrin receptors, it is deposited into an ubiquitous water soluble protein known as ferritin. Alternatively iron may be stored as haemosiderin, a water insoluble protein complex (Crichton 1991; Williams *et al.* 2006). If iron is not stored as ferritin or haemosiderin, it will form ROS. Ferritin is a heteropolymer, consisting of two subunits, a heavy (H-form) and a light (L-form) chain. It forms a hollow protein shell capable of storing as many as 4500 iron molecules (Smith 1990; Harrison and Arosio 1996, Theil 1998; Williams *et al.* 2006). In the reticuloendothelial storage cells ferritin occurs most frequently in the L-form, the form that is intrinsically associated with iron storage (Levi *et al.* 1994). As iron accumulates in the cell, more iron will be deposited as haemosiderin, as this protein is able to store more iron per unit volume in the cell than ferritin (Knutson and Wessling-Resnick 2003). Ferritin possess a detoxification mechanism which prevents iron from forming free radicals, thereby preventing cell damage (Lieu *et al.* 2001). It has been inferred that the presence of ROS may up-regulate the transcription of ferritin.

1.3.4 IRON RECYCLING

The total body iron content of a healthy adult is in the region of 3-4 g. The bone marrow requires about 20-24 mg per day (approximately 80% of the iron demand in humans) for the manufacture of around 200 billion erythrocytes (Knutson and Wessling-Resnick 2003; Hentze *et al.* 2004). Iron must therefore be recycled to meet the demand of haem production required for erythropoiesis (Figure 1.8). Macrophages of the reticuloendothelial system are responsible for the iron recycling process. Macrophages occur most abundantly in the liver, small and large intestine, bone marrow,

kidneys and spleen. The spleen, liver and bone marrow are the most active sites of erythrocyte destruction (erythrophagocytosis) and iron recycling (Lee *et al.* 1985; Rossi 2005). Macrophages phagocytose the red blood cells and the subsequent proteolytic digestion of haemoglobin releases haem. The haem molecules are either diffused or transported by an unknown mechanism from the phagolysome in the macrophage to a site where HMOX1 is present. HMOX1 catabolises haem to produce biliverdin, ferrous and carbon monoxide (Maines 1997). Iron molecules are then returned to the blood, where they bind to transferrin or are stored within the macrophage as ferritin.

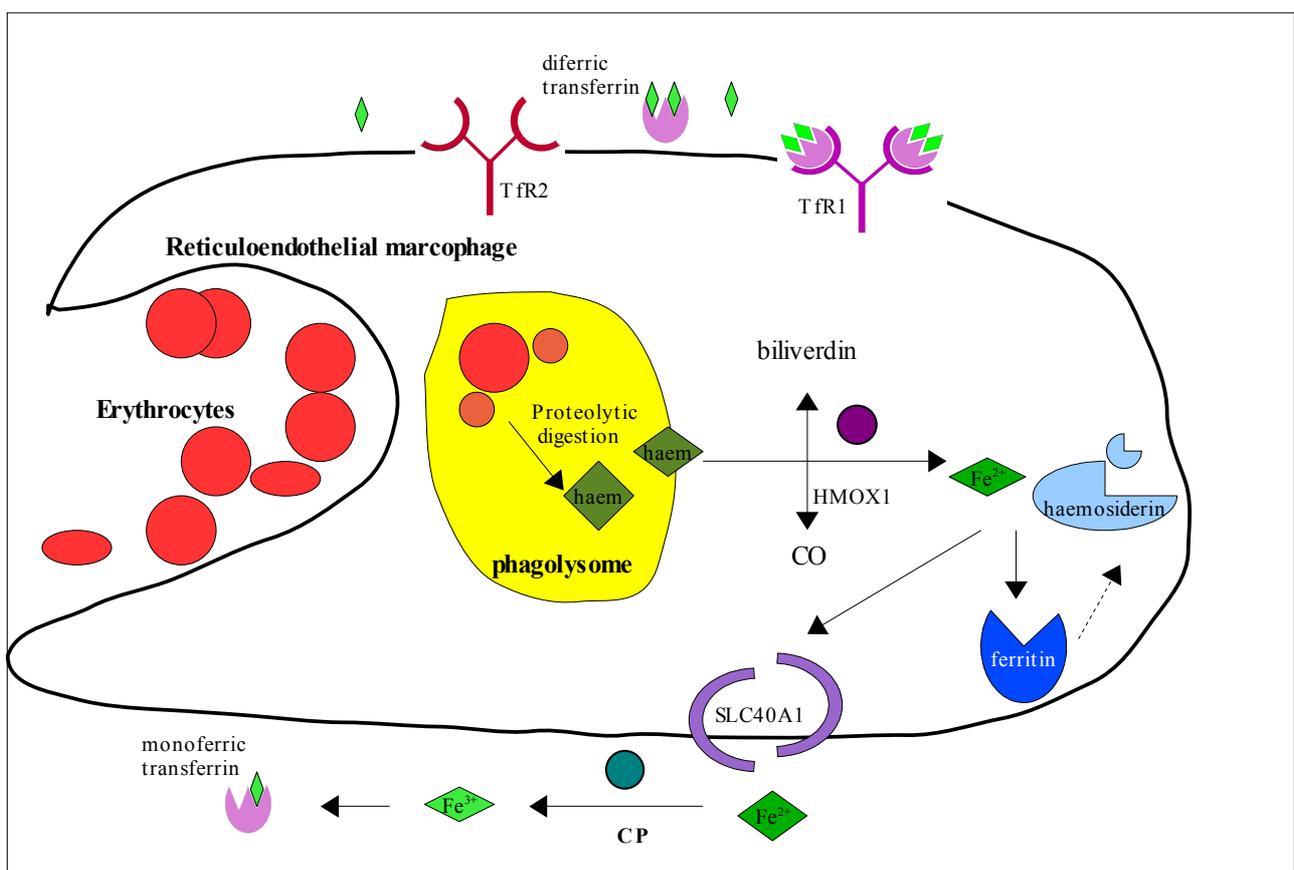


Figure 1.8 A schematic representation of iron recycling (Drawing not to scale). Senescent erythrocytes are phagocytosed by the reticuloendothelial macrophages and converted to haem. They are then either stored or transported to the bone marrow for erythropoiesis, as described (Section 1.4.4)

Abbreviations: CO, carbon monoxide; CP, ceruloplasmin protein; Fe²⁺, ferrous iron; Fe³⁺, ferric iron; HMOX1, haem oxygenase 1 protein; SLC40A1, solute carrier family 40 member 1 protein; TfR1, transferrin receptor 1; TfR2, transferrin receptor 2.

1.3.5 IRON EXCRETION

It has been mentioned that ferroportin is responsible for iron release from macrophages, enterocytes and hepatocytes (Knutson *et al.* 2005; Cairo *et al.* 2006). Thus, at the cellular level, iron excretion is tightly regulated. However, the human body does not possess a regulatory systemic mechanism for iron release. The only means by which loss of iron (approximately 1 mg per day) is achieved, is from bleeding (including menstruation), biliary excretion, epithelial cell desquamation and pregnancy (Smith 1990; Cairo *et al.* 2006).

1.3.6 IRON HOMEOSTASIS

Since the body lacks an efficient pathway for the excretion of iron, the fundamental means of regulating iron metabolism is *via* iron absorption. Iron homeostasis is maintained on both the cellular and systemic level *via* different regulatory molecules and processes.

1.3.6.1 Iron regulation at the cellular level

The complex regulation of iron metabolism at the cellular level requires co-ordinated expression of the genes involved in iron absorption, iron storage and iron release (Calzolari *et al.* 2007), especially between the iron storage (ferritin) and the iron uptake proteins (TfR1) (Cairo *et al.* 2006). The regulation mechanism involves the direct interaction between iron responsive/regulatory proteins 1 and 2 (IRP1 and IRP2) and IREs (Rouault *et al.* 1990). IREs are conserved stem loop structures found in the 5' and 3' UTRs of various messenger RNA (mRNA) coding for iron metabolism proteins (such as ferritin, TfR1, NRAMP2, and SLC40A1). IRPs bind to IREs with high affinity and high specificity (Hentze and Kühn 1996; Hanson and Leibold 1999; Cairo and Pietrangelo 2000).

IRP1 is regulated by the intracellular labile iron pool (Haile *et al.* 1992a; Haile *et al.* 1992b;

Rouault *et al.* 1992). When iron levels are high, an increase in IRP1 activity inhibits IRE binding through the formation of a [4Fe-4S] cluster; thereby causing efficient translation of ferritin mRNA and a decreased stability of TfR1 mRNA. This eventually leads to enhanced iron storage vs decreased iron uptake. In contrast, when iron levels are low, IRP1 activity is enhanced and binds to the IREs, causing a decrease in ferritin translation and a more stabilised form of TfR1 mRNA, resulting in enhanced iron uptake and availability within the cells (Cairo *et al.* 2006).

IRP2 does not possess a [Fe-S] cluster and is regulated by its own degradation in response to iron levels. IRP2 is degraded under conditions of high levels of iron, while it accumulates when iron levels are low (Guo *et al.* 1995; Pantopoulos *et al.* 1995a; Pantopoulos *et al.* 1995b; Hentze *et al.* 2004).

The amount of iron present in the labile iron pool of the enterocyte also appears to regulate several proteins involved in iron homeostasis. Transferrin receptor and ferritin expression respond to such changes in the labile iron pool (Anderson *et al.* 1990; Pietrangelo *et al.* 1992; McKie *et al.* 1996; Pountney *et al.* 1999). The presence of IREs in ferroportin and DMT1 have led to the claim that the translation of these mRNAs may be altered during periods of fluctuations in the labile iron pool. Ferroportin has an IRE in its 5' UTR, signifying that mRNA translation would occur more efficiently in the presence of a high labile iron pool (Gunshin *et al.* 1997; Abboud and Haile 2000; McKie *et al.* 2000). In contrast, DMT1 has an IRE in its 3' UTR, indicating the mRNA would be degraded in the presence of a high labile iron pool.

1.3.6.2 Iron regulation at the systemic level

The communication between iron storage cells (enterocytes, hepatocytes and macrophages) and iron consuming cells (erythroid precursors) needs to be highly effective to ensure that transfer between these tissues and cells maintains iron homeostasis (Hentze *et al.* 2004). As mentioned before, the

most effective means of controlling iron homeostasis is by regulating iron absorption, especially dietary iron absorption. The regulation of iron absorption is dependent on several factors, such as the rate of erythropoiesis, hypoxia, level of iron in storage cells and recent dietary iron uptake (Lieu *et al.* 2001).

The “store regulator” may modulate enterocytic iron absorption in response to the iron levels in the storage cells of the liver, skeletal muscle, blood and tissue macrophages (Finch 1994; Roy and Enns *et al.* 2000). It seems that the “store regulator” acts to increase the absorption of non-haem dietary iron at a slow rate until iron stores are replete (Sayers *et al.* 1994). The “store regulator” also functions to prevent iron overload (Finch 1994). Once the body's iron requirements have been satisfied, the absorptive cells of the intestinal epithelium are reconditioned, to reduce the subsequent uptake of iron (Finch 1994; Roy and Enns *et al.* 2000). It is hypothesised that a putative soluble plasma signal allows communication between these storage cells. The “store regulator” in effect helps to prevent iron overload (Finch 1994).

The rate of erythropoiesis has the greatest effect on iron absorption. The erythroid cells are the major consumers of iron. When the erythron's demand for iron is greater than the ability of the storage cells to mobilise and release iron for erythropoiesis, the “erythroid regulator” signals to the intestinal region to increase iron absorption (Andrews 1999; Hentze *et al.* 2004). This form of regulation is presumed to be totally independent of body iron stores and is thought to require a soluble plasma protein that signals between the bone marrow and the intestine (Cazzola *et al.* 1999). There is also evidence that intestinal iron absorption is induced in response to hypoxia (lack of oxygen in blood and tissue) by a humoral “hypoxia regulator” (McKie *et al.* 2000; Hentze *et al.* 2004).

Mucosal block is a phenomenon brought about by enterocytes resisting additional iron absorption for a few days after sufficient dietary iron has been consumed, thereby satisfying the body's intracellular iron requirements (Andrews 1999). The level of iron in storage cells can influence the uptake of dietary iron, especially in iron-deficient states (Finch 1994). Several dietary elements either inhibit or enhance iron absorption in the GIT. Ascorbic acid, meat and fermented vegetables are known to enhance the absorption of iron (Siegenberg 1991), while calcium, bran products and iron binding compounds such as tea and coffee, inhibit the absorption of iron (Gillooly 1983; Hallberg 1993).

During the course of inflammation and infection, iron is retained within the cells, possibly to prevent access of the metal to the microbes. This response is called the “inflammatory regulator”. The process causes iron retention and accumulation in the macrophages responsible for erythrophagocytosis, and as a result iron absorption is interrupted (Hentze *et al.* 2004). These regulators may not be independently controlled; the same regulating molecules may activate a specific regulator in a dose-dependant manner (Hentze *et al.* 2004).

Several studies have concluded that hepcidin antimicrobial peptide (HAMP) may be one of the vital iron regulating molecules (Pigeon *et al.* 2001; Rossi 2005). HAMP, also known as liver-expressed antimicrobial peptide (LEAP1) or hepcidin (HEPC), is a peptide hormone synthesised by the liver. The HFE-related HH hepcidin model clarifies the basic principals of iron regulation by HAMP (Pietrangelo 2004; Fleming 2005). This model suggests that the rate of iron efflux is initially dependent on the plasma levels of hepcidin. When iron plasma levels are high, HAMP synthesis is upregulated, and the release of iron from storage cells is diminished. Conversely, HAMP synthesis is down-regulated when iron plasma levels are low and this increases the release of iron from storage cells. HAMP maintains plasma iron levels by regulating the mobility of iron from

hepatocytes and the absorption of dietary iron from the GIT, as well as controlling the release of iron from macrophages (Nemeth *et al.* 2004). Hepcidin is secreted in response to inflammation, anaemia, hypoxia and liver iron levels. *HAMP* expression by hepatocytes is also regulated by haemojuvelin (*HJV*) via bone morphogenetic proteins (BMPs) (Babitt *et al.* 2006).

Ferroportin expression is controlled by a number of factors, including inflammation, iron status and hypoxia, which modulates transcriptional and posttranscriptional regulation (Pietrangelo 2004). Ferroportin is also negatively regulated by HAMP. HAMP and ferroportin, interact directly on a molecular level, causing ferroportin to be internalised and prevention of iron export from the enterocytes (Nemeth *et al.* 2004). Since ferroportin is also present in hepatic cells and macrophages, *HAMP* also modulates the release of iron from these cells. This post-translational regulation of *SLC40A1* by hepcidin may complete a loop/link in iron homeostasis.

1.4 IRON AND PCT

A tightly regulated body iron status is vital for maintaining good health. A disruption in iron homeostasis will lead to either iron overload or iron deficiency disorders, responsible for serious and even fatal health problems. Iron overload/hepatic siderosis has long been associated with PCT (Berlin 1962; Uys and Eales 1963; Lundvall *et al.* 1970), and is detected in approximately 80% of PCT patients (Nordmann and Puy 2002; Lambrecht *et al.* 2007). Iron overload is defined as the condition in which the rate of iron uptake exceeds the rate of iron release by hepatocytes (Trinder *et al.* 2000).

As iron accumulates, transferrin becomes saturated with iron. Excess iron is referred to as non transferrin-bound iron (NTBI). Under normal circumstances, NTBI is quickly removed from the plasma *via* a carrier mediated mechanism involving iron reduction (Wright *et al.* 1986; Randell *et al.* 1994). Under circumstances of iron overload, there is not a concomitant increase in iron reduction, allowing accumulation of NTBI and generation of ROS *via* the Fenton reaction. This leads to cellular toxicity and subsequently the impaired synthesis of proteins, membrane lipids and carbohydrates, DNA damage, induction of proteases, altered cell proliferation and ultimately cell death (Halliwell and Gutteridge 1992). These damaging effects imply that iron may play a significant role in the development of carcinogenesis, atherosclerosis and neurodegenerative disorders (Connor *et al.* 1992).

PCT patients often experience elevated serum aminotransferases and gamma-glutamyl transpeptidase levels in their hepatic tissues (Badminton and Elder 2005). Elevated levels of plasma ferritin, similar to that observed in HH patients, have been identified in PCT patients. Furthermore, 30 to 40% of PCT patients display inflammation, fibrosis, necrosis and fat deposition in the liver.

Patients may also develop liver cirrhosis and hepatic carcinomas (Francanzani *et al.* 2001). These symptoms highlight the effect of PCT on the liver or *vice versa*.

Various studies support the hypothesis that iron overload is a risk factor for PCT (Felsher 1977; Lundvall *et al.* 1970; Elder and Roberts 1995). Three distinct observations by several researchers are summarised by Kushner *et al.* (1972). Firstly, PCT patients invariably display hepatocytic iron overload. Secondly, clinical and biochemical remission is achieved in patients receiving regular phlebotomy. Thirdly, subsequent administration of iron causes relapse. The mechanism by which iron may potentially trigger the induction of PCT remains unclear. Iron has been suggested to play an indirect role in the inhibition of UROD activity (Elder and Roberts 1995), through an iron dependent process (Badminton and Elder 2005). It is thought that increased iron levels contributes to elevated oxidation of ROS. The excessive accumulation of these oxidised products causes inhibition of UROD activity, decreased decarboxylation of uroporphyrinogens and accumulation of porphyrins. Recently, a uroporphomethene molecule was identified as a byproduct of an iron dependant oxidation reaction (Phillips *et al.* 2007). This molecule may represent the sought after putative UROD inhibitor as the clinical manifestation of PCT requires the generation of a porphomethane inhibitor *via* the iron dependent oxidation reaction.

The hepatic iron overload observed in PCT patients, appears to be similar to that observed in HH patients. Edwards *et al.* (1998) were the first to suggest that *HFE* mutations associated with HH might be responsible for the iron overload in PCT patients. The frequencies of the two common HH associated *HFE* mutations (C282Y and H63D), were more prevalent in PCT patients than in control populations. HH patients are four times more likely to suffer from sporadic PCT than healthy individuals (Elder and Worwood 1998). However, most studies argue that iron overload of PCT patients is not severe, reaching only within the lower end of the HH range in 10% of patients

(Lundvall *et al.* 1970; Edwards *et al.* 1989). It has been noted that iron overload and consequently liver damage and cirrhosis are particularly prevalent in populations that consume large amounts of beer and wine (MacPhail *et al.* 1979; Fletcher *et al.* 1999)

There is a consensus that *HFE* gene mutations impose an elevated risk for the development of PCT. It has not yet been established whether other genes (possible modifier loci) involved in iron metabolism have a potential role in the pathogenesis of PCT. Abberations present in one or more of these genes or modifier loci, along with an inherited *UROD* mutation or acquired defect in *UROD* enzyme activity, may explain the clinical manifestations of this disease and elucidate the phenotype-genotype correlation of PCT. Furthermore, the positive identification of modifier genes involved in iron metabolism could represent an important development in the elucidation of the relationship between iron overload and PCT.

1.5 GENES INVOLVED IN IRON HOMEOSTASIS

Numerous genes have been implicated in iron homeostasis, several of which have only recently been identified. These genes include, amongst others, the haem oxygenase 1 (*HMOX1*) gene (Tenhunen *et al.* 1969); the transferrin receptor 1 (*TfR1*) gene (Enns *et al.* 1982; Goodfellow *et al.* 1982); the ferritin light chain (*FTL*) gene (Caskey *et al.* 1983; McGill *et al.* 1984); the ferritin heavy chain 1 (*FTH1*) gene (McGill *et al.* 1984); the lactotransferrin [*LTF*, also known as lactoferrin] gene (Naylor *et al.* 1987) the high iron (*HFE*) gene (Feder *et al.* 1996); the solute carrier family 11 (proton-coupled divalent metal iron transporter) member 2 gene [*SLC11A2*, also known as the natural resistance-associated macrophage protein 2 (*NRAMP2*) gene or divalent metal transporter 1 (*DMT1*) gene] (Gruenheid *et al.* 1997; Gunshin *et al.* 1997); the transferrin receptor 2 (*TfR2*) gene (Kawabata *et al.* 1999); the hephaestin (*HEPH*) gene (Vulpe *et al.* 1999), Kaplan and Kushner 2000, Anderson *et al.* 2002, Petrak and Vyrol 2005), the solute carrier family 40 (proton-coupled

divalent metal ion transporter) member 1 [*SLC40A1*, also known as the solute carrier family 11 (proton-coupled divalent metal ion transporter) member 3 (*SLC11A3*) gene, ferroportin 1 (*FPN1*) gene, iron regulated transporter 1 (*IREG1*) gene or metal transporter 1 (*MTP1*) gene (Abboud and Haile 2000, Donovan *et al.* 2000, McKie *et al.* 2000), the hepcidin antimicrobial peptide gene [*HAMP*, also known as liver expressed antimicrobial peptide (*LEAP1*) gene or hepcidin (*HEPC*) gene] (Krause *et al.* 2000, Nicolas *et al.* 2001, Park *et al.* 2001, Pigeon *et al.* 2001), the cytochrome b reductase 1 gene [*CYBRD1*, also known as the duodenal cytochrome b (*DCYTB*) gene] (McKie *et al.* 2001), the ceruloplasmin (*CP*) gene (Royle *et al.* 1987, Cairo *et al.* 2001) and the haemojuvelin (*HJV*) gene (Papanikolaou *et al.* 2004). It is suggested that one or more of these genes interact with one another to ensure tight regulation of iron metabolism (Nemeth *et al.* 2004). This current investigation focused on four genes, i.e. *CP*, *CYBRD1*, *HAMP* and *SLC40A1*. These genes are discussed in detail below.

1.5.1 CERULOPLASMIN (*CP*)

Ceruloplasmin (*CP*, OMIM *117700), also known as ferroxidase or iron (II):oxygen oxidoreductase, was mapped to chromosome 3q23-24 (Royle *et al.* 1987). The *CP* gene is 3 kb long and encompasses 19 exons. (Daimon *et al.* 1995). This α 2- glycoprotein protein consists of a single polypeptide of 1046 amino acids (Takahashi *et al.* 1984).

CP is synthesised in the liver and is the principal protein involved in copper transport (Frieden 1981) and plays a significant role in iron regulation. *CP* possesses ferroxidase activity responsible for the peroxidation of ferrous iron to ferric iron (Frieden and Hsieh 1976). This ferroxidase activity is important for the transferrin loading of iron (Roeser *et al.* 1970) and for iron uptake from non-transferrin sources (Harris *et al.* 1999). Evidence has been provided that *CP* aids the mobilisation and release of iron from storage cells (mainly hepatocytes) into the blood plasma, conceivably in

concert with *SLC40A1* (Roeser *et al.* 1970; Frieden and Hsieh 1976; Mukhopadhyay *et al.* 1998; Roy and Andrews 2001).

A murine knockout model of aceruloplasminemia (*CP*^{-/-}), demonstrated that iron parameters were normal at birth, but iron loading occurred from an early age in the hepatocytes and reticuloendothelial cells (Harris *et al.* 1999). The iron accumulation in these mice was shown to be as a result of impaired iron compartmentalisation and iron efflux and not due to defective iron absorption or iron turnover.

1.5.2 CYTOCHROME B REDUCTASE 1 (*CYBRD1*)

Cytochrome b reductase 1 (*CYBRD1*, *OMIM 605745) was mapped to the chromosomal region 2q31 by The international radiation hybrid mapping consortium (McKie *et al.* 2001). This 35.6 kb gene consists of five exons and codes for a 4254 bp long mRNA. This gene encodes a plasma membrane di-heme 286 amino acid protein, which shares approximately 50% homology with cytochrome b561, also a membrane reductase, but one that is apparently not involved in iron homeostasis (McKie *et al.* 2001; Gunshin *et al.* 2005). Cytochrome b651 possesses binding sites for two substrates, ascorbic acid and semidehydroascorbic acid and these binding sites are similar or partially conserved in *CYBRD1*, suggesting that these substrates may also bind to and react with *CYBRD1* (McKie *et al.* 2001). Three transcripts of different sizes (1, 4, and 5 kb) have been detected by northern Blot analysis.

CYBRD1 is expressed in the brush border membrane of duodenal enterocytes, where it is responsible for catalysing the reduction of ferric iron to ferrous iron during intestinal iron absorption. The duodenum is the fundamental active site of iron absorption. Studies performed in homozygous *hpx* (without circulating transferrin) mice revealed that *CYBRD1* mRNA and protein

levels were up-regulated in the duodenal mucosa, indicating that chronic anaemia up-regulated *CYBRD1* levels. In the same way, iron deficient mice, also showed elevated mRNA and protein expression levels. (McKie *et al.* 2001). Thus, *CYBRD1* expression appears to be down-regulated during states of iron overload. However, more recently mRNA expression of *CYBRD1* was shown to be elevated in mice with haemochromatosis, (Muckenthaler *et al.* 2003; Hermann *et al.* 2004) Hypoxia, iron deficiency and hypotransferrinemia are all modulators of *CYBRD1* expression.

1.5.3 HEPCIDIN ANTIMICROBIAL PEPTIDE (*HAMP*)

HAMP (OMIM *606464) is an antimicrobial peptide synthesised by the liver in response to inflammation and iron overloading (Park *et al.* 2001). The cDNA encoding *HAMP* was isolated by Krause *et al.* (2000). This 2.5 kb gene has been localised to chromosome region 19q13 and contains only three exons (Krause *et al.* 2000; Park *et al.* 2001; Pigeon *et al.* 2001). *HAMP* codes for an 84 amino acid propeptide which is subject to enzymatic cleavage to produce three mature peptides of 20, 22 and 25 amino acids (Park *et al.* 2001).

HAMP maintains plasma iron levels by regulating the mobility of iron from hepatocytes and the absorption of dietary iron from the GIT, and controlling the release of iron from macrophages (Nemeth *et al.* 2004). Therefore, under conditions of iron overload, *HAMP* mediates the cessation of iron absorption from enterocytes, iron recycling in macrophages and iron storage in hepatocytes (Rossi 2005). In response to hypoxia, anaemia and in HH patients, *HAMP* expression is down-regulated causing iron overload in hepatocytes, increased iron export by ferroportin and increased enterocytic iron absorption. Complete inhibition of *HAMP* expression was demonstrated in mice that exhibited iron overload, due to the targeted disruption of the upstream regulatory factor (*USF2*) gene (Nicolas *et al.* 2001). The iron overload seen in these mice resembled that observed in human HH patients and in *HFE* knock-out mice (Zhou *et al.* 1998; Bahram *et al.* 1999; Levy *et al.* 1999).

Studies performed by Nemeth *et al.* (2004) concluded that *HAMP* interacts on a molecular level with *SLC40A1* through a process involving internalisation and degradation. Mutations in *HAMP* have been associated with juvenile haemochromatosis (JH) (Roetto *et al.* 2003). Disruption of *HAMP* regulation causes iron overload and subsequent cellular toxicity, cirrhosis and endocrine failure in HH patients.

1.5.4 SOLUTE CARRIER FAMILY 40 MEMBER 1 (*SLC40A1*)

The *SLC40A1* gene (OMIM *604653) was first identified in the zebrafish (*danio rerio*) and subsequently in humans as a putative iron exporter (Donovan *et al.* 2000; McKie *et al.* 2000) and localised to chromosome 2q31 (Haile 2000). This gene is 20 kb long, consists of eight exons (Njajou *et al.* 2001) and encodes a 571 amino acid long transmembrane protein (McKie *et al.* 2000), that is most abundant in the reticuloendothelial macrophages of the bone marrow, liver, spleen, the placenta and enterocytes (Donovan *et al.* 2000).

The 5' UTR mRNA of *SLC40A1* contains a conserved sequence, that forms a hairpin loop characteristic of a functional IRE. It is speculated that the IRE(s) bind and/or release IRP(s) to alter the expression of *SLC40A1* (McKie *et al.* 2000). *SLC40A1* expression is regulated by several factors including, changes in body iron status, hypoxia, inflammation and erythropoiesis (Abboud and Haile 2000; Yang *et al.* 2002).

HH patients with *SLC40A1* mutations have larger reticuloendothelial iron stores compared to other patients with other forms of HH. Reticuloendothelial iron overload also occurs in patients with ferroportin disease. Partial or complete loss of function mutations in *SLC40A1* have been predicted to cause abnormal iron regulation and consequently iron overload, as well as retarded or non-existent reticuloendothelial iron release (Montosi *et al.* 2001; Cazzola *et al.* 2002).

1.6 REGULATION OF GENE EXPRESSION

Upon completion of the human genome project, man's entire DNA sequence became available to the scientific community (http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml; Collins et al. 2003). This enormous international undertaking spanning several years serves as a crucial foundation for molecular biologists striving to unravel the complexity of proteomics, transcriptomics, metabolomics and the regulation of eukaryotic gene expression.

The regulation of gene expression occurs at several levels including transcription (initiation, elongation and termination), splicing, translation and protein degradation. It is a complicated process, brought about by the intricate interaction of genes and their protein products in interconnecting molecular networks, including cells, various tissues and organs and the entire organism (Kel *et al.* 2005). In order to grasp the transcriptional control of gene expression, which is the predominant level of gene regulatory control, a great deal of research has concentrated on the functional or regulatory elements that may play a role in this process. The promoter region of a gene constitutes one such regulatory area, as it is able to regulate gene expression, most notably at the level of transcription (Pesole *et al.* 2000; Pesole *et al.* 2001). Promoter regions are non-coding sequences and are usually located upstream (5') of the first translation site. These regions contain the basic machinery necessary for transcription, and in concert with several transcription factors (TFs) can both initiate and regulate gene expression. Most eukaryotic promoter regions contain a preinitiation complex (PIC). This complex, which includes the general transcription factors (GTF) such as (TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH) and RNA polymerase II, recognises the core promoter region and is responsible for transcription initiation.

1.6.1 TRANSCRIPTION FACTORS

A TF can be defined as a protein that regulates transcription by interaction with DNA or with a protein that can be assembled into a sequence-specific DNA-protein complex (Wingender 1997). The specific sequences or motifs to which TFs bind are known as transcription factor binding sites (TFBSs) (van Deursen *et al.* 2007). TF motifs are usually between 5 and 20 bp in length and in most cases are degenerate (Burden and Weng 2005, Kel *et al.* 2005). It is presumed that the degeneracy has been selected for during evolution to allow different levels of activity within different promoters (Stormo 2000).

TFBSs also function to bind TF in a particular orientation, especially relative to other components of protein complexes (Kel *et al.* 2005). TFs can influence and regulate transcription mechanisms in the following ways. Most TFs are responsible for enhancing the formation of the PIC at the TATA-box or Initiator (Inr) element *via* the interaction of a *trans*-activating domain with the basal transcription complex. They may also act as auxiliary, repressive or altering factors. As auxiliary factors they manipulate DNA conformation to ensure optimal activity of other TFs. As repressive factors they act indirectly to disrupt a TF ensemble or directly, to prevent binding to an inhibiting domain. As altering factors they alter chromosomal conformation, enabling easier access for the RNA polymerases. TFs may be composed of modules or present in a single polypeptide chain. Depending on the specific function of a TF, its modules may contain one or more of the following domains: DNA-binding, *trans*-activating or *trans*-repressing, ligand binding, oligomerisation (enabling TFs to bind as dimers) or modulating (where modifying enzymes may bind) domains (Yuh and Davidson 1996; Kel *et al.* 2005).

The entire promoter region (the promoter sequence, all transcription factors and/or regulatory proteins) enables the tissue specific and biochemically relevant regulation of gene expression at the

correct developmental stage in each cell, tissue and organ (Tjian 1995). Genetic variation in the promoter region may disrupt or create a TFBS, thereby altering the regulation of gene expression. *In silico* analysis combines computationally generated information with biologically relevant information such as gene signal transduction, metabolic and physiological pathways, thus providing a key starting point in predicting putative TFBS. While the methodology may accurately anticipate the presence of a TFBS, it cannot foretell the *in vivo* reality, where intricate protein-DNA and/or protein-protein interactions are specific in different cell or tissue environments. Instead it merely provides relatively sound information on which molecular biologists can base follow-up functional experiments.

1.6.1.1 Transcription factors and iron metabolism

Several TFs have a characterised role in human hepatocytes (Cereghini 1996; Zaret 2002; Costa *et al.* 2003, Odom *et al.* 2006). Common TFs that play a role in iron metabolism and/or in the transcriptional regulation of genes expressed in the liver include hepatic nuclear factors (HNFs) (Kuo *et al.* 1972; Duncan *et al.* 1998), CCAAT enhancer binding proteins (C/EBPs) (Johnson *et al.* 1987), cyclic-AMP responsive enhancer binding protein (CREB) (Montminy *et al.* 2004), and upstream regulatory factor 1 and 2 (USF1, USF2) (Pajukanta *et al.* 2004).

The hepatic nuclear factors are classified as HNF1 α [also known as transcription factor 1 (TCF1)] (Ktistaki and Taliandis 1997; Tronche *et al.* 1997), HNF3 β [also referred to as forkhead box A1 (FOXA1)] (Pani *et al.* 1992; Lee *et al.* 2005) and HNF4 α [also known as transcription factor 14 (TCF14)] (Costa *et al.* 1989). They are members of the forkhead class of DNA-binding proteins. HNF1 α is predominantly expressed in the liver, where it regulates the expression of liver specific genes at several developmental stages, especially during liver organogenesis (Courtois *et al.* 1987). HNF1 α TFBSs in the liver-enriched genes tend to associate within clusters of other TFBSs and

these clusters are enriched in the promoter regions of such genes (Tronche *et al.* 1997). HNF3 β is responsible for the coordinated control of the expression of several hepatic genes (Pani *et al.* 1992) and is critical for the initiation of liver development (Lee *et al.* 2005). Furthermore, HNF3 β plays a role in ketogenesis and lipid metabolism (Wolfrum *et al.* 2003, 2004). HNF4 α , expressed in high levels in the liver, intestine and kidney (Sladek *et al.* 1990), is also responsible for regulating the expression of many hepatic genes, especially those involved in insulin homeostasis, during the differentiation of the liver ectoderm (Li *et al.* 2000; Parviz *et al.* 2003, Tirona *et al.* 2003). The HNFs may regulate the expression of their family members. For instance, HNF3 β up-regulates the expression of HNF1 α and HNF4 α regulates the expression of HNF1 α (Duncan *et al.* 1998).

The C/EBP family of transcription factors are not part of the PIC, but frequently interact with the PIC in order to activate the transcription of certain promoters (Cao *et al.* 1991, Fickett and Hatzigeorgiou 1997). This family contains bZIP domains, which mediates the dimerization required for DNA binding (Cereghini *et al.* 1996). C/EBP α has a palindromic consensus sequence (5' ATTGCGCAAT3') and is involved in hepatocytic gene expression and tissue differentiation (Umek *et al.* 1990; Niehof *et al.* 2001), as well as the maintenance of the differentiated nonproliferative state (Cereghini *et al.* 1996). CREB is a TF that responds to a soluble messenger called cAMP. CREB usually acts in concert with several other transcription factors to regulate the expression of numerous genes. One of the metabolic processes that CREB is known to be active in, is hepatic lipid metabolism (Herzig *et al.* 2003). The upstream stimulatory factors 1 and its isoform 2 (USF1, USF2) belong to the Myc protooncogene (Myc/cMyc) family of DNA-binding proteins (Gregor *et al.* 1990, Viollet *et al.* 1996). Various studies have implicated USF in hepatic gene regulation (Vallet *et al.* 1998, Casado *et al.* 1999).

1.7 GENETIC VARIATION

Ninety percent of variation at the DNA level is attributed to single nucleotide substitutions (Collins *et al.* 1998; Kruglyak 1999). Single nucleotide polymorphisms (SNPs) are described as single nucleotide positions in genomic DNA where alternate alleles may exist in certain normal population groups (Brookes 1999). SNPs occur in the human genome at least once in every 1000 bp (Lai *et al.* 1998; Taillon-Miller *et al.* 1998). Not all SNPs are functionally important, but several studies have provided evidence that single nucleotide substitutions occurring in fundamental genomic areas such as regulatory regions may impact on phenotype (Ober *et al.* 2003; Jiang *et al.* 2005; Masotti *et al.* 2005; Prior *et al.* 2006). Whilst a single SNP is not normally considered as disease-causing, these polymorphisms may be applied as indicators of relative risk for genetic disease, often in response to environmental fluctuations (Guo and Jamison 2005). Polymorphisms in genes that may increase susceptibility to diseases and disorders include *RET* (Hirschsprungs disease) (Puffenburger *et al.* 1994), *HLA* (autoimmunity and infection) (Lechler and Warrens 2005), *PTPN22* (rheumatoid arthritis and type 1 diabetes) (Begovich *et al.* 2004; Bottini *et al.* 2004), *NOD2* (inflammatory bowel disease) (Ogura *et al.* 2001; Hugot *et al.* 2001) and *APOE4* (Alzheimer's disease) (Strittmatter and Roses 1996). For this reason, SNPs have become a linchpin in biomedical research; eventually, a better insight into genomic variation may assist in the treatment of genetic disease.

Specific alleles of SNPs located in close proximity to one another are more often than not inherited together. These patterns of SNPs called haplotypes, are known to alter genomic susceptibility to disease, thereby influencing the relative risk of common diseases (Brookes 1999). Linkage disequilibrium (LD) describes the close association between two SNPs and arises due to the shared ancestry of a particular chromosome. The relationship between common polymorphisms and the haplotypes in which they arose, has facilitated genetic association studies. In this manner haplotypes

putatively associated with disease have been determined and subsequently the causative variant within the haplotype could be established. A classic example of a disease related-variant that was inferred by haplotype is the C282Y *HFE* gene mutation in the human leukocyte antigen (HLA) region associated with HH (Simon *et al.* 1987). This ancestral haplotype is very common in Northwestern European countries. It should be emphasised that the inheritance of a specific haplotype does not necessarily cause disease; it may increase genomic susceptibility to disease.

1.7.1 GENETIC VARIATION IN THE PROMOTER REGION

Single nucleotide variation is not evenly distributed in the promoter region between sites -2000 and -1; a higher frequency is found closer to the transcription initiation site (Guo and Jamison 2005). Overall the frequencies of transition and transversion substitutions are similar in the promoter vs the entire genome, except for an increased incidence of higher G/C substitutions in the promoter area. This finding is in accordance with the GC content of promoter regions. Furthermore, transition type substitutions account for approximately 66.67% of all substitutions in gene promoter regions. The proportion of transition substitutions identified in the current study (77%) was higher than the published average frequency (Guo and Jamison 2005).

1.7.1.1 Regulatory SNPs (rSNPs)

SNPs present in promoter regions that alter gene expression are termed regulatory SNPs (rSNPs) (Montgomery *et al.* 2007). Several rSNPs have been implicated in the aetiology of certain diseases (Bulyk *et al.* 2003; Montgomery *et al.* 2007). Examples include cancer (Miao *et al.* 2003; Bond *et al.* 2004), perinatal HIV-1 transmission (Kostrikis *et al.* 1999) and depression (Caspi *et al.* 2003). rSNPs can alter the function of a specific site (e.g. TFBS) essential for normal regulation, cause a change in the protein binding affinity at such a site, or create a site that is not required or generally associated with regulation.

1.8 HYPOTHESIS AND OBJECTIVES

PCT patients often display iron overload, the cause of which is not understood. The hypothesis to be tested was whether genes involved in iron metabolism are implicated in the aetiology of PCT. Given the fact that the regulatory regions of genes are known to influence gene expression, this preliminary study will investigate the 5' non-coding regulatory region of four genes (*CP*, *CYBRDI*, *HAMP* and *SLC40A1*). The aim of this study was to identify putative candidate genetic variants or modifier loci associated with PCT.

The objectives of this study include:

- i) The mutation analysis of the 5' non-coding regulatory promoter region (known to influence gene expression) of four genes (*CP*, *CYBRDI*, *HAMP* and *SLC40A1*) by:
 - Polymerase chain reaction (PCR) amplification of the promoter regions
 - Heteroduplex single stranded conformation polymorphism (HEX-SSCP) and restriction fragment length polymorphism (RFLP) analysis of the amplified products to detect novel and/or known variants
 - Semi-automated DNA sequencing to identify novel and/or known variants
- ii) Statistical and bioinformatic analysis of the identified variants to determine:
 - Allele and genotype frequencies of the patient and population-matched control groups
 - Association between patients and population-matched controls by means of Fisher's exact test and/or chi-squared (χ^2) analysis
 - Potential putative TFBSs in the promoter region that could explain the regulation of these genes

**CHAPTER 2:
MATERIALS AND
METHODS**



MATERIALS

2.1 STUDY COHORT

A total of 74 blood samples were collected from unrelated patients diagnosed with PCT. Due to the variable phenotypic expression of this disease, and to ensure conformity in diagnostic criteria, the patients in this study cohort were all diagnosed by the same clinician at Groote Schuur Hospital, Cape Town. The patient cohort consisted of 39 males and 35 females from the diverse South African population groups. These groups included, 15 Black (Xhosa) [eight males (mean age of onset: 49.88 ± 21 years) and seven females (mean age of onset: 37.38 ± 17 years)], 30 Caucasian [13 males (mean age of onset: 46.64 ± 18 years) and 17 females (mean age of onset: 51.21 ± 16 years)], and 29 Coloured [18 males (mean age of onset: 43.33 ± 15 years) and 11 females (mean age of onset: 39.70 ± 17 years)] patients. In this study, 'Black' refers to South Africans from central African descent, 'Caucasian' refers to individuals from European descent and 'Coloured' includes patients from San, Khoi, African Negro, Madagascan, Javanese and European origin (Nurse *et al.* 1985). Written informed consent was obtained from all participants. In addition, DNA samples from 132 population-matched-control individuals were obtained from Tygerberg Hospital, Cape Town. This group included 30 Black individuals (two males and 28 females), 60 Caucasians (25 males and 35 females), and 42 Coloured (14 males and 28 females).

Individual informed consent was not obtained from all participants. A written exclusion was obtained from the Research Ethics committee of the University of Cape Town on the basis that all subjects were clinically-affected persons who had already been subjected to *HFE* gene mutation analysis as part of their standard workup for PCT, and that this investigation could be viewed as an extension of their clinical investigation. Furthermore, all results were anonymised with the key retained only by the clinician in charge of the patients. This project has also been approved by the Research and Ethics/Biosafety committee of Stellenbosch University (Project Ethics number is

NO4/08/123).

DETAILED EXPERIMENTAL PROCEDURES

2.2 DNA ISOLATION AND PURIFICATION

2.2.1 TOTAL GENOMIC DNA (gDNA) ISOLATION FROM WHOLE BLOOD

An adaptation of the Miller *et al.* (1988) protocol was used to perform DNA extractions from whole blood. Whole blood (5 ml) was collected in an ethylene diamine tetra-acetic acid (EDTA) ($C_{10}H_{16}N_2O_8$) (Appendix 1) coated glass tube and stored at $-20^{\circ}C$. The thawed blood was transferred to a 50 ml polypropylene Falcon tube (Lasec). Thirty ml cold cell lysis buffer [155 mM ammonium chloride (NH_4Cl), 10 mM potassium hydrogen carbonate ($KHCO_3$), 0.1 mM EDTA, pH 7.4] was added and the tube was placed on ice for 30 min. The sample was briefly mixed by inversion at 10 min intervals to ensure lysis of the red blood cells and centrifuged at 1 500 rpm (Avanti™ 30 or GS-15R, Beckman or Hermle Z 200 A), at $4^{\circ}C$ for 10 min.

The supernatant was removed and the pellet was resuspended in 10 ml cold phosphate buffer saline (PBS) [27 mM potassium chloride (KCl), 137 mM sodium chloride (NaCl), 8 mM di-sodium hydrogen orthophosphate anhydrous (Na_2HPO_4), 1.5 mM potassium di-hydrogen orthophosphate (KH_2PO_4), pH 7.3] solution. The solution was centrifuged at 1 500 rpm, at $4^{\circ}C$ for 10 min. The supernatant was removed and the following solutions were added to the pellet: 3 ml nuclear lysis buffer [10 mM Tris (hydroxymethyl) aminomethane (Tris-HCl) ($(CH_2OH)_3CNH_2-Cl$), 400 mM NaCl, 2 mM EDTA, pH 8.2], 1.5 mg/ml Proteinase K (Roche Diagnostics) and 10% (w/v) sodium dodecyl sulphate (SDS). The tube was inverted to mix the solution and incubated in a water bath at $55^{\circ}C$ overnight.

Following overnight incubation, 1 ml 6 M NaCl was added, and the tube was shaken vigorously for

one min. The solution was centrifuged at 3 500 rpm at room temperature for 30 min. The supernatant was then transferred to a clean Falcon tube, shaken for 15 sec and centrifuged at 2 500 rpm for 15 min. The supernatant (without any foam or pellet) was transferred to a clean Falcon tube. Two volumes of ice-cold \pm 99.9% (v/v) ethanol (EtOH) was added to precipitate the DNA.

The precipitated DNA was removed and placed in a 1.5 ml tube (Eppendorf) containing 70% (v/v) EtOH. The solution was centrifuged at 14 000 rpm for 10 min (Eppendorf Centrifuge 4517C). The supernatant was removed and the pellet allowed to air dry. The dry pellet was dissolved in 200-800 μ l of double distilled water (ddH₂O) (SABAX, Adock Ingram) depending on the size of the pellet. The pellet was resuspended overnight on a shaker (Eppendorf Mixer 5432) at room temperature. DNA quantity and quality was calculated by means of spectrophotometric measurements based on the A_{260} and A_{260}/A_{280} ratios respectively (Nanodrop[®] ND-1000 Spectrophotometer, Nanodrop Technologies). The DNA was diluted to a concentration of 50 ng/ μ l and stored at 4°C.

2.2.2 gDNA PURIFICATION

gDNA was purified as described by Maniatis *et al.* 1989. Briefly, an equal volume of cold phenol/chloroform [24:1] was added to 100 μ l DNA in a 1.5 ml tube (Eppendorf). The solution was vortexed (Vortex Genie 2, Lasec) for 5 sec and centrifuged at 13 200 rpm (Eppendorf spectrafuge 5415D, Labnet) for 5 min. The supernatant was transferred to a fresh 1.5 ml tube and an equal volume of chloroform was added. The sample was vortexed for 5 sec and centrifuged at 13 200 rpm for 5 min. The supernatant was transferred to a fresh 1.5 ml tube, 5M sodium acetate (C₂H₃O₂Na) equal to one tenth of the volume of the DNA solution was added, and the tube mixed by inversion. Two volumes of \pm 99.9% (v/v) ice-cold EtOH was added and the vial incubated for 20 min at -20°C. The sample was centrifuged at 13 000 rpm for 1 hr at 4°C, the supernatant was removed, the pellet washed in 70% (v/v) ethanol and centrifuged at 13 000 rpm for 5 min. The supernatant was

removed carefully so as not to disturb the pellet. The pellet was allowed to air dry, suspended in ddH₂O and allowed to dissolve overnight on the shaker. The DNA concentration determinations, dilutions and storage were as described in Section 2.2.1.

2.3 POLYMERASE CHAIN REACTION (PCR) AMPLIFICATION

2.3.1 OLIGONUCLEOTIDE PRIMERS

PCR amplification of the promoter region (5'-UTR) of the ceruloplasmin (*CP*), cytochrome b reductase 1 (*CYBRD1*), hepcidin antimicrobial peptide (*HAMP*) and solute carrier family 40 (proton-coupled divalent metal ion transporter) member 1 (*SLC40A1*) genes was performed. Primer sets were designed to generate approximately 300 bp amplicons, each gene promoter region therefore having multiple primer sets. Primers sequences and size of the amplified products are shown in Table 2.1. The Primer3 software (Rozen and Skaletsky 2000) was used to design the primers. The primer design was such to allow for an overlap of at least 30 bp between amplicons. The melting temperature (T_m) of the primers, was calculated according to the formula $T_m = 2(nA + nT) + 4(nG + nC)$ (Thein and Wallace 1986).

Table 2.1 Oligonucleotide primer sequences, amplicon sizes and PCR conditions for the amplification of the promoter region of the *CP*, *CYBRD1*, *HAMP*, and *SLC40A1* genes.

◇Key to follow on pg 53

^o <i>CP</i> ^p ENSG00000047457 (ENSEMBL)							
Primer Name	Primer Sequence	Product size (bp)	T _m (°C)	T _{A1} (°C)	T _{A2} (°C)	MgCl ₂ (mM)	PCR Cycle
*CP1 F	5' CCTGTTAGGCTCTGCTAGTT 3'	267	60	55		1.5	2
*CP1 R	5' CTTATGGGACCACCATCACA 3'		60				
#CP2 F	5'AGAAA TAGTCATGCACCAC 3'	311	54	55	50	1.5	3
#CP2 R	5' TGCATTGTTAGGCTATTTTG 3'		54				
*CP3 F	5'CAGGTTTGTAGCCTAGGAGC 3'	282	62	56		1.5	2
*CP3 R	5' ATGCTCCCTTTGTTCTCTG 3'		60				
#CP4 F	5' CTCAGAACGTATCCCTGT CAT 3'	275	62	57		1.5	2
#CP4 R	5' TTGCCAGGCTTCTCTGACTG 3'		62				
*CP5 F	5' CAGAGGAACAAGGGAGCAT 3'	300	60	56		1.5	2
*CP5 R	5' GGAGCCTGAGAAGAAATGAAG 3'		62				
*CP6 F	5' CAGTCAGAGAAGCCTGGC 3'	366	58	52		1.5	2
*CP6 R	5' GACTTACGTGTCAACAG 3'		56				

Table 2.1 continued

^o <i>CYBRDI</i> ^p ENSG00000071967 (ENSEMBL)						
Primer Name	Primer Sequence	Product size (bp)	T _m (°C)	T _{A1} (°C)	MgCl ₂ (mM)	PCR Cycle
*BP1 F	5' CACACTAAACGGCAAGTCCA 3'	255	60	57	1.5	2
*BP1 R	5' GTCCCCATGAGGTGTCTCAC 3'		64			
*BP2 F	5' CAGTGCAAGTGGTGGATCAT 3'	310	62			
*BP2 R	5' GTGAAACCCTGTCTCTACTG 3'		60			
*BP3 F	5' GCTCAAGTGA TCCTCCTTCC 3'	275	62	57	1.0	2
*BP3 R	5' GCCTCCTTGCAAGGCTATAAC 3'		62			
*BP4 F	5' CTGCGCCTGGCCTTTACTGT 3'	300	64	55	1.5	2
*BP4 R	5' GAGAACATGTTACATGCATG 3'		56			
*BP5 F	5' AGGTTACAGGAGGAGCTATG 3'	294	60	56	1.5	2
*BP5 R	5' TGCAACAGCCAGAACTACAG 3'		62			
*BP6 F	5' CTGTACACTGGCCAGAACCG 3'	309	64	58	1.5	2
*BP6 R	5' AGGCGGACAGAACAGACTCT 3'		62			
*BP7 F	5' GTAGGAGTGTGTGACTTAGG 3'	312	60	56	1.5	2
*BP7 R	5' GCTGGAGAAATGGCTGTCTC 3'		62			
*BP8 F	5' GGTCCATTCAATCAGTGAGG 3'	246	62	58	1.5	2
*BP8 R	5' CACTTCCCGTGTCTCGTCTC 3'		64			
*BP9 F	5' CTACCCAACGGATCCCTCTC 3'	276	64	57	1.5	2
*BP9 R	5' CTGGCCCCAACTCAGAAATG 3'		62			
*BP10 F	5' AAGCCCTCTGCGAGCTTGG 3'	371	60	55	1.5	2
*BP10 R	5' GCGAAGATCACCGACAGGAA 3'		60			

Table 2.1 continued

^a <i>HAMP</i> ^b ENSG00000105697 (ENSEMBL)							
Primer Name	Primer Sequence	Product size (bp)	T _m (°C)	T _{A1} (°C)	T _{A2} (°C)	MgCl ₂ (mM)	PCR Cycle
*PP1 F	5' GACGGGGAGGAGGTTCTAA 3'	274	56	56		1.5	2
*PP1 R	5' AGCCTGGGTGACAGATGAG 3'		56				
#PP2 F	5' CATCGGACTGTAGATGTTAGC 3'	256	60	57		1.5	2
#PP2 R	5' TCAAGACTAGCCTGGGAAC 3'		62				
*PP3 F	5' CACGCCTGGCTAAATTTGTT 3'	326	58	55		2.0	2
*PP3 R	5' CACCACACGTGCATAGGTTTC 3'		62				
#PP4 F	5' TCAAGGGTCTGACACTGGG 3'	312	62	59	54	2.0	2
#PP4 R	5' CCATCACGATGTCA TTCTGC 3'		60				
*PP5 F	5' AAGTGAGTGGAGGAGAGCG 3'	285	62	55		2.0	2
*PP5 R	5' CTTTGCTCTGTCTCATTTCC 3'		58				
*PP6 F	5' CTGAGGGTGACACAACCCT 3'	291	60	56		1.0	2
*PP6 R	5' AGAGCCA CTGGTCAGGCTG 3'		62				

Table 2.1 continued

° <i>SLC40A1</i> ^b ENSG00000138449 (ENSEMBL)							
Primer Name	Primer Sequence	Product size (bp)	T _m (°C)	T _{A1} (°C)	T _{A2} (°C)	MgCl ₂ (mM)	PCR Cycle
*SP1 F	5' GTAGACCTTTGGGGCTCCTG 3'	288	64	60		1.0	1
*SP1 R	5' TGGAGGGTGAAGGTGAATGAC 3'		62				
*SP2 F	5' GCCACAGCAGGATTAACG 3'	289	60	55		1.5	2
*SP2 R	5' AATCCGTCCTGGAGACAGAA 3'		60				
*SP3 F	5' GGGAGAAGGAATGATGGTGA 3'	302	60	57.5		1.0	1
*SP3 R	5' ATGCCACAGAGGCCGCTTTC 3'		60				
*SP4 F	5' GAAGCCCTGCTATGCAGTCC 3'	255	64	60		1.0	1
*SP4 R	5' GTCAGGTGCTGGAAGAAAGC 3'		64				
*SP5 F	5' CTGGAGCTTTGCACTGCGAC 3'	357	64	58	54	1.5	3
#SP5 R	5' AGACGAGCTCCCGTCAACCT 3'		60				
SP6 F	5' TGGAACGCGTCGAGGCGAA 3'	369	62				
SP6 R	5' AGCTAACACTGTAGCTGAAG 3'		58				
*SP7 F	5' CTGAGCCTCCCAAACCGCTT 3'	315	64	60		1.0	1
*SP7 R	5' TCACCACAGCATCCTCTCTG 3'		62				

° gene name; ^b accession number for reference sequence; * primers manufactured by Inqaba Biotech; # primers manufactured by Intergrated DNA Technologies (IDT)

Abbreviations: 5', 5'-prime; 3', 3'-prime; °C, degrees Celcius; A, adenine; bp; base pairs; BP, *CYBRDI* promoter; C, Cytosine, *CP*; Ceruloplasmin gene; CP, Ceruloplasmin promoter; *CYBRDI*, cytochrome b reductase 1 gene; F, forward primer; G, guanine; *HAMP*, hepcidin antimicrobial peptide gene; MgCl₂, magnesium chloride; mM, millimolar; PCR, polymerase chain reaction; *PP*, *HAMP* promoter; R, reverse primer; *SLC40A1*, solute carrier family 40 member A1 gene; SP, *SLC40A1* promoter; T, thymine; T_{A1}, annealing temperature 1; T_{A2}, annealing temperature 2; T_m, melting temperature.

2.3.2 PCR AMPLIFICATION CONDITIONS

Amplification conditions were optimised individually for each PCR primer set as detailed in Table 2.1. Each 25 μ l reaction mixture consisted of 50 ng template DNA, 1 x ammonium sulphate $[(\text{NH}_4)_2\text{SO}_4]$ buffer (Fermentas), 10 pmol of each primer, and 0.5 U *Taq* polymerase (Fermentas). All samples received 0.1 mM of each 2'-deoxynucleotide triphosphate (dNTP) (dCTP, dATP, dGTP, dTTP) (Fermentas), except the *SLC40A1* amplicons SP1, SP4 and SP7, where 0.2 mM of each dNTP was used. MgCl_2 (Fermentas) concentration ranged from 1-2 mM, as outlined in Table 2.1.

The PCR amplification was performed using a GeneAmp® PCR cycler 2700 (Applied Biosystems). The amplicons were amplified using one of 3 PCR cycle options, as shown in Table 2.1 and explained below.

Cycle 1 included an initial denaturing step at 95°C for 5 min. This was followed by denaturation at 95°C for 1 min, annealing at the specified temperature (T_{A1} ; Table 2.1) for 2 min, for 35 cycles. Thereafter, elongation was performed at 72°C for 10 min.

Cycle 2 encompassed the following thermo-cycling conditions: initial denaturation at 95°C for 2 min, followed by 35 cycles of denaturing at 95°C for 30 sec, annealing at annealing temperature 1 (T_{A1} ; Table 2.1) for 45 sec and elongation at 72°C for 30 sec. Final elongation was at 72°C for 5 min.

Cycle 3 included an initial denaturing step at 95°C for 2 min. This was followed by a denaturing step at 95°C for 30 sec, an annealing step at the annealing temperature 1 (T_{A1} ; Table 2.1) for 2 min, and an elongation step at 72°C for 30 sec, for 10 cycles. The same cycle was repeated 30 times at

the decreased annealing temperature 2 (T_{A2} ; Table 2.1). Final extension was at 72°C for 5 min.

2.4 AGAROSE GEL ELECTROPHORESIS

All PCR products were electrophoresed on agarose gels to determine whether amplification occurred. A 2% (w/v) agarose gel [4 g agarose, 200 ml 1 x TBE (90 mM Tris-HCl, 90 mM H_3BO_3 and 0.1 mM EDTA, pH 8.0) and 0.01% (v/v) ethidium bromide (EtBr)] was prepared. The gel was allowed to set and transferred to a horizontal electrophoresis tank filled with 1 x TBE buffer. PCR product (5 μ l) was added to 5 μ l of cresol red loading buffer [0.02% (w/v) cresol red and 0.35% (w/v) sucrose], and loaded into the wells of the agarose gel. In addition, 5 μ l of a 100 bp O'GeneRuler™ ladder (Fermentas) size marker was loaded. The DNA samples were electrophoresed at 120 V for 1 hr. Visualisation of the PCR products was achieved by UV light transillumination and the image captured on the Multigenius Bio Imaging System (Syngene).

2.5 HETERODUPLEX SINGLE-STRANDED CONFORMATION

POLYMORPHISM (HEX-SSCP) ANALYSIS

Successfully amplified PCR products were subjected to (HEX-SSCP) analysis (Kotze *et al.* 1995) on the Hoefer SE 600 Basic slab 30 cm vertical gel apparatus. Polyacrylamide gels [12% (w/v) supplemented with urea] were used [12% (v/v) PAA (1% C of a 40% (w/w) stock (99 acrylamide (AA):1 N, N' methylenebisacrylamide (BAA), 7.5% (w/v) urea, 1.5 x TBE (135 mM Tris-HCl, 135 mM H_3BO_3 and 2 mM EDTA, pH 8.0), 0.1% (w/v) ammonium persulphate (APS) ($(NH_4)_2S_2O_8$) and 0.1% (v/v) N, N, N', N'- Tetramethylethylenediamine (TEMED)]. Gels were allowed to polymerise and the wells were cleaned with 1.5 x TBE buffer. The gels were assembled to the upper chambers and positioned in the electrophoresis tanks filled with 1 x TBE buffer before the former was filled with 1.5 x TBE buffer. The gel system was left to acclimatise at 4°C for at least 1 hr.

PCR samples (20 μ l) were mixed with 15 μ l bromophenol blue loading buffer [95% (v/v) de-ionised

formamide (H_2NCHO), 20 mM EDTA, 0.05% (w/v) bromophenol blue ($\text{C}_{19}\text{H}_{10}\text{Br}_4\text{O}_5\text{S}$), 0.05% (w/v) xylene cyanol ($\text{C}_{31}\text{H}_{28}\text{N}_2\text{Na}_4\text{O}_{13}\text{S}$), denatured at 95°C for 10 min and immediately placed on ice prior to loading the gel. An amount of 15 μl of the denatured product and dye mixture was loaded. Electrophoresis was performed at 350 V, at 4°C for 16 hr.

After completion of electrophoresis, the gels were removed from the tanks and dismantled. Gels were stained in 0.01% EtBr in 1 x TBE buffer for 10 min and destained in dH_2O for 3 min. Visualisation of the DNA banding patterns and capture of the image was as described in Section 2.4.

2.6 SEMI-AUTOMATED DNA SEQUENCING

2.6.1 PURIFICATION OF PCR PRODUCTS

Samples displaying aberrant banding patterns upon HEX-SSCP analysis were subjected to bi-directional semi-automated DNA sequencing once the PCR products had been purified using the GenElute™ PCR Clean-up Kit (Sigma). In brief, column preparation solution (500 μl) was added to a collection tube, containing a GenElute Miniprep Binding Column, and centrifuged at 11 400 rpm for 30 sec (Eppendorf spectrafuge 5415D, Labnet). The elute was discarded and a mixture comprising 5 volumes of binding solution and 1 volume of PCR product was added to the vial containing the binding column. This solution was centrifuged at 13 200 rpm for 1 min. Again, the elute was discarded. An amount of 500 μl of wash solution was applied to the binding column and the solution was centrifuged with the lid of the tube open at 13 200 rpm for 3 min. The binding column was then transferred to a new 2 ml collection tube, 50 μl of ddH_2O (SABAX, Adcock Ingram) was added and the sample incubated at room temperature for 1 min. Following a final centrifugation at 13 200 rpm for 1 min, the binding column was removed and the purified DNA stored at 4°C . The constituents of the Sigma GenElute™ PCR Clean-up Kit column preparation,

binding and wash solutions are not provided by the manufacturer.

2.6.2 CYCLE SEQUENCING

Purified PCR product (9.9 ng) to be sequenced was added to 1 µl termination ready reaction mix [(BigDye[®] Terminator v3.1 cycle sequencing kit (Applied Biosystems))] and 3.3 pmol primer. The primers used for sequencing were the same as those used for PCR amplification (Table 2.1). Cycle sequencing was performed on a GeneAmp[®] PCR 2700 thermocycler (Applied Biosystems). Initial denaturation was at 96°C for 10 sec. This was followed by 25 cycles of denaturing at 96°C for 10 sec, annealing at 55°C for 10 sec and elongation at 60°C for 4 min. Following cycle sequencing, the products were analysed by electrophoresis on an ABI Prism 3130X1 Genetic Analyzer automated sequencer (Applied Biosystems). The resulting electrophoregram were examined visually for variation. BioEdit Sequence Alignment Editor v 7.0.5.2 (Hall 1999) was applied to align the sequences, with the published reference sequence as indicated in Table 2.1.

2.7 RESTRICTION FRAGMENT LENGTH POLYMORPHISM (RFLP)

ANALYSIS

Semi-automated DNA sequence analysis of aberrant banding patterns seen on HEX-SCCP gels revealed four variants that were suitable for detection by restriction enzyme recognition sites. The population-matched controls were screened for these variants using RFLP analysis. These variants are depicted in Table 2.2. The specific nucleotide change, restriction enzyme used for variant detection and allele sizes are also indicated. For each variant, Allele *1 refers to the common allele as identified in the reference sequences (Ensembl <http://www.ensembl.org/>). Allele *2 refers to the nucleotide substitution as identified in this study.

Table 2.2 Four RFLPs detected in the promoter region of the *CYBRDI*, *HAMP* and *SLC40A1* genes. The specific base pair change, relevant restriction enzyme and its recognition site and digestion requirements are shown. Fragment sizes of common (Allele *1) and variant (Allele *2) alleles are also indicated.

Gene	Variant	Enzyme	Recognition Site (5' – 3')	NEB Buffer	Incubation Temperature (°C)	Allele *1 fragment(s) size (bp)	Allele *2 fragment(s) size (bp)
<i>CYBRDI</i>	-244G/A	<i>BsrDI</i>	GCAATGNN [~]	2	65	276	216 and 60
<i>HAMP</i>	-1010C/T	<i>RsaI</i>	GT [~] AC	1	37	192 and 82	274
<i>HAMP</i>	-582A/G	<i>BstUI</i>	CG [~] CG	2	60	326	245 and 81
<i>SLC40A1</i>	-1355C/G	<i>BfaI</i>	C [~] TAG	4	37	285	255 and 30

5', 5-prime; 3', 3-prime; °C, Degrees celcius; [~], restriction enzyme cutting site, *1, known allele *2, variant allele identified in study

Abbreviations: A, adenine bp, basepair; *BfaI*, *Bacteroides fragilis* I restriction enzyme; *BsrDI*, *Bacillus stearothermophilus* D70 restriction enzyme; *BstUI*, *Bacillus stearothermophilus* restriction enzyme UI; C, cytosine; CP, ceruloplasmin; *CYBRDI*, cytochrome b reductase 1; G, guanine, *HAMP*, hepcidin antimicrobial peptide; n, adenine or cytosine or guanine or thymine; NEB buffer, New England Biolab buffer; R, adenine or guanine; *RsaI*, *Rhodopseudomonas sphaeroides* I restriction enzyme; *SLC40A1*, solute carrier family 40 member 1; T, thymine.

The PCR product (10 µl) was used in a total volume of 20 µl containing appropriate restriction enzyme (2U) and 1 x NEB (New England Biolab) buffer (Table 2.2). Three types of NEB buffer were used depending on the specific restriction enzyme and as recommended by the manufacturer. NEB buffer 1 consists of 10 mM Bis-Tris-propane-HCl, 10 mM MgCl₂, 1mM dithiothreitol, (pH 7.0); NEB buffer 2 consists of 10 mM Tris-HCl, 50 mM NaCl, 10 mM MgCl₂, 1 mM dithiothreitol, (pH 7.9) and NEB buffer 4 consists of 20 mM Tris-acetate, 50 mM potassium acetate, 10 mM magnesium acetate and 1 mM dithiothreitol (pH 7.9). For the restriction enzyme digestion of the *CYBRDI* variant (-238G/A) bovine serum albumin (BSA) at a final concentration of 100 µg/ml was also added to the reaction. DNA samples were digested for 16 hr in a water bath at the relevant temperature indicated in Table 2.2. Electrophoresis conditions and UV visualisation of the digested

PCR products was as described in Section 2.4. Electrophoresis was performed on a 2% (w/v) agarose gel at 80 V for 90 min.

2.8 STATISTICAL ANALYSIS

Allele and genotype frequencies were estimated by allele counting. Statistically significant differences between patient and control groups were assessed by the Fisher's exact test and/or Chi-Squared (χ^2) analysis. The Fisher's exact test was applied when the total sample size (patients and controls) was less than 20, or when the total sample size was between 20 and 40 but with one of the values in the contingency table less than 5 (Cochran 1954). Probability values less than 0.05 ($p < 0.05$) were regarded as statistically significant. These analysis were performed using STATISTICA [StatSoft, Inc. (2007) STAT (data analysis software system), version 8] and 2x3 contingency tables (Microsoft excel software). In addition, patients and controls in each population group were divided according to gender (female patients *vs* female controls and male patients *vs* male controls) and the χ^2 and/or Fisher's exact test was applied. For each variant identified, the patient and control groups were tested for departure from Hardy-Weinberg equilibrium (HWE) by means of a Chi-squared (χ^2) goodness-of-fit test using the Tools for population genetics association studies (TFPGA) program version 1.3 (Miller 1997). Haplotype analysis and linkage disequilibrium analysis was performed on the polymorphic loci of each gene (Haploview 4.0, Barrett *et al.* 2005). Default parameters were applied to test for LD (the coefficient of association (D'): where $D'=1$ for perfect linkage disequilibrium; logarithm of the likelihood odds ratio, measure of confidence (LOD): $LOD > 3$; correlation coefficient between two loci (r^2): $r^2 > 0.8$ where $r^2 = 1$ for perfect linkage disequilibrium). The default block definition was applied (Gabriel *et al.* 2002) upon haplotype analysis.

2.9 BIOINFORMATIC ANALYSIS

In silico analysis was performed on each of the variants identified in the study. Several bioinformatic databases are available for *in silico* analysis, of which the following three were used in this study: JASPAR CORE (<http://jaspar.genereg.net/>) (Sandelin *et al.* 2004), TRANSFAC^{®7} (www.gene-regulation.com/pub/databases.html#transfac) (Wingender *et al.* 2001), and rVISTA (<http://genome.lbl.gov/vista/rvista/submit.shtml>) (Loots *et al.* 2002). JASPAR CORE is a non-redundant curated database based on position weight matrices (PWM). The analytical method employed in this study involved the comparison of a putative TFBS in the reference sequence (Genatlas, <http://www.genatlas.org/>) with the variant sequence. In this way, the creation or abolishment of a TFBS due to a particular nucleotide substitution/deletion/insertion can be anticipated. The relative profile score threshold was set to 80%. From the second database, TRANSFAC^{®7}, two programs were used, namely MATCH[™] (v1.0) (Kel *et al.* 2005) and ALIBABA2 (v.21). MATCH[™] employs the combination of mononucleotide weight matrices and pattern matching, ensuring a higher accuracy compared to using the methods independently. In this study, the parameters were set to 0.8 (80%) for matrix similarity and 0.85 (85%) for core similarity. Utilising the TRANSFAC^{®7} database, ALIBABA2 predicts putative TFBSs in unknown sequences. The default parameters were used upon analysis in ALIBABA2. The final program used was rVISTA, a database for comparative analysis of genomic sequences. The same parameters were used for this program as was for MATCH[™].

CHAPTER 3: RESULTS AND DISCUSSION



MUTATION ANALYSIS OF THE PROMOTER REGION OF FOUR GENES IMPLICATED IN IRON HOMEOSTASIS (*CP*, *CYBRD1*, *HAMP* and *SLC40A1*) IN PORPHYRIA CUTANEA TARDA (PCT) PATIENTS.

ABSTRACT

Porphyria cutanea tarda (PCT) is a disorder characterised by the accumulation of abnormally high levels of porphyrin in the liver, skin, urine and faeces. This condition is a result of defective uroporphyrinogen decarboxylase (UROD) enzyme activity in the haem biosynthesis pathway and presents in two forms: familial PCT (fPCT) and sporadic PCT (sPCT). A number of precipitating factors have been associated with both forms of PCT. These include iron overload, high iron (*HFE*) gene mutations, oestrogen exposure, alcohol abuse and liver disease. The nature of the association between iron overload and PCT is unknown. The aim of this study was to investigate whether variation in any of the four genes involved in iron metabolism [ceruloplasmin (*CP*), cytochrome b reductase 1 (*CYBRD1*), hepcidin antimicrobial peptide (*HAMP*), and solute carrier family 40 member 1 (*SLC40A1*)] contribute to the development of PCT. The study group included 74 PCT patients from three diverse South African population groups and 132 population-matched controls. The promoter regions of the selected genes were analysed for evidence of DNA sequence variation using the techniques of polymerase chain reaction (PCR) amplification, heteroduplex single-stranded conformation polymorphism analysis (HEX-SSCP) and bi-directional semi-automated DNA sequencing. A total of 34 variants (11 novel and 23 known) were identified in all four genes. Two of the polymorphisms represented nucleotide expansions i.e. repeat units; the remaining were all single nucleotide substitutions. Statistical analysis determined that several variants were significantly associated with PCT. These polymorphic alleles were either predicted to be related to increased susceptibility for disease or protective against disease. *In silico* analysis revealed a number of putative transcription factor binding sites (TFBSs) in the DNA regions spanning the

variants in all four genes. It was determined that in nine instances, presence of the significant variant allele had the ability to either disrupt an existing or create a new relevant TFBS. The loss or gain of the TFBS's integrity potentially proposes a mechanism of gene dysregulation.

INTRODUCTION

Porphyria cutanea tarda (PCT), the most common form of hepatic porphyria, is characterised by urinary excretion and the hepatic synthesis of abnormally large amounts of porphyrins (Nordmann and Puy 2002). Porphyrin accumulation occurs as a result of a deficiency in the uroporphyrinogen decarboxylase (UROD) enzyme during the fifth step of the haem biosynthesis pathway (Sampietro *et al.* 1999).

Whilst familial PCT (fPCT) is associated with an *UROD* gene mutation, both sporadic PCT (sPCT) and fPCT are associated with several additional precipitating factors including alcohol abuse, oestrogen and chemical exposure, mutations in the high iron (*HFE*) gene and hepatitis C virus (HCV) virus and human immunodeficiency virus (HIV) infections (Mendez *et al.* 2005). Iron overload has also been implicated in the pathogenesis of PCT. An *in vitro* porcine model demonstrated that iron overload affects hepatic porphyrin metabolism in a manner indicative of the abnormalities associated with PCT (Kushner *et al.* 1972). Many genes are involved in iron homeostasis, some of which have been identified only recently, indicating that several of these genes interact with one another to ensure the strict regulation of iron metabolism (Nemeth *et al.* 2004).

The current study focused on the regulatory regions of four genes, ceruloplasmin (*CP*), cytochrome b reductase 1 (*CYBRD1*), hepcidin antimicrobial peptide (*HAMP*) and solute carrier family 40 member A1 (*SLC40A1*), that play a significant role in iron homeostasis. The promoter regions of

genes are responsible for regulating gene expression at the transcriptional level (Pesole *et al.* 2002). The purpose of this pilot study was to identify candidate loci within the promoter region of *CP*, *CYBRD1*, *HAMP* and *SLC40A1* appropriate for follow-up experimental functional studies encompassing gene expression assays. It is anticipated that findings obtained from the proposed future research will elucidate the association between iron overload and PCT.

MATERIALS AND METHODS

The study was approved by the Ethics Review committee of the University of Stellenbosch (Project number NO4/08/123). The experimental procedures employed for this study were as described in the Materials and Methods section of this thesis (Chapter 2). In this chapter details are also provided on the South African population groups selected for the investigation.

RESULTS

The promoter region of four genes (*CP*, *CYBRD1*, *HAMP* and *SLC40A1*) implicated in iron metabolism were screened in PCT patients and in unaffected, unrelated control individuals of three South African populations, namely Blacks, Caucasians and Coloureds. Mutation analysis of the regulatory region of the genes mentioned above revealed 11 novel and 23 previously described variants. The variants were initially detected by HEX-SSCP analysis and confirmed by bi-directional semi-automated DNA sequencing (Appendix 3). Four variants (one in *CYBRD1*, two in *HAMP*, one in *SLC40A1*) were suited for RFLP analysis (Appendix 3). PCR conditions for amplicon 2 of *CYBRD1* and amplicon 6 of *SLC40A1* could not be optimised utilising their designated primers (refer to Table 2.1). Nevertheless, for both amplicons PCR product was obtained using the adjacent forward and reverse primers. Thus the BP1F and BP3R primer pair was used to analyse the fragment originally spanned by the BP2 primer pair and the SP5F and SP7R primer pair was used to analyse the fragment originally spanned by the SP6 primer pair.

The *CYBRDI* variants, -849C/G, -167C/G and -163G/A, were detected by aberrant banding patterns on HEX-SSCP gels and confirmed by semi-automated DNA sequencing. Due to technical difficulties encountered with HEX-SSCP methodology, the entire study cohort could not be screened successfully. These variants were omitted from further standard analytical procedures and were not included in Tables 3.1-3.5 but are listed in Appendix 3.

The PCR products amplified using the aforementioned BP1F, BP3R and SP5F, SP7R primer pairs were sequenced following amplification and not analysed by means of HEX-SSCP analysis. The *CYBRDI* variants identified on the sequencing electropherograms included the -1849T/G, -1844G/C, -1834G/A, -1813C/T, -1540G/A, -1540G/A, -1459T/C and -1467T/C single nucleotide substitutions. The *SLC40A1* variants detected on the sequencing electropherograms include the -622T/C, -524C/T, -501C/T, 98G/C and -8C/G. In all instances, sequence variation was numbered from the first translation site (ATG). For most cases, the nucleotide that was identified upon inspection of the reference sequence (www.ensembl.org) and cross referenced with a sequence available on another database (www.genatlas.org) was considered the common allele. However, for the following single base pair changes at nucleotide positions -439T/C (*CP*), -1844C/G, -238A/G (*CYBRDI*), -1355G/C, -622C/T, -501T/C (*SLC40A1*), the allele on the reference sequence had a lower incidence than the polymorphic allele in this study. After comparing the international incidences of these alleles and/or taking the ancestral allele into consideration (HapMap), the variants were annotated as such that the allele with a higher frequency (usually the ancestral allele) was accepted as the common allele. These variants were therefore annotated as follows: -439C/T (*CP*), -1844G/C, -238G/A, (*CYBRDI*) and -1355C/G, -622T/C, -501C/T (*SLC40A1*).

Variants identified by the above means were subjected to statistical analysis. Genotype frequencies observed in patients and unaffected control individuals of all three racial groups are tabulated in

Table 3.1. The frequencies of the polymorphic alleles were compared between patients and controls within each population group and are also indicated in Table 3.1. The P values were determined by a Chi-squared (χ^2) goodness-of-fit test or the Fisher's exact test using 2x2 (STATISTICA 8.0) or 2x3 contingency tables (Microsoft Excel software). The Fisher's exact test was applied when the total sample size (patients and controls) was less than 20, or when the total sample size was between 20 and 40 but with one of the values in the contingency table less than 5 (Cochran 1954). A P value smaller than 0.05 ($P < 0.05$) was considered statistically significant. Furthermore, in each population group, the patient and control groups were divided according to gender and the χ^2 and/or Fisher's exact test was applied (data not shown).

Patient and control groups were tested for deviation from Hardy-Weinberg equilibrium (HWE) by means of a χ^2 goodness-of-fit test using the Tools for population genetics association studies (TFPGA) program v1.3 (Miller 1997). A probability (P) value greater than 0.05 ($P > 0.05$) determined that the group was in HWE. The P values at each locus for the three population groups (patients and controls) are indicated in Table 3.2. A statistically significant association between variants at one locus in *CYBRDI* and one locus in *HAMP* and *PCT* were demonstrated to be gender specific in the current study and are summarised in Table 3.3.

Table 3.1 Genotypic and polymorphic allele frequencies of the *CP*, *CYBRD1*, *HAMP* and *SLC40A1* variants identified in patients and control individuals in South African Blacks, Caucasians and Coloureds.

◇ key to follow on pg 71

		<i>CP</i>																							
		Blacks						Caucasians						Coloureds											
Nucleotide position and change	Study cohort	n	Genotype			P	2n	Variant Allele	P	n	Genotype			P	2n	Variant Allele	P	n	Genotype			P	2n	Variant Allele	P
			CC	CG	GG						CC	CG	GG						CC	CG	GG				
-567C/G #rs34053109	Patients	11	0.91	0.09	0.00	0.78	22	G	0.79	23	0.91	0.00	0.09	0.06	46	0.09	0.01	17	1.00	0.00	0.00	-	34	0	-
	Controls	16	0.94	0.06	0.00		32	0.03		41	1.00	0.00	0.00		82	0.00		-	17	0.82	0.12		0.06	0.93	
-563T/C #rs17838834	Patients	11	0.82	0.18	0.00	0.44	22	C	0.48	23	1.00	0.00	0.00	-	46	0.00	-	23	0.78	0.13	0.09	0.50	46		0.15
	Controls	16	0.69	0.31	0.00		32	0.16		51	0.84	0.14	0.02		102	0.09		26	0.68	0.32	0.00		56	0.16	
-439C/T #rs701749	Patients	14	0.71	0.29	0.00	0.11	28	T	0.05	28	0.82	0.18	0.00	0.68	56	0.09	0.98	28	0.68	0.32	0.00	0.50	56	0.16	0.50
	Controls	22	0.95	0.05	0.00		44	0.02		51	0.84	0.14	0.02		102	0.09		26	0.77	0.23	0.00		52	0.12	
-364delT #rs17838833	Patients	15	0.87	0.13	0.00	0.08	30	-	0.08	27	0.74	0.26	0.00	0.42	54	0.13	0.45	28	0.68	0.32	0.00	0.05	56	0.16	0.06
	Controls	22	1.00	0.00	0.00		44	0.00		55	0.82	0.18	0.00		110	0.09		40	0.88	0.12	0.00		80	0.06	
-354T/C #rs17838832	Patients	14	0.86	0.07	0.07	0.18	28	C	0.02	26	0.92	0.08	0.00	0.35	52	0.04	0.35	25	0.92	0.08	0.00	0.59	50	0.04	0.28
	Controls	23	1.00	0.00	0.00		46	0.00		38	0.97	0.03	0.00		76	0.01		33	0.85	0.12	0.02		66	0.09	
-350C/T #rs3433174	Patients	14	0.93	0.07	0.00	0.87	28	T	0.87	26	1.00	0.00	0.00	-	52	0.00	-	25	1.00	0.00	0.00	-	50	0.00	-
	Controls	23	0.91	0.09	0.00		46	0.04		26	1.00	0.00	0.00		52	0.00		25	1.00	0.00	0.00		50	0.00	
-282A/G #17838831	Patients	16	0.81	0.13	0.06	0.05	32	T	0.001	29	1.00	0.00	0.00	-	58	0.00	-	23	0.93	0.07	0.00	0.39	46	0.04	0.40
	Controls	30	1.00	0.00	0.00		60	0.00		29	1.00	0.00	0.00		58	0.00		36	0.86	0.14	0.00		72	0.07	

Table 3.1 continued

		<i>CYBRD1</i>																							
		Blacks			Caucasians			Coloureds																	
Nucleotide position and change	Study cohort	<i>n</i>	Genotype			<i>P</i>	<i>2n</i>	Variant Allele	<i>P</i>	<i>n</i>	Genotype			<i>P</i>	<i>2n</i>	Variant Allele	<i>P</i>								
			TT	TG	GG						TT	TG	GG					TT	TG	GG					
-1849T/G #rs12692964	Patients	12	0.58	0.42	0.00	^a 0.87	24	G	0.89	27	0.59	0.37	0.04	0.97	54	0.22	0.80	24	0.67	0.25	0.05	0.63	48	0.21	0.36
	Controls	13	0.31	0.39	0.00		26	0.19		32	0.72	0.26	0.02		64	0.20		28	0.54	0.36	0.11		56	0.29	
-1844G/C #rs7585974	Patients	13	0.69	0.23	0.08	^a 0.59	26	C	0.44	24	0.58	0.25	0.17	0.14	48	0.29	0.04	27	0.60	0.33	0.07	0.17	54	0.24	0.06
	Controls	13	0.77	0.23	0.00		26	0.12		36	0.75	0.22	0.03		72	0.14		28	0.82	0.14	0.04		56	0.11	
-1834G/A #rs7586174	Patients	15	0.69	0.25	0.06	^a 0.86	30	A	0.84	28	0.64	0.29	0.07	0.15	56	0.21	0.03	26	0.62	0.31	0.08	0.32	52	0.23	0.10
	Controls	12	0.75	0.17	0.08		24	0.17		37	0.43	0.35	0.22		74	0.39		28	0.46	0.32	0.21		56	0.38	
-1813C/T #rs12692965	Patients	12	0.50	0.42	0.08	^a 0.30	24	T	^a 0.12	22	0.59	0.36	0.04	0.59	44	0.23	0.31	24	0.71	0.21	0.08	0.66	48	0.19	0.58
	Controls	10	0.80	0.20	0.00		20	0.10		32	0.72	0.26	0.02		64	0.16		28	0.61	0.32	0.07		56	0.23	
-1540G/A This study	Patients	14	0.43	0.50	0.07	^a 0.51	28	A	1.00	28	0.79	0.18	0.03	0.30	56	0.13	0.70	26	0.54	0.38	0.08	0.80	52	0.27	0.57
	Controls	14	0.36	0.64	0.00		28	0.31		37	0.70	0.30	0.00		74	0.15		27	0.59	0.37	0.04		54	0.22	
-1477G/A This study	Patients	14	0.43	0.50	0.07	^a 0.38	28	A	0.78	28	0.82	0.18	0.00	0.14	56	0.06	0.04	26	0.46	0.42	0.12	0.32	52	0.33	0.51
	Controls	14	0.29	0.71	0.00		28	0.36		36	0.61	0.33	0.06		72	0.22		27	0.30	0.63	0.07		54	0.39	
-1459T/C #rs10199858	Patients	14	0.50	0.21	0.29	^a 0.28	28	C	0.59	28	0.50	0.32	0.18	0.07	56	0.29	0.14	26	0.38	0.24	0.38	0.03	52	0.50	0.01
	Controls	14	0.29	0.50	0.21		28	0.46		34	0.23	0.59	0.18		68	0.47		26	0.07	0.35	0.58		52	0.75	
-1452T/C This study	Patients	9	0.67	0.22	0.11	^a 0.50	18	C	0.40	19	0.79	0.16	0.05	0.02	38	0.13	0.11	12	0.92	0.08	0.00	^a 0.01	24	0.04	0.02
	Controls	12	0.75	0.25	0.00		24	0.13		28	0.46	0.54	0.00		56	0.27		24	0.46	0.54	0.00		48	0.27	

Table 3.1 continued

		<i>CYBRD1</i>																							
		Blacks			Caucasians			Coloureds																	
Nucleotide position and change	Study cohort	n	Genotype			P	2n	Variant Allele	P	n	Genotype			P	2n	Variant Allele	P								
			TT	TC	CC						TT	TC	CC					TT	TC	CC					
-1346T/C This study	Patients	14	0.64	0.21	0.14	*0.19	28	0.25	0.59	29	0.86	0.14	0.00	0.36	58	0.07	0.38	22	0.59	0.27	0.14	0.003	44	0.27	0.00
	Controls	19	0.61	0.39	0.00		38	0.19		35	0.77	0.23	0.00		70	0.11		23	1.00	0.00	0.00		46	0.00	
-1272T/C This study	Patients	14	0.79	0.21	0.00	*0.37	28	0.11	0.45	29	0.79	0.21	0.00	0.83	58	0.10	0.84	22	0.73	0.27	0.00	0.01	44	0.14	0.01
	Controls	18	0.89	0.11	0.00		36	0.05		35	0.77	0.23	0.00		70	0.11		23	1.00	0.00	0.00		46	0.00	
-645T/C This study	Patients	11	0.91	0.09	0.00	*0.19	22	0.22	0.09	30	0.90	0.07	0.03	0.23	60	0.07	0.22	25	0.92	0.00	0.08	0.15	50	0.15	0.46
	Controls	9	0.78	0.00	0.22		18	0.05		42	0.98	0.00	0.02		84	0.02		23	0.96	0.00	0.04		46	0.04	
-624G/A #rs884408	Patients	11	0.64	0.09	0.27	*0.24	22	0.32	0.04	27	0.55	0.40	0.07	0.20	54	0.26	0.34	25	0.72	0.20	0.08	0.28	50	0.18	0.09
	Controls	9	0.89	0.11	0.00		18	0.06		42	0.62	0.38	0.00		84	0.19		23	0.87	0.13	0.00		46	0.07	
-238G/A #rs868106	Patients	8	0.63	0.12	0.25	*0.18	16	0.31	0.15	20	0.50	0.10	0.40	0.00	40	0.43	0.88	19	0.42	0.05	0.53	0.0001	38	0.55	0.00
	Controls	12	0.75	0.25	0.00		24	0.12		35	0.23	0.71	0.06		70	0.41		23	0.74	0.26	0.00		46	0.13	
G(T)₈G(T)_n G(T)_nG(T)₉ This study	Patients	16	0.50	0.19	0.31	0.08				32	0.06	0.34	0.60	0.21				32	0.56	0.16	0.28	0.67			
	Controls	24	0.17	0.33	0.50					38	0.18	0.38	0.44		58	0.58	0.21	0.21							

*1, G(T)₈G(T)₆G(T)₆G(T)₉; *2, G(T)₈G(T)₈G(T)₆G(T)₉; *3, G(T)₈G(T)₆G(T)₉

Table 3.1 continued

		<i>HAMP</i>																							
		Blacks			Caucasians			Coloureds																	
Nucleotide position and change	Study cohort	<i>n</i>	Genotype			<i>P</i>	<i>2n</i>	Variant Allele	<i>P</i>	<i>n</i>	Genotype			<i>P</i>	<i>2n</i>	Variant Allele	<i>P</i>								
			CC	CT	TT						T	CC	CT					TT	T	CC	CT	TT	T		
-1010C/T #rs10414846	Patients	12	0.83	0.17	0.00	*0.64	24	0.08	0.87	27	0.88	0.12	0.00	0.95	54	0.06	0.62	21	0.91	0.09	0.00	0.06	42	0.05	0.09
	Controls	14	0.86	0.14	0.00		28	0.07		46	0.85	0.15	0.00		92	0.08		33	0.70	0.30	0.00		66	0.15	
-582A/G #rs10421768	Patients	13	0.39	0.53	0.08	*0.09	24	0.35	0.04	25	0.84	0.16	0.00	0.44	50	0.08	0.21	26	0.65	0.27	0.08	0.02	52	0.21	0.00
	Controls	22	0.73	0.27	0.00		44	0.14		56	0.71	0.27	0.02		112	0.15		43	0.93	0.02	0.00		86	0.01	
-429 G/T This study	Patients	13	0.77	0.23	0.00	*0.59	26	0.12	0.85	29	1.00	0.00	0.00	-	29	0.00	-	26	0.80	0.12	0.08	0.58	52	0.14	0.92
	Controls	19	0.74	0.26	0.00		38	0.13		31	0.77	0.20	0.03	62	0.13										

		<i>SLC40A1</i>																							
		Blacks			Caucasians			Coloureds																	
Nucleotide position and change	Study cohort	<i>n</i>	Genotype			<i>P</i>	<i>2n</i>	Variant Allele	<i>P</i>	<i>n</i>	Genotype			<i>P</i>	<i>2n</i>	Variant Allele	<i>P</i>								
			TT	TC	CC						C	TT	TC					CC	C	TT	TC	CC	C		
-1461T/C This study	Patients	13	0.69	0.15	0.15	0.37	26	0.23	0.09	22	0.96	0.04	0.00	0.66	44	0.02	0.44	19	0.74	0.21	0.05	0.73	38	0.16	0.46
	Controls	5	1.00	0.00	0.00		10	0.00		28	0.92	0.04	0.04		56	0.05		24	0.83	0.13	0.04		48	0.10	
-1399G/A This study	Patients	14	0.86	0.14	0.00	*0.003	28	0.07	*0.03	23	0.87	0.13	0.00	0.20	46	0.07	0.21	19	0.79	0.21	0.10	0.54	38	0.11	0.54
	Controls	5	0.60	0.00	0.40		10	0.40		29	0.97	0.03	0.00		58	0.02		28	0.96	0.04	0.00		56	0.02	
-1355C/G #rs3811621	Patients	11	0.36	0.64	0.00	*0.00	22	0.31	0.94	22	0.36	0.59	0.05	0.02	44	0.34	0.45	20	0.45	0.35	0.20	0.12	40	0.37	1.00
	Controls	21	0.67	0.05	0.28		42	0.32		44	0.46	0.27	0.27		88	0.41		28	0.57	0.11	0.32		56	0.37	
-750G/A #rs13015236	Patients	11	0.46	0.45	0.09	*0.01	22	0.32	0.003	29	0.21	0.59	0.20	0.52	58	0.50	0.34	26	0.31	0.54	0.15	0.31	52	0.42	0.42
	Controls	17	0.94	0.06	0.00		34	0.03		43	0.28	0.60	0.12		86	0.42		28	0.14	0.72	0.14		56	0.50	

Table 3.1 continued

		<i>SLC40A1</i>													
		Blacks			Caucasians			Coloureds							
Nucleotide position and change	Study cohort	n	Genotypes		P	2n	Variant Allele	P	n	Genotype		P	2n	Variant Allele	P
-622T/C #rs12693542	Patients	3	TT	TC	CC	0.33	C	0.66	16	TT	TC	CC	0.56	C	0.86
	Controls	4	0.33	0.67	0.00		6		0.67	32	0.50	0.38		0.12	
-524C/T This study	Patients	3	CC	CT	TT	-	T	-	16	CC	CT	TT	-	T	-
	Controls	4	0.50	0.25	0.25		8		0.62	18	0.55	0.22		0.22	
-501C/T #rs6728200	Patients	3	CC	CT	TT	-	T	-	16	CC	CT	TT	0.30	T	0.30
	Controls	4	1.00	0.00	0.00		8		0.00	17	1.00	0.00		0.00	
-98G/C #rs13008848	Patients	6	GG	GC	CC	0.20	C	0.22	18	GG	GC	CC	0.65	C	0.48
	Controls	4	0.67	0.33	0.00		12		0.17	36	0.67	0.22		0.11	
-8C/G #rs11568351	Patients	6	CC	CG	GG	0.67	G	0.66	18	CC	CG	GG	0.00	G	0.003
	Controls	4	0.75	0.25	0.00		8		0.13	39	0.13	0.64		0.23	
(CGG) _n #rs3833570	Patients	3	*1*1	*1*2	*2*2	0.41	*2	0.51	13	*1*1	*1*2	*2*2	0.03	*2	0.001
	Controls	3	0.33	0.67	0.00		6		0.33	26	0.31	0.69		0.00	
									10	0.50	0.30	0.20		20	0.14
									14	0.71	0.29	0.00		28	0.65
									15	0.53	0.40	0.07		30	0.27
									15	0.33	0.40	0.27		30	0.47
									15	0.93	0.07	0.00		30	0.03
									14	0.86	0.14	0.00		28	0.07
									15	0.86	0.07	0.07		30	0.10
									12	0.83	0.00	0.17		24	0.17
									21	0.52	0.33	0.14		42	0.31
									9	0.78	0.11	0.11		18	0.17
									21	0.57	0.29	0.14		42	0.29
									7	0.86	0.14	0.00		14	0.07
									16	0.25	0.63	0.12		32	0.65
									14	0.71	0.29	0.00		28	0.65

*1, (CGG)₇; *2, (CGG)₈

◇ Key to **Table 3.1**

* RefSNP identification number; *Fisher's exact test; ζ polymorphic allele frequency.

Abbreviations: A, adenine; C, cystosine; *CP*, ceruloplasmin; *CYBRDI*, cytochrome b reductase 1; del, deletion, G, guanine. *HAMP*, hepcidin antimicrobial peptide; *n*, number of individuals; *2n*, number of alleles; *P*, Probability value calculated using Chi squared (χ^2) exact test or Fisher's exact; *SLC40A1*, solute carrier family 40 member 1; T, thymine.

purple	novel variant identified in the current study
red	within a particular population group, variant was identified in patients only
yellow	variant identified in heterozygote state only
blue	variant identified in homozygote state only
pink	variant identified in one population group only
green	statistically significant difference between patients and controls ($P < 0.05$)

Table 3.2 The Probability (*P*) values at variant loci in the *CP*, *CYBRD1*, *HAMP* and *SLC40A1* genes tested for departure from HWE. PCT patients and control individuals from three South African groups (Blacks, Caucasians and Coloureds) were analysed.

◇ key to follow on page 74

Nucleotide position and change	Study cohort	Blacks	Caucasians	Coloureds
		<i>P</i>	<i>P</i>	<i>P</i>
<i>CP</i>				
-567C/G	Patients	0.87	0.00	-
	Controls	0.90	-	-
-563T/C	Patients	0.74	-	0.07
	Controls	0.46	-	0.02
-439C/T	Patients	0.53	0.60	0.31
	Controls	0.91	0.29	0.51
-364DelT	Patients	0.78	0.44	0.31
	Controls	-	0.46	0.67
-354T/C	Patients	0.02	0.84	0.84
	Controls	-	0.93	0.13
-350C/T	Patients	0.89	-	-
	Controls	0.83	-	-
-282A/G	Patients	0.09	-	0.84
	Controls	-	-	0.65
<i>CYBRD1</i>				
-1849T/G	Patients	-	0.71	0.51
	Controls	0.39	0.73	0.24
-1844G/C	Patients	0.35	0.05	0.65
	Controls	0.63	0.67	0.18
-1834G/A	Patients	0.47	0.42	0.50
	Controls	0.17	0.11	0.09
-1813C/T	Patients	0.98	0.89	0.12
	Controls	0.73	0.99	0.60
-1540G/A	Patients	0.58	0.33	0.91
	Controls	0.08	0.29	0.71
-1477G/A	Patients	0.58	0.60	0.84
	Controls	0.04	0.83	0.09
-1459T/C	Patients	0.04	0.80	0.01
	Controls	0.98	0.29	0.69

Table 3.2 continued

		<i>CYBRD1</i>		
-1452T/C	Patients	0.28	0.36	0.88
	Controls	0.62	0.05	0.07
-1346G/A	Patients	0.11	0.69	0.14
	Controls	0.31	0.45	-
-1272T/C	Patients	0.65	-	0.54
	Controls	0.80	0.45	-
-645T/C	Patients			
	Controls			
-624G/A	Patients	0.01	-	0.11
	Controls	0.86	0.13	0.74
-238G/A	Patients	0.04	0.32	0.00
	Controls	0.62	0.01	0.47
		<i>HAMP</i>		
-1010C/T	Patients	0.75	0.76	0.82
	Controls	0.77	0.58	0.31
-582A/G	Patients	0.40	0.66	0.33
	Controls	0.46	0.71	0.93
-429G/T	Patients	0.64	-	0.01
	Controls	0.51	-	0.62

Table 3.2 continued

Nucleotide position and change	Study cohort	<i>P</i>	<i>P</i>	<i>P</i>
<i>SLC40A1</i>				
-1461T/C	Patients	0.04	0.91	0.36
	Controls	-	0.06	0.11
-1399G/A	Patients	0.77	0.74	0.61
	Controls	0.03	0.92	0.92
-1355C/G	Patients	0.12	0.74	0.26
	Controls	0.01	0.92	0.00
-750G/A	Patients	0.87	0.35	0.60
	Controls	0.90	0.11	0.02
-622T/C	Patients	0.39	0.61	0.93
	Controls	0.35	0.14	0.45
-524C/T	Patients	-	-	0.89
	Controls	-	-	0.90
-501C/T	Patients	-	0.90	0.00
	Controls	-	-	0.01
-98G/C	Patients	0.62	0.00	0.31
	Controls	-	0.13	0.07
-8C/G	Patients	0.62	0.02	0.17
	Controls	0.78	0.06	0.17

-, HWE not calculated as polymorphic allele not detected in respective population group

Abbreviations: A, adenosine; C, cysteine; *CP*, ceruloplasmin gene; *CYBRDI*, cytochrome b reductase 1 gene; del; deletion; G, guanine; *HAMP*, hepcidin antimicrobial peptide gene; *P*, Probability value; *SLC40A1*, solute carrier family 40 member 1 gene; T, thymidine

Table 3.3. Allele frequency probability values (*P*) of variants demonstrating gender specificity in the current study.

Nucleotide position and change	Population group	Gender	<i>P</i>
<i>CYBRDI</i>			
-1459T/C	Coloured	Female	0.02
<i>HAMP</i>			
-582A/G	Coloured	Female	0.00

; recurring

Abbreviations: *CP*, ceruloplasmin; *CYBRDI*, cytochrome b reductase 1; *HAMP*, hepcidin antimicrobial peptide; *P*, probability value, del, deletion, T, thymine; C, cytosine; A, adenine; G, guanine.

The polymorphic allele frequencies of the South African Black and Caucasian unaffected control groups were correlated with Sub-Saharan and European unaffected control groups available on HapMap. Incidence comparison was only carried out when possible i.e. when incidence data was available on HapMap. Although South African Blacks do not display exact genetic similarity to Sub-Saharan groups as Caucasians do not to Europeans, the polymorphic allele frequency correlations provided minimal validation of the accuracy of the results (allele frequencies) obtained in the current study. The South African Coloured group is a more recently founded population which arose through admixture. It is thus not possible to compare allele frequencies with mixed ancestry incidence data available on HapMap. The polymorphic allele frequencies in the South African unaffected control group screened in the current study as well as the international incidence available on HapMap is tabulated in Table 3.4. Only loci where incidence data were available on HapMap for comparison were included. The variant allele frequencies determined in the South African patient group are also presented in Table 3.4 for reference, but should not be compared with the HapMap incidence data as the allele incidences on this database refer to the general population and not to a selected group of diseased patients.

Table 3.4. Comparison of variant (polymorphic) allele frequencies.

Nucleotide position and change	Ancestral allele	SA population group	Polymorphic allele frequencies			HapMap population group
			SA PCT patients	SA unaffected individuals	HapMap unaffected individuals	
<i>CP</i>						
-567C/G	NA	Blacks	0.05	0.03	0.02	SS African
		Caucasians	0.09	0.00	0.00	European
-563T/C	C	Blacks	0.09	0.16	0.07	SS African
		Caucasians	0.00	-	-	European
-439C/T	C	Blacks	0.14	0.02	0.00	SS African
		Caucasians	0.09	0.09	0.11	European
-364delT	NA	Blacks	0.07	0.00	0.03	SS African
		Caucasians	0.13	0.09	0.09	European
-354T/C	T	Blacks	0.11	0.00	0.05	SS African
		Caucasians	0.04	0.01	0.06	European
-350C/T	NA	Blacks	0.04	0.04	0.03	SS African
		Caucasians	0.00	-	-	European
-282A/G	NA	Blacks	0.13	0.11	0.08-0.13	SS African
		Caucasians	0.00	-	-	European
<i>CYBRDI</i>						
-238G/A	NA	Blacks	0.31	0.12	0.31	SS African
		Caucasians	0.43	0.41	0.37-0.50	European
<i>HAMP</i>						
-1010C/T	C	Blacks	0.08	0.07	0.13	SS African
		Caucasians	0.06	0.08	0.18	European
-582A/G	A	Blacks	0.35	0.14	0.13	SS African
		Caucasians	0.08	0.15	0.16	European
<i>SLC40A1</i>						
-750G/A	G	Blacks	0.32	0.03	0.23	SS African
		Caucasians	0.50	0.42	0.60	European
-622T/C	T	Blacks	0.67	0.62	0.56	SS African
		Caucasians	0.31	0.31	0.29	European

-, polymorphic allele not identified in the current study's control group; 0.00, polymorphic allele not identified in population group.

Abbreviations: A, adenine; C, cytosine; *CP*, ceruloplasmin gene; *CYBRDI*, cytochrome b reductase 1 gene; G, guanine; *HAMP*, hepcidin antimicrobial peptide gene; NA, not available; PCT, porphyria cutanea tarda; SA, South African; *SLC40A1*, solute carrier 40 member 1 gene; SS African, Sub-Saharan African; T, thymine.

Linkage disequilibrium (LD) and haplotype analysis using the Haploview 4.0 software were performed on the identified polymorphic loci in each of the four genes (Barrett *et al.* 2005). Regarding the complete patient cohort, a total of 44.59% of the patients possessed one of the seven variants in *CP* and 16.62% of the patients possessed two or more variants in the *CP* gene. Seventy two percent (72%) of patients heterozygous for -364delT were also homozygous for -439T/C in *CP*. The LD test was applied, and these two SNPS were not found to be in LD (D' : 0.17, LOD: 0.95, r^2 : 0.02) in the study cohort. In addition, LD analysis and haplotype analysis was performed for all loci in *CP*, and no significant results were obtained.

The LD plot for the *CYBRD1* loci revealed that two novel *CYBRD1* variants (-1540G/A and -1477G/A) were potentially in LD (D' = 0.95, LOD= 26.25, r^2 = 0.69). Haplotype analysis revealed that the -1492 variant A allele and the -1555 common G allele were most likely to be inherited as a haplotype (P =0.03). No further investigations were performed in this study.

In *HAMP*, 28 patients (37.84%) contained one of the three variants in the *HAMP* gene. One patient (Coloured female) was observed to be heterozygous for all three variants. Five of eight patients (62.5%) heterozygous for the -1010C/T variant were also heterozygous for the -582A/G variant. The two variants were tested for LD and there was no evidence of the -1010C/T and -582A/G variants being in linkage disequilibrium in the study cohort (D' : 0.249, LOD: 0.78, r^2 : 0.031).

Haplotype analysis did not yield any significant results. Every patient carried at least one variant in *SLC40A1*. LD and haplotype analysis revealed no significant linkage for any of the loci in *SLC40A1*. The small sample size available for the current study might have impaired accurate LD and haplotype analysis for all four of the genes, as sample size may contribute to the accuracy of these methodologies (Wang *et al.* 2005).

In silico analysis was performed on all variants identified in this study (Table 3.1). The bioinformatic programs were applied as follows: the disruption of an existing or creation of a new of TFBS was investigated by comparing the reference sequence with the variant sequence. Identification of the same TFBS by more than one program, increases the likelihood that the motif exists. *In silico* analysis revealed several putative TFBSs located in regions adjacent to and spanning the variants detected in the promoter regions of *CP*, *CYBRD1*, *HAMP* and *SLC40A1*. The putative TBFSs predicted by JASPAR, TRANSFAC^{®7} and rVISTA, are tabulated in Table 3.5. Only TFs considered to be biologically relevant to the current investigation, i.e. involved in iron metabolism pathways, liver-specific or known to regulate one of the genes analysed here, were included in Table 3.5. Only the putative TFBS that span the region of the statistically significant variants identified in this study will be discussed.

Table 3.5 Predicted TFBSs in the promoter region of the *CP*, *CYBRD1*, *HAMP*, and *SLC40A1* genes, obtained from data generated by *in silico* analysis. The databases used were JASPAR CORE, TRANSFAC[®]7 (comprising 2 programs, i.e. MATCH[™] and ALIBABA2) and rVISTA. TFBSs listed were expected to (i) remain intact, (ii) be disrupted or (iii) newly created in the presence of the polymorphic allele.

◇ key to follow on pg 82

<i>CP</i>												
Nucleotide position and change	JASPAR			MATCH			ALIBABA			rVISTA		
	Intact	Disrupted	Created	Intact	Disrupted	Created	Intact	Disrupted	Created	Intact	Disrupted	Created
-567C/G	GATA2	YY1	GATA2	GATA	-	GATA	-	-	-	HNF1 α HNF3 β HNF4 α	C/EBP α YY1	GATA 1 GATA 2 GATA 3
-563T/C	YY1	-	GATA2	GATA	-	-	-	-	-	HNF1 α HNF3 β HNF4 α	C/EBP α YY1	GATA 1 GATA 2 GATA 3
-439C/T	-	-	-	-	AP1 C/EBP α	-	-	C/EBP α , ER	-	HNF1 α HNF3 β HNF4 α	AP1 YY1	C/EBP α
-364delT	PU.1	-	-	-	-	-	-	-	-	-	-	C/EBP α
-354T/C	PU.1	-	-	-	AP1 HNF3 β	-	-	-	-	C/EBP α HNF1 α HNF3 β HNF4 α	C/EBP α	C/EBP α
-350C/T	PU.1	-	-	-	AP1	-	-	-	-	HNF1 α HNF3 β HNF4 α	AP1, YY1	-
-282A/G	-	-	GATA2	STAT1	C/EBP α	-	-	C/EBP α	-	C/EBP α AP1 USF	-	-

Table 3.5 continued

Nucleotide position and change	<i>CYBRD1</i>											
	JASPAR			MATCH™			ALIBABA2			rVISTA		
	Intact	Disrupted	Created	Intact	Disrupted	Created	Intact	Disrupted	Created	Intact	Disrupted	Created
-1849T/G	-	GATA 2	GATA 2	-	GATA 2 GATA 3	-	-	-	-	NF-Y	-	-
-1844G/C	-	-	-	-	-	-	-	-	-	-	-	-
-1834G/A	USF1	-	PU.1	-	-	GfiI	-	SP1	-	AP1, C/EBP α , USF1	GATA 1	-
-1813C/T	-	-	PU.1	-	-	USF	-	-	-	C/EBP α , SP1	-	-
-1540G/A	-	GATA 3 PU.1	-	-	-	-	-	-	-	STAT1, C/EBP α , SP1	-	-
-1477G/A	GATA 2 GATA 3 USF1	-	-	-	GATA 2	-	-	-	-	C/EBP α , GATA1, USF	-	-
-1459T/C	-	-	-	-	-	-	-	-	-	SP1	-	-
-1452T/C	-	-	-	-	-	-	-	-	-	C/EBP α	-	-
-1346T/C	GATA 2 GATA 3	-	PU.1	-	-	-	-	GATA 1	-	C/EBP α , GATA 1	-	SP1

Table 3.5 continued

CYBRD1

Nucleotide position and change	JASPAR			MATCH™			ALIBABA2			rVISTA		
	Intact	Disrupted	Created	Intact	Disrupted	Created	Intact	Disrupted	Created	Intact	Disrupted	Created
-1272T/C	PU.1	-	C/EBP α	-	-	-	-	-	-	C/EBP α	-	-
-645T/C	-	-	-	-	NF-E2	-	-	SP1	-	-	-	-
-624G/A	-	RUNX1	-	-	-	-	-	-	SP1	C/EBP α	-	-
-238G/A	-	SP1	NF-Y	-	-	-	-	-	-	AP1	-	-

HAMP

Nucleotide position and change	JASPAR			MATCH™			ALIBABA2			rVISTA		
	Intact	Disrupted	Created	Intact	Disrupted	Created	Intact	Disrupted	Created	Intact	Disrupted	Created
-1010C/T	GATA1, GATA2 GATA3	-	-	CREB	-	-	-	-	-	AP1, GATA1, GATA2, YY1, USF	-	-
-582A/G	-	USF1	-	NFKB	-	-	-	-	-	HNF3 β , USF	-	-
-429G/T	YY1, USF1	-	-	-	-	-	-	-	-	YY1, USF	-	-

Table 3.5 continued

Variant	JASPAR CORE			MATCH™			ALIBABA2			rVISTA		
	Intact	Disrupted	Created	Intact	Disrupted	Created	Intact	Disrupted	Created	Intact	Disrupted	Created
-1461T/C	-	-	PU.1	-	Gfi1	-	-	-	-	C/EBP α , YY1, USF, NFKB, STAT1	-	-
-1399G/A	-	SP1	GATA 2 GATA 3 NF-Y	-	Gfi1 NF-Y	-	SP1	-	-	C/EBP α , STAT1, SP1, YY1, USF, HNF1 α	-	-
-1355C/G	-	-	-	-	-	-	-	-	-	ER, YY1, USF, HNF1 α	-	-
-750G/A	PU.1 SP1B	PU.1	-	-	-	-	-	-	-	C/EBP α , STAT1, YY1, HNF1 α , NFKB	-	-
-622T/C	USF1	CREB PU.1	-	USF1	-	-	-	NFKB	SP1	STAT1, SP1, YY1	-	-
-524C/T	-	-	SP1	-	-	-	-	-	YY1	C/EBP α , SP1, YY1, NFKB	-	-
-501C/T	-	SP1	-	-	-	-	SP1, AP2	-	-	USF, HNF1 α , NFKB	-	-
-98G/C	GATA2 GATA 3	-	RUNX1	-	-	-	-	-	-	C/EBP α , GATA 1, GATA 2,	-	-
-8C/G	-	-	-	-	-	-	-	-	-	YY1, USF, HNF1 α	GATA 2	-

Abbreviations: AP1, activator protein 1, C/EBP α , CCAAT enhancer binding protein; CREB, cAMP response element-binding protein 1; GATA 1, GATA binding protein 1; GATA 2, GATA binding protein 2; GATA 3, GATA binding protein 3; Gfi1, growth factor independent 1; HNF1 α , hepatic nucleic factor 1 alpha; HNF3 β , hepatic nuclear factor 3 beta; NF-E2, nuclear factor erythroid 2; NFKB, nuclear factor kappa-B; NF-Y, nuclear factor -Y, PU.1, spleen focus forming virus proviral integration oncogene SPI1; RUNX1, runt-related transcription factor 1; SP1, specificity protein 1; STAT1, signal transducer and activator of transcription 1; USF, upstream stimulatory factor 1; YY1, Ying Yang-1

DISCUSSION

The current study employed various techniques to analyse variation in the promoter region of *CP*, *CYBRD1*, *HAMP*, and *SLC40A1* and to perform statistical and bioinformatic analysis. Conventionally, variants were initially detected by HEX-SSCP analysis and confirmed by bi-directional semi-automated DNA sequencing. RFLP analysis was employed where feasible. Statistical analysis determined which variants were associated with PCT (Tables 3.1). All loci were tested for departure from HWE (Table 3.2). Males generally show an increased incidence in PCT. For this reason, males and females were analysed separately for genotype and allele frequencies (data not shown) and the significant results were summarised in Table 3.4. Putative TFBSs were predicted by *in silico* analysis (Table 3.5) and their potential implications were investigated.

When comparing allele frequencies in patients and controls it was noticed that in some cases the polymorphic allele was more frequent in controls. This may be interpreted as an indication that the inheritance of the polymorphic allele confers a protective effect. Alternatively, the polymorphic allele may have no effect at all in the development of PCT. Inheritance of a particular allele or haplotype may provide protection against a disease either by preventing its manifestation or more likely decreasing the phenotypic severity. Protective alleles have been identified in amongst others diabetes type 2 (Motala *et al.* 2005), Alzheimer's disease (Qiu *et al.* 2004), malaria (Hill 1999) and rheumatoid arthritis (Gourraud *et al.* 2006). Due to the intricate interaction and inheritance of alleles (susceptibility, neutral or protective) the presence of a protective allele does not necessarily negate onset of disease.

Five loci showed departure from HWE (Table 3.2). There are a number of reasons causing departure from HWE. These include genotyping errors, random genetic drift, inbreeding, admixture of ethnic groups (population stratification), assortative mating, mutation and selection (Klug and

Cummings 2003). Genotyping errors cannot be entirely eradicated as molecular techniques (sampling, DNA extraction, molecular analysis, technical artefacts) and manual handling (scoring, data analysis) are not 100% reliable (Bonin *et al.* 2004).

In *CP*, the -563T/C locus deviated from HWE. The patient and control groups were screened for this variant using HEX-SSCP analysis. Another variant, -567C/G, was also identified four bp upstream from the -563T/C variant. The close proximity of these two variants may have altered the predicted conformational banding pattern, leading to inaccurate genotyping of the -563T/C variant.

The -1477G/A (*CYBRDI*) locus in the Black population was not in HWE. All samples were subjected to semi-automated DNA sequencing directly after PCR amplification. After the identification and confirmation of this variant upon bi-directional DNA sequencing, the remaining samples were sequenced in only one direction. It is thus possible that sequencing errors occurred. Two DNA sequencing problems that may be encountered include the failure to detect heterozygotes unambiguously and compression in GC rich areas (e.g. promoter regions of genes) resulting in sequences that are not 100% accurate (Li *et al.* 2003). The small sample size of the Black population may also have contributed to the observed departure from HWE.

The -238G/A locus in the Caucasian control population was not in HWE. The control individuals were genotyped using RFLP analysis and the patients were genotyped upon HEX-SSCP analysis. In all three patient groups, a low number of heterozygotes were detected. This is in contrast to the high heterozygosity detected in the control groups. Due to DNA sample shortages, it was not possible to genotype the patients using RFLP analysis. It is possible that the HEX-SSCP banding patterns were not clear enough to accurately distinguish between the heterozygous and homozygous form of the variant. It is therefore highly recommended that the patient group be genotyped using a

more suitable method (bi-directional semi-automated DNA sequencing or RLFP analysis) before the results in this study are considered for the -238G/A locus.

Upon investigation of the -238G/A variant incidence in the European population (HapMap), it was noted that that the G and A allele had equal frequencies (0.5). In the European population, the GG and AA genotypes each had a frequency of 0.29 and the GA genotype had a frequency of 0.42. In contrast, in the Black and Coloured control group, the polymorphic allele frequency was considerably lower (0.12 and 0.13). This trend was consistently observed in the African American and Mixed Ancestry control population groups (HapMap). The unusual distribution of alleles in the Caucasian control group, as is also observed in the European population, could account for the deviation from HWE at this locus.

In *SLC40A1*, the -1399G/A locus showed departure from HWE. Only 5 control individuals were screened for this variant. It is thus most likely that the small sample size contributed to the deviation from HWE. It is imperative to increase the sample size before the results obtained in this study can be accurately interpreted.

The -501T/C variant in the Coloured population was not in HWE. This variant was genotyped using semi-automated DNA sequencing. The variant was identified and confirmed by bi-directional sequencing but subsequently the remaining patients and controls were sequenced in only one direction. No heterozygotes were detected in this population group but TT and CC homozygotes were detected. It is possible that these results are due to allele dropout. Allele dropout refers to the amplification of only one of the two alleles at a specific locus, causing false homozygotes (Taberlet *et al.* 1996).

It is possible that deviation from HWE in the Coloured population (-1355G/C; -501T/C (*SLC40A1*) investigated in the current study is as a result of admixture. Admixture is an evolutionary force and arises from the intermixture of two or more previously isolated populations (Halder and Shriver 2003). This concept should be kept in mind during the comparison of genomic data with complex disease traits. According to Wahlund's principle, admixture can be demonstrated by the relative high homozygosity and low heterozygosity observed in a population (Hartl and Clark 2007). In recently admixed populations, false-positive associations are often not replicated in subsequent studies (Thomas and Witte 2002). Type-I errors (false-positives) are under-estimated under circumstances of HWE deviation. The impact of Type-I errors is often ignored, but the implications of this error in association studies in particular have not been entirely investigated (Schaid and Jacobsen 1999).

The importance of population stratification in association studies and the fact that frequencies of polymorphisms may be highly population-dependent needs to be highlighted. In this study, the -1477G/A variant was detected in all three population groups. A statistically significant association was noted in Caucasians only, where homozygosity for the variant allele was present in controls and not in patients. In contrast, the homozygous variant form was present in Black patients and not in controls. Examples of disorders where disease-related polymorphism frequencies vary in populations include sarcoidosis pulmonary disease (McGrath *et al.* 2001), alcoholism (Blum *et al.* 1990; Gelernter *et al.* 1993) and prostate cancer (Khoury and Yang 1998).

If a true association exists between a variant and disease, the expected genotypes in patients may depart from HWE (Schaid and Jacobsen 1999). The magnitude of deviation from HWE depends on the disease-related genetic mechanisms. This may account for the deviation in HWE noticed in some patient groups in the current study. Factors causing deviations from HWE, were taken into

consideration and avoided where possible during the selection of patient samples.

UROD AND PCT

The production of haem is an essential process for the survival of man. Haem is required for incorporation into several molecules, most notably haemoglobin; the molecule that in essence enables man to breathe. A decrease in activity of one of the haem biosynthesis enzymes, due to a disruption in the tightly regulated pathway leads to an accumulation of precursors upstream from the site of enzyme defect. Inadequate activity of the fifth enzyme in this pathway, UROD, leads to an accumulation of uroporphyrinogens and is associated with the development of PCT. Upon diagnosis, patients in the current study had not been screened for mutations in the *UROD* gene. It is therefore not clear whether the patients in the present study suffer(ed) from fPCT or sPCT. However, the exonic regions of *UROD* were recently analysed in these patients (Carstens *et al.* 2006; Vervalle J, personal communication). Two novel missense variants were identified in the South African Caucasian group (Carstens *et al.* 2006). A statistically significant association between these variants and PCT was not noted. Preliminary investigations of the *UROD* gene in Blacks and Coloureds have not revealed any significant variation.

The lack of significant *UROD* mutations revealed in the initial studies implies that these patients suffer from sPCT and that at least one exacerbating factor is likely to contribute to decreased UROD activity in the liver. In the event that a patient possesses a *UROD* mutation; the clinical manifestation of PCT almost always only occurs in the presence of a precipitating factor, demonstrating that a *UROD* mutation alone does not cause the onset of PCT. Thus, in both sPCT and fPCT patients, one or more precipitating factors contribute to the manifestation of this disease.

Whilst UROD protein levels are halved in fPCT patients; levels are normal in sPCT patients (Elder *et al.* 1985; Bulaj *et al.* 2000). Furthermore, UROD enzyme activity is reduced by up to 50% in the

liver in sPCT. This observation leads to the speculation that a UROD inhibitor intrinsic to hepatocytes, where haem biosynthesis occurs, must exist. Recently, a porphomethene inhibitor molecule was identified in two murine models of PCT (Phillips *et al.* 2007). This molecule is produced by the incomplete oxidation of uroporphyrinogen to uroporphyrins. In PCT patients, uroporphyrinogen urinary excretion levels (especially uroporphyrinogen I) are greatly increased due to a decrease in UROD activity, usually exceeding 1,000µg/day compared to the average <50µg/day (Smith *et al.* 1980). The accumulation of uroporphyrinogens causes the excessive generation of the porphomethene inhibitor leading to a further decrease in UROD activity and ultimately the clinical manifestation of PCT.

The mechanism of UROD inhibition by the porphomethene molecule is strongly suggested to be iron-dependent (Elder *et al.* 1985; Phillips *et al.* 2007). This speculation is validated by the presence of normal UROD activity upon phlebotomy-induced clinical remission (Elder *et al.* 1985). Iron overload is invariably observed in PCT patients (Felsher 1977; Lundvall *et al.* 1970; Elder and Roberts 1995). Thus, the generation of a molecule which inhibits UROD activity *via* an iron-dependent pathway is substantiated.

The mechanism by which iron accumulates in PCT patients is not clear. Approximately 20% of PCT patients are homozygous for the C282Y *HFE* mutation and a further 15% are heterozygous for the mutation (Bulaj *et al.* 2000). In these patients, the iron overload can be attributed to the presence of these mutations. In the remaining cases, the cause of iron overload is unknown. It thus seemed viable to investigate the role of other genes involved in iron metabolism in PCT. Should variants identified in these genes be positively correlated with iron overload in PCT patients, it may provide a platform for the elucidation of iron overload associated with PCT.

GENES INVOLVED IN IRON HOMEOSTASIS INVESTIGATED IN PCT

CERULOPLASMIN (CP)

Mutation analysis of the 1130 bp *CP* promoter region revealed seven previously described SNPs (-567C/G, -563T/C, -439C/T -364delT, -354T/C, -350C/T, -282A/G) (Table 3.1) (Appendix 3).

The entire *CP* promoter region of the rat (*Rattus norvegicus*) has also been characterised (Fleming and Gitlin 1992). A critical region of approximately 300 bp is necessary for the activation of transcription and is highly conserved in humans. In the rat, possible repressor activity was revealed through the identification of negative *cis* elements in the region -5200 to -732. In contrast, regions -198 to -103, -308 to -269 and -398 to -348 had a positive effect on gene expression. The region between -398 and -348 shows 80% homology to the albumin D site, a site known to bind C/EPB α . Interestingly, three of the seven variants (-364delT, -354T/C, -350T/C) identified in this study lie within these known C/EPB α binding motifs.

Apart from C/EPB α , many putative TFBSs were also predicted for the HNF1 α , HNF3 β , and HNF4 α motifs in the *CP* promoter region (Table 3.5). Competitive binding of the HNFs to the critical region of the rat promoter has not been observed (Fleming and Gitlin 1992). This does not rule out the possibility that the HNF motifs may exist in the human promoter region as the two promoter regions are not 100% homologous. Although these TFs are known regulators of gene expression in the human liver, their relative abundance in the liver does not allow an accurate forecast of the potential role of TF binding at one motif.

The -567C/G variant was predicted to be situated within a putative YinYang-1 (YY1) binding site. A C to G transversion at this site causes a predicted disruption of the YY1 TFBS. YY1 is a zinc finger protein and is unusual in that it can regulate transcription in three ways: by repression,

activation or initiation (Seto *et al.* 1991; Shrivastava and Calame 1994). YY1 has been implicated in iron metabolism, especially because of its known interaction with transferrin (Adrian *et al.* 1996). YY1 binding capabilities, presumed to increase with age, have a negative regulatory effect on transferrin transcription causing decreased serum transferrin levels associated with the aged. YY1 also has a role in the liver, where it binds to the acid alpha-glucosidase (*GAA*) gene, and may cause transcriptional silencing (Yan *et al.* 2001). The putative core recognition sequence (5'-CCATNTT-3') (positioned at -568 to -562) of YY1 in *CP* spans the -567 nucleotide position (highlighted in bold) with only one nucleotide change (highlighted in *italics*) (5'-CCATNTA-3'). The flanking sequences of the core sequence generally determine the nature of regulation (repression, initiation or activation) by YY1, although this is not clearly defined (Seto *et al.* 1991).

A study performed on the rat (*Rattus norvegicus*) demonstrated that the *CP* promoter region (-1800 bp to -1 bp) contained a TBFS for YY1 and that the core sequence was characteristic of such sequences of genes that are expressed in the liver (Lee *et al.* 2001, Gyulikhandanova *et al.* 2004). Since the *CP* gene is highly conserved (97,8%) between the rat and human (Nishiyana *et al.* 2003), and has an elucidated involvement in iron metabolism, it reiterates the possibility of a YY1 TBFS in *CP*. A C/EBP α TFBS was also anticipated to be disrupted in the presence of the variant allele. Although this particular C/EBP α site does not lie within the regions where these TFBSs have been determined to lie within the rat *CP* promoter, the role of this TF should not be underestimated.

The -567C/G polymorphism was also expected to create a novel putative binding site for the GATA transcription factor. The -282A/G polymorphism is also expected to create a new binding site for the GATA transcription factor. As mentioned previously, the GATA binding proteins have defined roles in haematopoiesis (Tsai *et al.* 1994). GATA 1 plays a key role in erythropoiesis and is known to target genes in the haem biosynthesis pathway (Ferreira *et al.* 2005), whilst GATA 2 controls the

proliferate capabilities of haematopoietic progenitor cells (Tsai *et al.* 1994). GATA 3 is expressed in the haemopoietic system, but its function is less well defined (Ferreira *et al.* 2005). The zinc finger DNA-binding domains of each GATA TF (GATA-1, 2, 3) recognise a slight variation of the core sequence (5'-AGATAA-3') (Ko and Engel 1993), hence the various software programs were unable to distinguish precisely which GATA binding site was created. Although the potential role of the GATAs in *CP* expression is not known, these TFs do play a key role in genes expressed in the liver. It is therefore viable to research the potential implication of GATA binding to *CP* mRNA. A putative oestrogen receptor (ER) binding site was expected at nucleotide position -439. The core binding sequence, referred to as an oestrogen responsive element (ERE), consists of a 13 bp inverted repeat (5'- GGTCANNNTGACC-3') (Driscolli *et al.* 1998). The flanking sequences also play an important role in the interaction and binding capabilities of ERE and ER. A single nucleotide deviation from the core sequence may prevent binding, whilst in contrast two nucleotide changes may have no detrimental effect on TF binding. It is therefore clear that ER-ERE binding is a highly intricate event.

There is no documentation supporting the notion that *CP* possesses an ERE element. However, a recent murine *in vitro* study showed that *CP* expression is oestrogen-responsive during uterine growth and differentiation (Stuckey *et al.* 2006), increasing the likelihood that *CP* may possess an ER TFBS. Since oestrogen is a known risk factor in PCT, it may seem feasible to perform follow up expression studies. If it is found that PCT patients carrying a *CP* gene mutation display iron overload and are known oestrogen consumers, it could imply that oestrogen induced dysregulation of *CP* causes iron overload frequently observed in PCT patients.

An activator protein 1 (AP1) site in the region spanning -439 was also predicted by *in silico* analysis. The substitution at position -439 (C to T) abolishes the putative AP1 TFBS. AP1 is not a

single protein; instead it complexes to basic region-leucine zipper proteins that belong to the JUN, FOS, MAF and ATF oncogene subfamilies (Shaulian and Karin 2002). AP1 controls cell proliferation, differentiation, apoptosis and tumorigenesis. In murine studies it was shown that C-jun complexes with AP1 to control hepatogenesis (Hilberg *et al.* 1993). A novel AP1 site was detected 3.7 kb upstream of the translation initiation site in the human *CP* promoter region (Das *et al.* 2007), indicating a degree of functional importance between AP1 and *CP* regulation. Upon mutational analysis (functional studies), the putative AP1 sites spanning the -439C/T SNP identified in this study may therefore reveal an underlying importance in the regulation of *CP* in PCT patients.

The -439 C to T change results in either the creation or disruption of a C/EBP α recognition site in the immediate region of the -439C/T variant. Putative C/EBP α binding sites were also predicted in the regions of -364 and -354. In the presence of the -364delT and -354T/C variants, these TFBSs are expected to remain intact. All three mentioned variants lie within predicted C/EBP α binding motifs in the rat. This is noteworthy, since C/EBP α has a positive regulatory effect on *CP*. Should expression studies verify that the -439C/T disrupts a C/EBP α TFBS, it could potentially provide an explanation for the mechanism of iron overload under conditions of defective regulation of *CP*.

The -354 T to C nucleotide change was expected to abolish an AP1 site. Since AP1 binding has been implicated in *CP* gene regulation, it is recommended that mutational analysis be performed in future. A HNF3 β core motif is expected to be disrupted in the presence of the C allele. Considering that this TF is required for the co-ordinated expression of several hepatic genes (Pani *et al.* 1992), it seems feasible to investigate the possible role of this TF in *CP* expression.

A putative C/EBP α binding site was predicted at position -282. This site was shown to be disrupted in the presence of the -282A/G transition.

CYBRDI

Seven novel (-1540G/A, -1477G/A, -1452T/C, -1346T/C, -1272T/C, -645T/C and the [G(T)₈ G(T)_n G(T)_n G(T)₉] repeat unit) and ten previously described variants (-1849T/G, -1844G/C, -1834G/A, -1813C/T, -1459T/C, -849C/G -633 G/A, -238G/A, -167C/G, -163G/A) were identified in the 2122 bp region of the *CYBRDI* promoter (Table 3.1).

While the function of *CYBRDI* has recently been established, there is no literature providing possible regulatory mechanisms of *CYBRDI* expression.

The -1834G/A variant is located in the region of the following putative TBFSs: spleen focus forming integration oncogene SPI1 (PU.1), specificity protein 1 (SP1) and growth factor-independent 1 (Gfi1) (Table 3.3). The variant A allele was predicted to create a new PU.1 (5'-GGAA-3') and Gfi1 (5'-GGAA-3') binding motif. PU.1 is a member of the tissue-specific ETS family and the strict regulation of this TF is crucial for the generation of haematopoietic lineages (DeKoter 1998, Rosenbauer *et al.* 2004). Gfi1 represses proliferation and maintains the integrity of haematopoietic stem cells (Hock *et al.* 2004). Since the variant allele was more frequent in controls than in patients, the function of this putative TFBS may be necessary for well-being and may even have a protective effect. Alternatively, the polymorphic allele is expected to abolish an existing SP1 and GATA1 binding site. SP1 contains zinc finger motifs and is a positive regulator of transcription. It has been implicated in iron metabolism; a decrease in its binding activity to the transferrin promoter causes a noticeable downregulation of transferrin expression (Adrian *et al.* 1996). Thus the role of these TFs spanning the region of this significant variant should be investigated in future studies; in particular in the Caucasian population where a statistically significant association was noted.

Two putative TFBSs, PU.1 and USF, were predicted to lie within the region of the -1813C/T variant. Upon visual comparison of the PU.1 core sequence with the sequence spanning the variant, it was noted that the TFBS lies in reverse orientation (present on the lagging strand) and that in the presence of this variant a recognition site was created. A USF binding motif was also predicted to be created in the presence of the T allele. As USF has an active role in hepatic regulation (Vallet *et al.* 1998, Casado *et al.* 1999) including *HAMP*, it is most certainly feasible to investigate the functionality of this TF in *CYBRDI* expression, especially in the Black population, as a statistically significant association was found in this ethnic group.

The sequence spanning the -1540G/A variant contains a core motif for the GATA 2 transcription factor. The variant allele is predicted to abolish the recognition motif. Considering the role of GATA 2 in the regulation of genes expressed in the liver, it is deemed feasible to investigate the functionality of this TF.

The disruption of the putative GATA 1 and SP1 recognition motif is predicted in the presence of the -1346T/C variant. Alternatively, the C allele is expected to create a PU.1 site. Given that these TFs have a proposed function in liver regulation, it would be feasible to pursue an investigation into the role of these TFs in *CYBRDI* gene regulation in especially the Coloured population where a statistically significant association was determined. The variant C allele of -1272T/C was predicted to create a new recognition site for the C/EBP α , a TF well characterised for its function in regulating genes expressed in the liver.

It is significant that the -645T/C transition is predicted to disrupt putative SP1 and NF-E2 binding motifs. NF-E2 plays a role in haem synthesis and iron procurement and has documented recognition sites in the promoter region of two haem biosynthesis genes, ferrochelatase (*FECH*) and

porphobilinogen deaminase (*PBGD*) (Peters *et al.* 1993). This variant could prove to be significant in populations that do not depart from HWE. Under these conditions, future studies encompassing functional analysis investigating the possible role of NF-E2 and SP1 in *CYBRDI* expression is highly recommended.

The presence of the -624G/A variant causes disruption of the core motif of a putative runt-related protein 1 (RUNX1) TFBS. RUNX1 plays a critical role in haematopoiesis; it is primarily responsible for the generation of the first definitive haematopoietic stem cells in the vasculature of the mouse embryo (Ono *et al.* 2007). RUNX1 recognises the core sequence 5'-TGT/cGGT-3' which closely resembles the sequence spanning the variant -636 to -631 (5'-TGGGGT-3') (normal allele highlighted in bold). Since the variant was shown to be significant in Blacks, it would be worthwhile to further investigate the relationship between RUNX1 and the effect of the AA genotype on *CYBRDI* expression. The variant A allele is also predicted to create a novel SP1 binding motif.

The -238G/A variant is located within a nuclear factor Y (NF-Y) TFBS. This ubiquitous TF which recognizes the CCAAT-box, has a role in haemoglobin synthesis (Liberati *et al.* 1998) and enterocyte differentiation (Bevilacqua *et al.* 2002). The variant A allele creates the core sequence, promoting NF-Y binding (Bucher 1990). NF-Y binding motifs often occur in proximal promoter regions (Bellorini *et al.* 1997; Frontini *et al.* 2002), near the sites where the basal transcription machinery associates. NF-Y may also interact with nucleosomes, thereby manipulating chromatin structure and enhancing binding of transcriptional activators (Motta *et al.* 1999, Coustry *et al.* 2001; Romier *et al.* 2003). Therefore, it is noteworthy that a putative NF-Y binding site has been predicted at position -238, fairly close to the transcription initiation site. Furthermore, several lines of evidence show that NF-Y may interact with several transcription factors involved in liver

regulation, such as C/EBP (Milos and Zaret 1992), HNF4 (Ueda *et al.* 1998), SP1 (Yamada *et al.* 2000) and CREB (Eggers *et al.* 1998).

The A allele, required for the putative binding of NF-Y, is present in the minority of the general populations, but has an increased frequency in the patient groups of the current study, indicating that the allele may have a protective effect. Since the precise role of *CYBRDI* is not clear, the potential influence of this variant on expression of the gene and in iron metabolism remains elusive but warrants further investigation.

A repeat unit, 1752 bp upstream from the translation initiation ATG codon, was identified in all three population groups. No literature could be found describing this repeat unit and it was not identified on HapMap. For this reason, it is proposed as a novel repeat unit in this study. Three alleles were distinguished: allele *1 G(T)₈ G(T)₆ G(T)₆ G(T)₉, allele *2 G(T)₈ G(T)₆ G(T)₉ and allele *3 G(T)₈ G(T)₈ G(T)₆ G(T)₉. Individuals were either homozygous for the G(T)₈ G(T)₆ G(T)₆ G(T)₉ or G(T)₈ G(T)₆ G(T)₉ repeat units or heterozygous for the G(T)₈ G(T)₈ G(T)₆ G(T)₉/G(T)₈ G(T)₆ G(T)₉. The following configurations were observed: *1*1, *2*3 and *2*2. Repeat units have long been associated with disease (Frontali *et al.* 1999). An unstable expansion of repeat units above the average threshold may lead to disease. In the regulatory regions of genes, repeat units may influence mRNA transcription (Usdin and Grabczyk 2000). A (GT)_n repeat unit in the promoter region of the solute carrier family 11 member 1 (*SLC11 A1*) (*NRAMP1*) gene has been implicated in the susceptibility to infectious and autoimmune conditions (Blackwell *et al.* 1995). Thus, considering that repeat units in genes involved in iron metabolism have been implicated in other disease states, further investigation into the potential role of this *CYBRDI* repeat unit in PCT is warranted.

Several of the aforementioned TFs play a role in iron homeostasis, haematopoiesis and liver

regulation. It is therefore suggested that they may have a potential role in the regulation of *CYBRDI* expression. Until the transcriptional regulation of *CYBRDI* has been clarified, no substantial conclusions can be drawn at this point in time.

HAMP

Two previously described (-1010C/T, -582A/G) and one novel variant (-429G/T) were identified in the 1084 bp region of the *HAMP* promoter analysed in the present study (Table 3.2). A single variant, -582A/G, demonstrated a potential statistically significant association.

Experimental studies on 950 bp of the human *HAMP* promoter demonstrated that the promoter region (Courselaud *et al.* 2002) contains a sequence displaying TATA box homology (Courselaud *et al.* 2002). Several putative motifs for liver-enriched TFs were predicted in the regulatory promoter area, namely HNF4 α (-75 to -62 and -602 to -589), C/EBP α (-249 to -236, -298 to -285 and -90 to -77) and a STAT binding site, which overlaps the -90 to -70 C/EBP α binding site. Furthermore, negative *cis*-acting elements in the region -960 to -633 were revealed. The murine *HAMP* promoter region was also characterised and shown to contain the same regulatory motifs. It was shown that HNF4 α does not possess any activating capabilities but rather plays a role in the negative regulation of hepcidin. More interestingly, C/EBP α binds to the promoter and *trans*-activates the human and murine *HAMP* promoter *in vivo*. Mice lacking C/EBP α TFs showed decreased levels of hepcidin mRNA accompanied by iron overload in hepatocytes. Iron overload in turn induces C/EBP α expression. The link between C/EBP α and *HAMP* expression and hepatic iron overload indicates that cross talk occurs between the C/EBP α signalling pathway and iron metabolism.

More recently, several E-box motifs were identified in the *HAMP* promoter region (Bayele *et al.*

2006). E-boxes have a hexanucleotide consensus sequence (5'-CANNTG-3') that is recognised by members of the basic helix-loop-helix leucine zipper (bHLH-ZIP) family of transcription factors (such as USF1, USF2, c-Myc/Max) (Atchley and Fitch 1997). In particular, two canonical E-boxes (5'-CACGTG-3') were identified at positions -529 to -524 and at -457 to -452 and two E-boxes (5'-CACATG-3') at -230 to -225 and at -786 to -781. These E-box motifs lie within the regions of the C/EBP α binding sites.

In vitro experiments suggested that the human and murine *HAMP* promoters are strongly regulated by the USF1/USF2 heterodimers (Bayele *et al.* 2006). The study demonstrated that USF1 may *trans*-activate *HAMP* expression whilst USF2 may *cis*-activate *HAMP* expression. Although USF1 may activate the promoter independently of USF2, or *vice versa*, it is more likely that the dimers operate as heterodimers. This is supported by the fact that heterodimers represent more than two thirds of USF binding activity (Violett *et al.* 1996). USF2 association with *HAMP* has already been elucidated (Nicolas *et al.* 2001). Absolute inhibition of *HAMP* expression and resulting iron overload was demonstrated in USF2^{-/-} knock-out mice. The iron overload seen in these mice resembled that observed in human HH patients and in *HFE* knock-out mice (Zhou *et al.* 1998; Bahram *et al.* 1999; Levy *et al.* 1999).

A USF1 binding site was predicted in the region spanning the normal A allele of the -582 A/G variant. The variant G allele is expected to abolish the USF1 recognition motif (5'-CACGTG-3') (normal A allele highlighted in bold). The core sequence was identical to the reference sequence at nucleotide positions -583 to -578. Although a functional USF1 binding site has not been identified in this region by previous studies, it does not exclude the possibility that USF1 may bind to this region. It is recommended that mutational analysis be performed to investigate the functional status of the predicted abolished USF1 binding site. Further research on this promising find is encouraged,

particularly in Blacks and in Coloured females where statistically significant differences between patients and controls were demonstrated, in order to determine a possible association between the polymorphism and iron overload in PCT patients.

SLC40A1

The current study screened the 1653 bp region of the *SLC40A1* promoter for the presence of variation. Seven previously described (-1355C/G, -750G/A, -622T/C, -501C/T, -98G/C, -8C/G, (CGG)_n) and three novel (-1461T/C, -1399G/A, -524C/T) variants were identified (Table 3.2). There are no records in the literature indicating potential functional implications of the previously described variants on the IRE of *SLC40A1*.

The -1461T/C transition was identified in patients only in the Black population and this variant is predicted to disrupt two TFBSs: PU.1 and Gfil. Both these TFs recognise a (5'-GGAA-3') sequence. These TFs are involved in the generation and maintenance of haematopoietic lineages (DeKoter 1998, Hock *et al.* 2004; Rosenhauer *et al.* 2004). However, upon visual inspection, the sequence spanning the variant did not contain a sequence at all homologous to the core recognition sequence. It is thus unclear what the role of this TF in *SLC40A1* may be, and if the TF binds at all.

A PU.1 TFBS was predicted to be abolished in the presence of the -750G/A variant. It is recommended that a larger sample size for the Black population be screened for this variant; if the variant still displays a statistical significant association with PCT, it would be feasible to investigate the possibility of *SLC40A1* dysregulation by PU.1.

The -501 C to T transition was predicted to disrupt the core sequence of the SP1 transcription factor (JASPAR CORE). In contrast, according to the ALIBABA2 program, the TFBS would remain

intact in the presence of the variant. SP1 recognises a GC rich sequence (Tamaki *et al.* 1995). The presence of the T nucleotide could therefore be predicted to disrupt the GC rich area spanning the variant. However, whether the presence of the T nucleotide is sufficient to abolish the core recognition sequence is not certain. SP1 has been implicated in iron metabolism and has been described to bind to the transferrin promoter region (Adrian *et al.* 1996). The possibility that the potential disruption of the SP1 TFBS could cause *SLC40A1* dysregulation, should be further investigated, particularly in the Caucasian population.

The -98G/C variant has been recorded earlier (Douabin-Gicquel *et al.* 2001) and was detected in all three population groups in the current study. In a previous study, Zaahl *et al.* (2004) identified the -98G/C variant in the South African Black and Caucasian groups. The variant allele frequency was higher in Caucasian patients presenting with primary iron overload (0.20) than in unaffected controls (0.15). In Blacks, the variant allele was only detected in unaffected controls (0.08). The present study noted a similar frequency in Caucasian controls (0.17). In contrast, the variant was not detected in Black controls, but seen in patients. The homozygous form of the variant was not reported in the earlier study on South African subjects (Zaahl *et al.* 2004). Here, the homozygous state was detected in Caucasians and Coloureds, in patients as well as controls. A novel RUNX1 TFBS was expected to be created in the presence of the -98G/C variant. RUNX1 is the DNA binding element of the RUNX1/CBF β complex and binds to core enhancer sequences present in the promoter regions of several genes involved in haematopoietic cell differentiation (Cleary 1999). Furthermore, reciprocal chromosomal translocations involving both RUNX1 and CBF β have been implicated in leukemogenesis (Zhang *et al.* 1997). Although there is no evidence supporting the role for RUNX1 in *SLC40A1* expression, a possible involvement cannot be ruled out.

The previously recognised -8C/G variant (Douabin-Gicquel *et al.* 2001) was identified in all three population groups. In a former study on South African Blacks and Caucasians, the -8 C/G variant was detected in only the heterozygous form in patients with primary iron overload and controls (Zaahl *et al.* 2004). In the current study, the variant homozygous state was not seen in Black patients and controls and Coloured patients, but it was detected in Caucasian patients and controls, males as well as females. The G allele was predicted to cause the disruption of a GATA 2 TFBS. This TF controls the proliferation of haematopoietic progenitor cells (Tsai *et al.* 1994). The potential role of GATA 2 in *SLC40A1* expression should be investigated, especially in the Caucasian population as the variant was noted to display a statistically significant difference between patients and controls.

A previously described (CGG)_n repeat unit (Lee *et al.* 2001) was identified in all population groups analysed here. The repeat unit was located 330 bp upstream from the translation initiation ATG codon. Allele *1 represents seven triplet repeat units [(CGG)₇] and allele *2 (the variant allele) represents eight triplet repeats [(CGG)₈]. Heterozygous individuals and homozygotes for both configurations were observed. In the Caucasian population group the variant allele was more prevalent amongst patients than controls, suggesting that this repeat unit [(CGG)₈] may be associated with PCT. A single Coloured male patient showed (CGG)₇ on one allele and (CGG)₉ on the other allele. A previous study reported that this microsatellite (CGG)₉ could possibly contribute to iron overload in the Black South African population (Zaahl *et al.* 2004). Repeat units, particularly trinucleotide repeats, have been implicated in genetic diseases (Frontali *et al.* 1999). One example is the expansion of the (CGG) repeat in the 5'UTR of the fragile X mental retardation 1 (*FMRI*) gene responsible for Martin-Bell (fragile-X) syndrome (Devys *et al.* 1993). It is strongly recommended that the implication of this CGG repeat unit in iron metabolism be investigated in larger populations.

GENERAL DISCUSSION

In the current study an initial investigation of four genes implicated in iron homeostasis, revealed several promising findings. The C282Y mutation status of the patients in the current study was not available. Although no association could be made between iron overload and any of the identified variants due to lack of clinical information i.e. iron parameters, several statistically significant associations were determined between variants and PCT. In *CP*, variation was identified in all three population groups, but was predominantly confined to Blacks. Variation in *CYBRDI* was detected in all three ethnic groups, with Coloureds and Caucasians representing the population groups containing the most and Blacks the least amount of variation. A statistically significant association was noted for one *HAMP* variant (-582A/G) in a single population group (Blacks). Variation in *SLC40A1* was detected in all three population groups.

Only three variants were identified in the *HAMP* promoter; a relatively low frequency compared to the remaining three genes. Considering the role of *HAMP* is critical in iron metabolism, and that the gene sequence is fairly conserved in vertebrates (Chen *et al.* 2005; <http://www.ncbi.nlm.nih.gov/BLAST/>) it can be suggested that in the event that variation does occur, it will have a detrimental effect on expression. Of all the types of haemochromatosis, juvenile haemochromatosis (JH), characterised by the presence of *HAMP* or *HJV* mutations (Papanikolaou *et al.* 2004; Roetto *et al.* 2003), has the earliest disease onset. JH patients have severe iron overload and symptoms such as hypogonadotropic hypogonadism, cardiac arrhythmias and intractable heart failure (Cazzola *et al.* 1983). This demonstrates the crucial role of *HAMP*, and that variation in this gene may have serious health consequences.

Several environmental and precipitating factors have been implicated in the development of PCT. Different populations may be susceptible to the effects of alternate precipitating factors. For

example, the *HFE* C282Y gene mutation is more prevalent in the northern European countries (Merryweather-Clarke *et al.* 2000), whilst HCV infection is more prevalent in southern Europe (Thomas *et al.* 2000). Therefore the respective populations are more likely to be exposed to those specific exacerbating factors; both of these conditions are known risk factors for PCT. In comprehending the complex aetiology of disease, two integral components, i.e. gene-gene and gene-environment interactions must be grasped first. Gene-gene interaction generally describes the relationship between two or more genes or modifier loci, where the inheritance of an allele at one locus may modify the action of another allele at another locus. The interactions between these modifying loci are responsible for variable phenotypic expression, often characteristic of complex diseases (Rutter 2007).

PCT patients are often C282Y carriers, which in conjunction with an inherited *UROD* gene mutation may increase an individual's susceptibility to PCT. This is in essence an example of gene-gene interaction. Other well described diseases where gene-gene interaction plays a fundamental role include atrial fibrillation, diabetes type 2 and hypertension (Moore 2004).

Documented evidence depicts interactions between genes investigated in this study. In particular, *HAMP* interacts on a molecular level with *SLC40A1* (Nemeth *et al.* 2004). *HAMP* has been identified as a key regulator of iron homeostasis. *HAMP* expression in the liver is developmentally regulated; the highest levels of expression occurring shortly before birth and during adulthood (Courselaud *et al.* 2002). This seems to be in accordance with the functional status of the liver at various developmental stages. *HAMP* expression is positively regulated by iron overload, inflammation and high dietary iron. Conversely, *HAMP* expression is attenuated under conditions of hypoxia and anaemia.

It was recently determined that hepcidin expression is also affected by alcohol (Harrison-Findik *et al.* 2006); one of the most frequent precipitating factors implicated in PCT. *In vivo* studies demonstrated that mice treated with ethanol displayed significant down regulation of *HAMP*. Ethanol metabolism in the cell results in the generation of ROS. Alcohol mediated down-regulation of *HAMP* expression was alleviated with the subsequent administration of Vitamin E, which abolished the effect of alcohol on the binding activity of C/EBP α to *HAMP*. For these reasons it would seem feasible to correlate the iron parameters with the alcohol consumption of each patient that possess a *HAMP* gene variant identified in the current study.

The *SLC40A1* gene codes for ferroportin, the sole iron exporting protein identified to date. Ferroportin is responsible for releasing iron at the basolateral membrane of enterocytes, in the liver and macrophages. Ferroportin has also recently been described to manipulate iron uptake from the enterocytic apical membrane (Thomas and Oates 2002). *SLC40A1* expression is regulated by changes in body iron status, hypoxia, inflammation and erythropoiesis. *SLC40A1* gene mutations have been implicated in primary iron overload disorders such as Type IV HH or ferroportin disease (Montosi *et al.* 2001; Njajou *et al.* 2001; Cazzola *et al.* 2002; Camaschella 2005). Patients with this disease display reticuloendothelial iron overload and retarded or non-existent reticuloendothelial iron release (Knutson *et al.* 2003). The HH phenotype resulting from *SLC40A1* mutations, is inherited in an autosomal dominant manner (Schimanski *et al.* 2005). It is, however, not clear how these mutations lead to the development of HH. One model suggests that an increase in *SLC40A1* expression as a result of a gain-of-function mutation leads to an increase in iron transport out of the enterocytes, causing an iron overload phenotype characterised by high transferrin saturation and high serum ferritin levels. These findings reveal that *SLC40A1* plays a fundamental role in iron metabolism.

In response to iron loading, HAMP causes SLC40A1 to be internalised and degraded, thereby preventing iron export from the enterocytic and hepatic cells and iron release from macrophages post erythrophagocytosis (Pigeon *et al.* 2001; Nemeth *et al.* 2004). This inverse and complex relationship between the *SLC40A1* and *HAMP* gene and protein products is thought to involve redox signalling mechanisms (Yeh *et al.* 2004). The post-translational regulation of *SLC40A1* by hepcidin may complete a loop/link in iron homeostasis. *HAMP* may also interact with other genes involved in iron metabolism. Modifier loci in *HAMP* and *HFE* (C282Y modifier locus) have been implicated in HH (Merryweather-Clarke *et al.* 2003).

It has been put forward that CP functions in concert with ferroportin, the only protein responsible for iron export from the liver (Roeser *et al.* 1970; Frieden and Hsieh 1976; Mukhopadhyay *et al.* 1998; Roy and Andrews 2001). CP is a liver expressed protein but exerts its function in the plasma surrounding storage cells, i.e. hepatocytes and macrophages, but not enterocytes (Roeser *et al.* 1970). Hepatocyte and reticuloendothelial cell iron overload occurs in *CP*^{-/-} knock out mice (Harris *et al.* 1999). Following iron export by SLC40A1, CP oxidises ferrous iron to ferric iron required for transferrin loading. In the event of *CP* dysregulation, it can be hypothesised that this protein will not oxidise ferrous iron efficiently, resulting in an increase in serum ferritin iron. Saturated ferritin levels have been observed in PCT patients, and is characteristic of HH. Due to the increase in serum ferritin levels and the decrease in oxidation of iron to the ferric form, it is postulated that a signalling or feedback mechanism exists that prevents ferroportin from releasing iron from storage cells, causing iron accumulation in storage cells (hepatocytes, macrophages). This would account for the hepatic iron overload in PCT patients. This hypothesis is supported by evidence of hepatic iron overload observed in a patient who was a compound heterozygote for two novel *CP* mutations (Bosio *et al.* 2002).

To complicate matters, HH associated dominant *SLC40A1* mutations can create a loss-of-function or gain-of-function effect. In the event of a gain-of-function mutation, the signalling mechanism between HAMP and SLC40A1 is disrupted, preventing HAMP from degrading ferroportin. This results in continued iron export out of the cell and causing elevated transferrin saturation levels. On the other hand, the loss-of-function mutation causes a disruption in ferroportin's basal activity, preventing iron export and causing iron accumulation in storage cells. This study revealed the following conceivably important candidate loci that are likely to be involved in *CP* and *SLC40A1* repression. There were two *CP* variants: -567C/G disrupts a YY1 putative TFBS and -439T/C disrupts putative ER, AP1 and C/EBP α binding motifs. In *SLC40A1* there were also two variants of note: -501C/T disrupts a SP1 TFBS and -8C/G disrupts a GATA 2 binding motif. Functional analysis are needed to prove whether one or more of these variants may be accountable (either solely or in conjunction with each other) for the repression of *CP* and *SLC40A1* expression. This research may explain how iron overload comes about in PCT patients.

CYBRD1 is a membrane protein responsible for the reduction of ferric iron to ferrous iron, enabling absorption of dietary iron into enterocytes, although the precise mechanism remains unclear (McKie *et al.* 2001; Frazer *et al.* 2005; Gunshin *et al.* 2005). When the body requires more iron and when serum iron levels are low, *CYBRD1* regulation is increased. This is supported by the knowledge that hypoxia, iron deficiency and hypotransferrinaemia are all modulators of *CYBRD1* expression. Thus, during states of iron overload, *CYBRD1* expression appears to be down-regulated. It is thought that other ferrireductases may exist to reduce ferric iron to ferrous iron. This may complicate the elucidation of *CYBRD1* expression patterns in states of iron overload and iron deficiency. However, it is strongly hypothesised that *CYBRD1* expression is diminished under conditions of iron overload. It is speculated that the promoter variants identified in this study may disrupt a TFBS, imperative for TF binding to decrease expression under the iron overload

conditions observed in PCT patients. In this way *CYBRDI* expression is not diminished and subsequently dietary iron absorption is not impaired. This leads to excessive iron accumulation. In essence the variant could disrupt the negative feedback regulatory circuit that exists between *CYBRDI* expression and iron levels. This may explain the iron overload observed in PCT patients who possess *CYBRDI* variants. Alternatively, a variant may create a novel TFBS, which promotes increased *CYBRDI* expression, so that dietary iron absorption continues even though the body's iron requirements have been substantiated.

Single nucleotide substitutions may abolish or create a TF binding motif and therefore potentially alter gene regulation and expression. Defects or malfunctions in the intricate DNA-protein interaction between a TF and its binding motif can lead to human disease (Bulyk 2003). Examples include the -463G/A SNP in the Myeloperoxidase (*MPO*) gene associated with the atherosclerotic process of end-stage renal disease (ESRD) (Maruyama *et al.* 2004; Nordfors *et al.* 2005) and the -1082G/A SNP in interleukin-10 (*IL-10*) gene associated with inflammation in infectious diseases and systemic lupus erythematosus (SLE) (Reuss *et al.* 2002). Transcriptional regulation and initiation is a complex process involving many transcription factors working alone or in concert with one another. Thus the ability or inability of one TF to bind to a motif may or may not cause an alteration in basal transcription levels. If however a binding motif central to the initiation or regulation of transcription is disrupted, gene expression will be dysregulated, resulting in disease.

Computational analysis alone cannot conclusively determine the potential effect of a single nucleotide change on the integrity of a binding motif (Guo and Jamison 2005). Only experimental functional studies can verify or reject the predicted outcome. The *in silico* analysis performed in this study has highlighted several putative binding motifs, which a few have previously described functional importance in the regulation of *CP*, *CYBRDI*, *HAMP*, *SLC40A1* and of other genes

involved in iron metabolism or in liver specific genes. Although only a small porportion of these presumed binding sites predicted by *in silico* analysis prove to be functionally important (Fickett and Wassermann 2000), the methodology represents an appropriate predictive basis suitable for the design of follow-up studies.

Whilst bioinformatic methodologies for the investigation of gene-gene interactions have not been optimised for the large scale analysis of unlinked markers, statistical tools are currently being developed. This will be a groundbreaking advancement, expected to enrich the current limited understanding of gene-gene interactions and to enable a more accurate interpretation of how these gene interactions contribute to the aetiology of complex diseases. It is highly recommended that gene-gene-interaction analysis be performed in future PCT investigations, particularly between *HAMP* and *SLC40A1* due to their documented interplay and between *SLC40A1* and *CYBRD1* due to their juxtaposed chromosomal position (2q31-32).

The study presented here highlighted several tentative significant associations between PCT and a number of polymorphic loci in all four genes analysed. Future studies investigating the genotype-phenotype correlation between iron levels and these variants will undoubtedly expand the current comprehension of this multifactoral disease. In various instances presence of the polymorphic allele was considered to affect TF binding, in the event producing altered gene expression. A more thorough understanding of TFs and their binding motifs as well as their interactions with one another will provide a more accurate interpretation of regulatory pathways within biological systems. An enhanced comprehension of all these aspects is central in characterising complex diseases.

CHAPTER 4:
CONCLUSIONS AND
FUTURE PROSPECTS



4. CONCLUSIONS AND FUTURE PROSPECTS

PCT is characterised by a defect in uroporphyrinogen decarboxylase activity (UROD) in the fifth step of the haem biosynthesis pathway, causing a hepatic and cutaneous accumulation of porphyrins. A number of genetic and environmental precipitating factors contribute to the clinical manifestation of this complex disease, which displays phenotypes of varying intensities. Acute cutaneous appearances include blisters and lesions, few in number; whilst chronic cutaneous appearances may include scars, hyperpigmentation and alopecia. Severe skin disease in HEP may manifest as disfigured, blotchy and coarse scars with eventual sclerodactyly. Iron overload, a significant exacerbating factor observed in approximately 80% of PCT patients, is to some extent characteristic of the iron overload seen in HH patients. The aim of the study presented here was to establish whether variants in genes involved in iron metabolism are associated with PCT. Variants that demonstrate a positive correlation with PCT or affect TFBSs, thereby altering gene expression, may conceivably be implicated in the aetiology of this disease.

The first objective of this study was to perform mutation analysis in order to identify novel and/or previously described variants in the promoter region of four genes (*CP*, *CYBRDI*, *HAMP* and *SLC40A1*) known to be associated with iron accumulation in a variety of iron loading diseases. The methods included PCR amplification, HEX-SSCP and RFLP analysis, and bi-directional semi-automated DNA sequencing analysis. The following results were obtained: 11 novel variants (*CYBRDI*: -1540G/A, -1477G/A, -1467T/C -1346T/C, -1272T/C,-645T/C; *HAMP*: -429G/T; *SLC40A1*: -1461T/C, -1399G/A, -750G/A, -524C/T) and 23 previously described variants (*CP*: -567C/G, -563C/T, -439C/T, -364delT, -354T/C, -350C/T, -282G/A; *CYBRDI*: -1849T/G, -1844G/C, -1834G/A -1813C/T, -1459T/C, -624G/A, -238G/A, [G(T)₈G(T)_nG(T)_nG(T)₉]); *HAMP*: -1010C/T, -582A/G; *SLC40A1*: -1355C/G, -622T/C, -501C/T, -98G/C, -8C/G, [(CGG)_n]) were identified.

Clinical samples comprised genomic DNA extracted from whole blood from PCT patients and normal control individuals of three population groups native to South Africa, i.e. Blacks, Caucasians and Coloureds. PCR amplification of the patient samples was successful in the majority of the samples. However, due to restrictions in DNA quantity, not all patients and non-patients could be amplified for each amplicon. This limitation is indicated by the varying number of individuals (*n*) scored for each locus. Not all known variants in the promoter region of the four genes were identified. The failure to detect these polymorphisms could be as a result of the low sensitivity of the chosen HEX-SSCP methodology. Studies report that SSCP analysis is only approximately 70% sensitive (Hayashi and Yandell 1993). Several factors, such as temperature, electrophoresis time, gel composition, pH and DNA fragment size are known to influence the sensitivity of the method. (Fujita and Silver 1994). DNA fragments smaller than 250-300bp are most suitable to reveal single basepair changes (Hayashi 1991), with optimal results achieved when using 150-200 bp fragments (Sheffield *et al.* 1993) Since only one of the fragments analysed here was smaller than 250 bp, and a significant proportion of the fragments were larger than 300 bp, it may account for some of the known polymorphisms going undetected in this study.

Combining the SSCP methodology with HEX analysis (HEX-SSCP) improves the sensitivity of polymorphism detection. Furthermore, HEX-SSCP analysis performed under two or three different conditions increases the sensitivity to between 70 and 95% (Fujita and Silver 1994). The HEX-SSCP detection method employed in this study (ethidium bromide staining combined with UV transillumination) is less sensitive than silver staining or radiolabelling methods, which also provides clearer banding patterns. This factor may also have contributed to variation being overlooked or genotyping errors. All things considered, a substantial number of polymorphisms were identified. Due to time constraints it was not considered feasible to further optimise HEX-SSCP conditions in order to locate additional variants. Nevertheless, this should be considered in

future. Due to the above mentioned limitations of the HEX-SSCP methodology, it is suggested that other methods be employed in follow-up studies. A time efficient method for the detection of polymorphisms is denaturing high performance liquid chromatography (dHPLC) (Underhill *et al.* 1997). Additional advantages of this method include a greater sensitivity level (92%-100%) and a more efficient screening of larger sized fragments (198-732 bp). DNA sequencing is the most accurate method for the successful detection of variation (Kristensen *et al.* 2001). However, this method is not a cost effective approach for the initial screening of a large number of samples. DNA chip technology (e.g. allele specific oligonucleotide (ASO) hybridisation also offers a highly sensitive and specific method for the detection of single base pair mutations (Kozal *et al.* 1996; Nollau and Wagener 1997).

The second objective encompassed statistical and *in silico* analysis, aimed at determining whether statistically significant associations exist between any variants and PCT. Not all controls groups were in HWE. It is imperative to ensure that all control populations conform to HWE specifications. This allows for the accurate determination of the significance of the study's findings. It is therefore of utmost importance to repeat the mutation analysis and even expand the sample size of loci that displayed deviation from HWE in controls, before an association between disease and variant can be regarded as valid. It is also vital to consider population stratification during the study design of a project. In recently admixed populations, false-positive associations are often not replicated in subsequent studies (Thomas and Witte 2002). It is essential to gain as much information as possible on the study subjects' (patients and controls) ethnicity. This will enable suitable project design and will inform researchers whether statistical corrections need to be performed to account for population stratification.

Allele and genotype frequencies were estimated. In several instances, polymorphic incidences

varied significantly between populations and between patients and controls. This trend has occurred in other studies (Perez-Lezaun *et al.* 1997) and may be related to the genetic distance between populations. It is beneficial to perform the Power calculation before the initiation of a project. By revealing the number of subjects required for a study, the Power calculation ensures that the study will have sufficient power to detect a selected effect (Goodman and Berlin 1994; Hoenig and Heisey 2002). Should the Power calculation be performed *post-hoc*, it should never be used as an explanation for failing to detect association in a study. Bonferroni's correction, often applied in association studies where more than one statistical test is necessary, was not applied in this study as there has been much debate regarding the significance of this test (Perneger 1998). The major concern of the Bonferroni's correction is that it is concerned with the general null hypothesis (all hypotheses are valid simultaneously); which in most cases does not provide valuable information to the researcher (Rothman 1990). A fundamental purpose of the Bonferroni's adjustment, is to minimize Type I errors; this cannot decrease without inflating Type II errors.

In silico analysis sought to identify putative TFBSs that may regulate the expression of the four genes investigated, as well as other genes considered biologically relevant to the study, i.e. liver specific or involved in iron metabolism pathways. *In silico* examination predicted several putative TFBSs spanning the variants in the promoter regions of the four genes. The programs used different logarithms and criteria to determine the presence or abolishment of a TFBS. For this reason, *in silico* analysis is merely a predictive method, which aims to highlight a putative binding motif. In this manner *in silico* analysis forms a suitable and cost free basis on which future functional studies may be designed. The relevance of a motif with regards to the specifics of the particular study, needs to be evaluated before further experimental analysis are strategised. Functional studies may include transfection experiments using a luciferase reporter assay to compare expression levels between normal and variant promoters cloned in suitable vectors.

Of the several variants identified in this study, nine warrant further investigation and are tabulated in Table 4.1. These include loci where 1) a statistically significant association was determined in a specific population and 2) the presence of the polymorphic allele possibly alters gene expression by disrupting an existing or creating a novel TF binding motif. Polymorphic loci were shown to be present in elevated incidences in patients compared to controls and *vice versa*. The following TBFSs spanning the DNA sequence in which two significant variants were detected, should be highlighted: the putative USF1 recognition sequence spanning the *HAMP* -582G/A variant and the putative AP1 and C/EBP α binding motifs spanning the *CP* -439T/C variant. All three of these TFs have described roles in the regulation of the respective genes. More importantly, the presence of a C/EBP α binding motif spanning the nucleotide position -439 in *CP* has been documented. It is encouraged that future research initially focusses on the aforementioned variants.

It is suggested that future research into the nature of iron overload in PCT be expanded to include the screening of all genes involved in iron homeostasis, e.g. *HEPH*, *LTF*, *TfR1*, *FTHC* and *FTLC* for potential disease-related variants. To this extent, the promoter regions of *HFE*, *HJV* and *HMOX1* are currently being screened (Stellenbosch University). It is recommended that the entire gene (exons, introns and 3'UTR) be subjected to mutation screening. Additional clinical information of each patient (including smoking, oestrogen therapy, alcohol consumption, *HFE* C282Y, and chemical exposure status) should be obtained to enable insightful investigation into the association of these variants and PCT. Furthermore, the patients should be screened for mutations in the *UROD* gene, to determine if the patient is presenting with fPCT or sPCT. There is also a need to correlate iron parameters (serum iron and ferritin levels and transferrin saturation) and genotypes between patient and non-patient groups in order to ascertain which variants (if any) are associated with iron overload. In depth gene-gene interaction, linkage disequilibrium and haplotype analysis will elucidate the complex relationship of genes involved in iron homeostasis potentially implicated in

the aetiology of PCT.

Table 4.1 Variants identified in the current study that warrant further investigation.

Gene	Nucleotide position and change	* South African population group	Allele frequency <i>P</i>	#Disruption/creation of putative TFBS	Role of TF/relevance
<i>CP</i>	-567C/G	Caucasians	^a 0.01	#YY1	Increases transferrin expression in the aged, but may repress or initiate transcription of several genes
	-439T/C	Blacks	^a 0.05	#ER, #AP1 #C/EBP α	Unknown role in <i>CP</i> , precipitating factor in PCT Previously reported in <i>CP</i> regulation, unknown role at this nucleotide location Positive regulatory effect on <i>CP</i> expression
<i>CYBRD1</i>	-1477G/A	Caucasians	^b 0.04	#GATA 2	Controls proliferative capabilities of haematopoietic progenitor cells
	-1459T/C	Coloureds	^b 0.01	#USF1	Active role in the regulation of genes expressed in the liver
<i>HAMP</i>	-582A/G	Blacks	^a 0.04	#USF1	<i>Trans</i> -activates <i>HAMP</i> expression
<i>SLC40A1</i>	-8C/G	Caucasians	^b 0.003	#GATA 2	Controls proliferative capabilities of haematopoietic progenitor cells
	(CGG) _n	Caucasians	^A 0.001	N/A	

Bold text, novel variants identified in the study; *, South African population in whom the variant was determined to be statistically significant; ^a, variant allele more prevalent amongst patients than controls; ^b, variant allele more prevalent amongst controls than patients; #, variant disrupts putative TFBS; ^c, variant creates novel TFBS

Abbreviations: AP1, activator protein 1; C/EBP α , CCAAT enhancer binding protein α ; *CP*, ceruloplasmin; *CYBRD1*, cytochrome b reductase 1; ER, estrogen receptor; GATA2, GATA binding protein 2; *HAMP*, hepcidin antimicrobial peptide; N/A, not applicable; *P*, Probability value, PU.1, spleen focus-forming intergration oncogene SP11; SP1, specificity protein 1, *SLC40A1*, solute carrier family 40 member 1 TFBS, transcription factor binding site; TF, transcription factor, USF1, upstream stimulatory factor 1, YY1, Yin Yang-1.

This study represents only an initial phase of a broader long-term study. It demonstrated conclusively that several polymorphic loci in genes involved in iron homeostasis occur in higher frequencies in PCT patients than in the general population, suggesting that such loci predispose carriers to disease. Conversely, higher frequencies of some alleles in controls compared to patients were also noticed, indicating that this category of loci act to provide a possible protective effect against disease. It thus reasoned feasible to extend the investigation to include the remaining genes that play a role in the iron metabolic pathway. It is envisaged that data from the completed large study could yield specific haplotypes associated with iron overload, thereby clarifying the basis of iron overload associated with PCT.

An in-depth understanding of the effect of DNA sequence variation on biologic traits is envisaged to improve the diagnosis, treatment and ultimately prevention of human genetic disorders. Only once insight is gained into the combined complications arising from gene-gene and gene-environment interactions, does genotype-phenotype correlation becomes possible. Promising research data from the study presented here provide a solid foundation for future investigation into the environmental and genetic factors (in particular the associated iron overload), implicated in the complex aetiology of PCT.

CHAPTER 5: REFERENCES



5.1 GENERAL REFERENCES

Abboud S and Haile D (2000) A novel mammalian iron-regulated protein involved in intracellular iron metabolism. *J Biol Chem* 26: 19906-19912

Adrian GS, Seto E, Fischbach KS, Rivera EV, Adrian EK, Herbert DC, Walter FJ, Weaker FJ, Bowman BH (1996) YY1 and SP1 transcription factors bind the human transferrin gene in an age-related manner. *J Gerontol A Biol Sci Med Sci* 51: B66-B75

Anderson GJ, Deepak D, Wilkins SJ, Frazer DM (2006) Regulation of systemic iron homeostasis: how the body responds to changes in iron demand. *Biometals* 26: 16046-16053

Anderson GJ, Frazer DM, Wilkins SJ, Becker EM, Millard KN, Murphy TL, McKie AT, Vulpe CD (2002) Relationship between intestinal iron-transporter expression, hepatic hepcidin levels and the control of iron absorption. *Biochem Soc Trans* 30: 724-726

Anderson GJ, Powell LW, Halliday JW (1990) Transferrin receptor distribution and regulation in the rat small intestine: effect of iron stores and erythropoiesis. *gastroenterol* 98: 576-585

Anderson KE, Sassa S, Bishop DF, Desnick RJ (2001) Disorders of heme biosynthesis: X-linked sideroblastic anaemia and the porphyrias. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds), *The metabolic and molecular bases of inherited disease*. McGraw-Hill, New York, pp2961-3062

Andrew SE (1999) Muddy waters: looking beyond the disease-causing mutation, DNA mutations at second sites influence phenotype. *Clin Genet* 55: 144-148

Andrews NC (1999) Disorders of iron metabolism. *N Engl J Med* 341: 1986-1995

Andrews NC, Fleming MD, Gunshin H (1999) Iron transport across biological membranes. *Nutr Rev* 57: 114-123

Atchley WR and Fitch WM (1997) A natural classification of the basic helix-loop-helix class of transcription factors. *Proc Natl Acad Sci USA* 94: 5172-5176

Babitt JL, Huang FW, Wrighting DM, Xia Y, Sidis Y, Samad TA, Campagna JA, Chung RT, Schneyer AL, Woolf AL, Andrews NC, Lin HY (2006) Bone morphogenetic protein signalling by hemojuvelin regulates hepcidin expression. *Nature Genet* 38: 531-539

Badminton MN and Elder GH (2005) Molecular mechanisms of dominant expression in porphyria. *J Inher Metab Dis* 28: 277-286

- Bahram S, Gilfillan S, Kuhn LC, Moret RM, Schulze JB, Lebau A, Schumann K (1999) Experimental hemochromatosis due to MHC class I HFE deficiency: Immune status and iron metabolism. *Proc Natl Acad Sci USA* 96: 13312-13317
- Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21: 263-265
- Bayele HK, McArdle H, Srai SK (2006) Cis and trans regulation of hepcidin expression by upstream stimulatory factor. *Blood* 15: 4237-4245
- Begovich AB, Carlton VEH, Honigberg LE, Schrodi SJ, Chokkalingam AP, Alexander HC, Ardlie KG, Huang Q, Smith AM, Spoerke JM, Conn MT, Chang M, Chang SHP, Saiki RK, Catanese JJ, Leong DU, Garcia VE, McAllister LB, Jeffrey DA, Lee AT, Batliwalla F, Remmers E, Criswell LA, Seldin MF, Kastner DL, Amos CI, Sninsky JJ, Gregersen PK (2004) A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. *Am J Hum Genet* 75: 330-337
- Bellorini M, Lee DK, Dantonel JC, Zenzoumi K, Roeder RG, Tora L and Mantovani R (1997) CCAAT binding NF-Y TBP interactions: NF-YB and NF-YC require short domains adjacent to their histone fold motifs for association with TBP basic residues. *Nucleic Acids Res* 25: 2174-2181
- Benedetto AV, Kushner JP, Taylor JS (1978) Porphyria cutanea tarda in three generations of a single family. *New Eng J Med* 298: 358-361
- Berlin SO and Brante G (1962) Iron metabolism in porphyria and haemochromatosis. *Lancet* 2: 729
- Beutler E, Felitti V, Gelbart T, Ho N (2001) Genetics of iron storage and hemochromatosis. *Drug Metab Dispos* 29: 495-499
- Bevilacqua MA, Faniello MC, Iovine B, Russo T, Cimino F, Costanzo F (2002) Transcription factor NF-Y regulates differentiation of CaCo-2 cells. *Arch Biochem Biophys* 407: 39-44
- Blackwell JM, Barton CH, White JK, Searle S, Baker AM, Williams H, Shaw MA (1995) Genomic organization and sequence of the human NRAMP gene: identification and mapping of a promoter region polymorphism. *Mol Med* 1: 194-205
- Blum K, Noble EP, Sheridan PJ, Montgomery A, Ritchie T, Jagadeeswaran P, Nogami H, Briggs AH, Cohn JB (1990) Allelic association of human dopamine D2 receptor gene in alcoholism. *JAMA* 263: 2055-2060
- Bond GL, Hu W, Bond EE, Robins H, Lutzker SG, Arva NC, Bargonetti J, Bartel F, Taubert H, Wuerl P, Onel K, Yip L, Hwang SJ, Strong LC, Lozano G, Levine AJ (2004) A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans. *Cell* 119: 591-602

- Bonin A, Bellemain E, Eidesin PB, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetic studies. *Mol Ecol* 13: 3261-3273
- Boisseau AM, Couzigou P, Forestier JF, Legrain V, Aubertin J, Doutre MS, Maleville J, Geniaux M, Beylot C (1991) Porphyria cutanea tarda associated with human immunodeficiency virus infection. A study of four cases and review of the literature. *Dermatologica* 182: 155-159
- Boldt DH (1999) New Perspectives on iron metabolism: An introduction. *Am J Med Sci* 318: 207-212
- Bosio S, De Gobbi M, Roetto A, Zecchina, Leonardo E, Rizzetto M, Lucetti C, Petrozzi U, Binyccelli U, Camaschella C (2002) Anemia and iron overload due to compound heterozygosity for novel ceruloplasmin mutations. *Blood* 100: 2246-2248
- Bothwell TH, Charlton RW, Cook JD, Finch CA (1979) Iron metabolism in man. Blackwell Scientific Publications, Oxford, pp113
- Bothwell TH, Charlton RW, Motulski AG (1995) Hemochromatosis. In: Scriver CR, Beaudet AL, Sly WS, Valee D (eds), *The metabolic and molecular bases of inherited disease*. McGraw Hill, New York, pp2237-2269
- Bottini N, Musumeci L, Alonso A (2004) A functional variant of lymphoid tyrosine phosphatase associated with type I diabetes. *Nat Genet* 36: 337-338
- Brodie MJ, Moore MR, Goldberg A (1977) Enzyme abnormalities in the porphyrias. *Lancet* 2: 699-701
- Brookes AJ (1999) The essence of SNPs. *Gene* 234: 177-186
- Bucher P (1990) Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. *J Mol Biol* 212: 563-578
- Bulaj ZJ, Phillips JD, Ajioka RS, Franklin MR, Griffen LM, Guinee DJ, Edwards CQ, Kushner JP (2000) Hemochromatosis genes and other factors contributing to the pathogenesis of porphyria cutanea tarda. *Blood* 95: 1565-1571
- Bulyk ML (2003) Computational prediction of transcription-factor binding site locations. *Genome Biol* 5: 201.1-201.7
- Burden HE and Weng Z (2005) Identification of conserved structural features at sequentially degenerate locations in transcription factor binding sites. *Genome Informatics* 16: 49-58
- Cairo G, Bernuzzi F, Recalcati S (2006) A precious metal: Iron, an essential nutrient for all cells. *Genes Nutr* 1: 25-40

- Cairo G, Conte D, Bianchi L, Fraquelli M, Recalcati S (2001) Reduced serum ceruloplasmin levels in hereditary haemochromatosis. *Br J Haematol* 114: 226-229
- Cairo G and Pietrangelo A (2000) Iron regulatory proteins in pathobiology. *Biochem J* 352: 241-250
- Calzolari A, Oliviero I, Testa U (2007) Transferrin receptor 2 is emerging as a major player in the control of iron metabolism. *Cent Euro J Biol* 2: 34-35
- Camaschella C (2005) Understanding iron homeostasis through genetic analysis of haemochromatosis and related disorders. *Blood* 106: 3710-3717
- Camaschella C, Roetto A, Cali A, De Gobbi M, Garozzo G, Carella M, Majorano N, Totaro A, Gasparini P (2000) The gene *TFR2* is mutated in a new type of haemochromatosis mapping to 7q22. *Nat Genet* 25: 14-15
- Cao Z, Umnek RM, McKnight SL (1991) Regulated expression of three C/EBP isoforms during adipose conversion of 3T3-L1 cells. *Genes Dev* 5: 1538-1552
- Carstens JJ, Hift RJ, Zaahl MG (2006) Mutation analysis of the *UROD* gene in patients with Porphyria cutanea tarda. Hons diss. Univerisity of Stellenbosch
- Casado M, Vallet VS, Kahn A, Vaulont S (1999) Essential role *in vivo* of upstream stimulatory factors for a normal dietary response of the fatty acid synthase gene in the liver. *J Biol Chem* 274: v2009-v2013
- Caskey JH, Jones C, Miller YE, Seligman PA (1983) Human ferritin gene is assigned to chromosome 19. *Proc Nat Acad Sci USA* 80: 482-486
- Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H, McClay J, Mill J, Martin J, Braithwaite A, Poulton R (2003) Influence of Life Stress on Depression: Moderation in the 5-HTT Gene. *Science* 301: 386-389
- Cazzola M, Cremonesi L, Papaioannou M, Soriani N, Kioumi A, Charalambidou A, Paroni R, Romstou K, Levi S, Ferrari M, Arosio P, Christakis J (2002) Genetic hyperferritinaemia and reticuloendothelial iron overload associated with a three base pair deletion in the coding region of the ferroportin gene (SLC11A3). *Br J Haematol* 119: 539-546
- Cazzola M, Beguin Y, Bergamaschi, Guornone R, Cerani P, Barella S, Cao A, Galanello R (1999) Soluble transferrin receptor as a potential determinant of iron loading in congenital anaemias due to ineffective erythropoiesis. *Br J Haematol* 106: 752-755
- Cazzola M, Ascari E, Barosi G, Claudiani G, Dacco M, Kaltwasser JP, Panaiotopoulos N, Schalk KP, Werner EE (1983) Juvenile idiopathic haemochromatosis: a life-threatening disorder presenting as hypogonadotropic hypogonadism. *Hum Genet* 65: 149-154

- Cereghini J (1996) Liver-enriched transcription factors and hepatocyte differentiation. *FASEB J* 10: 267-282
- Chen SL, Xu MY, Ji XS, Yu GC, Liu Y (2005) Cloning, characterisation, and expression analysis of hepcidin from red sea bream (*Chrysophyrus major*). *Antimicrob Agents Chemother* 49: 1608-1612
- Christiansen L, Bygum A, Thomsen K, Brandup F, Hørder M, Petersen NE (2000) Association between CYP1A2 polymorphisms and susceptibility to porphyria cutanea tarda. *Hum Genet* 40: 1-25
- Christiansen L, Bygum A, Thomsen K, Brandup F, Hørder M, Petersen NE (1999a) Denaturing gradient gel electrophoresis analysis of the Haemochromatosis (HFE) gene: Impact of HFE gene mutations on the manifestation of porphyria cutanea tarda. *Clin Chem* 45: 2025-2026
- Christiansen L, Ged C, Hombrados I, Brons-Paulsen J, Fontanellas A, de Verneuil H, Horder M, Peterson NE (1999b) Screening for mutations in the uroporphyrinogen decarboxylase gene using denaturing gradient gel electrophoresis. Identification and characterization of six novel mutations associated with familial PCT. *Hum Mut* 14: 222-232
- Ciehanover A, Schwartz AL, Lodish HF (1983) The asialoglycoprotein receptor internalizes and recycles independently of the transferrin and insulin receptors. *Cell* 32: 267-275
- Cleary ML (1999) A new angle on a pervasive oncogene. *Nature Genet* 23: 134-135
- Cochran WG (1954) Some methods for strengthening the common χ^2 tests. *Biometrics* 10: 417-451
- Collins FS, Green E, Guttmacher A, Guyer M (2003) A vision for the future of genomics and research. *Nature* 422: 835-847
- Collins FS, Brooks LD, Chakravarti S (1998) A DNA polymorphism discovery resource for research on human genetic variation. *Genome Res* 8: 1229-1231
- Connor SL, Gustafson JR, Sexton G, Becker N, Artuad-Wild S, Connor WE (1992) The diet habit survey: a new method of dietary assessment that relates plasma cholesterol changes. *J Am Dietetic Assoc* 92: 41-47
- Conrad ME, Umbreit JN, Moore EG (1999) Iron absorption and transport. *Am J Med Sci* 318: 213-329
- Conrad ME, Parmley RT, Osterloh K (1987) Small intestinal regulation of iron absorption in the rat. *J Lab Clin Med* 110: 418-426
- Cortes JM, Oliva H, Paradinas FJ, Fernandes-Guio C (1980) The pathology of the liver in porphyria cutanea tarda. *Histopathol* 21: 1500-1503

- Costa RH, Kalinichenko VV, Holterman AX, Wang X (2003) Transcription factors in liver development, differentiation, and regeneration. *Hepatology* 38: 1331–1347
- Costa RH, Grayson DR, Darnell JE Jr (1989) Multiple hepatocyte-enriched nuclear factors function in the regulation of transferrin and α 1-antitrypsin genes. *Mol Cell Biol* 9: 1415-1425
- Courselaud B, Pigeon C, Inoue Y, Inoue J, Gonzalez PL, Leroyer P, Gilot D, Boudjema K, Guguen-Guillouzo C, Brissot P, Loreal O, Ilyin G (2002) C/EBP α regulates hepatic transcription of hepcidin, an antimicrobial peptide and regulator of iron metabolism. *J Biol Chem* 277: 41163-41170
- Courtois G, Morgan GJ, Campbell LA, Fourel G, Crabtree GR (1987) Interaction of a liver-specific nuclear factor with the fibrinogen and alpha 1-antitrypsin promoters. *Science* 238: 688-692
- Coustry F, Hu Q, de Crombrughe B, Maity SN (2001) CBF/NF-Y functions both in nucleosomal disruption and transcription activation of the chromatin-assembled topoisomerase IIa promoter. Transcription activation by CBF/NF-Y in chromatin is dependent on the promoter structure. *J Biol Chem* 276: 40621-40630
- Crichton RR (1991) *Inorganic Biochemistry of iron metabolism*. Chichester: Ellis Horwood, pp263
- Cripps DJ (1987) Diet and alcohol effects on the manifestation of hepatic porphyrias. *Federation Proceedings* 46:1894-900
- Daimon M, Yamatani K, Igarashi M, Fukase N, Kawanami T, Kato T, Tominaga M, Sasaki H (1995) Fine structure of the human ceruloplasmin gene. *Biochem Biophys Res Commun* 208: 1028-1035
- Daniell WE, Stockbridge HL, Labbe RF, Woods JS, Anderson KE, Bissell DM, Bloomer JR, Ellefson RD, Moore MR, Pierach CA, Schreiber WE, Tefferi A, Franklin GM (1997) Environmental chemical exposures and disturbances of heme synthesis. *Environ Health Perspect* 105: 37-43
- Das D, Tapryal N, Goswami SK, Fox PL, Mukhopadhyay CK (2007) Regulation of ceruloplasmin in human hepatic cells by redox active copper: identification of a novel AP-1 site in the ceruloplasmin gene. *Biochem J* 402: 135-141
- DeKoter RP, Walsh JC, Singh H (1998) PU.1 regulates both cytokine dependent proliferation and differentiation of granulocyte/macrophage progenitors. *EMBO J* 17: 4456-4468
- Dereure O, Aguilar-Martinez P, Bessis D, Perney P, Vallat C, Guillot B, Blanc F, Guilhou JJ (2001) HFE mutations and transferrin receptor polymorphism analysis in porphyria cutanea tarda: a prospective study of 36 cases from southern France. *Br J Dermatol* 144: 533-539

- de Verneuil H, Bourgeois F, de Rooij F, Siersema PD, Wilson JHP, Grandchamp B, Nordmann Y (1992) Characterisation of a new mutation (R292G) and a deletion at the human uroporphyrinogen decarboxylase locus in two patients with hepatoerythropoietic porphyria. *Hum Genet* 89: 548-552
- de Verneuil H, Grandchamp B, Beaumont C, Picat C, Nordmann Y (1986) Uroporphyrinogen decarboxylase structural mutant (Gly 281--> Glu) in a case of porphyria. *Science* 234: 732-734
- de Verneuil H, Beaumont C, Deybach JC, Nordmann Y, Sfar Z, Kastally R (1984) Enzymatic and immunological studies of uroporphyrinogen decarboxylase in familial porphyria cutanea tarda and hepatoerythropoietic porphyria. *Am J Hum Genet* 36: 613-622
- de Verneuil H, Aitken G, Normann Y (1978) Familial and sporadic porphyria cutanea: two different diseases. *Hum Genet* 44: 145-151
- Devys D, Lutz Y, Rouyer N, Bellocq JP, Mandel JL (1993) The FMR-1 protein is cytoplasmic, most abundant in neurons and appears normal in carriers of a fragile X premutation. *Nat Genet* 4: 335-340
- Dickson PW, Aldred AR, Marley PD, Tu GF, Howlett GJ, Shreiber G (1985) High prealbumin and transferrin mRNA levels in the choroid plexus of rat brain. *Biochem Biophys Res Commun* 127: 890-895
- Donovan A, Roy CN, Andrews NC (2006) The ins and outs of iron homeostasis. *Physiology* 21: 115-123
- Donovan A, Brownlie A, Zhou Y, Shepard J, Pratt SJ, Moynihan J, Paw B, Drejer A, Barut B, Zapata Z, Law TC, Brugnara C, Lux SE, Pinkus GS, Pinkus JL, Kingsley PD, Palis J, Fleming MD, Andrews NC, Zon LI (2000) Positional cloning of zebrafish ferroportin1 identifies a conserved vertebrate iron exporter. *Nature* 403: 776-781
- Douabin-Gicquel V, Soriano N, Ferran H, Wojcik F, Palierne E, Tamim S, Jovelin T, McKie AT, Le Gall JY, Davind V, Mosser J (2001) Identification of 96 single nucleotide polymorphisms in eight genes involved in iron metabolism: efficiency of bioinformatic extraction compared with a systematic sequencing approach. *Hum Genet* 106: 393-401
- Driscoll M, Sathya DG, Muyan M, Klinge CM, Hilf R, Bambara RA (1998) Sequence requirements for estrogen binding to estrogen responsive elements. *Biol Chem* 273: 29321-29330
- Drobacheff C, Derancourt C, Van Landyt H, Devred D, de Waziens B, Cribier B, Rey D, Lang JM, Grosieux C, Kalis B, Laurent R (1998) Porphyria cutanea tarda associated with human immunodeficiency virus infection. *Eur J Dermatol* 8: 492-296
- Dubart A, Mattei MG, Raich N, Beaupain D, Romeo PH, Mattei JF, Goossens M (1986) Assignment of human uroporphyrinogen decarboxylase (URO-D) to the p34 band of chromosome 1. *Hum Genet* 73: 277-279
- Duncan SA, Navas MA, Dufort D, Rossant J, Stoffel M (1998) Regulation of a transcription factor network required for

differentiation and metabolism. *Science* 281: 692-695

Dunn LL, Rahmanto YS, Richardson DR (2006) Iron uptake and metabolism in the new millennium. *Trends Cell Biol* 17: 93-100

Edwards CQ, Griffen LM, Goldgar D, Drummond C, Skolnick MH, Kushner JP (1998) Prevalence of hemochromatosis among 11, 065 presumably healthy blood donors. *N Engl J Med* 318: 1355-1362

Edwards CQ, Griffen LM, Goldgar DE, Skolnick MH, Kushner JP (1989) HLA-linked hemochromatosis alleles in sporadic porphyria cutanea tarda. *Gastroenterol* 97: 972-981

Egger NG, Goeger DE, Payne DA, Miskovsky EP, Weinman SA, Anderson KE (2002) Porphyria cutanea tarda: Multiplicity of Risk factors including *HFE* mutations, Hepatitis C, and inherited uroporphyrinogen decarboxylase deficiency. *Digest Dis Sci* 47: 419-426

Eggers A, Siemann G, Blume R, Knepel W (1998) Gene-specific transcriptional activity of the insulin cAMP responsive element is conferred by NF-Y in combination with cAMP response element-binding protein. *J Biol Chem* 273: 18499-18508

Elder GH (1998) Porphyria cutanea tarda. *Semin Liver Dis* 18: 67-75

Elder GH (1993) Molecular genetics of disorders of haem biosynthesis. *J Clin Pathol* 46: 977-981

Elder GH, Urquhart AJ, De Salamaca RE, Munoz JJ, Bonkovsky HL (1985) Immunoreactive uroporphyrinogen decarboxylase in the liver in porphyria cutanea tarda. *Lancet* 2: 229-233

Elder GH, Lee GB, Tovey JA (1978) Decreased activity of hepatic uroporphyrinogen decarboxylase in sporadic porphyria cutanea tarda. *New Eng J Med* 299: 274-278

Elder GH, Gray CH, Nicholson DC (1972) The porphyrias: a review. *J Clin Path* 25: 1013-1033

Elder GH and Roberts AG (1995) Uroporphyrinogen decarboxylase. *J Biomembr Bioeng* 27: 207-214

Elder GH and Worwood M (1998) Mutations in the haemochromatosis gene, porphyria cutanea tarda and iron overload. *Hepatology* 27: 289-291

Enns CA, Suomalainen HA, Gebhardt JE, Schroder J, Sussman HH (1982) Human transferrin receptor: expression of the receptor is assigned to chromosome 3. *Proc Natl Acad Sci USA* 79: 3241-3245

Fargion S, Fracanzani AL, Romano R, Cappellini D, Fare M, Mattioli M, Piperno A (1996) Genetic hemochromatosis in Italian patients with porphyria cutanea tarda: possible explanation for iron overload. *J Hepatol* 24: 564-569

- Feder JN, Penny DM, Irrinki A, Lee VK, Lebron JA, Watson N, Tsuchihashi Z, Sigal E, Bjorkman PJ, Schatzman RC (1998) The haemochromatosis gene product complexes with the transferrin receptor and lowers its affinity for ligand binding. *Proc Natl Acad Sci USA* 95: 1472-1477
- 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 (1996) A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nat Genet* 13: 399-408.
- Felsher BF, Carpio NM, Engleking DW, Nunn AT (1982) Decreased hepatic uroporphyrinogen decarboxylase activity in porphyria cutanea tarda. *New Eng J Med* 306: 766-369
- Felsher BF and Kushner JP (1977) Hepatic siderosis and porphyria cutanea tarda. Relation of iron excess to the metabolic defect. *Semin Hematol* 14: 243-251
- Felsher BF and Redeker AG (1966) Effect of chloroquine on hepatic uroporphyrin metabolism in patients with porphyria cutanea tarda. *Medicine* 45: 575-583
- Ferreira R, Ohneda K, Yamamoto M, Philipsen S (2005) GATA1 function, a paradigm for transcription factors in haematopoiesis. *Mol Cell Biol* 25: 1215-1227
- Fickett JW and Wassermann JJ (2000) Discovery and modeling of transcriptional regulatory regions. *Curr Opin Biotech* 11: 19-24
- Fickett JW and Hatzigeorgiou AG (1997) Eukaryotic promoter recognition. *Genome Res* 7: 861-878
- Finch C (1994) Regulators of iron balance in humans. *Blood* 84: 1697-1702
- Fleming MD, Trenor CC III, Su MA, Foernzler D, Beier DR, Dietrich WF, Andrew NC (1997) Microcytic anaemia mice have a mutation in Nramp2, a candidate iron transporter gene. *Nat Genet* 16: 383-386
- Fleming RE, Britton RS, Waheed A, Sly WS, Bacon BR (2005) Pathophysiology of hereditary hemochromatosis. *Semin Liver Dis* 25: 411-419
- Fleming RE and Gitlin (1992) Structural and functional analysis of the 5'-flanking region of the rat ceruloplasmin gene. *Biol Chem* 267: 479-486
- Fletcher L, Halliday J, Powell L (1999) Interrelationships of alcohol and iron in liver disease with particular reference to the iron-binding proteins, ferritin and transferrin *J Gastroenterol Hepatol* 14: 202-214

- Francauzani AL, Taioli E, Sampietro M (2001) Liver cancer risk is increased in patients with porphyria cutanea tarda in comparison to matched control patients with chronic liver disease. *J Hepatol* 35: 498-503
- Frazer DM, Wilkins SJ, Vulpe CD, Anderson GJ (2005) The role of duodenal cytochrome b in intestinal iron absorption remains unclear. *Blood* 106: 4413
- Frazer DM and Anderson GJ (2003) The orchestration of body iron uptake: how and where do enterocytes receive their cues? *Blood Cells Mol Dis* 30: 288-297
- Frieden E and Hsieh HS (1976) Ceruloplasmin: the copper transport protein with essential oxidase activity. *Adv Enzymol Relat Areas Mol Biol* 44: 187-236
- Frontali M, Novelletto A, Annesi GM, Jodice C (1999) CAG repeat instability, cryptic sequence variation and pathogenicity: evidence from different loci. *Philos T Biol Sci* 354: 1089-1094
- Frontini M, Imbriano C, diSilvio A, Bell B, Bogni A, Romier C, Moras D, Tora L, Davidson I, Mantovani R (2002) NF-Y recruitment of TFIID, multiple interactions with histone fold TAF(II)s. *J Biol Chem* 277: 5841-5848
- Fujita K and Silver J (1994) Single-strand conformational polymorphism. *PCR Methods Appl* 4: 137-140
- Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D (2002) The structure of haplotype blocks in the human genome. *Science* 296: 2225-2229
- Gale E, Torrance J, Bothwell T (1963) The quantitative estimation of total iron stores in human bone marrow. *J Clin Invest* 42: 1076-1082
- Garey JR, Franklin KF, Brown DA, Harrison LM, Metcalf KM, Kushner JP (1993) Analysis of uroporphyrinogen decarboxylase complementary DNAs in sporadic porphyria cutanea tarda. *gastroenterol* 105: 165-169
- Garey JR, Harrison LM, Franklin KF, Metcalf KM, Radisky ES, Kushner JP (1990) Uroporphyrinogen decarboxylase: a splice site mutation causes the deletion of exon 6 in multiple families with porphyria cutanea tarda. *J Clin Invest* 86: 1416-1422
- Garey JR, Hansen JL, Harrison LM, Kennedy JB, Kushner JP (1989) A point mutation in the coding region of uroporphyrinogen decarboxylase associated with familial PCT. *Blood* 73: 892-895
- Gelernter J, Goldman D, Risch N (1993) The A1 allele at the D2 dopamine receptor gene and alcoholism. *JAMA* 269: 1673-1677

- Gillooly M (1983) The effect of organic acids, phylates and polyphenols on absorption of iron from vegetables. *Br J Nutr* 49: 331-342
- Goodman SN and Berlin JA (1994) The use of predicted confidence levels when planning experiments and the missuse of power when intpreting results. *Ann Intern Med* 121: 200-206
- Goodfellow PN, Banting G, Sutherland R, Greaves M, Solomon E, Povey S (1982) Expression of human transferrin receptor is controlled by a gene on chromosome 3: assignment using species specificity of a monoclonal antibody. *Somat Cell Genet* 8: 197-206
- Gorman N, Zaharia A, Trask HS, Szakacs JG, Jacobs NJ, Jacobs JM, Balestra D, Sinclair JF, Sinclair PR (2007) Effect of iron and ascorbate on uroporphyrin in ascorbate-requiring mice as a model for porphyria cutanea tarda. *Hepato* 45: 187-194
- Gourraud PA, Boyer JF, Barnetche T, Abbal M, Cambon-Thomsen A, Cantagrel A, Constantin A (2006) A new classification of HLA-DRB1 alleles differentiates predisposing and protective alleles for rheumatoid arthritis structural severity. *Arthritis Rheum* 54: 593–599
- Gregor PD, Sawadogo M, Roeder RG (1990) The adenovirus major late transcription factor USF is a member of the helix-loop-helix group of regulatory proteins and binds to DNA as a dimer. *Genes Dev* 4: 1730-1740
- Grossman ME, Bickers DR, Poh-Fitzpatrick MB, Delco VA, Harber LC (1979) Porphyria cutanea tarda: clinical features and laboratory findings in 40 patients. *Am J Med* 67: 277-286
- Gruenheid S, Cellier M, Vidal S, Gros P (1997) Identification and characterisation of a second mouse *Nramp* gene. *Genomics* 25: 514-525
- Gunshin H, Starr CN, DiRenzo C, Fleming MD, Jin J, Greer EL, Sellers VM, Galica SM, Andrews NC (2005) *Cybrd1* (duodenal cytochrome b) is not necessary for dietary iron absorption in mice. *Blood* 106: 2879-2883
- Gunshin H, Mackenzie B, Berger UV, Gunshin Y, Romero MF, Boron WF, Nussberger S, Gollan JL, Hediger MA (1997) Cloning and characterization of a mammalian proton-coupled metal-ion transporter. *Nature* 388: 482-488
- Guo B, Brown FM, Phillips JD, Yu Y, Leibold EA (1995) Characterization and expression of iron regulatory protein 2 IRP2. Presence of multiple IRP2 transcripts regulated by intracellular iron levels. *J Biol Chem* 270: 16529-16535
- Gou Y and Jamison DC (2005) The distribution of SNPs in human gene regulatory regions. *Genomics* 6: 140-151
- Gyulikhandanova NE, Tsybalenko NV, Platonova NA, Babich VS, Puchkova LV (2004) Regulation of Ceruloplasmin Gene in Mammals. *B Exp Biol Med* 137: 485-489

- Haile DJ, Rouault TA, Harford JB, Kennedy MC, Blondin GA, Beinert H, Klausner RD (1992a) Cellular regulation of the iron-responsive element binding protein: disassembly of the cubane iron-sulfur cluster results in high-affinity RNA binding. *Proc Natl Acad Sci USA* 89: 11735-11739
- Haile DJ, Rouault TA, Tang CK, Chin J, Harford JB, Klausner RD (1992b) Reciprocal control of RNA-binding and aconitase activity in the regulation of the iron-responsive element binding protein: role of the iron sulphur cluster. *Proc Natl Acad Sci USA* 89: 7536-7540
- Halder I and Shriver MD (2003) Measuring and using admixture to study the genetics of complex diseases. *Hum Genome* 1: 52-62
- Hall TA (1999) Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/95/NT. *Nucl Acids Symp Ser* 41: 95-98
- Hallberg L (1993) Inhibition of haem-iron absorption in man by calcium. *Br J Nutr* 69: 533-540
- Halliwell B (1992) Reactive oxygen species and the central nervous system. *J Neuro Chem* 59: 609-623
- Halliwell B and Gutteridge JM (1992) Biologically relevant metal iron-dependent hydroxyl radical generation. An update. *FEBS letter* 27: 108-112
- Hanson ES and Leibold EA (1999) Regulation of the iron regulatory proteins by reactive nitrogen and oxygen species. *Gene Expr* 7: 367-376
- Harris ZI, Durley AP, Man TK, Gitlin JD (1999) Targeted gene disruption reveals an essential role for ceruloplasmin in cellular iron efflux. *Proc Natl Acad Sci USA* 96: 10812-10817
- Harris ZL, Klomp LW, Gitlin JD (1998) Aceruloplasminemia: an inherited neurodegenerative disease with impairment of iron homeostasis. *Am J Clin Nutr* 67: 927S-977S
- Harrison PM and Arosio P (1996) The ferritins: molecular properties, iron storage function, and cellular regulation. *Biochim Biophys Acta* 1275: 161-203
- Harrison-Findik D, Schafer D, Klein E, Timchenko NA, Kulaksiz H, Clemens D, Fein E, Andriopoulos B, Pantopoulos K, Gollan J (2006) Alcohol metabolism-mediated oxidative stress down-regulates hepcidin transcription and leads to increased duodenal iron transported expression. *J Biol Chem* 281: 22974-22982
- Hartl DL and Clark AG (2007) *Principles of Population Genetics*. Sinauer Associates, Sunderland

- Hayashi K (1991) PCR-SSCP: A simple and sensitive method for detection of mutations in genomic DNA. *PCR Methods Applic* 1: 34-38
- Hayashi K and Yandell DW (1993) How sensitive is a PCR-SSCP? *Hum Mutat* 2: 338-346
- Hentze MW, Muckenthaler MU, Andrews NC (2004) Balancing acts: Molecular control of mammalian iron metabolism. *Cell* 117: 285-297
- Hentze MW and Kühn LC (1996) Molecular control of vertebrate iron metabolism: mRNA-based regulatory circuits operated by iron, nitric oxide, and oxidative stress. *Proc Natl Acad Sci USA* 93: 8175-8182
- Hermann T, Muckenthaler MU, van der Hoeven F, Brennan K, Gehrke SG, Hubert N, Sergi C, Gröne HJ, Kaiser I, Gosch I, Volkmann, Reidel HD, Hentze MW, Stewart AF, Stremmel W (2004) Iron overload in adult HFE-deficient mice independent of changes in the steady-state expression of the duodenal iron transporters DMT1 and Ireg1/ferroportin. *J Mol Med* 82: 39-48
- Herzig S, Hedrick S, Morantte I, Koo SH, Galimi F, Montminy F (2003) CREB controls hepatic lipid metabolism through nuclear hormone receptor PPAR- γ . *Nature* 426:190-193
- Hift RJ, Corrigan AV, Hancock V, Kannemeyer J, Meissner PN, Kirsch RE (2002) Porphyria cutanea tarda: The etiological importance of mutations in the HFE gene and viral infection is population dependent. *Cell Mol Biol* 48: 1-7
- Hift RJ (1999) The diagnosis of porphyria. *S Afr Med J* 89: 611-614
- Hilberg F, Aguzzi A, Howells N, Wagner EF (1993) c-Jun is essential for normal mouse development and hepatogenesis. *Nature* 365: 179-181
- Hill AV (1999) The immunogenetics of resistance to malaria. *Proc Assoc Am Physicians* 111: 272-277
- Hock H, Hambleton MJ, Rooke HM, Schindler JW, Saleque S, Fujiwara Y, Orkin SH (2004) Gfi-1 restricts proliferation and preserves functional integrity of haematopoietic stem cells. *Nature* 431: 1002-1007.
- Hoenig JM and Heisey DM (2001) The abuse of power: the pervasive fallacy of Power calculations for data analysis. *Am Stat* 55: 19-24
- Holti G, Rimington C, Tate BC, Thomas G (1958) An investigation of 'porphyria cutanea tarda'. *Quart J Med* 27: 1-19

- Hugot JP, Chamaillard M, Zouali H, Lesage S, Cézard JP, Belaiche J, Almer S, Tysak C, O'Marain CO, Gassull M, Binder V, Finkel Y, Cortot A, Modigliani R, Laurent-Puig P, Gower-Rousseau C, Macry J, Colombel JF, Sahbatou M, Thomas G (2001) Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 434: 599-603
- Ippen H (1961) Allgemeine symptome. Der spätèn hautporphyrie (Porphyria cutanea tarda) als Hinweise für deren Behandlung. *Dtsch Med Wochenschr* 86: 127-133
- Jackson HA, Whatley SD, Roberts AG, Morgan RR, Worwood M, Elder GH (1997) Haemochromatosis gene mutations in familial and sporadic porphyria cutanea tarda. *Blood* 90 (Suppl 1): 604a
- James MFM and Hift RJ (2000) Porphyrins. *Brit J Anaesth* 85: 143-153
- Jiang X, Xu K, Hoberman J, Tain F, Marko AJ, Waheed JF, Harris CR, Marini A, Enoch MA, Lipsky RH (2005) BDNF Variation and Mood Disorders: A novel functional promoter polymorphism and Val66Met are associated with anxiety but have opposing effects. *Neuropsychopharmacol* 30: 1353–1361
- Johnson PF (1987) Transcriptional activators in hepatocytes. *Cell Growth Differ* 1: 47-52
- Kaplan J and Kushner JP (2000) Mining the genome for iron. *Nature* 403: 711-713
- Kappas A, Sassa A, Galbraith RA, Nordmann Y (1985) The porphyrias. In: Scriver CR, Beaudet AL, Sly WS, Valle WD (eds), *The Molecular and Metabolic Basis of Inherited Disease*, McGraw-Hill, New York, pp2103-2159
- Kawabata H, Nakamaki T, Ikonomi P, Smith RD, Germain RS, Koeffler HP (2001) Expression of transferrin receptor 2 in normal and neoplastic hematopoietic cells. *Blood* 98: 2714–2719.
- Kawabata H, Yang R, Hiramata T, Vuong PT, Kawano S, Gombart AF, Koeffler HP (1999) Molecular cloning of transferrin receptor 2. A new member of the transferrin receptor-like family. *J Biol Chem* 274: 20826–20832
- Kel A, Kel-Margoulis O, Borlak J, Tchekmenev D, Wingender E (2005) Databases and tools for *in silico* analysis of regulation of gene expression. *Handbook of Toxiogenomics: Strategies and Applications*. Ed J Borlak, WILEY-VCH p253-290
- Khoury MJ and Yang Q (1998) The future of genetic studies of complex human diseases: an epidemiologic perspective. *Epidemiol* 9: 350-354
- Klug WS and Cummings MR (2003) Population genetics. In: *Concepts of Genetics*, Pearson Education LTD, London, pp 631-650

- Knutson MD, Oukka M, Koss LM, Aydemir F, Wessling-Resnick M. Iron release from macrophages after erythrophagocytosis is up-regulated by ferroportin 1 overexpression and down-regulated by hepcidin (2005) *Proc Natl Acad Sci USA* 102: 1324–1328
- Knutson M and Wessling-Resnick M (2003) Iron metabolism in the reticuloendothelial system. *Crit Rev Biochem Mol Biol* 38: 61-88
- Ko LJ and Engel JD (1993) DNA binding specificities of the GATA transcription factor family. *Mol Cell Biol* 13: 4011-4022
- Kostrikis LG, Neumann AU, Thomson B, Korber BT, McHardy P, Karanickolas R, Deutsch L, Huang Y, Lew JF, McIntosh K, Pollack H, Borkowsky W, Spiegel HM, Palumbo P, Oleske J, Bardeguet A, Luzuriaga K, Sullivan J, Wolinsky SM, Koup RA, Ho DD, Moore JP (1999) A polymorphism in the regulatory region of the CC-chemokine receptor 5 gene influences perinatal transmission of human immunodeficiency virus 1 type to African-American infants. *J Virol* 73: 10264-10271
- Kotze MJ, Theart L, Callis M, Peeters A, Thiart R, Langenhoven E (1995) Nonradioactive multiplex PCR screening strategy for the simultaneous detection of multiple low-density lipoprotein receptor gene mutations. *PCR methods Applic* 4: 352-356
- Kozal MJ, Shah N, Shen N, Yang R, Fucini R, Merigan TC, Richman DD, Morris D, Hubbell E, Chee M, Gringeras TR (1997) Extensive polymorphisms observed in HIV-1 clade B protease gene using high-density oligonucleotide arrays. *Proc Natl Acad Sci USA* 2: 753-759
- Krause A, Neitz S, Magert HJ, Schulz A, Forssmann WG, Schulz-Knappe P, Adermann K (2000) LEAP-1, a novel highly disulfide-bonded human peptide, exhibits antimicrobial activity. *FEBS Lett* 480: 147-150
- Kristensen VN, Kelefiotis D, Kristensen T, Borrensens-Dale AL (2001) High-throughput methods for detection of genetic variation. *Biotechniques* 30: 318-322
- Kruglyak L (1999) Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat Genet* 22: 139-144
- Ktistaki E and Taliandis I (1997) Modulation of hepatic gene expression by hepatocyte nuclear Factor 1. *Science* 277: 109-112
- Kuo CJ, Conley PB, Chen L, Sladek FM, Darnell JE, Crabtree GR (1972) A transcriptional hierarchy involved in mammalian cell-type specification. *Nature* 355: 457-461
- Kushner JP, Barbuto AJ, Lee GR (1976) An inherited enzyme defect in porphyria cutanea tarda. *J Clin Invest* 58: 1089-1097

- Kushner JP, Lee GR, Nacht S (1972) The role of iron in the pathogenesis of porphyria cutanea tarda. *J Clin Invest* 51: 3044-3051
- Lai CL, Wu PC, Lin HJ, Wong KL (1984) Case report of symptomatic porphyria cutanea tarda and histiocytic lymphoma. *Cancer* 53: 573-576
- Lai E, Riley J, Purvis I, Roses A (1998) A 4-Mb high-density single nucleotide polymorphism-based map around human APOE. *Genomics* 54: 31-38
- Lambrecht RW, Thapar M, Bonkovsky HL (2007) Genetic aspects of porphyria cutanea tarda. *Semin Liver Dis* 27: 99-108
- Lamoril J, Andant C, Gouya L, Malonova E, Grandchamp B, Martásek P, Deybac hJC, Puy H (2002) Hemochromatosis (HFE) and transferrin receptor-1 (TFRC1) genes in sporadic porphyria cutanea tarda (sPCT). *Cell Mol Biol* 48: 33-41
- Latunde-Dada G, Simpson RJ, McKie AT (2006) Recent advances in mammalian haem transport. *Trends Biochem Sci* 31: 182-184
- Lechler R and Warrens A (2005) HLA in health and disease. Academic press, London
- Lee CS, Friedman JR, Fulmer JT, Kaestner KH (2005) The initiation of liver development is dependent on FOXa transcription factors. *Nature* 435: 944-947
- Lee PL, Gelbart T, West C, Halloran C, Felitti V, Beutler E (2001) A study of genes that may modulate the expression of hereditary hemochromatosis: transferrin receptor-1, ferroportin, ceruloplasmin, ferritin light and heavy chains, iron regulatory proteins (IRP)-1 and -2, and hepcidin. *Blood Cells Mol Dis* 27: 783-802
- Lee SH, Starkey PM, Gordon S (1985) Quantitative analysis of total macrophage content in adult mouse tissues. Immunchemical studies with monoclonal antibody F4/80. *J Exp Med* 161: 475-489
- Levi S, Santambrogio P, Cozzi A, Rovida E, Corsi B, Tamborini E, Spada S, Albertini A, Arosio P (1994) The role of the L-chain in ferritin iron incorporation. Studies of homo and heteropolymers. *J Mol Biol* 238: 649-654
- Levy JE, Montross LK, Cohen DE, Fleming MD, Andrews NC (1999) The C282Y mutation causing hereditary haemochromatosis does not produce a null allele. *Blood* 94: 9-11
- Li J, Ning G, Duncan SA (2000) Mammalian hepatocyte differentiation requires the transcription factor HNF-4-alpha. *Genes Dev* 14: 464-474
- Li JY, Ram G, Gast K, Chen X, Barasch K, Mori K, Schmidt-Ott K, Wang J, Kuo HC, Savage-Dunn C, Garrick MD, Barasch J (2004) Detection of intracellular iron by its regulatory effect. *Am J Physiol Cell Physiol* 287: 1547-1559

- Li Z, Bai X, Ruparel H, Kim S, Turro NJ, Ju J (2003) A photocleavable fluorescent nucleotide for DNA sequencing and analysis. *Proc Natl Acad Sci USA* 100: 414-419
- Liberati C, Ronchi A, Lievens P, Ottolenghi S and Mantovani R (1998) NF-Y organizes the gamma-globin CCAAT boxes region. *J Biol Chem* 273: 16880-16889
- Lieu PT, Heiskala M, Peterson PA, Yang Y (2001) The roles of iron in health and disease. *Mol Aspects of Med* 22: 1-87
- Lombard M, Chua E, O'Toole P (1997) Regulation of intestinal non-haem iron absorption. *Gut* 40: 435-439
- Loots GG, Ovcharenko I, Patcher L, Dubchak I, Rubin EM (2002) rVISTA for comparative sequence-based discovery of functional transcription factor binding sites. *Genome Res* 12: 832-839
- Lundvall O, Weinfield A, Lundin P (1970) Iron storage in porphyria cutanea tarda. *Acta Med Scand* 188: 37-53
- MacPhail AP, Simon MO, Torrance JD, Charlton RW, Bothwell TH, Isaacson C (1979) *Am J Clin Nutr* 32: 1272-1278
- Maines MD (1997) The heme oxygenase system: a regulator of second messenger gases. *Annu Rev Pharmacol Toxicol* 37: 517-544
- Majuri R and Grasbeck R (1987) A rosette receptor assay with haem-microbeads. Demonstration of a haem receptor on J562 cells. *Eur J Haematol* 38: 21-25
- Maniatis T, Fritsch EF, Sambrook (1982) *Molecular cloning: A laboratory Manual*. Cold Spring Harbor Lab, New York.
- Maruyama Y, Lindholm B, Stenvinkel P (2004) Inflammation and oxidative stress in ESRD- the role of myeloperoxidase. *J Nephrol* 17: S72-S76
- Masotti C, Armelin-Correa LM, Splendore A, Lin CJ, Barbosa A, Sogayor MC, Passos-Beuno MR (2005) A functional SNP in the promoter region of *TCOF1* is associated with reduced gene expression and YY1 DNA-protein interaction. *Gene* 359: 44-52
- Mattei MG, Dubart A, Beaupain D, Goossens M, Mattei JF (1985) Localization of uroporphyrinogen decarboxylase gene to 1p34, by in situ hybridization. *Cytogenet Cell Genet* 40: 692
- Mauzerall DC (1998) Evolution of porphyrins. *Clin Dermatol* 16: 195-201
- Mauzerall DC and Granick S (1958) Porphyrin biosynthesis in erythrocytes III. Uroporphyrinogen and its decarboxylation. *J Biol Chem* 232: 1141-1162

- McGill JR, Boyd D, Barrett KJ, Drysdale JW, Moore CM (1984) Localization of human ferritin H (heavy) and L (light) subunits by in situ hybridization. *Am J Hum Genet* 36: 146S
- McGrath DS, Goh N, Foley PJ, du Bois RM (2001) Sarcoidosis: genes and microbes-soil or seed. *Sarcoidosis Vasc Diffuse Lung Dis* 18: 149-164
- 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 (2001) An iron-regulated ferric reductase associated with the absorption of dietary iron. *Science* 291: 1755-1759
- 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 (2000) A novel duodenal iron-regulated transporter, IREG1, implicated in the basolateral transfer of iron to the circulation. *Mol Cell* 5: 299-309
- McKie AT, Raja KB, Peters TJ, Farzaneh F, Simpson RJ (1996) Expression of genes involved in iron metabolism in mouse intestine. *Am J Physiol* 271: G772-G779
- McManus JF, Begley CG, Sassa S, Ratnaik (1996). Five new mutations in the uroporphyrinogen decarboxylase gene identified in families with cutaneous porphyria. *Blood* 88: 3589-3600
- Mehrany K, Drage LA, Branhagen DJ, Pittelkow MR (2003) Association of porphyria cutanea tarda with hereditary haemochromatosis. *J Am Acad Dermatol* 51: 205-211
- Mendez M, Rosetti MV, Batlle AM, Parera VE (2005) The role of inherited and acquired factors in the development of porphyria cutanea tarda in the argentinean population. *J Am Acad Dermatol* 52: 417-424
- Mendez M, Sorkin L, Rossetti MV, Astrin KH, Batlle AMC, Parera VE, Aizencang G, Desnick RJ (1998) Familial Porphyria Cutanea Tarda: Characterization of Seven Novel Uroporphyrinogen Decarboxylase Mutations and Frequency of Common Hemochromatosis Alleles. *Am J Hum Genet* 63: 1363-1375
- Merryweather-Clarke AT, Cadet E, Bomford A, Capron D, Viprakasi V, Miller A, McHugh PJ, Chapman RW, Pointon JJ, Wilmhurst, VLC, Livesey KJ, Tanphaichitr V, Rochette J, Robson KJH (2003) Digenic inheritance of mutations in HAMP and HFE results in different types of haemochromatosis. *Hum Molec Genet* 12: 2241-2247
- Merryweather-Clarke AT, Pointon JJ, Jouanolle AM, Rochette J, Robson KJH (2000) Geography of HFE C282Y and H63D mutations. *Genet Test* 4: 183-198
- Merryweather-Clarke AT, Pointon JJ, Shearman JD, Robson KJH (1997) Global prevalence of putative haemochromatosis mutations. *J Med Genet* 34: 275-278

- Miao X, Yu C, Tan W, Xiong P, Liang G, Lu W, Lin D (2003) A functional polymorphism in the matrix metalloproteinase-2 gene promoter (-1306C/T) is associated with the risk of development but not metastasis of gastric cardia adenocarcinoma. *Cancer Res* 63: 3987-3990
- Miller MP (1997) Tools for population genetic analysis (TFPGA) 1.3. A windows program for the analysis of allozyme and molecular population genetic data.
- Miller S, Dykes D, Polesky H (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucl Acids Res* 16: 1215
- Mills M and Payne SM (1995) Genetics and regulation of heme iron transport in *Shigella dysenteriae* and detection of an analogous system in *Escherichia coli* O157:H7. *J Bacteriol* 177: 3004-3009
- Milos PM and Zaret KS (1992) A ubiquitous factor is required for C/EBP-related proteins to form stable transcription complexes on an albumin promoter segment in vitro. *Genes Dev* 6: 991-1004
- Montminy M, Koo SH, Zhang X (2004) The CREB family: key regulators of hepatic metabolism. *Ann Endocrinol (Paris)* 65: 73-75
- Montosi G, Donovan A, Totaro A, Garuti C, Pignatti E, Cassanelli S, Trenor CC, Gasparini P, Andrews NC, Pietrangelo A (2001) Autosomal-dominant hemochromatosis is associated with a mutation in the ferroportin (SLC11A3) gene. *J Clin Invest* 108: 619-623
- Montgomery SB, Griffith OL, Schuetz JM, Brooks-Wilson A, Jones SJM (2007) A survey of genomic properties for the detection of regulatory elements. *Plos Comput Biol* 3: e106
- Moore JH (2004) Computational analysis of gene-gene interactions using multifactor dimensionality reduction. *Expert Rev Mol Diagn* 4: 795-803
- Mor Z and Capsi E (1997) Cutaneous complications of hormonal replacement therapy. *Clin Dermatol* 15: 147-154
- Moran-Jimenez MJ, Ged C, Romana M, Enriquez De Salamance R, Taieb G, Topi G, D'Alessandro L, de Verneuil H (1996) Uroporphyrinogen decarboxylase: complete human gene sequence and molecular study of three families with hepatoerythropoetic porphyria. *Am J Hum Genet* 58: 712-121
- Morgan EH (1983) Effect of pH and iron content of transferrin on its binding to reticulocyte receptors. *Biochim Biophys Acta* 762: 498-502
- Moss GM (1987) Recommendations on organic & biochemical nomenclature, symbols & terminology. *Pure Appl Chem* 59: 779-832

- Motala AA, Busson M, Al-Harbi EM, Khuzam MA, Al-Amari EM, Arekat MR, Almawi WY (2005) Susceptible and protective human leukocyte antigen class II alleles and haplotypes in bahraini type 2 (Non-Insulin-Dependent) diabetes mellitus patients. *Clin Diagn Lab Immunol* 12: 213-217
- Motta MC, Caretti G, Badaracco GF, Mantovani R (1999) Interactions of the CCAAT-binding trimer NF-Y with nucleosomes. *J Biol Chem* 274: 1326-1333
- Muckenthaler M, Roy CN, Custodio AO, Minana B, deGraaf J, Montross LK, Andrews NC, Hentze MW (2003) Regulatory defects in liver and intestine implicate abnormal hepcidin and *Cybrd1* expression in mouse hemochromatosis. *Nat Genet* 34: 102-107
- Mukhopadhyay CK, Attieh ZK, Fox L (1998) Role of ceruloplasmin in cellular iron uptake. *Science* 279: 714-717
- Murphy GH (1993) The cutaneous porphyrias: a review of diagnosis and management. *J Eur Acad Derm Venereol* 2: 11-17
- Naylor SL, Marshall A, Solomon A, McGill JR, McCombs J, Magnuson VL, Moore CM, Lalley PA, Pentecost BT, Teng C (1987) Lactoferrin maps to human chromosome 3(q21-q23) and mouse chromosome 9. *Cytogenet Cell Genet* 46: 669
- Nemeth E, Tuttle MS, Powelson J, Vaughn MB, Donovan A, Ward DM, Ganz, Kaplan J (2004) Hepcidin regulates cellular iron efflux by binding to ferroportin and inducing its internalization. *Science* 306: 2090-2093
- Nicolas G, Bennoun M, Devaux I, Beaumont C, Grandchamp B, Kahn A, Vaulont S (2001) Lack of hepcidin gene expression and severe tissue iron overload in upstream regulatory factor 2 (USF2) knockout mice. *Proc Natl Acad Sci USA* 98: 8780-8785
- Niehof M, Kubicka S, Zender L, Manns MP, Trautwein C (2001) Autoregulation enables different pathways to control CCAAT/enhancer binding protein (C/EBP) transcription. *J Mol Biol* 309: 855-868
- Nishiyama C, Yokota T, Nishiyama M, Ra C, Okumura K, Ogawa H (2003) Molecular cloning of rat transcription factor YY1. *Biosci Biotechnol Biochem* 67: 654-658
- Njajou OT, Vaessen N, Joosse M, Berghuis B, van Dongen JW, Breuning MH, Snijders PJ, Rutten WP, Sandkuijl LA, Nollau P and Wagener C (1997) Methods for detection of point mutations: performance and quality assessment. *Clin Chem* 43: 1114-1128
- Nordfors L, Lindholm, Stenvinkel P (2005) End-stage renal disease-not an equal opportunity disease: the role of genetic polymorphisms. *J Intern Med* 258: 1-12
- Nordman Y and Puy H (2002) Human hereditary hepatic porphyrias. *Clin Chim Acta* 325: 17-37

- Nurse GT, Weiner JS, Jenkins T (1985) The growth of hybrid communities. In: The peoples of southern Africa and their affinities. Oxford Clarendon, Oxford, pp218
- Ober C, Aldrich CL, Chervoneva I, Billstrand C, Rahimov F, Gray HL, Hyslop T (2003) Variation in the HLA-G promoter region influences miscarriage rates. *Am J Hum Genet* 72: 1425–1435
- Odom DT, Dowell R, Jacobsen ES, Nekludova L, Rolfe PA, Danford TW, Gifford DK, Fraenkel E, Bell GI, Young RA (2006) Core transcriptional regulatory circuitry in human hepatocytes. *Mol Sys Biol* 27: 1868-1880
- Ogura IN, Bonen DK, Inohara N, Nicolae DL, Chen FF, Ramos R, Britton H, Moran T, Kargliuskus R, Duerr RH, Achkar JP, Brant SR, Bayless TM, Kirschner BSM, Hanaver SB, Nunez G, Cho JH (2001) A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* 411: 603-606
- Ono M, Yaguchi H, Ohkura N, Kitabayashi I, Nagamura Y, Nomura T, Miyachi Y, Tsukada T, Sakaguchi S (2007) Foxp3 controls regulatory T-cell function by interacting with AML1/Runx1. *Nature* 446: 685-689
- Oostra BA, van Duijn CM, Heutink P (2001) A mutation in SLC11A3 is associated with autosomal dominant hemochromatosis. *Nat Genet* 28: 213-214
- Pajukanta P, Lilja HE, Sinsheimer JSM, Cantoe RM, Lusa AJ, Gentile M, Duan XJ, Soro-Paavonen A, Naukkarinen J, Saarela J, Laakso M, Ehnholm C, Taskinen MR (2004) Familial combined hyperlipidemia is associated with upstream transcription factor 1 (USF1). *Nature Genet* 36: 371-376
- Pani L, Quian XB, Clevidence D, Costa RH (1992) The restricted activity of the liver transcription factor hepatocyte nuclear factor 3 beta involves a cell-specific factor and positive autoactivation. *Mol Cell Biol* 12: 552-562
- Pantopoulos K, Hentze MW (2005a) Rapid responses to oxidative stress mediated by iron regulatory protein. *EMBO J* 14: 2917–2924
- Pantopoulos K, Hentze MW (2005b) Nitric oxide signaling to iron-regulatory protein: direct control of ferritin mRNA translation and transferrin receptor mRNA stability in transfected fibroblasts. *Proc Natl Acad Sci U S A* 92: 1267–1271
- Papanikolaou G, Samuels ME, Ludwig EH, MacDonald MLE, Franchini PL, Dube MP, Andres L, MacFarlane J, Sakellaropoulos N, Politou M, Nemeth E, Thompson J, Risler JK, Zaborowska C, Babakiaiff R, Radomski CC, Pape TD, Davidas O, Christakis J, Brissot P, Lockitch G, Ganz T, Hayden MR, Goldberg YP (2004) Mutations in HFE2 cause iron overload in chromosome 1q-linked juvenile hemochromatosis. *Nature Genet* 36: 77-82
- Park CH, Valore EV, Waring AJ, Ganz T (2001) Hepsidin, a urinary antimicrobial peptide synthesized in the liver. *J Biol Chem* 276: 7806-7810

- Parkkila S, Waheed A, Britton RS, Bacon BR, Zhou XY, Tomatsu S, Fleming BR (1997) Association of the transferrin receptor in human placenta with HFE, the protein defective in hereditary hemochromatosis. *Proc Natl Acad Sci* 94: 13198-13202
- Parviz , Matullo C, Garrison WD, Savatski L, Adamson JW, Ning G, Kaestner KH, Rossi JM, Zaret KS, Duncan SA (2003) Hepatocyte nuclear factor 4-alpha controls the development of a hepatic epithelium and liver morphogenesis. *Nature Genet.* 34: 292-296
- Pérez-Lezaun A, Calafell F, Seielstad M, Mateu E, Comas D, Bosch E, Bertranpetit J (1997) Population genetics of Y-chromosome short tandem repeats in humans. *J Mol Evol* 45: 265–270
- Perneger TV (1998) What's wrong with Bon Ferroni adjustments. *BMJ* 316: 1236-1238
- Pesole G, Liuni S, Grillo G, Licciulli F, Mignone F, Gissi C, Sacconne C (2002) UTRdb and UTRsite: specialized database of sequences and functional elements of 5' and 3' untranslated regions of eukaryotic mRNAs. Update 2002. *Nucleic Acids Res* 30: 335-340
- Pesole G, Mignone F, Gissi C, Grillo G, Licciulli F, Liuni S (2001) Structural and functional features of eukaryotic mRNA untranslated regions . *Gene* 276: 73-81
- Pesole G, Grillo G, Larizza A, Liuni (2000) The untranslated regions of eukaryotic mRNAs: Structure, function, evolution and bioinformatic tools for their analysis. *Brief Bioinform* 1: 236-249
- Peters LL, Bishop TR, Andrews NC (1993) Globin-enhancer binding factor NF-E2 is implicated in the regulation of heme biosynthesis and iron uptake in mk/mk mice. *Blood* 82: 179a
- Petrak J and Vyoral D (2005) Hephaestin-a ferroxidase of cellular iron export. *Int J Biochem Cell Biol* 37: 1173-1178
- Phillips JD, Bergonia HA, Reilly CA, Franklin MR, Kushner JP (2007) A porphomethene inhibitor of uroporphyrinogen decarboxylase causes porphyria cutanea tarda. *Proc Natl Acad Sci USA* 104: 2079-2084
- Phillips JD, Jackson LK, Bunting M, Franklin MR, Thomas KR, Levy JE, Andrews NC, Kushner JP (2001) A mouse model of familial porphyria cutanea tarda. *Proc Natl Acad Sci USA* 98: 259-264
- Pietrangelo A (2004) Hereditary Haemochromatosis- a new look at an old disease. *New Engl J Med* 350: 2383-2397
- Pietrangelo A, Rocchi E, Casalgrandi G, Rigo G, Ferrari A, Perini M, Ventura E, Cairo G (1992) Regulation of transferrin, transferrin receptor, and ferritin genes in human duodenum. *gastroenterol* 102: 802–809
- Pigeon C, Ilyin G, Courselaud B (2001) A new mouse liver-specific gene, encoding a protein homologous to human antimicrobial peptide hepcidin, is overexpressed during iron overload. *J Biol Chem* 276: 7 811-7819

- Ponka P (1999) Iron metabolism: physiology and pathophysiology. *Trace Elem Res Hum* 5: 55-57
- Pountney DJ, Konijn AM, McKie AT, Peters TJ, Raja KB, Salisbury JR, Simpson RJ (1999) Iron proteins of duodenal enterocytes isolated from mice with genetically and experimentally altered iron metabolism. *Br J Haematol* 105: 1066-1073
- Prior SJ, Hagberg JM, Paton CM, Douglass LW, Brown MD, McLenithan JC, Roth SM (2006) DNA sequence variation in the promoter region of the VEGF gene impacts VEGF gene expression and maximal oxygen consumption. *Am J Physiol Heart Circ Physiol* 290: H1848-H1855
- Puffenberger EG, Hosoda K, Washington SS, Nakao K, deWit D, Yanagisawa M, Chakravarti A (1994) A missense mutation of the endothelial-B receptor gene in multigenic Hirschsprungs disease. *Cell* 79: 1257-1266
- Qiu C, Fratiglioni L, Karp A, Winblad B, Bellander T (2004) Occupational exposure to electromagnetic fields and risk of Alzheimer's disease. *Epidemiol* 15: 687-694
- Randell EW, Parkes JG, Olivieri NF, Templeton DM (1994) Uptake of non-transferrin-bound iron by both reductive and non reductive processes is modulated by intracellular iron. *J Biol Chem* 269: 16046-16053
- Reuss E, Fimeers R, Kruger A, Becker C, Rittner C, Hohler T (2002) Differential regulation of interleukin-10 production by genetic and environmental factors - a twin study. *Genes Immun* 3: 407-413
- Rich MW (1999) Porphyria cutanea tarda. Don't forget to look at the urine. *Postgrad Med* 105: 208-214
- Ridley M (2004) *Nature via nature*. Harper Perennial, London, pp248
- Roberts AG, Elder GH, Newcombe RG, Enriquez de Salamanca R, Munoz JJ (1988) Heterogeneity of familial porphyria cutanea tarda. *J Med Genet* 25: 669-676
- Roberts AG, Whatley SD, Morgan RR, Worwood M, Elder GH (1997) Increased frequency of the haemochromatosis Cys282Tyr mutation in sporadic porphyria cutanea tarda. *Lancet* 349: 321-323
- Robinson SW, Clothier B, Akhtar RA, Yang AL, Latour I, Van Ijperen C, Festing MF, Smith AG (2002) Non-ahr gene susceptibility loci for porphyria and liver injury induced by the interaction of 'dioxin' with iron overload in mice. *Mol Pharmacol* 61: 674-681
- Rocchi E, Stella AM, Cassanelli M, Borghi A, Nardella N, Seium Y, Casalgrandi G (1995) Liposoluble vitamin and naturally occurring caretonoids in porphyria cutanea tarda. *Eur J Clin Invest* 25: 510-514
- Roeser HP, Lee GR, Nacht S, Cartwright GE (1970) Role of ceruloplasmin in iron metabolism. *J Clin Invest* 49: 2408

- Roetto A, Papanikolaou G, Politou M, Alberti F, Girelli D, Christakis J, Loukopoulos D, Camaschella C (2003) Mutant antimicrobial peptide hepcidin is associated with severe juvenile hemochromatosis. *Nature Genet* 33: 21-22
- Roetto A, Totaro A, Piperno A, Piga A, Longo F, Garozzo G, Cali A, De Gobbi M, Gasparini P, Camaschella C (2001) New mutations inactivating transferrin receptor 2 in hemochromatosis type 3. *Blood* 97: 2555-2560
- Romana M, Grandchamp B, Dubart A, Amselem S, Chabret C, Nordmann Y, Goossens M, Romeo P-H (1991) Identification of a new mutation responsible for hepatoerythropoietic porphyria. *Eur J Clin Invest* 21: 225-229
- Romana M, Dubart A, Beaupain D, Chabret C, Goossens M, Romeo PH (1987) Structure of the gene for human uroporphyrinogen decarboxylase. *Nucleic Acids Res* 15: 7343-7356
- Romier C, Cocchiarella F, Mantovani R, Moras D (2003) The NF-YB/NF-YC structure gives insight into DNA binding and transcription regulation by CCAAT factor NF-Y. *J Biol Chem* 278: 1336-1345
- Rosenbauer F, Wagner K, Kutok JL, Iwasaki H, Le Beau MM, Okuno Y, Akashi K, Fiering S, Tenen DG (2004) Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. *Nature Genet* 36: 624-630
- Rossi E (2005) Hepcidin-the iron regulatory hormone. *Clin Biochem Rev* 26: 47-49
- Rothman KJ (1990) No Adjustments are needed for multiple comparisons. *Epidemiology* 1: 43-46
- Rouault TA, Haile DJ, Downey WE, Philpott CC, Tang C, Samaniego F, Chin J, Paul I, Orloff D, Harford JB, Klausner RD (1992) An iron-sulfur cluster plays a novel regulatory role in the iron-responsive element binding protein. *Biomaterials* 5: 131-140
- Rouault TA, Tang CK, Kaptain S, Burgess WH, Haile DJ, Samaniego F, McBride OW, Harford JB, Klausner RD (1990) Cloning of the cDNA encoding an RNA regulatory protein-the human iron-responsive element-binding protein. *Proc Natl Acad Sci USA* 87: 7958-7962
- Roy CN and Andrews NC (2001) Recent advances in disorders of iron metabolism: mutations, mechanisms and modifiers. *Hum Mol Genet* 10: 2181-2186
- Roy CN and Enns CA (2000) Iron homeostasis: new tales from the crypt. *Blood* 96: 4020-4027
- Royle NJ, Irwin DM, Koschinsky ML, MacGillivray RTA, Hamerton JL (1987) Human genes encoding prothrombin and ceruloplasmin map to 11p11-q12 and 3q21-24, respectively. *Somat. Cell Molec. Genet* 13: 285-292

- Rozen S and Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds), *Bioinformatics methods and protocols: methods in molecular biology*. Human Press, Totowa, NJ, pp 365-386
- Rutter M (2007) Gene-environment interdependence. *Developmental Sci* 10: 12-18
- Sampietro M, Fiorelli G, Fargion S (1999) Iron overload in porphyria cutanea tarda. *Haematologica* 84: 248-253
- Sampietro M, Piperno A, Lupica L, Arosio C, Vergani A, Corbetta N, Malosio I, Mattioli M, Fracanzani AL, Cappellini MD, Fiorelli G, Fargion S (1998) High prevalence of the His63Asp HFE mutation in Italian patients with porphyria cutanea tarda. *Hepatology* 27: 181-184
- Sandelin A, Alkema W, Engstrom P, Wasserman WW, Lenhard B (2004) JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res* 32: 91-94
- Sarkany RPE (2001) The management of porphyria cutanea tarda. *Clin Exp Dermatol* 26: 225-232
- Sassa S (2002) The porphyrias. *Photodermatol Photoimmunol Photomed* 18: 56-67
- Sayers MH, English G, Finch C (1994) Capacity of the store-regulator in maintaining iron balance. *Am J Hematol* 47: 194-197
- Schaid DJ and Jacobsen SJ (1999) Biased tests of association: comparisons of allele frequencies when departing from Hardy-Weinberg proportions. *Am J Epidemiol* 149: 706-711
- Schimanski LM, Drakesmith H, Merryweather-Clarke AT, Viprakasit V, Edwards JP, Sweetland E, Bastin JM, Cowley D, Chinthammitr Y, Robson KJH, Townsend ARM (2005) *In vitro* functional analysis of human ferroportin (FPN) and hemochromatosis-associated FPN mutations. *Blood* 105: 4096-4102
- Seto E, Shi Y, Shenk T (1991) YY1 is an initiator sequence-binding protein that directs and activates transcription *in vitro*. *Nature* 354: 241-245
- Shaulian E and Karin M (2002) AP-1 as a regulator of cell life and death. *Nat Cell Biol* 4: E131-E136
- Shrivastava A and Calame K (1994) An analysis of genes regulated by the multi-functional transcriptional regulator Yin Yang-1. *Nucleic Acid Res* 22: 5151-5155
- Sheffield VC, Beck JS, Kwitek AE, Sandstrom DW, Stone EM (1993) The sensitivity of single-stranded conformation analysis for the detection of single base substitutions. *Genomics* 16: 325-332

- Siah CW, Ombiga J, Adams LA, Trinder D, Olynyk JK (2006) Normal iron metabolism and the pathophysiology of iron overload disorders. *Clin Biochem Rev* 27: 5-16
- Siegenberg D (1991) Ascorbic acid prevents the dose-dependent inhibitory effects of polyphenols and phylates on non-haem iron absorption. *Am J Clin Nutr* 53: 537-541
- Simon M, Le Mignon L, Fauchet R, Yaouanq J, David V, Edan G, Bourel M (1987) A study of 609 HLA haplotypes marking for the hemochromatosis gene: (1) mapping of the gene near the HLA-A locus and characters required to define a homozygous population and (2) hypothesis concerning the underlying cause of hemochromatosis-HLA association. *Am J Hum Genet* 41 :89-105
- Sinclair PR, Gorman N, Shedlofsky SI, Honsinger CP, Sinclair JF, Kavagas MR, Anderson KE (1997) Ascorbic acid deficiency in porphyria cutanea tarda. *J Lab Clin Med* 130: 197-20
- Sinha A, Dixon N, O'Sullivan MM, Sowden JM (1999) Porphyria cutanea tarda in a patient with systemic lupus erythematosus. *Rheumatol* 38: 1166-1168
- Sladek FM, Zhong W, Lai E, Darnell JE Jr (1990) Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. *Genes and Dev* 4: 2353-5365
- Smith AG, Francis JE (1983) Synergism of iron and hexachlorobenzene inhibits hepatic uroporphyrinogen decarboxylase in inbred mice. *Biochem J* 214: 909-913
- Smith LH (1990) Overview of haemochromatosis. *West J Med* 153: 296-308
- Smith SG, Rao KR, Jackson AH (1980) The porphyrins of normal human urine, with a comparison of the excretion pattern in porphyria cutanea tarda. *Int J Biochem* 12: 1081-1084
- Stormo G (2000) DNA binding sites: representation and discovery. *Bioinformatics* 16: 16-23
- Strachan T and Read AP (2004) *Human molecular genetics* 3. Garland Science, London, pp 101-112
- Strittmatter WJ and Roses AD (1996) Apolipoprotein E and Alzheimer's disease. *Ann Rev Neurosci* 19: 53-77
- Stuckey R (2006) Induction of iron homeostasis genes during estrogen-induced uterine growth and differentiation *Mol Cell Endocrinol* 253: 22-29
- Taberlet P, Griffin S, Goossens B, Questiau S, Manceau V, Escaravage N, Waits LP, Bouvet J (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acid Res* 24: 3189-3194
- Taddeini L and Watson CJ (1968) The clinical porphyrias. *Semin Hematol* 5: 335-369

- Taillon-Miller P, Gu Z, Li Q, Hiliier L, Kwok PY (1998) Overlapping genomic sequences: A treasure trove of single nucleotide polymorphisms. *Genome Res* 8: 748-754
- Takahashi N, Ortel TL, Putnam FW (1984) Single-chain structure of human ceruloplasmin: the complete amino acid sequence of the whole molecule. *Proc Natl Acad Sci USA* 81: 390-394
- Takeda A, Devenyi A, Connor JR (1998) Evidence for non-transferrin-mediated uptake and release of iron and manganese in glial cell cultures from hypotransferrinemic mice. *J Neurosci Res* 51: 454-462
- Taljaard JFM, Shanley BC, Stewart-Wynne E, Deppe WM, Joubert SM (1972) Studies of low dose chloroquine therapy and the action of chloroquine in symptomatic porphyria. *Brit J Derm* 87: 261-269
- Tamaki T, Ohnishi K, Hartl C, LeRoy EC, Trojanowska M (1995) Characterization of a GC-rich region containing Sp1 binding site(s) as a constitutive responsive element of the alpha 2(I) collagen gene in human fibroblasts. *J Biol Chem* 270: 4299-4304
- Tapiero H, Gate L, Tew KD (2001) Iron: deficiencies and requirements. *Biomed Pharmacother* 55: 324-332
- Tenhunen R, Marver HS, Schmid R (1969) Microsomal heme oxygenase. Characterization of the enzyme. *J Biol Chem* 244: 6388-6394
- Thadani H, Deacon A, Peters T (2000) Diagnosis and management of porphyria. *BMJ* 320: 1647-1651
- Theil EC (1998) The iron responsive element (IRE) family of mRNA regulators. Regulation of iron transport and uptake compared in animals, plants and microorganisms. *Metal Ions Biol Syst* 35: 403-434
- Thein SL and Wallace RB (1986) The use of synthetic oligonucleotides as specific hybridization probes in the diagnosis of genetic diseases. In: Davies KE (ed), *Human genetic diseases: a practical approach*. IRL Press, Oxford, pp33-50
- Thomas C and Oates PS (2004) Ferroportin/IREG-1/MTP-1/SLC40A1 modulates the uptake of iron at the apical membrane of enterocytes. *Gut* 53: 44-49
- Thomas DC and Witte JS (2002) Population stratification: A problem for case-control studies of candidate-gene associations. *Cancer Epidemiol Biomarkers Prev* 11: 505-512
- Thomas DL, Astemborski J, Rai RM, Anania FA, Schaeffer M, Galai N, Nolt K, Nelson KE, Strathdee SA, Johnson MS, Laeyendecker O, Boitnott J, Wilson LE, Vlahov D (2000) The natural history of hepatitis C virus infection. *JAMA* 284: 450-456
- Thorstensen K and Romslo I (1990) The role of transferrin in the mechanism of cellular iron uptake. *Biochem J* 271: 1-10

- Tjian R (1995) Molecular machines that control genes. *Sci Am* 272: 54-61
- Tirona RG, Lee W, Leake BF, Lan LB, Cline CB, Lamba V, Parviz F, Duncan SA, Inoue Y, Gonzalez F J, Schuetz E G, Kim RB (2003) The orphan nuclear receptor HNF4-alpha determines PXR- and CAR-mediated xenobiotic induction of CYP3A4. *Nature Med* 9: 220-224
- Toll A, Cells R, Ozalla MD, Bruguera M, Herrero C, Ercilla MG (2006) The prevalence of HFE C282Y gene mutation is increased in Spanish patients with porphyria cutanea tarda without hepatitis C virus infection. *J Eur Acad Dermatol Venereol* 20: 1201-1206
- Trinder D, Fox C, Vautier G, Olynyk JK (2002) Molecular pathogenesis of iron overload. *Gut* 51: 290-295
- Trinder D, Pates PS, Thomas C, Sadleir J, Morgan EH (2000) Localisation of divalent metal transporter 1 (DMT1) to the microvillus membrane of rat duodenal enterocytes in iron deficiency, but to hepatocytes in iron overload. *Gut* 46 270:276
- Tronche F, Ringeisen F, Blumenfeld M, Yaniv M, Pontoglio M (1997) Analysis of the distribution of binding sites for a tissue-specific transcription factor in the vertebrate genome. *J Mol Biol* 266: 231-245
- Tsai FY, Keller G, Kuo FG, Weiss M, Chen J, Rosenblatt M, Alt FW, Orkin SH (1994) An early haemopoietic defect in mice lacking transcription factor GATA 2. *Nature* 371: 221-226
- Tsiftoglou AS, Tsamadou AI, Papadopoulou LC (2006) Heme as a key regulator of major mammalian cellular functions: Molecular, cellular, and pharmacological aspects. *Pharmacol Ther* 111: 327-345
- Ueda A, Takeshita F, Yamashiro S, Yoshimura T (1998) Positive regulation of the human macrophage stimulating protein gene transcription. Identification of a new hepatocyte nuclear factor-4 (HNF-4) binding element and evidence that indicates direct association between NF-Y and HNF-4. *J Biol Chem* 273: 19339-19347
- Umek RM, Friedman AD, McKnight SL (1990) CCAAT-enhancer binding protein: a component of a differentiation switch. *Science* 251: 288-292
- Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, Davis RW, Cavalli-Sforza LL, Oefner PJ (1997) Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res* 7: 996-1005
- Usdin K and Grabczyk E (2000) DNA repeat expansions and human disease. *Cell Mol Life Sci* 57: 914-931
- Uys CJ and Eales L (1963) The histopathology of the liver in acquired (symptomatic) porphyria. *S Afr J Lab Clin Med* 9: 190-197

- Vallet VS, Casado M, Henrion AA, Bucchini D, Raymondjean M, Kahn A, Vaulont S (1998) Differential roles of upstream stimulatory factors 1 and 2 in the transcriptional response to liver genes. *J Biol Chem* 273: 20175-20179
- van Deursen D, Botma GJ, Jansen H, Verhoeven (2007) Comparative genomics and experimental promoter analysis reveal liver-specific elements in mammalian hepatic lipase genes. *Genomics* 8: 99
- Viollet B, Lefrancois-Martinez AM, Henrion A, Kahn A, Raymondjean M, Martinez A (1996) Immunochemical characterization and transacting properties of upstream stimulatory factor isoforms. *J Biol Chem* 271: 1405-1415
- Vulpe CD, Kuo YM, Murphy TL, Cowley L, Askwith C, Libina N, Gitschier J, Anderson GJ (1999) Hephaestin, a ceruloplasmin homologue implicated in intestinal iron transport, is defective in the sla mouse. *Nature Genet* 21: 195-199
- Waldenstrom J (1937) Studien über Porphyrrie. *Acta Med Scan* 82: 1-524
- Waldenstrom J and Haeger-Aronsen B (1960) The liver in porphyria cutanea tarda. *Ann Intern Med* 53: 286-292
- Wang J, Liao G, Usuka J, Peltz G (2005) Computational genetics: from mouse to human? *Trends Genet* 21: 526-532
- Warin RP (1963) Porphyria cutanea tarda associated with oestrogen therapy for prostatic carcinoma. *Brit J Dermat* 75: 298-299
- Watson CJ and Schwartz S (1941) A simple test for urinary porphobilinogen. *Proc Soc Exp Biol Med* 47: 393-394
- Weiling F (1991) Historical study: Johann Gregor Mendel 1822-1884. *Am J Med Genet* 107: 612-614
- Weizer-Stern O, Adamsky K, Margalit O, Ashur-Fabian O, Givol D, Amariglio N, Gideon R (2007) Hcpidin, a key regulator of iron metabolism, is transcriptionally activated by p53. *Brit J Haematol* 138: 253-263
- Wessling-Resnick M (1999) Biochemistry of iron uptake. *Crit Rev Biochem Mol Biol* 34: 285-314
- West AP, Giannetti AM, Herr AB, Bennett MJ, Nangiana JS, Pierce JR, Weiner LP, Snow PM, Bjorkman PJ (2001) Mutational analysis of the transferrin receptor binding sites. *J Mol Biol* 313: 385-387
- Whitby FG, Phillips JD, Kushner JP, Hill CP (1998) Crystal structure of human uroporphyrinogen decarboxylase. *EMBO J* 17: 2463-2471
- Williams MH, Andrews SC, Treffry A, Harrison PF (2006) The relationship between ferritin and haemosiderin. *Hyperfine Interact* 29: 1447-1450

- Wingender E, Chen X, Fricke E, Geffers R, Hehl R, Liebich I, Krull M, Matys V, Michael H, Ohnauer R, Pruss M, Schacherer F, Thiele S, Urbach S (2001) The TRANSFAC system on gene expression regulation. *Nucleic Acid Res* 29: 281-283
- Wingender E (1997) Classification scheme of eukaryotic transcription factors. *Mol Biol* 31: 483-497
- Wolfram S (1952) Uber Porphyrinkolik ein Beitrag zur Symptomatologie des Erythematodes Acutus. *Hautarzt* 3: 298-300
- Wolfrum C, Asilmaz E, Luca E, Friedman J M, Stoffel M (2004) Foxa2 regulates lipid metabolism and ketogenesis in the liver during fasting and in diabetes. *Nature* 432: 1027-1032
- Wolfrum C, Shih DQ, Kuwajima S, Norris AW, Kahn CR, Stoffel M (2003) Role of Foxa-2 in adipocyte metabolism and differentiation. *J Clin Invest* 112: 345-356
- Wood RJ and Han O (1998) Recently identified molecular aspects of intestinal iron absorption. *J Nutr* 128: 1841-1844
- Wright TL, Brissot P, Ma WL, Weisiger RA (1986) Characterization of non-transferrin bound iron clearance by rat liver. *J Biol Chem* 261: 10909-10914
- Yamada Y, Tanaka T, Miyamoto K, Noguchi T (2000) Sp family members and nuclear factor-Y cooperatively stimulate transcription from the rat pyruvate kinase M gene distal promoter region via their direct interactions. *J Biol Chem* 275: 18129-18137
- Yan B, Heus J, Lu N, Nichols RC, Raben N, Plotz P (2001) Transcriptional regulation of the human acid-glucosidase gene: identification of a repressor element and its transcription factors Hes-1 and YY1. *J Biol Chem* 276: 1789-1793
- Yang F, Lum JB, McGill JR, Moore CM, Naylor SL, van Bragt PH, Baldwin WD (1984) Human transferrin: cDNA characterization and chromosomal localization. *Proc Natl Acad Sci USA* 81: 2752-2756
- Yang F, Liu XB, Quinones M, Melby PH, Ghio A, Haile DJ (2002). Regulation of reticuloendothelial iron transporter MTP1 (Slc11a3) by inflammation. *J Biol Chem* 277: 39786-39791
- Yeh KY, Yeh M, Glass J (2004) Hepcidin regulation of ferroportin 1 expression in the liver and intestine of the rat. *Am J Physiol Gastrointest Liver Physiol* 286: G385-G394
- Yuh CH and Davidson EH (1996) Modular cis-regulatory organization of Endo16, a gut-specific gene of the sea urchin embryo. *Development* 122: 1069-1082
- Zaahl MG, Merryweather-Clarke AT, Kotze MJ, van der Merwe S, Warnich L, Robson KJH (2004) Analysis of genes implicated in iron regulation in individuals presenting with primary iron overload. *Hum Genet* 115: 409-417

- Zaret KS (2002) Regulatory phases of early liver development: paradigms of organogenesis. *Nat Rev Genet* 3: 499-512
- Zhang YW, Bae SC, Huang G, Fu YX, Lu J, Ahn MY, Kanno Y, Kanno T, Ito Y (1997) A novel transcript encoding an N-terminally truncated AML1/PEBP2-alpha-B protein interferes with transactivation and blocks granulocytic differentiation of 32Dcl3 myeloid cells. *Mol Cell Biol* 17: 4133-4145
- Zhou XY, Tomatsu S, Fleming RE, Parkila S, Waheed A, Jiang J, Fei Y, Brunt EM, Ruddy DA, Prass CE, Schatzman RC, O'Neill R, Britton RS, Bacon BR, Sly WS (1998) HFE gene knockout produces mouse model of hereditary hemochromatosis. *Proc Natl Sci USA* 95: 2492-2497
- Ziprkowski L, Krakowski A, Crispin M, Szeinberg A (1966) Porphyria cutanea tarda hereditaria. *Israel J Med Sci* 2: 338-343

5.2 ELECTRONIC REFERENCES

ALIBABA2: <http://www.gene-regulation.com/pub/programs/alibaba2>

Basic Local Alignment Search Tool (BLAST): <http://www.ncbi.nlm.nih.gov/blast>

Dermatology information services (DermIS): <http://www.dermis.net>

Ensembl: <http://www.ensembl.org>

European Porphyria Initiative (EPI): <http://www.porphyria-europe.com/>

GenAtlas: <http://www.genatlas.org/>

Haploview: <http://www.broad.mit.edu/mpg/haploview/>

Human Gene Mutation database: <http://www.hgmd.org>

Human Genome Project Information:

http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml

International HapMap Project: <http://www.hapmap.org>

JASPAR CORE: <http://jaspar.genereg.net/>

MATCHTM: <http://gene-regulation.com/pub/programs.html#match>

National Center for Biotechnology Information (NCBI): <http://www.ncbi.nlm.nih.gov>

Online Mendelian Inheritance in ManTM (OMIMTM): <http://www.ncbi.nlm.nih.gov/omim>

Porphyria South Africa online: <http://www.porphyria.uct.ac.za/>

Primer3 [online]: <http://frodo.wi.mit.edu/cgi-bin/primer3>

RefSNP: <http://www.ncbi/blast/dbSNP>

rVISTA: <http://genome.lbl.gov/vista/rvista/submit.shtml>

STATISTICA (StatSoft, Inc. (2007) STAT (data analysis software system), version 8

TRANSFAC^{®7}: <http://www.gene-regulation.com/pub/databases.html#transfac>

APPENDICES



APPENDIX 1: LIST OF REAGENTS AND CHEMICALS USED IN THIS STUDY

List of reagents and chemicals and their suppliers.

Reagent/Chemical	Supplier
Acrylamide (AA)	Fluka
Agarose	Laboratory Specialist Services
Ammonium chloride (NH ₄ Cl)	Seabreeze Suppliers
Ammonium persulphate (APS) (NH ₄) ₂ S ₂ O ₈)	Merck
Bis Acrylamide (BAA)	Sigma
Boric Acid	Seabreeze Suppliers
Bromophenol blue	Seabreeze Suppliers
Cresol Red	Merck
Di-sodium hydrogen orthophosphate anhydrous (Na ₂ HPO ₄)	Seabreeze Suppliers
Ethylene diamine tetra-acetic acid (EDTA) (C ₁₀ H ₁₆ N ₂ O ₈)	Seabreeze Suppliers
Ethanol (EtOH)	Seabreeze Suppliers
Ethidium Bromide (EtBr)	Fluka
Formaldehyde (HCHO)	Calbiochem (Merck)
Formamide (de-ionised) (HCONH ₂)	Merck
Phosphate buffered saline pellets (PBS)	Fluka
Potassium chloride (KCl)	Roche Diagnostics
Potassium di-hydrogen orthophosphate (KH ₂ PO ₄)	Seabreeze Suppliers
Potassium hydrogen carbonate (KHCO ₃)	Merck
Proteinase K	Roche Diagnostics
SABAX water (ddH ₂ O)	Adock Ingram
Sodium acetate (NaClO ₄)	Merck
Sodium chloride (NaCl)	Fluka
Sodium dodecyl sulphate (SDS) (CH ₃ (CH ₂) ₁₁ OSO ₃ Na)	Seabreeze Suppliers
Sodium hydroxide (NaOH)	Merck
Surcrose	Seabreeze Suppliers

Reagent/Chemical	Supplier
N, N, N', N' Tetramethylethylenediamine (TEMED)	Fluka
Tris-HCl	Fluka
Urea (NH ₂) ₂ CO	Sigma
Xylene cyanol	Fluka

APPENDIX 2: PROMOTER SEQUENCES OF *CP*, *CYBRD1*, *HAMP* and *SLC40A1* INDICATING THE POSITION OF DESIGNED PRIMER SETS AND VARIANTS IDENTIFIED IN THE CURRENT STUDY.

Promoter sequences (5'UTR) were selected from Ensembl (www.ensembl.org) and the reference number is indicated above each sequence. The designed primer sets for each gene promoter region (as indicated in Table 2.1) are highlighted using various colours (listed below) on the sequence. For each primer set, the arrow pointing right indicates the forward primer and the arrow pointing left indicates the reverse primer. The ATG corresponding to the translation initiation site is highlighted in bold text. Variation detected in the current study is highlighted in grey and the superscript number indicates the nucleotide position (listed below).

CP PROMOTER REGION

[ENSG00000047457 (ENSEMBL)]

5' CTAAAAAAAAAGCCATTTTCATACTTTTAGACATATACTTCCTGACTAAAAACAGGGTTTTAAG
 AAAAAAGTCTTCCACCTTAAAATTCATATCCACTTCAAAGAGAAAAATTTTCAGTCTAATGATAAACT
 GGGGTACTTTCTATTTACTAAGACTAATCTTTGCCTCTGAAGACAGGTTTAGTTATGGTTTGTTTA
 TTTGATGATTAATAACAAAGTTATTAGCC **CCTGTTAGGCTCTGCTAGTT** ATTTGCATATAGCAATAC
 AGAATGATCTGGTAAAGTTTTTATGCAGATTTGGTCCAAGCTCATAAAACAGATAAAATTCCTCCAA
 ATTAAATTCCTATTTTTGTTTACTTAAATTTGGCATGCTAGATATCATGTTTCATATCATAAAGCT
 GCAATTAATCATTGTAGACTTTTT **AGAAATAGTCATGCACCAC** ATAATAATGTTTGGAGTAAATGA
 CAACTGCATA **TGTGATGGTGGTCCCATAAG** ATTATAATAACATATTTTTACTGTACATTTTTCTAT
 GTTTGCATATGTATAGATACAGAAATATCATTGTGTTACAATTGTCTATAGCATTTCAGCATAGTTA
 TAGACTGTA **CAGGTTTGTAGCCTAGGAGC** AATAGGTATACC¹ATA²AGCCTACCTAGATGTATAGT
 AGGCTATATCATATAAGTCTGTGTAAGTACACTCTGTGATAGTCACACAATGA **CAAAATAGCCTAA**
CAATGCATTT **CTCAGAACGTATCCCTGTCAT** TAAGAAT³GACTGTAGTAAAAAGCTACATGTTTTGA
 TTCATTTTACATAAATCTTATCCTGATTTACTAAAAGTGCCTGTTT⁴GCAGGACACT⁵GTGC⁶TGTG
 ACATATTGT **CAGAGGAACAAAGGGAGCAT** TTAGACACGTTCTCTGCCCTCCTGGAATTTAC
CAGAGGAACAAAGGGAGCAT

ACA⁷ATGAATGGAAAAGGCACAGAGTTATGCACACCCTAATGCCTCCAACAATAACTGTTGACTTT
 TTATTTTT **CAGTCAGAGAAGCCTGGCA** ACCAAGAAGTGTTTTTTTGGTGGTTTACGA
CAGTCAGAGAAGCCTGGCAA

GAACCTAACTGAATTGGAAAATATTTGCTTTAATGAAACAATTTACTCTTGTGCAACACTAAATTG
 TGTCAATCAAGCAAATAAGGAAGAAAGTCTTATTTATAAAATTGCCTGCTCCTGATTTTTA **CTTCAT**
TTCTTCTCAGGCTCC AAGAAGGGGAAAAAA **ATG** AAGATTTTGATACTTGGTATTTTTCTGTTTTTA
 TGTAGTACCCCAGCCTGGGCGAAAAGAAAAGCATTATTACATTGGAATTATTGAAACGACTTGGGAT
 TATGCCTCTGACCATGGGGAAGAAACTTATTT **CTGTTGACACGTAAGTC** ACTATTTTTTATTGTT
 TATAGACCAAATTTTAAAGTTATTTTTTAAAGGATAG 3'

CP promoter primer sets

Variants identified in the current study

CP1**CP4**¹ -567C/G⁴ -364delT**CP2****CP5**² -563T/C⁵ -354T/C**CP3****CP6**³ -439C/T⁶ -350C/T⁷ -282A/G

CYBRD1 PROMOTER REGION

[ENSG0000007196 (ENSEMBL)]

5' TTTGGGTGGGCGGGCACTTTGAAAATCCATTTGT **CACACTAAACGGCAAGTCC** AGGTCCAGGAG
 GTTCCTGTCTTCCCTCTCTCAAGAGCAAAGTCAAGTAGTTTCATTGCAGGATAAGGCCAAGCCCA
 CCGCTGCCCATGCTGTTTTTGTGTTTCCAGACAGAGTCTAGCTGTCACCAGGCTG **CAGTGCAG**
TGGTGCAT¹ **CATAC**² CTCTCCTGAG³ CTCAAGTGATTCTCTTGCCTC⁴ CACCTCCCAAAGCCCTGGG
 ATTACAGCT **GTGAGACACCTCATGGGGAC** CCGGTTTACTGGGTTTTTTTTGTGTTTTTTGTGTTTTTTGT
 TTTTTTTT⁵ GAGACAGAGTCTCACTTTGTCAACCAAGCTGAAGTGCAGTGGTGAACCTCAGCTCAC
 TGCAGCCTCGACCTTCTGG **GCTCAAGTGATCCTCCTTCC** TCAGTCCCCCAAGTAGCTGGGGCTACA
 GGTGCATGCATTTGTATTTT **CAGTAGAGACAGGGTTTCAC** CTTGTTGCCCAGGCTG⁶ TTCTCAAAGT
 CCTGGACTCAAGTGATCTGCCCGCCTTAGCCTCCCAAAGTGCTAAGATTACAG⁷ GTGTGAGGCA **CTG**
CGCC⁸ **GGCCTT**⁹ **ACTGT** TAACTTAAAACAAAAAATTATAAATTTGAAAAAGAGGAGACTTTATT
 TCTTATAAAAG **GTTATAGCCTGCAAGGAGGC** CATTCCATAGGCTGATAAACAT¹⁰ AGCCTCTGGCCT
 AAGACCAGAGACAGGCACCTTGGAAAGGCAGAGGGGTTGGGGTAGGAGCTTTATGCTGAACAGT¹¹ TTG
 GCCAAACATACATACGTAAC **AGGTTACAGGAGGACTATG** AATATTAATGGAGGTGGTCCTTACAC
 ATGCATATTGAACAAA **CATGCATGTAACATGTTCTC** TTTGGGGTGGAGACTTAACATTTAATTGTA
 TTAATTCAAATACATTTAAATGTATTACTTCAAACCTACACTTCAAAGGTCTTTTTCAGGACGTGA
 ATGCATAAAGTGCACAATCC **CTGTACACTGGCCAGAACC** GTCCATGGTCCGTCTTCTTATCATG
 AAAAAGTTTCTGAAATCAGCCCAGTGAAAG **CTGTAGTTCTGGCTGGTGC** ACAGGGGTTTCCAGCTGG
 TCAGCATCTGTGAACTGATTAAGTTGTAATTGTTTTAATATTGCTTATCTCAAGCCAGTGCTTGT
 TAGCCTCTAGAGGAAAAGAAAACC¹² TTTGTGGCAGTTAGACCATAGTTTATTCTTAAGT **GTAGGA**
GTGTGTGACTTAGG TCCTTTTTATAATTTGATGTCTTATTGCTACAA **AGAGTCTGTTCTGTCCGCC**
TTATGATTTCTATTTTAAACATTAATGCTAGT CAGCTGTTGAGTCTAAATTTCCAAAATGGAGGGGG
 TAGACTTCCCTTCCGGCTGTAGCTAGAAACT¹³ CAGCTTAAAGGTTTTTCTGGG¹⁴ GTCTGCTTGGCC
 AAGGA **GGTCCATTTCAGTCAGTGAGG** GGCTTAGGATTTTATTTTTAGTTTACACCACCTTACAGGTT
 CCTCCTTAACACCCTTTGATGTCA **GAGACAGCCAATTCTCCAGC** ACTGTGGCTAAGGAGGCTTCGT
 TAGTAGTTAAAAGCACACACATAACC **CTACCCAACGGATCCCTCTC** TTCTCTTAACCAAGAGGGCG
 CACCCTGCCCCTACTCAACCTCCCCACAAATA **AGGACGAGACACGGGAAGTG** TAAAAGCAAGTGG
 TGAGTATAAAATCCTCGGAATTTTCTCTTTGAGCAACTAAGTCTGTTACTG **AAGCCCTCTGC**
GAGCTTGC ATGCTGGACGAGGGAGAGGGTGAGGCCACCAGGCA¹⁵ TGAGCGCCCTCGGCCGGCCGGC
 TGATCCCAGGGGGTGGGGCC **CATTTCTGAGTTGGGGCCAG** CTCCCACCC¹⁶ CAAG¹⁷ AGGCCCCACA
 TTCCGGGCCAGCAGCCAGAAAGTCCCTCCCCGCAGGCGGAGACAGCCCAAGAAGTGCAGCCCC
 GGTCCCAGCCGCCCCGGCCACTACCCAGAGGGCTGCCGCCGCTCTCCAAGTTCTTGTGGCCCCCGCG
 GTGCGGAGTATGGGGCGCT **ATG** GCCAGGAGGGCTACTGGCGCTTCTGGCGCTGCTGGGGTCCGGC
 ACTGCTCGTCGGC **TTCTCTGTCGGTGTCTTCGC** CCTCGT 3'

CYBRD1 promoter primer sets**Variants identified in the current study**

BP1	BP6	1-1849T/G	6-1540G/A	10-1346T/C	15-238G/A
BP2	BP7	2-1844G/C	7-1477C/T	11-1272T/C	16-167C/G
BP3	BP8	3-1834G/A	8-1459T/C	12-849C/G	17-163G/A
BP4	BP9	4-1813C/T	9-1452T/C	13-645T/C	
BP5	BP10	5-1752G(T),G(T),G(T),G(T),		14-624G/A	

HAMP PROMOTER REGION

[ENSG00000105697 (ENSEMBL)]

5' AGTGCCTTTTCTGTAAAGTGAAGGAAATGAGTGTCC **GACGGGGAGGAGGTTCCCTAA** AAGGGAGC
 AGGGTCTGGGGAGCCCAGGCCTCTGGGGTTGGGTGACTGAGAAGGCAGCCCCCTGAATACAGAGCAG
 AGCTGAAGGTGGGGCAGTAAGTGTCTGCTGGGAGAACAGGCAGCACAGGCTGAGTTGGTGCAGAAGT
 GAGTCAACATATGTGCCATCGTATAAAAATGTAC¹T **CATCGGACTGTAGATGTTAGC** TATTACTATTA
 CTGCTATTTTATGTTTTATAGACAGGGT **CTCACTCTGTCAACCAGGCT** GGAGTGCAGTCACACAAT
 CATAGCTCACTGCAACCTCAGCCTCCTGGGCTTAAGCGATCTGCCTCAGCCTCCCAAGTAGCTGGG
 ACTACAGATGTGTGCCAC **CACGCCTGGCTAAATTTGTT** TAAAATTTTTTTTTGTAGAGATGGGGTCT
 CCCTAT **GTTGCCCAGGCTAGTCTTGA** ACTTCTGGGCTCAAGCGACCCTCCTGCCTTGGCCTCCCAA
 ATTGCTGGGATTACAGGCATAAGCCACTGTGCTGGGCCATATTACTGCTGTCAATTTATGGCCAAAA
 GTTTGCTCAAACATTTTCCAGTTACCAGAGCCACATC **TCAAGGGTCTGACACTGGG** AAAACACCA²C
 GTGCGGATCGGGCACACGCTGATGCTTGCCCTGCTCAGGGCTATCTAGTGTTCCCTGCCA **GAACCT**
ATGCACGTGTGG TGAGAGCTTAAAGCAATGGATGCTTCCCCAACATGCCAGACACTCCTGAGGAG
 CCTGGCGGCTGCTGGCCATG³CCCCGTGTGCATGTAGGCGATGGGG **AAGTGAGTGGAGGAGAGCG** GA
 ACCTTGATTCTGCTCATCAAACCTGCTTAACCGCTGAAGCAAAGGGGGAACCTTTTTTCCCAGATCAG
CAGAATGACATCGTGATGG GGAAAGGGCTCCCCAGATGGCTGGTGAGCAGTGTGTGTCTGTGACCC
 CGTCTGCCCCACCCCTGAACACACCTCTGCCGG **CTGAGGGTGACACAACCCT** GTTCCCTGTGCT
 CTGTTCCCGCTTATCTCTCCCGCCTTTTCGGCGCCACCACCTTCTT **GGAAATGAGACAGAGCAAAG**
 GGGAGGGGGCTCAGACCACCGCCTCCCCTGGCAGGCCCCATAAAAGCGACTGTCACTCGGTCCCAG
 ACACCAGAGCAAGCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACG**ATG**GCACACTGAGCTC
 CCAGATCTGGGCCGCTTGCCTCCTGCTCCTCCTCCTCGC **CAGCCTGACCAGTGGCTCT** GTTTT
 CCCACAACAGGTGAGAGCCCAGTG 3'

HAMP promoter primer sets

Variants identified in the current study

PP1

PP4

¹ -1010C/T

PP2

PP5

² -582A/G

PP3

PP6

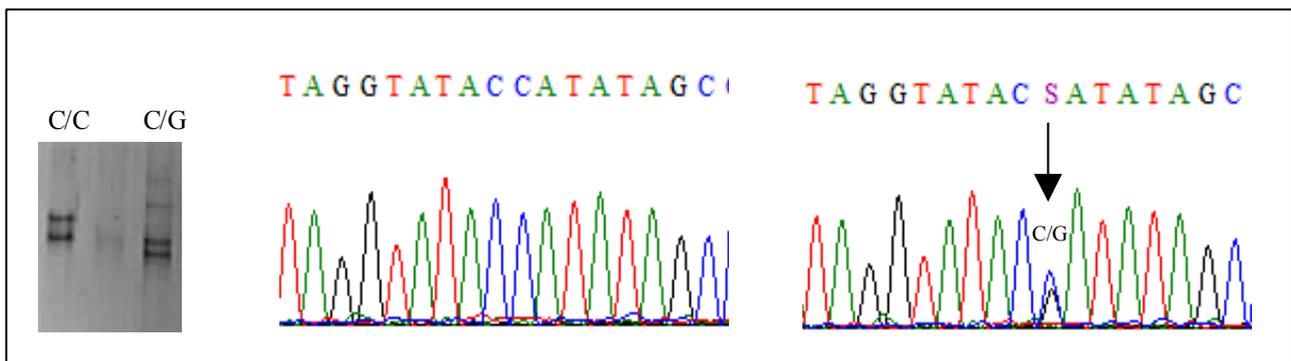
³ -429G/T

APPENDIX 3: HEX-SCCP GEL IMAGES AND SEQUENCING CHROMATOGRAMS OF VARIANTS IDENTIFIED IN THIS STUDY

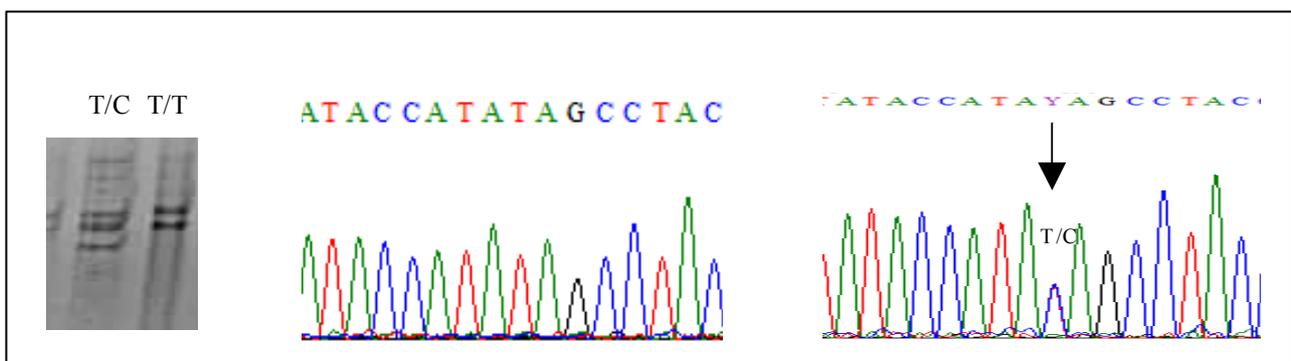
Variants were conventionally detected by HEX-SCCP analysis and confirmed by bi-directional semi-automated DNA sequencing. 1-7: *CP* gene variants; 8-24: *CYBRD1* gene variants; 25-27 : *HAMP* gene variants, 28-37 : *SLC40A1* gene variants. Images illustrate variants detected by the following means: A) abberant ^aSSCP banding pattern, B) abberant ^bheteroduplex formation or C) ^cRFLP analysis and D) the polymorphic allele on the sequencing chromatogram. Horizontal arrows indicate fragment sizes by RFLP analysis; vertical arrows depict specific basepair change identified on sequencing chromatograms. Variants 12, 13, 15,16, 17,18, 24, 28, 29 and 33 represent novel variants described in the current study. Variants 21, 22, and 23 were not included in statistical analysis.

CP variants

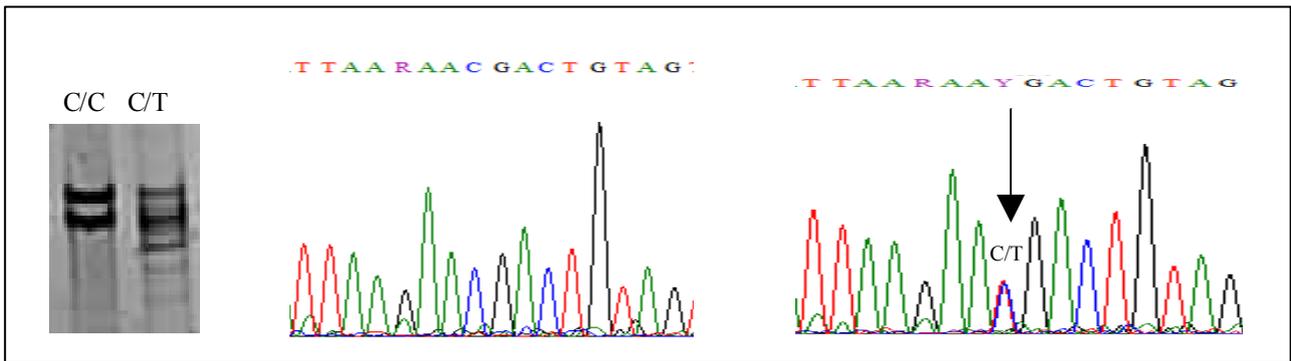
1) -567C/G^a



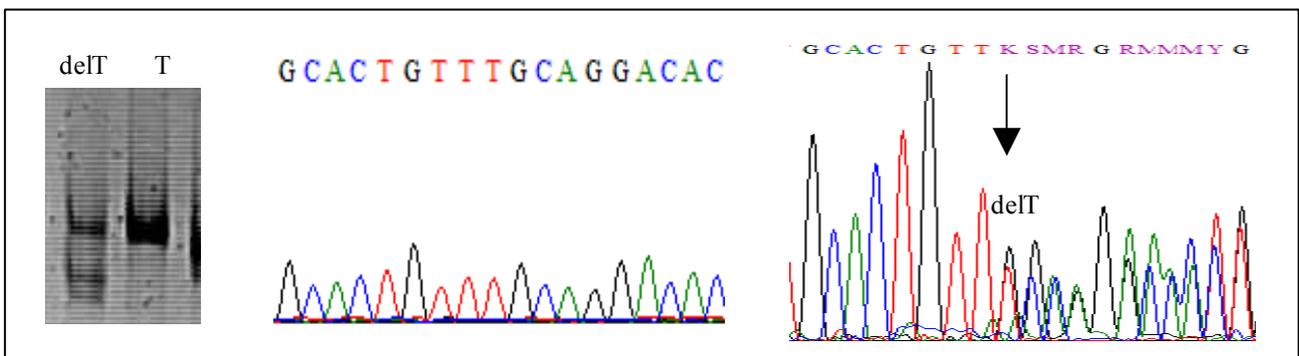
2) -563T/C^a



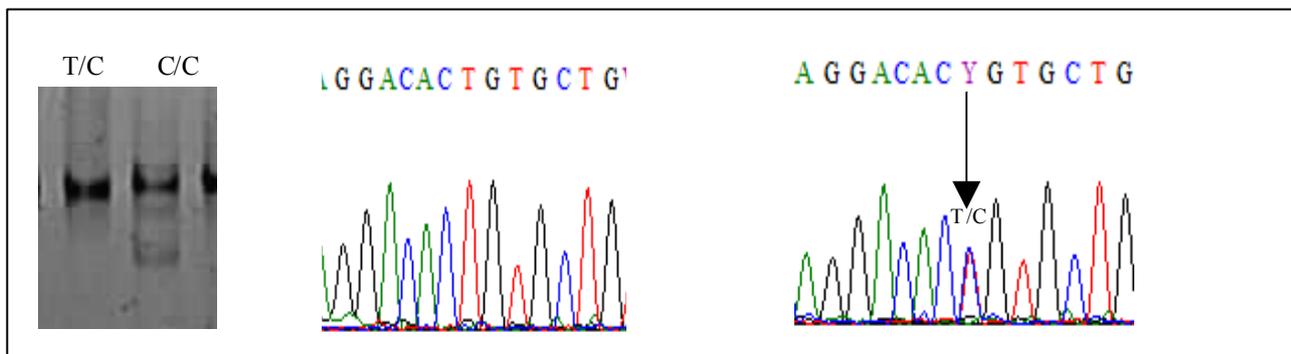
3) -439C/T^a



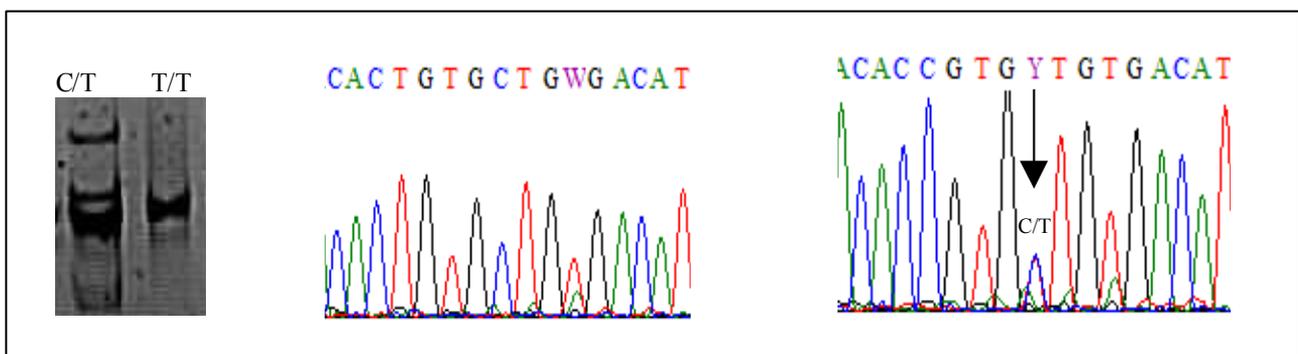
4) -364delT^b

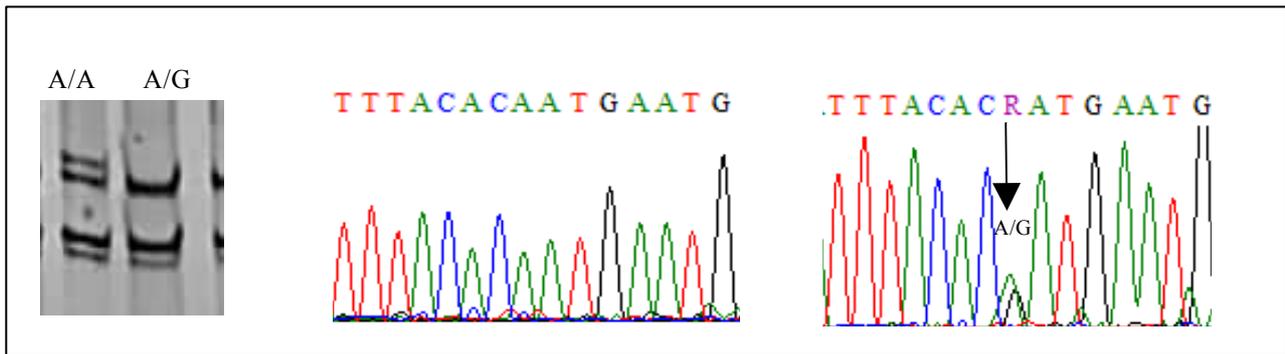
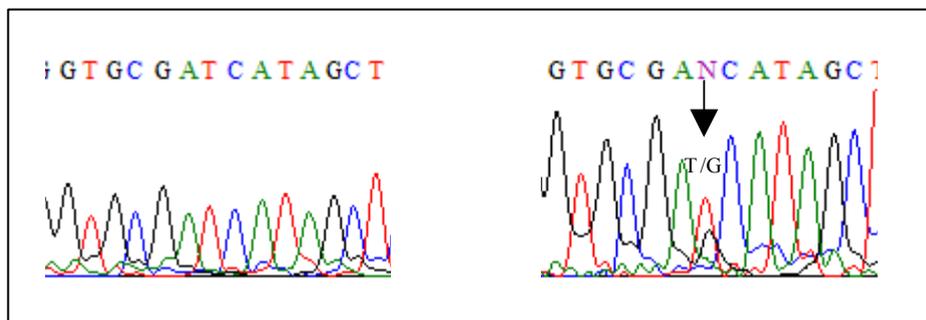
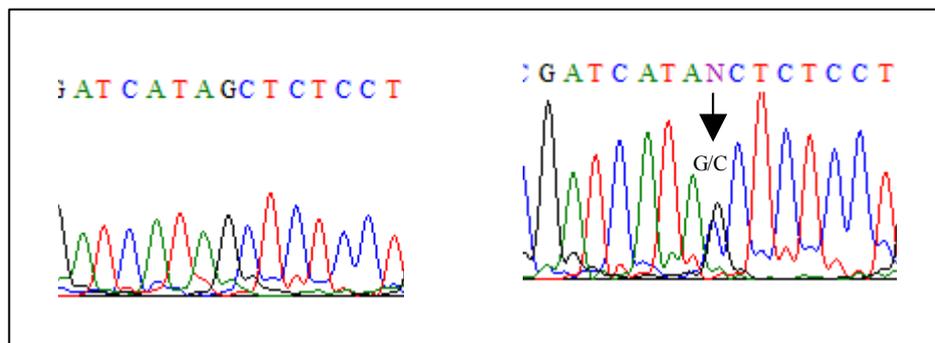


5) -354T/C^b

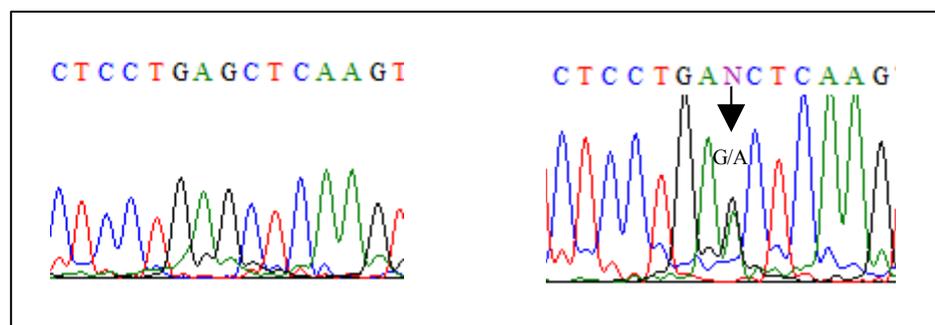


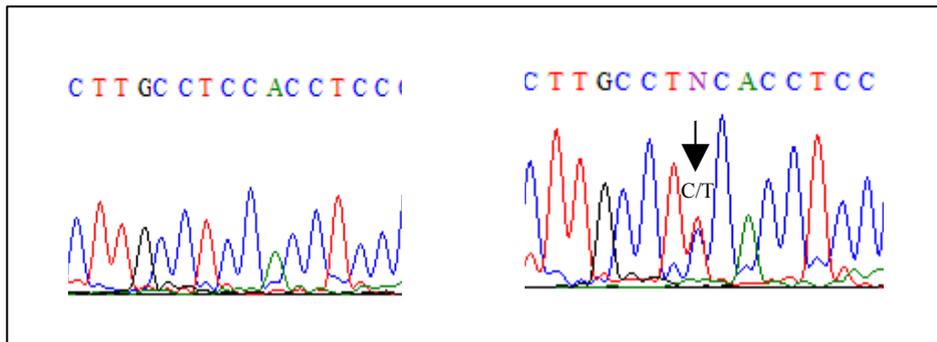
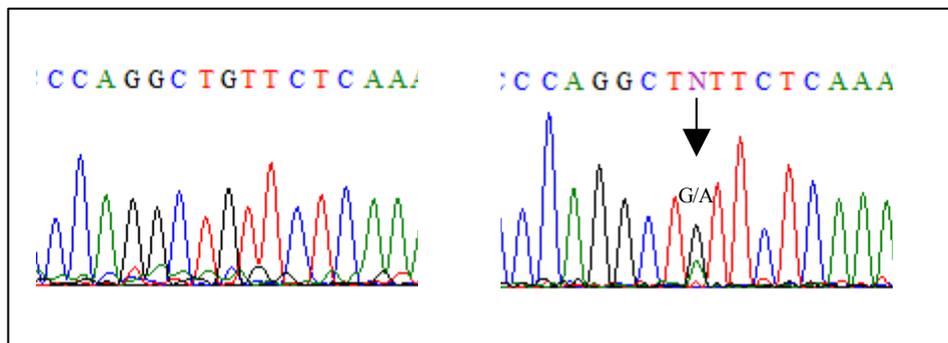
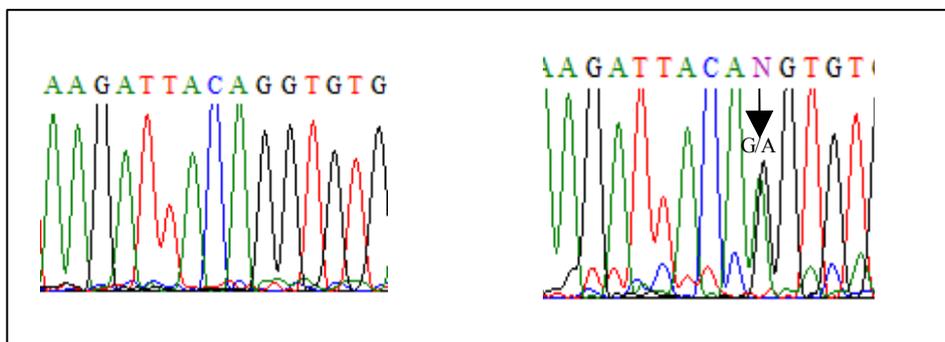
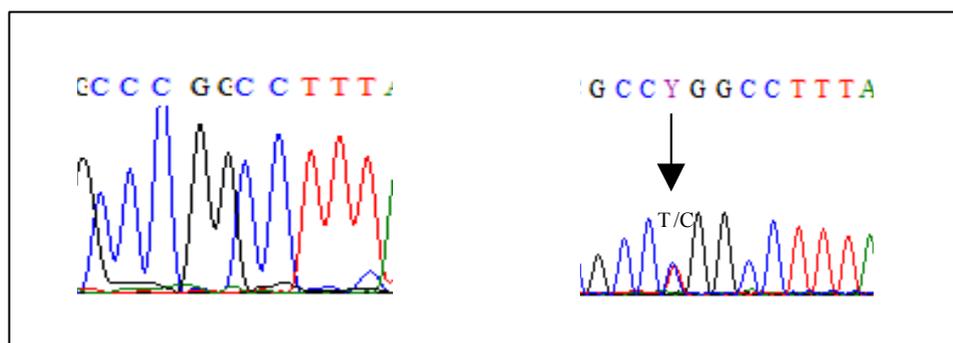
6) -350C/T^a (and -354C/C)



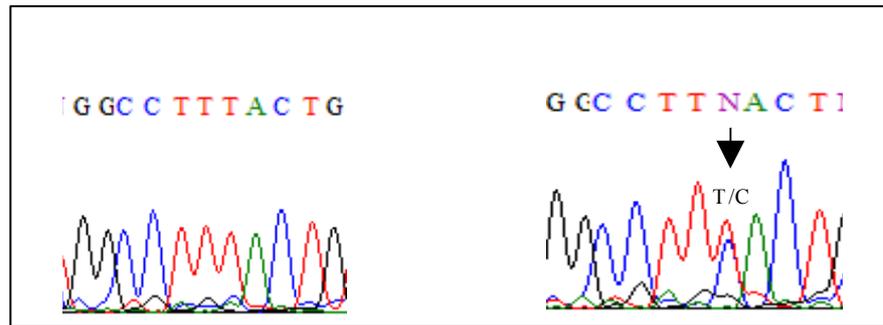
7)-282A/G^a*CYBRD1* variants8) -1849T/G^a9) -1844G/C^{a,e}

10) -1834G/A

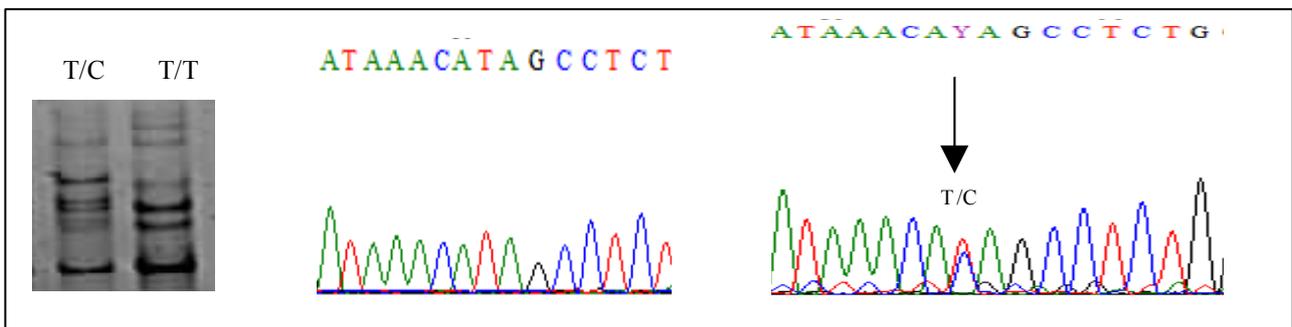


11) -1813C/T^d12) -1540G/A^d13) -1477G/A^d14) -1459T/C^d

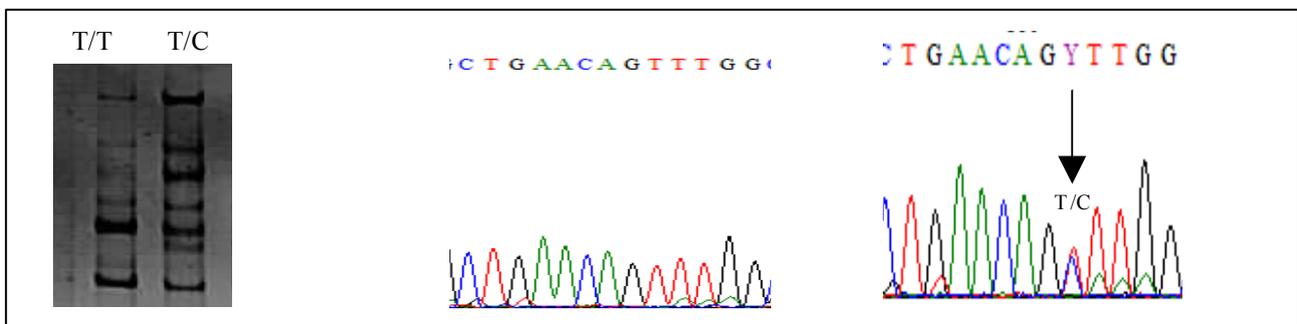
15) -1452T/C^d



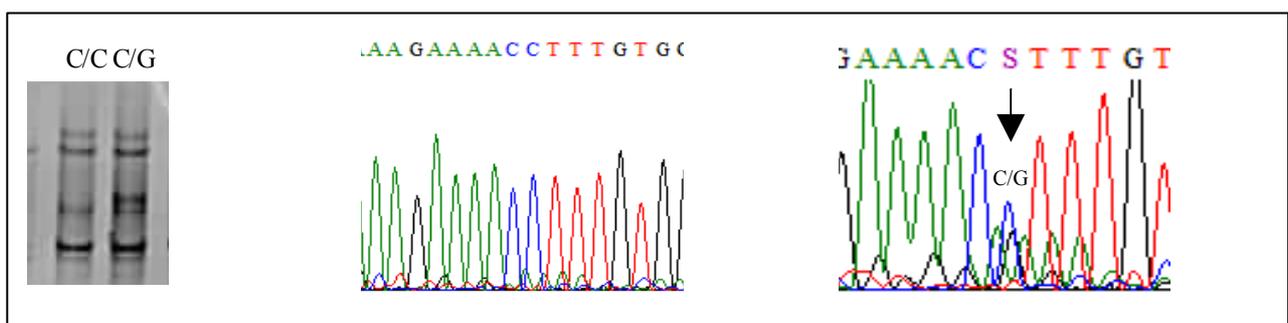
16) -1346T/C^a



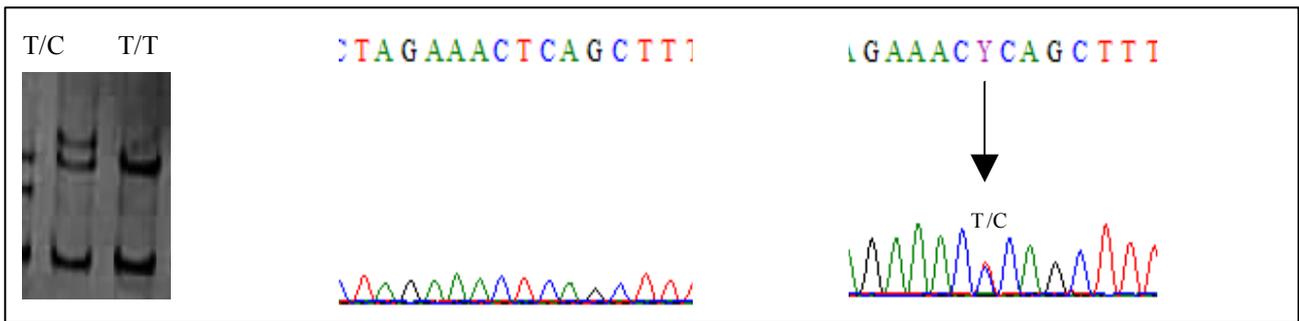
17) -1272T/C^a



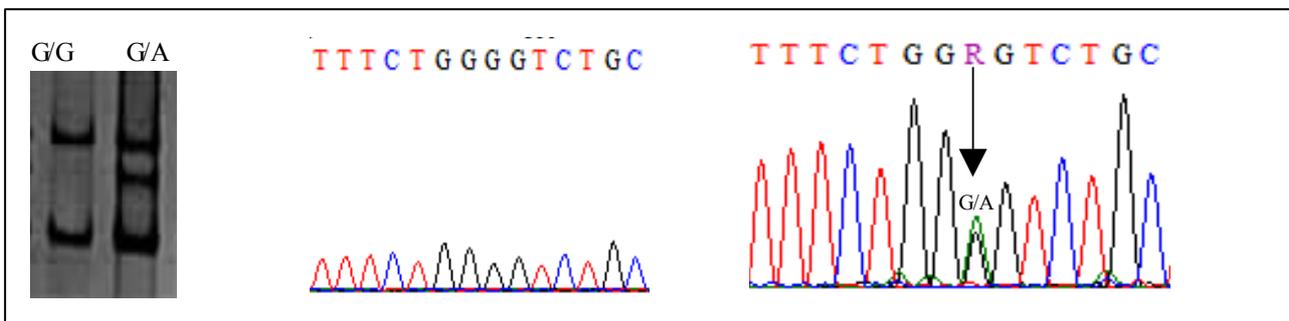
18) -849C/G^a



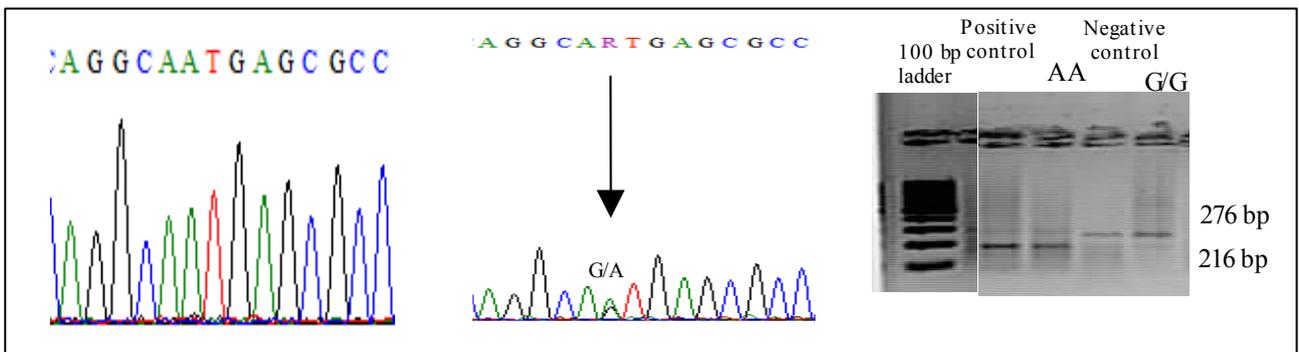
19) -645T/C^a



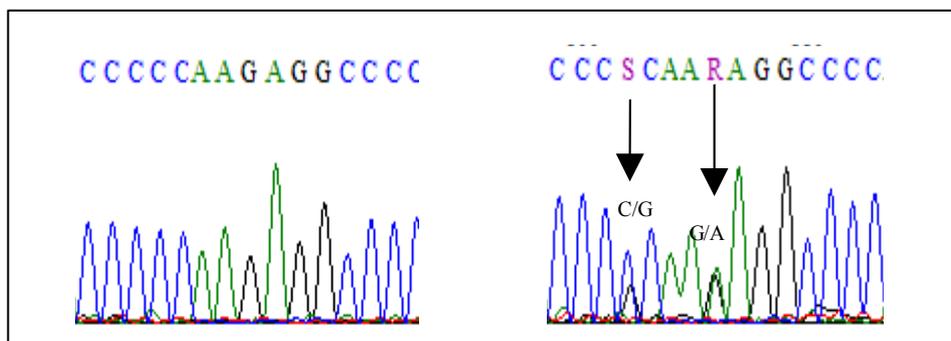
20) -624G/A^a



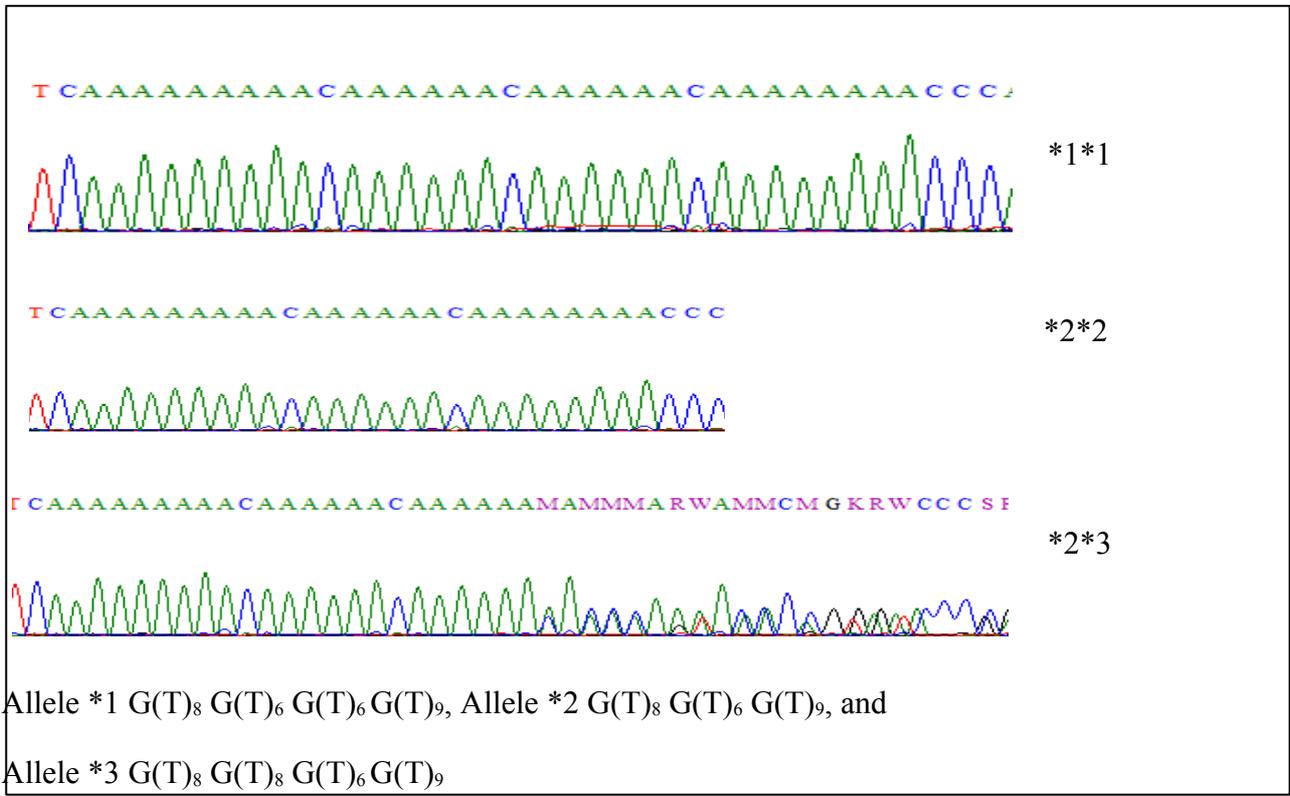
21) -238G/A^c



22) -167C/G^d and 23) -163G/A^d

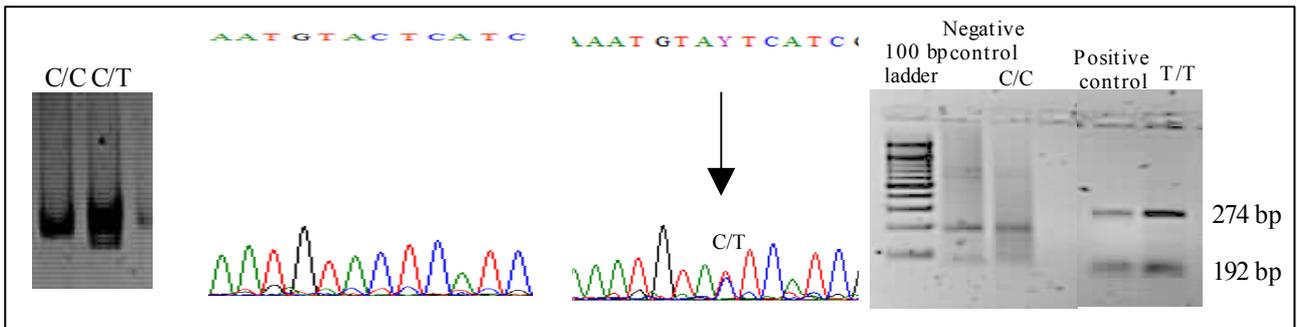


24) [G(T)₈G(T)_nG(T)_nG(T)₉] repeat ^e

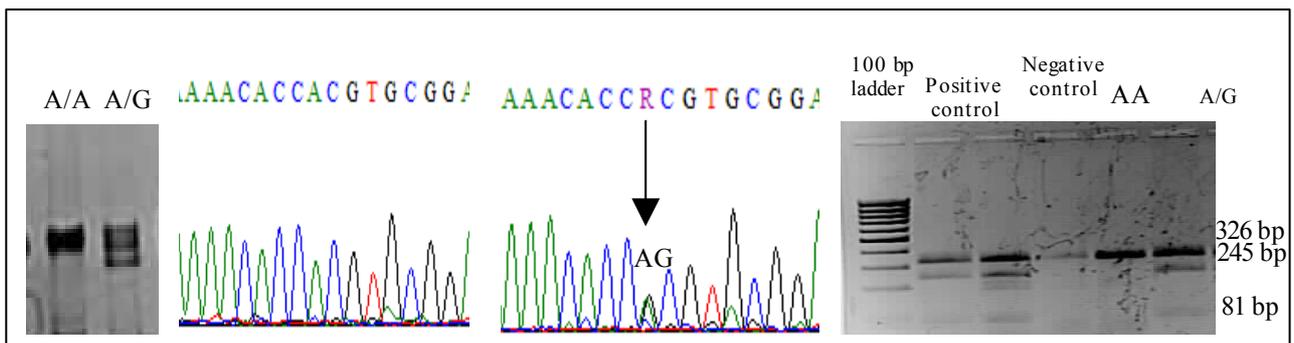


HAMP variants

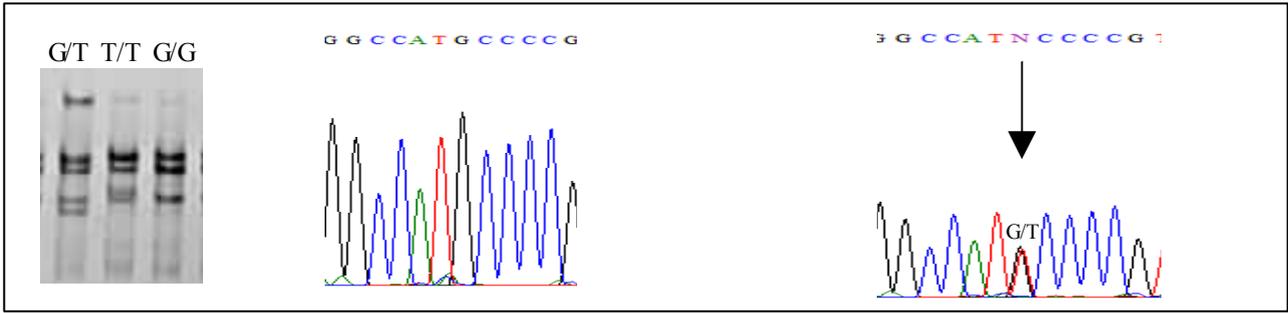
25) -1010C/T^{b,c}



26) -582A/G^a

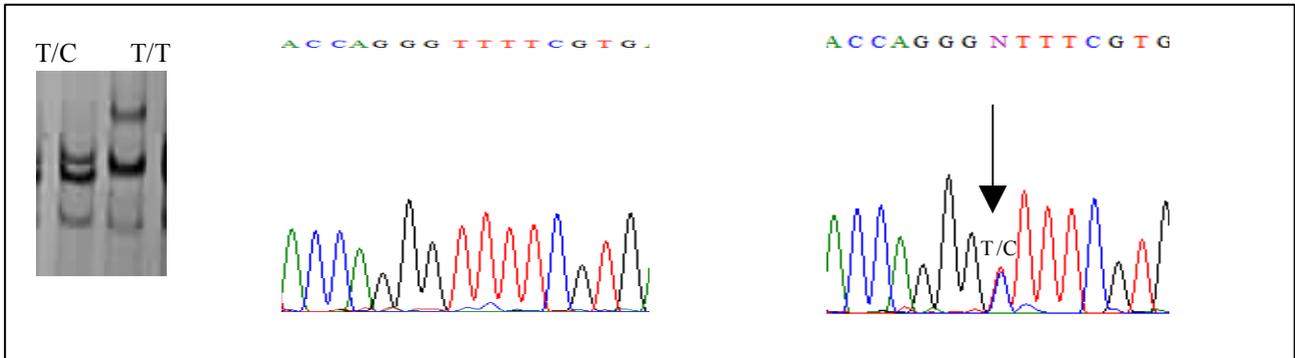


27) -429G/T^{a,c}

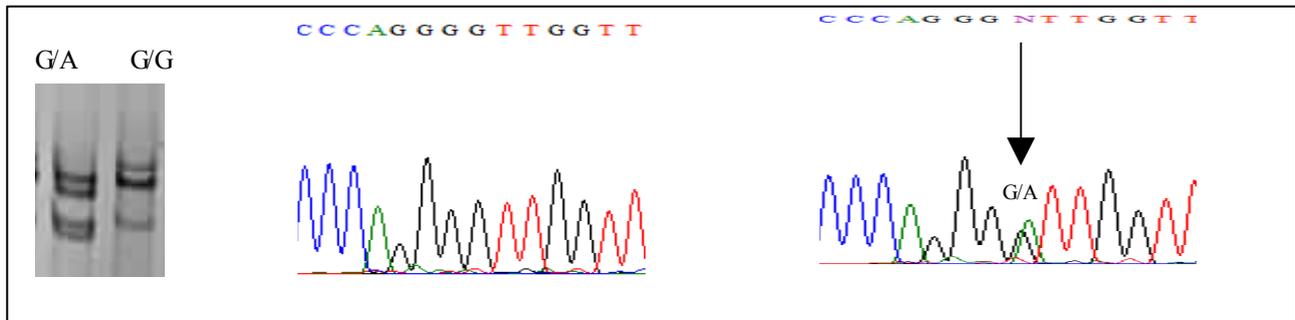


SLC40A1 variants

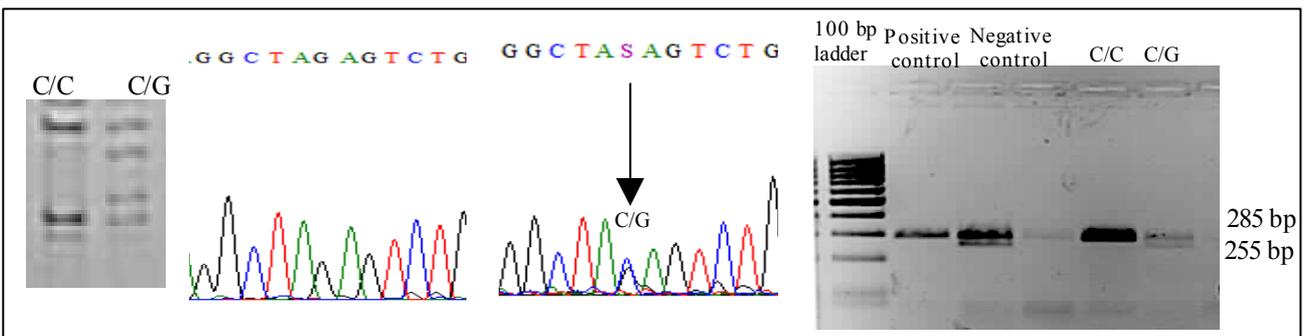
28) -1461T/C^a

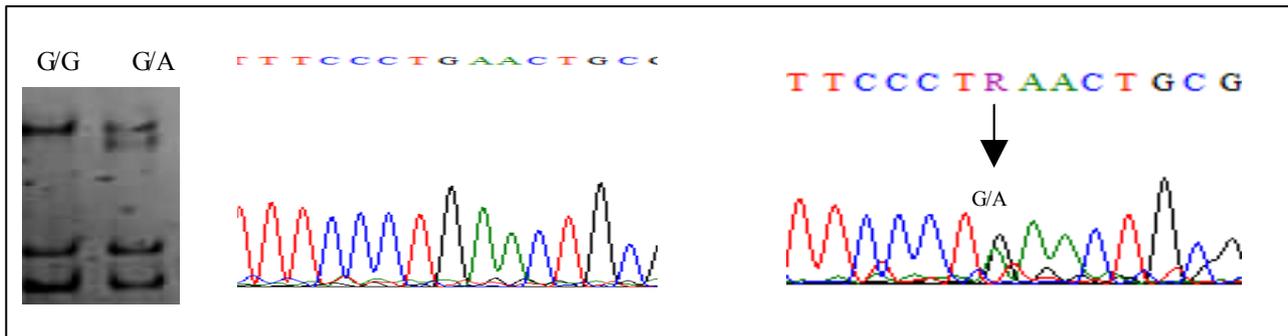
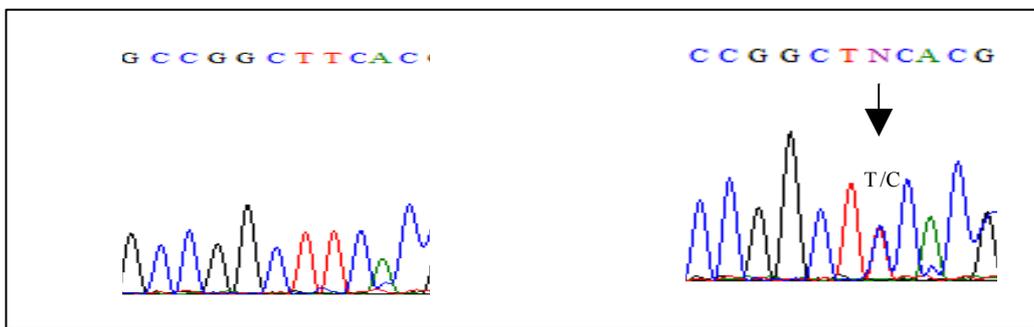
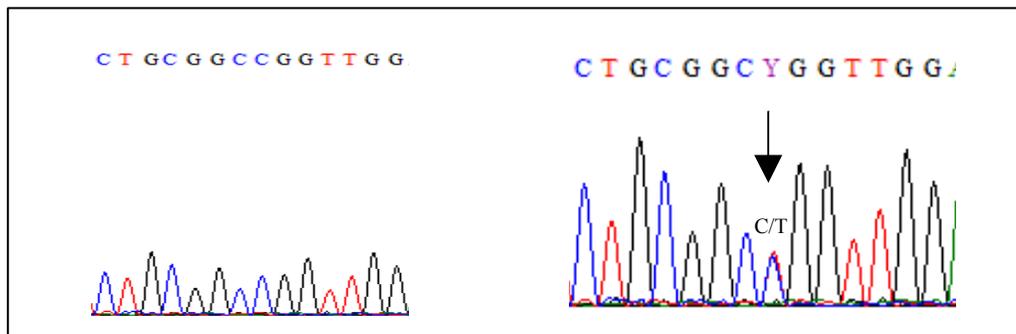
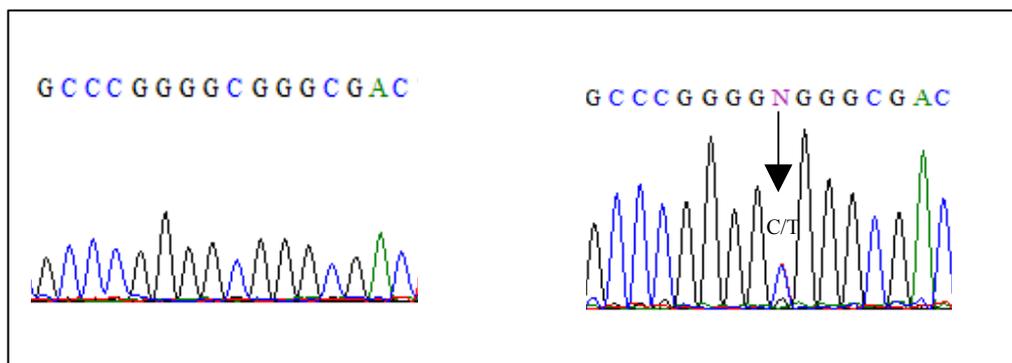


29) -1399G/A^a (and -1355G/C)

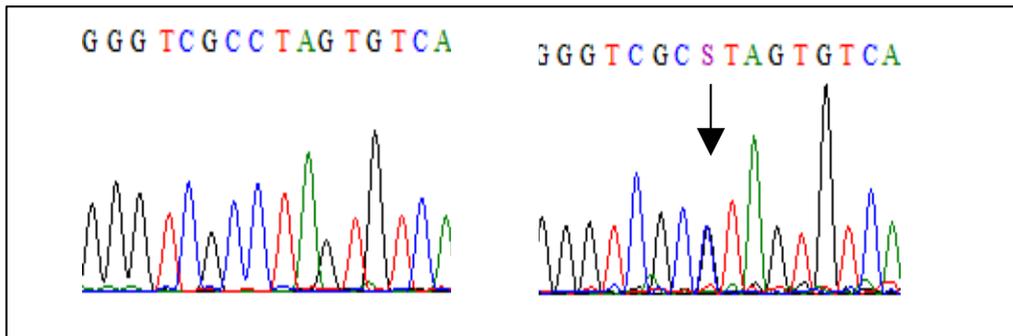


30) -1355C/G^{a,c}

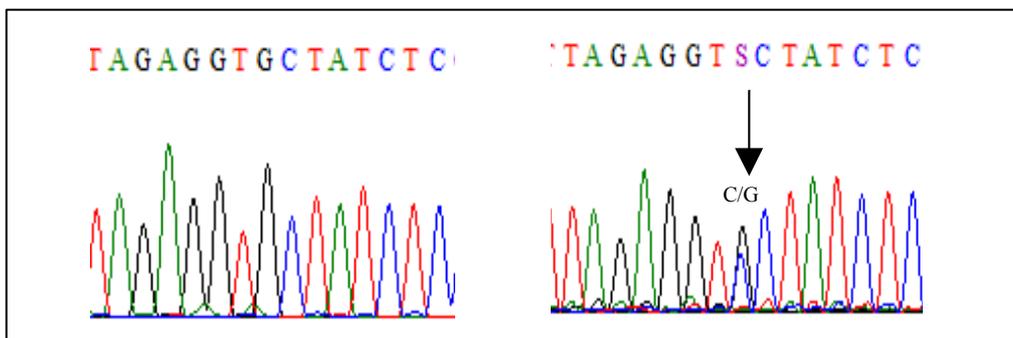


31) -750 G/A^a32) -622T/C^d33) -524C/T^d34) -501C/T^d

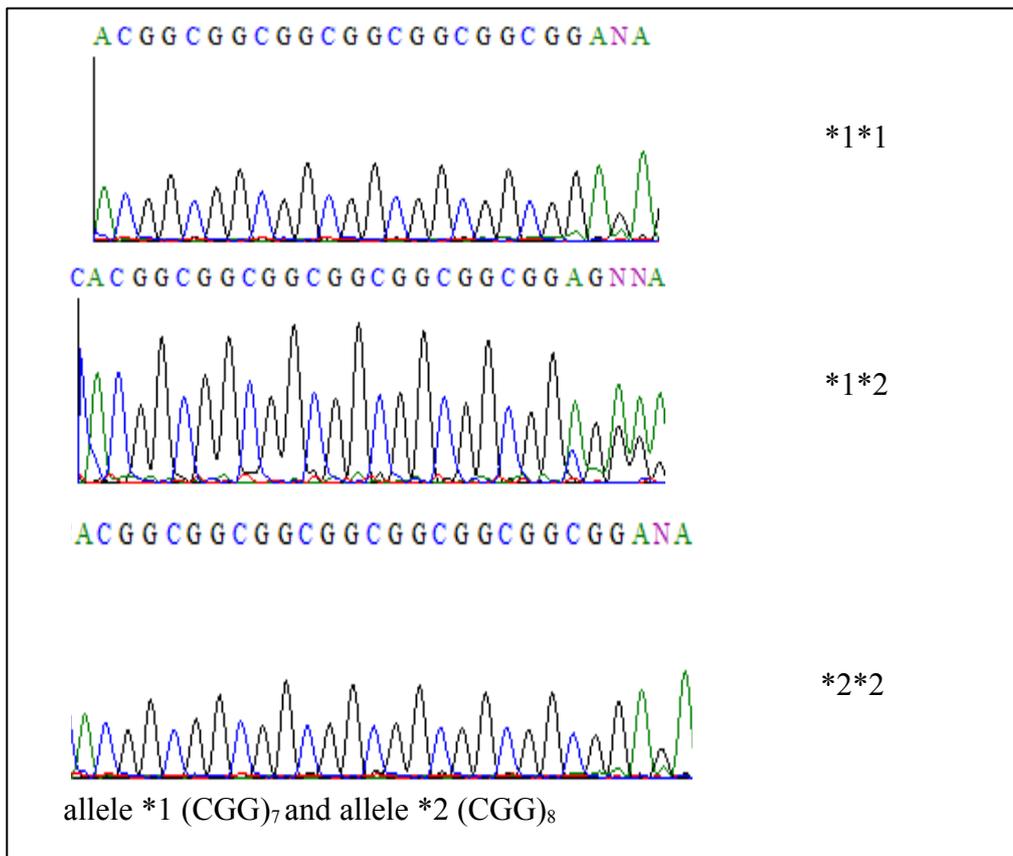
35) -98G/C^d



36) -8C/G^d



37) [(CGG)_n] repeat



APPENDIX 4: ABSTRACT OF ORAL PRESENTATION AT 11th SOUTH AFRICAN SOCIETY OF HUMAN GENETICS (SASHG) CONGRESS, GOLDEN GATE, SOUTH AFRICA, 1-3 MARCH 2007

PORPHYRIA CUTANEA TARDA (PCT) AND IRON OVERLOAD: AN ASSOCIATION STUDY

NA Panton¹, RJ Hift², L Warnich¹, MG Zaahl¹

NA Panton, oral presenter

¹Department of Genetics, University of Stellenbosch, South Africa

²Department of Medicine, University of Cape Town, South Africa

Porphyria cutanea tarda (PCT) is classified as a disorder of porphyrin metabolism affecting the biosynthesis of protohaem. This disorder is characterized clinically by light-sensitive dermatitis, appearing especially on the hands and face and biochemically by an increased excretion of urinary porphyrins, particularly uroporphyrin 1. The onset of PCT may be associated with alcohol use, oestrogen and chemical exposure, Human immunodeficiency virus (HIV) and Hepatitis C virus (HCV) infection. Iron overload is often present in PCT patients. Approximately 50% of Caucasian PCT patients carry the common C282Y mutation in the *HFE* gene. In the remainder, and in non-Caucasian patients, the cause for the iron overload is unknown. The promoter region of seven genes involved in iron metabolism, including, Haemojuvelin (*HJV*), Solute Carrier family 40 member 1 (*SLC40A1*), Cytochrome B reductase 1 (*CYBRD1*), Hepcidin antimicrobial peptide (*HAMP*), Ceruloplasmin (*CP*), High iron (*HFE*) and Haem oxygenase 1 (*HMOX1*), was investigated to determine their potential role in the development of PCT. DNA extracted from whole blood followed by Polymerase Chain Reaction (PCR) amplification of the various promoter regions was performed. Mutational analysis was performed by Heteroduplex single-stranded conformational polymorphism (HEX-SSCP) analysis. Semi-automated DNA sequencing was performed on the HEX-SSCP products demonstrating mobility shifts upon HEX-SSCP analysis. A large number of variations were identified by HEX SSCP analysis. The variants identified in *CP* were all novel and include -364 DelT (15 out of 80 patients-19%), -439 T/C (48 out of 80 patients- 60%), -354 T/C (6 out of 80 patients- 7.5%) and -282 A/G (2 out of 80 patients- 2.5%). All other variants identified still need to be confirmed by semi-automated DNA sequencing. These novel variants and any other variants found in these seven genes could potentially elucidate the association between iron overload and PCT. Recognition of other potential iron-loading influences in PCT patients would present an important development in the understanding of how the iron-loading associated with PCT, comes about.