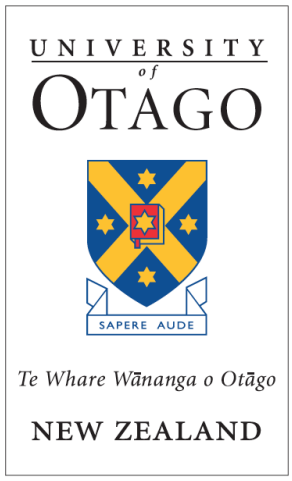
Gout and Metabolic Disease

Investigation of Potential Relationship in the New Zealand Population

Tahzeeb Fatima

Supervisor: Professor Tony R Merriman



A thesis submitted for the degree of
Doctor of Philosophy (PhD) in Biochemistry

University of Otago, Dunedin, New Zealand

March 2017

ABSTRACT

Hyperuricaemia, pathologically defined as the presence of higher levels of serum urate, results from a compromise in the delicate balance between the production and excretion of urate primarily in the liver and the kidneys, respectively. Hyperuricaemia is a prerequisite for gout, a painful inflammatory arthritis. The symptoms of gout arise from the body's immune response to monosodium urate crystals that accumulate in the synovial fluid of the joints. Hyperuricaemia and gout are complex traits. A number of genetic loci confer risk to develop hyperuricaemia. Genome-wide association studies (GWAS), an indispensable tool in population genetics, has identified at least twenty eight genomic loci that contain variants affecting serum urate concentration. Gene-environment interactions also play a significant role in this context. Exogenous factors such as the intake of purine-rich foods increase the frequency of gout flares. Population-specific genetic effects on gout are as evident, if not more, as for other complex phenotypes.

The prevalence of gout is much higher in the New Zealand Polynesian population compared to other populations. Approximately 7% of New Zealand Māori and Pacific Island people and 3% of New Zealand Europeans are affected by gout. The coexistence of metabolic conditions with gout, usually called gout-comorbidities, adds another level of complexity. However, not many studies have attempted to address the causal relationship between these traits. In fact, my research project was instigated as an attempt to study the causal associations between gout and its comorbidities and fill in some gap in the scientific literature. The research was, however, limited to three metabolic conditions/comorbidities of gout – imbalanced iron homeostasis, metabolic syndrome and disrupted lipid metabolism.

My study shows an association between increased serum ferritin and the risk of gout and seeded an idea that the consumption of iron-rich diet may play a role in increasing the frequency and severity of gout flares. Genetic association analysis using two variants in the *HFE* gene was done to confirm the association between ferritin and urate, which showed positive association in a smaller dataset and provoked the idea to investigate the causality, if it exists, between gout and iron metabolism. Using the robust 'Two-sample Mendelian randomisation' approach and exploiting summary statistics data from two large GWA studies, I was able to find an evidence of a causal effect of iron on urate metabolism, but not urate on iron metabolism.

In the context of the metabolic syndrome, the role of variants within/near the *ADRB3*, *MC3R*, *MC4R* and *ADTRP* genes were investigated. The positive effects identified for these variants supported the possible involvement of obesity and insulin resistance-related genes in gout pathophysiology.

With the help of gene sequencing-based rare variant analyses, several novel population-specific association signals were found within the coding regions of two lipid-related genes, *LRP2* and *A1CF*. Polynesian-specific novel genetic effects were identified to be predictive for gout for common variants within the *LRP2* gene. Rare variants within the *LRP2* gene were also identified and a higher prevalence of non-synonymous polymorphisms that can increase the risk of hyperuricaemia was observed in European individuals compared to Polynesians. These results indicated *LRP2* to contribute to the difference in gout prevalence between Māori and Pacific Island individuals compared to the New Zealand European population.

Collectively, my study reports a causal role of iron and ferritin in increasing serum urate concentration and the involvement of imbalanced iron homeostasis in hyperuricaemia. Also, positive genetic associations indicated that genes contributing to metabolic syndrome and lipid metabolism can increase the risk of gout, and also have population-specific effects for the Polynesian and European ancestral groups in New Zealand

PICTORIAL ABSTRACT

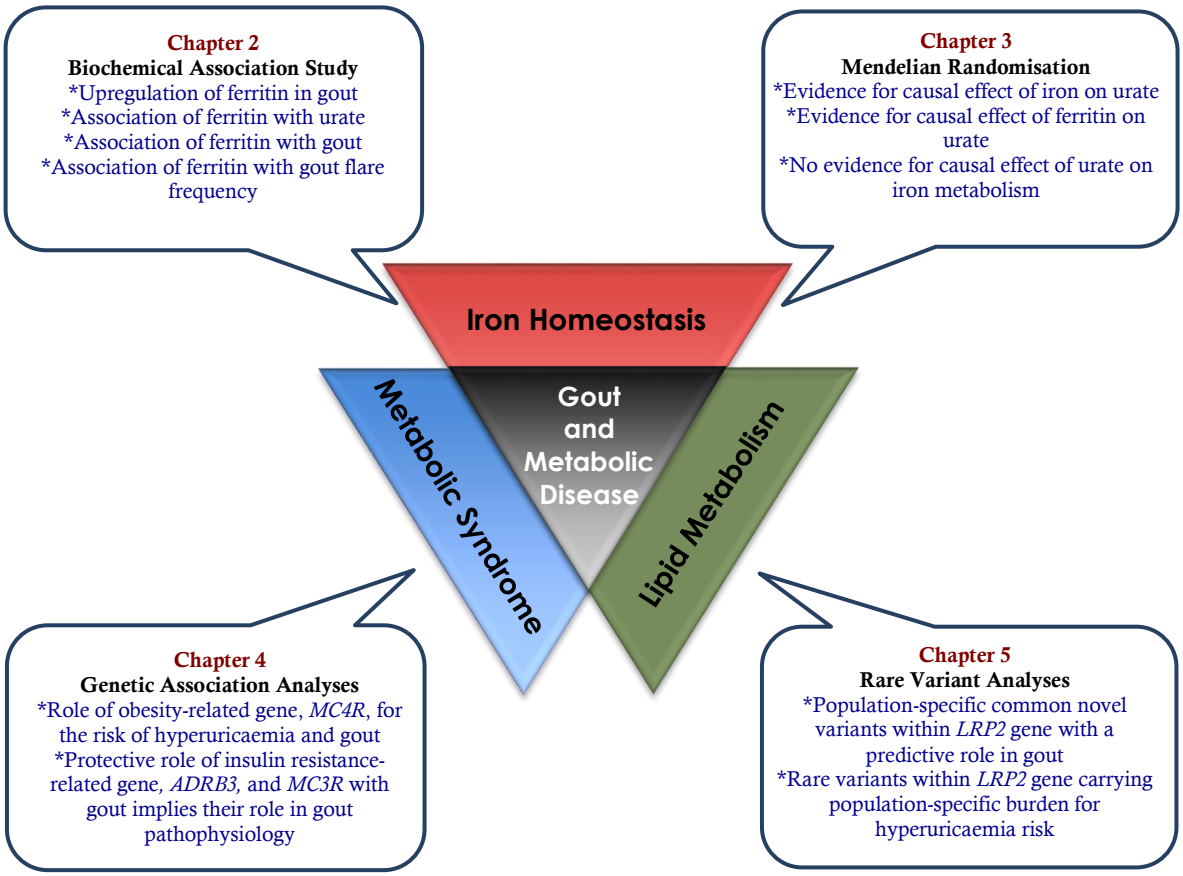


Figure 1: Pictorial abstract of discrete sections and major findings of this thesis

ACKNOWLEDGEMENTS

All praise to the Almighty ALLAH, the benevolent, the propitious and the sovereign whose blessings and glory flourished my thoughts and thrived my ambitions and who bestowed me with sound health and courage to complete my research.

It is my privilege and honor to express my heartiest gratitude to my dignified and respectable supervisor, Professor Tony Raymond Merriman, for his advice, inspiring guidance, continued patience, valuable comments, and constant encouragement. It would not have been possible for me to complete this task without his support. No words can express the depth of my heartiest feelings for the help that he graciously offered at all times. His efforts towards the inculcation of the spirit of hard work and maintenance of professional integrity besides other valuable suggestions will serve as a beacon of light through the cores of my life. Also, I would never have been able to complete my thesis without the guidance of my committee members Professor Warren Tate, Dr Torsten Kleffman and Dr Gregory T Jones, who continually and persuasively guided me throughout this project.

I have no adequate words to express my deepest gratitude towards Marilyn Merriman and my supportive lab fellows, Dr Tanya Major, Dr Anna Gosling, Amanda Phipps-Green, Ruth Topless, Murray Cadzow, James Boocock, Dr Hoang Tan Nguyen, Dr Humaira Rasheed, Mansour Zamaanpoor, Dr Sara Zeeshan, Edana Lord, Jaye Moors, Callum Tanner, Keresoma Leaupepe, Sepideh Sherif and Dr Cushla McKinney. They have been always caring and helpful throughout my stay here and guiding me for any help taht I needed, both within and outside the academic zone.

Tremendous love and special thanks to my lovely flatmates/friends/lab/country fellows Zeinab Mohammadi, Ranjeet Sohi, JiffinVir Singh Khosa, Amara Umair, Nuradhi Kalpani Jayasiri, Angel Aguirre and Alia Zulfahan for their continuous guidance, encouragement and remarkable cooperation. The precious moments of their cheerful company will remain in my mind for the rest of my life.

Special acknowledgments are due for my best friends, in none other than alphabetical order, Abida Mahmood, Hakimeh Azimi, Neelam Nazir, Padmini Parthasarathy and Sidra Liaqat to whom I will never be able to payback the way they always stood by me whenever life gave me harsh times. Special thanks to them for their consistent academic and moral support and to Padmini Parthasarathy for proof reading my thesis and for her 'nutritional' support through the hard times of my research/thesis write-up.

I would like to thank the administrative staff of the Biochemistry Department, Teena Joyce, Robyn Thompson, Julie Farquhar and Mel Kennedy for their support throughout my stay here.

Special thanks to the University of Otago for not only providing me a place to groom and polish myself but also a doctoral scholarship to financially support my studies. I am also thankful to the Department of Biochemistry, University of Otago, Division of Health Sciences, NZFGW and NZSBMB for providing travel grants for my trip to France and USA to attend the ECN and ACR conferences.

I have no sufficient words to acknowledge the appreciable efforts, love, care, and every possible kind of help provided to me by my very innocent and helpful parents and my dear siblings Ather, Ummara, Farah, Usman, and Laraib. Special thanks to my brave little star Laraib, who is none less than my first child.

Last but by no means the least my thanks to the real strength of my life, my best friend and my husband, Awais Rauf. Thank you, Awais for making everything possible.

Tahzeeb Fatima

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A	Alanine	G	Guanine (Purine)
A	Adenine (Purine)	GCK	Glucokinase
A1CF	Apolipoprotein B mRNA-editing enzyme 1 (APOBEC1) Complementation Factor	*GCKR*	Glucokinase Regulatory Protein
ABCG2	ATP Binding Cassette Subfamily G Member 2	GMP	Guanine Monophosphate
ACR	American College of Rheumatology	GRCh37	Genome Reference Consortium human genome build 37
ACTH	Adrenocorticotropic Hormone	GWAS	Genome-Wide Association Study
ADP	Adenosine Diphosphate	H	Histidine (His)
ADRB3	Beta Adrenergic Receptor	*HFE*	Haemochromatosis gene
ADTRP	Androgen-dependent Tissue Factor Pathway Inhibitor	HPFS	Health Professionals Follow-Up Study
AMP	Adenosine Monophosphate	HR	Hazard Ratio
AMPD	AMP-Deaminase	HTS	High-throughput Sequencing
AMPK	AMP-activated protein Kinase	HU	Hyperuricaemia
ApoA	A Apolipoprotein A	I	Isoleucine (Ile)
APOBEC1	Apolipoprotein B mRNA-editing enzyme 1	IHD	Ischemic Heart Disease
Arg	Arginine	IL1B	Interleukin 1 Beta
ARIC	Atherosclerosis Risk in Communities study	Ile	Isoleucinee
ASD	Autism Spectrum Disorders	IMP	Inosine Monophosphate
ATP	Adenosine Triphosphate	IR	Insulin Resistance
B	Asparagine (Asx)	JHS	Jackson Heart Study
BMI	Body Mass Index	K	Lysine (Lys)
C	Cysteine (Cys)	L	Leucine (Leu)
C	Cytosine (Pyramidine)	LD	Linkage Disequilibrium
CARDIA	Coronary Artery Risk Development in Young Adults study	*LRP2*	Low-density Lipoprotein Receptor-Related Protein 2
CETP	Cholesteryl Ester Transferase Protein	M	Methionine (Met)
CHD	Coronary Heart Disease	MCR	Melanocortin Receptor
CHS	Cardiovascular Heart study	MR	Mendelian Randomisation
CI	Confidence Interval	mRNA	messenger Ribonucleic Acid
CNS	Central Nervous System	MSH	Melanocyte-Stimulating Hormone
CVD	Cardiovascular Disease	MSU	Monosodium Urate
D	Aspartic Acid (Asp)	N	Asparagine (Asn)
DART	Direct Analysis in Real Time	NHANES III	Third National Health and Nutrition Examination Survey
dbGaP	The database of Genotypes and Phenotypes	NLRP3	NACHT, LRR and PYD domains-containing protein 3
DNA	Deoxyribonucleic Acid	NU	Normouricaemia
E	Glutamic Acid (Glu)	NZ	New Zealand
EDTA	Ethylene diamine tetra acetic Acid	OR	Odds Ratio
ELISA	Enzyme-Linked Immunosorbent Assay	P	Proline (Pro)
EP	Eastern Polynesian	PCR	Polymerase Chain Reaction
EPN	Eastern Polynesian with low Polynesian ancestry	*PCSK9*	Proprotein Convertase Subtilison Kexin type 9
EPWP	Mixed Eastern and Western Polynesian	POMC	Pro-opiomelanocortin
EPZ	Eastern Polynesian with high Polynesian ancestry	Q	Glutamine (Gln)
EUR	European	R	Arginine (Arg)
F	Phenylalanine (Phe)	ROS	Reactive Oxygen Species
FEUA	Fractional Excretion of Uric Acid	RR	Risk Ratio
FHS	Framingham Heart Study	RVAS	Rare Variant Association Studies
FPLC	Fast Protein Liquid Chromatography	S	Serine (Ser)
FTO	Fat Mass Obesity-associated gene	SCL	Southern Community Laboratories
G	Glycine (Gly)	SD	Standard Deviation

Continued

SE	Standard Error
SLC16A9	Solute Carrier family 16, member 9
SLC17A1	Solute Carrier family 17, member 1
SLC17A2	Solute Carrier family 17, member 2
SLC17A3	Solute Carrier family 17, member 3
SLC22A11	Solute Carrier family 22, member 11
SLC22A12	Solute Carrier family 22, member 12
SLC2A9	Solute Carrier family 2, member 9
SNP	Single Nucleotide Polymorphism
SU	Serum Urate
T	Threonine (Thr)
T2DM	Type 2 Diabetes Mellitus
TFF3	Trefoil Factor-3
TFPI	Tissue Factor Pathway Inhibitor
TG	Triglycerides
TLR	Toll-Like Receptor
Trp	Tryptophan
UK	United Kingdom
ULT	Urate-lowering Therapy
USA/US	United States of America
UTR	Untranslated Region
V	Valine (Val)
W	Tryptophan (Trp)
WP	Western Polynesian
X	No amino acid (Stop)
XO	Xanthine Oxidase
Y	Tyrosine (Tyr)
Z	Glutamic acid (Glx)
β	Beta value

Chapter 1

Introduction

SECTION 1.1 INTRODUCTION TO URATE, GOUT AND METABOLIC DISEASE

1.1.1 Background

The uric acid story began several centuries ago when the German-Swedish chemist, Karl Wilheim Scheele (in 1776), isolated the substance from a bladder stone. He called it 'lithic acid' by virtue of its acidic properties (reviewed in Richet (1995)). The name was later changed to 'uric acid' by George Pearson and Antoine Fourcroy, reflecting its presence in normal urine (Fourcroy, 1804). Since then, researchers across the globe have been intrigued to establish the physiological and pathophysiological roles of uric acid in the human body, especially following the scientific speculation of the crucial significance of uric acid homeostasis in the prevention and/or management of several disorders including gout. The story particularly relevant to disorders other than gout, however, remains incomplete and merits further exploration.

1.1.1.1 Basic chemical properties of uric acid

Uric acid ($C_5H_4N_4O_3$; 2,6,8-trihydroxypurine) is a weak diprotic acid with pK_{a1} ~5.4 and pK_{a2} ~10.3. Under physiological pH (~7.4), it is mostly present in the form of monovalent urate anion formed by the dissociation of a proton from the molecule (Figure 1.1). Therefore, the term urate is often used in literature to denote the dissociated monovalent anionic state of uric acid in the extracellular fluid. The terms uric acid and urate are interchangeably used to refer to the total uric acid pool in the body (Bobulescu and Moe, 2012).

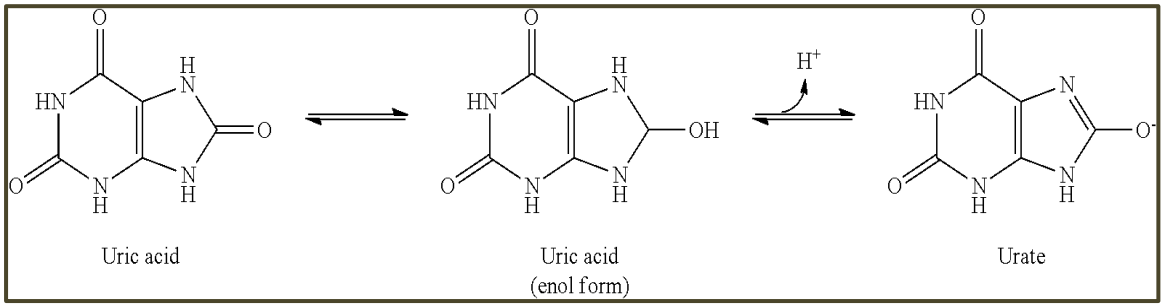


Figure 1.1: Formation of urate from uric acid via dissociation of one proton. Uric acid and urate are used interchangeably in the literature. Urate is the anionic form that circulates in blood at the normal physiological pH.

1.1.1.2 Urate handling - evolutionary insight

Although purine degradation is conserved across species, the end product of degradation is species dependent. In several organisms including most other mammals, urate is only an intermediary product of metabolism and is further subjected to oxidative degradation by the enzyme uricase (urate oxidase) in the peroxisomes of hepatocytes to form allantoin. Allantoin, being water soluble, is relatively easily excreted by the kidneys (Briggs *et al.*, 1977). However, birds, reptiles and some primates including humans, lack the uricase enzyme as a consequence of mutational silencing of the uricase gene. Evolutionary evidence suggests the occurrence of multiple missense and frameshift mutations in the promoter and coding regions of the gene during hominoid evolution, gradually decreasing the expression of the uricase gene and rendering the protein non-functional (Kratzer *et al.*, 2014; Masako *et al.*, 2013; Oda *et al.*, 2002). As a result, urate remains the end product of purine metabolism in these species and it is excreted as such by the kidneys (Rafey *et al.*, 2003). Figure 1.2 illustrates the uricase mutations during the early hominoid evolution.

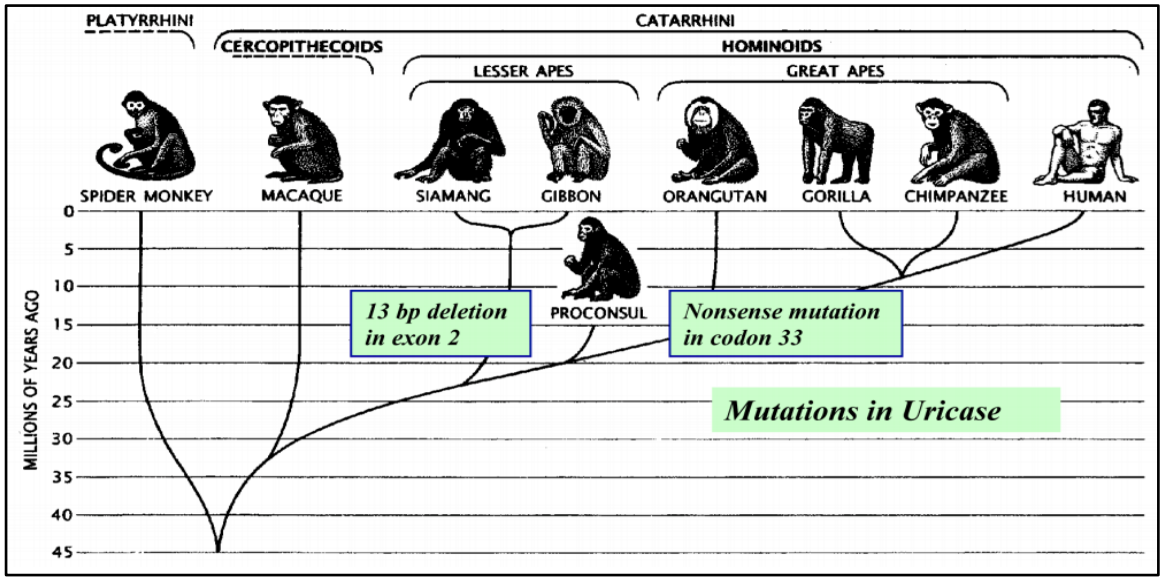


Figure 1.2: Uricase mutation in early hominoid evolution. Figure illustrates the process of loss of uricase enzyme through space and time. Source: (Johnson et al., 2008).

The mechanism for the clearance of urate from the body largely depends on the presence or absence of functional catabolic enzymes in the pathway and varies from species to species, as outlined in Figure 1.3.

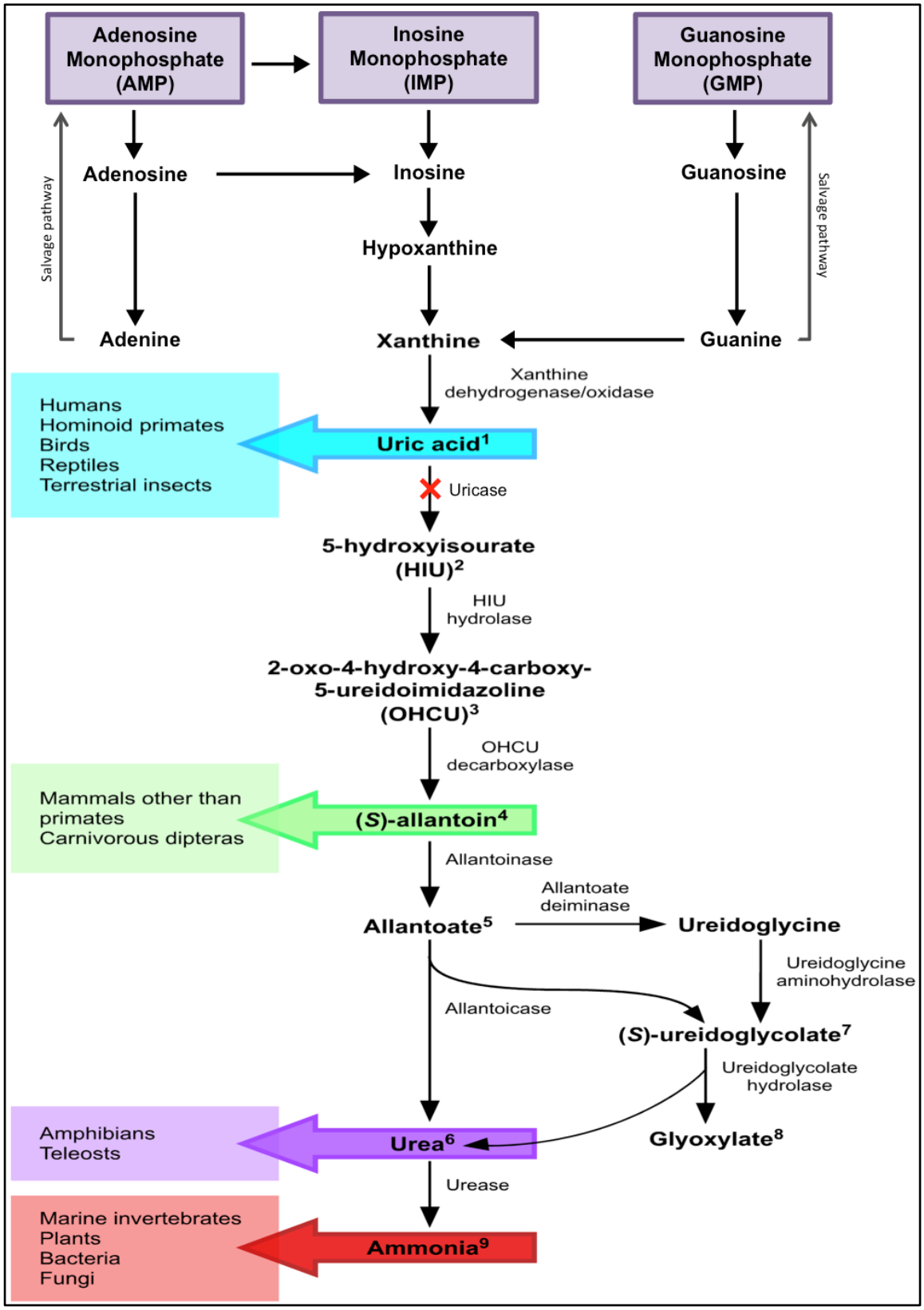


Figure 1.3: Flow sheet for conserved purine degradation and its variable end product metabolism among animal taxa. 1 indicates the end product of purine (AMP, IMP, GMP) metabolism in humans and other higher primates due to the evolutionary loss of uricase (indicated by red cross). Steps 2 to 9 illustrate the variation in end excretory products of purine metabolism between different species. Redrawn from Johnson et al. (2009b) and Lee et al. (2013).

1.1.1.3 Urate homeostasis in humans

In the human body, urate is the final breakdown product of purine catabolism. Degradation of endogenous (derived from nucleic acid metabolism) and exogenous purines (derived from dietary sources) result in the intracellular production of urate. In order to maintain a relatively constant (desirable) level of circulating urate in the body, a homeostatic mechanism functions to create a balance between the net production and excretion of urate. Urate homeostasis involves three determinants- its production in the liver, renal excretion and intestinal secretion (Rafey *et al.*, 2003). The production of uric acid occurs primarily in the liver by the action of a molybdenum metalloenzyme called xanthine oxidase. This poorly soluble intracellular product enters the body's circulation (as urate) by cellular efflux and is delivered to the kidney for excretion (Brondino *et al.*, 2006). Renal excretion of urate contributes to two-third of total urate excretion. A small proportion of urate that gains entrance to the intestine is degraded by the colonic bacteria and eliminated (Bobulescu and Moe, 2012; Rafey *et al.*, 2003). This process, termed intestinal/extra-renal uricolysis, is especially significant when renal handling of urate is compromised (Sorensen and Levinson, 1975).

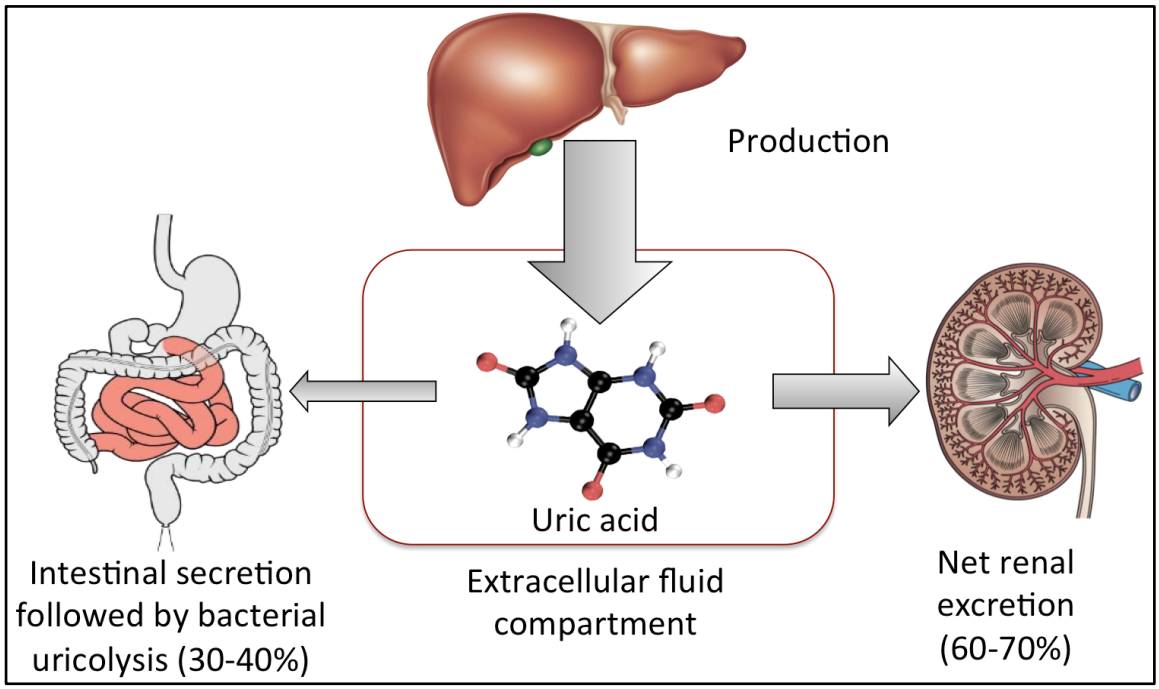


Figure 1.4: Schematic representation of the key determinants of urate homeostasis in humans. Redrawn from Bobulescu and Moe (2012).

1.1.2 Renal handling of urate and hyperuricaemia

Once urate in circulation reaches the kidney, it is freely filtered. However, under normal conditions, bidirectional transport of urate occurs in the renal proximal tubules, resulting in > 90% of the filtered urate being reabsorbed and < 10% of the filtered urate being excreted. The proportion of filtered urate eventually excreted in the urine (as uric acid) is termed as fractional excretion of uric acid (FEUA) (Fushimi *et al.*, 1990). The physiological value of FEUA under normal conditions is approximately 10% and is higher in females than males (Bobulescu and Moe, 2012). Fractional excretion of uric acid is indicative of renal function.

The serum urate level in humans is about ten times higher than in other mammals (Mandal and Mount, 2015). This can be attributed to the absence of the uricase enzyme and the urate reabsorption system in the kidney. The fact that the kidney invests resources in reabsorption of a large proportion of filtered urate compels us to speculate that urate is beyond an inert metabolic product in humans. Indeed, urate does have more roles to play in the human body apart from the undoubtedly important role of shuttling purine waste to the exterior (Richet, 1995). There is overwhelming evidence for its role as a respiratory and circulatory antioxidant (Carocho and Ferreira, 2013; Kang and Ha, 2014; Peden *et al.*, 1990).

Since the urate reabsorption system operates in the kidney, renal handling of urate is critical to maintain the desirable level of serum urate. The presence of a lower than normal level of serum urate is termed hypouricaemia. Hypouricaemia is defined as a serum urate concentration of < 0.2 mmol L^{-1} (Martín and Nieto, 2011). Defective urate reabsorption in the kidney has been shown to cause hereditary renal hypouricaemia characterised by hypouricaemia and increased FEUA (Windpessl *et al.*, 2016). Hereditary renal hypouricaemia is caused by mutations in two genes encoding renal transporters in the proximal renal tubule, *SLC22A12* (Solute carrier family 22 member 12) and *SLC2A9* (Solute carrier family 2 member 9) (Dinour *et al.*, 2010; Komoda *et al.*, 2004; Windpessl *et al.*, 2016). Conversely, hyperuricaemia is described as the presence of a greater than normal level of circulating urate in the blood stream.

Hyperuricaemia can be caused by either increased production of uric acid in the liver and/or reduced excretion by the kidney and gut. This in turn may be a consequence

of excessive intake of purine rich foods, and/or defective endogenous purine metabolism and/or endogenous defects in the urate reabsorption/secretion system. A lower FEUA is indicative of hyperuricaemia. The commonly used threshold values to define hyperuricaemia in adults include serum urate concentration of > 0.41 mmol L^{-1} in males and > 0.35 mmol L^{-1} in females (Dalbeth *et al.*, 2016).

Hyperuricaemia has been linked to diseases such as kidney stones and chronic uric acid nephropathy, evidence for causation being limited (Bobulescu and Moe, 2012). The relevance of hyperuricaemia in disease is most evident from its causal link with gout. Hyperuricaemia has also been described to be associated with gout-related comorbidities especially hypertension and diabetic renal disease (Doria and Krolewski, 2011; Perlstein *et al.*, 2006) and found to be a concomitant of the metabolic syndrome (MetS) that encompasses hypertension, obesity, insulin resistance and dyslipidaemia (Mandal and Mount, 2015). Hyperuricaemia observed in non-symptomatic patients is described as asymptomatic hyperuricaemia (aHU) (Richardson, 1991).

1.1.3 Gout

Gout being a complex disease is influenced by multiple genetic, environmental and interlinking risk factors (Robinson and Horsburgh, 2014). It is a painful inflammatory arthritic disease characterised by hyperuricaemia, hyperuricaemia being a necessary, although not sufficient, parameter for gout. Gout is the most common form of inflammatory arthritis in the elderly population. The global prevalence of gout is gradually increasing, which may be attributed partly to population aging, a rise in obesity and insulin resistance as well as changing lifestyle and dietary practices (Dalbeth *et al.*, 2016). The factors determining gout in the presence of hyperuricaemia are diverse and not completely understood (Dalbeth *et al.*, 2016).

The basic pathophysiological feature of gout is the deposition of monosodium urate (MSU) crystals in the synovial fluid of the joints following longstanding hyperuricaemia. Gout is therefore referred to as a urate crystal deposition disease. Joints at the extremities/periphery are commonly affected (Underwood, 2006), especially the metatarsal phalangeal joint of the big toe (Dalbeth *et al.*, 2016). The clinical features of gout arise due to the body's innate immune response to the accumulated crystals. Diagnosis of gout includes an inspection of clinical history, microscopic examination,

imaging and laboratory testing (Peláez-Ballestas *et al.*, 2010). The presence of tophus, defined as chronic inflammatory granulomatous lesions formed in response to crystal deposition (Dalbeth *et al.*, 2016) and elevated serum urate concentration, are strongly suggestive of gout (Janssens *et al.*, 2010).

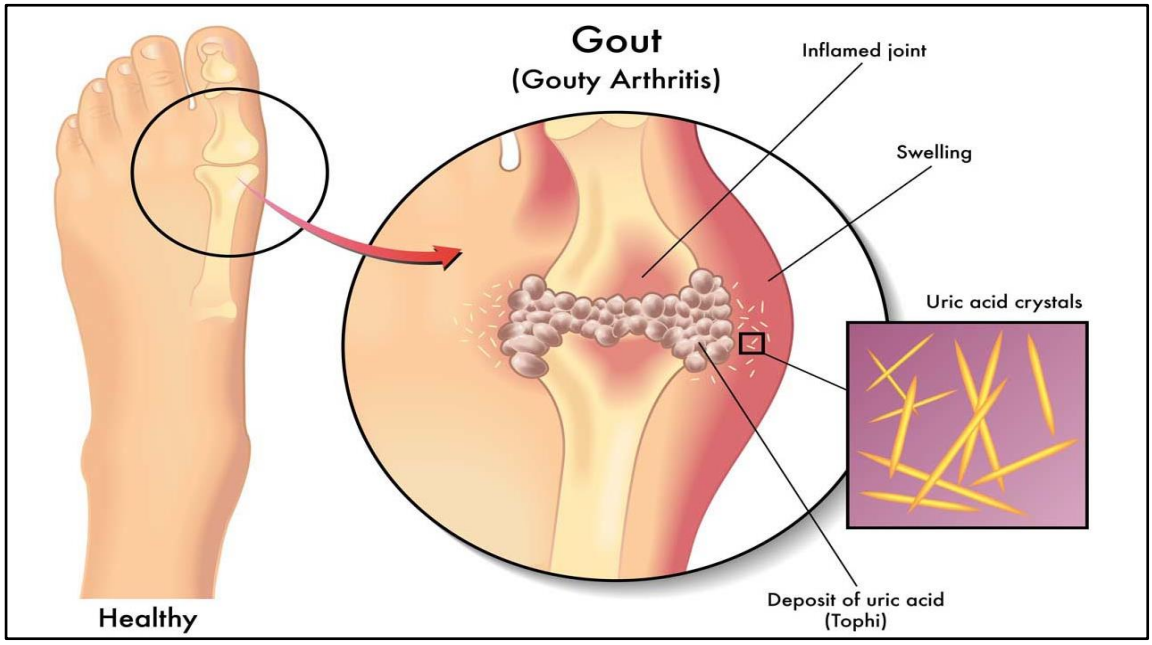


Figure 1.5: Pathophysiology of gout: Acute gout is characterised by redness, swelling and severe pain due to formation of monosodium urate (MSU) crystals within the synovial spaces of joints and an inflammatory response, while longstanding untreated hyperuricaemia can lead to deposition of MSU crystals (tophi) and subsequent permanent damage to the joints.

Acute gouty arthritis, the most common clinical presentation of gout, is characterised by severe pain, redness, tenderness, heat and swelling of the affected joint and causes restricted joint movement (Rome, 2012). Exogenous factors that trigger gout flares are diverse and include dehydration, starvation, alcohol consumption and intake of purine rich foods (Choi *et al.*, 2004b; Zhang *et al.*, 2012a)An acute gout flare peaks within the first day of attack and typically lasts for about a week, in the absence of treatment (Bellamy *et al.*, 1987; Lindsay *et al.*, 2011; Snaith, 2003).

If left untreated, acute gout may progress to advanced stage gout commonly called ‘chronic gout’. Chronic gout is essentially characterised by continuous low-level pain and frequent intermittent attacks (flares) of severe pain. Despite having low pain, the severity of gout may keep increasing over time (Choi *et al.*, 2005b; Lindsay *et al.*, 2011). This can lead to joint disability due to the formation of large deposits of MSU crystals in the joints called ‘tophi’ (singular: tophus) (Figure 1.5). Formation of tophi not only may be visually

distressing but can also cause irreparable damage to the joints in addition to restricted mobility (Choi *et al.*, 2005b).

1.1.3.1 Immunological perspective

Monosodium urate crystals are extremely potent elicitors of inflammation. A local immune response is evoked at the site of MSU crystal deposition via the recruitment of neutrophils, production of inflammatory mediators and proinflammatory cytokines (So, 2008). This process involves components of the innate immune system that can detect cellular products released by damaged cells in the body. The initial inflammatory response is triggered by two components, the toll-like receptors (TLRs) namely TLR2 and TLR4 (Akira and Takeda, 2004; Lim and Staudt, 2013), and the pattern recognition molecule called CD14 (Cluster of differentiation 14) (Fujihara *et al.*, 2003). Multiple intracellular processes lead to the formation of an inflammasome complex that triggers the maturation of interleukin 1-beta (IL1-ß) from pro-IL1-ß, which is the major mediator of the inflammatory effect (So, 2008) (Figure 1.6).

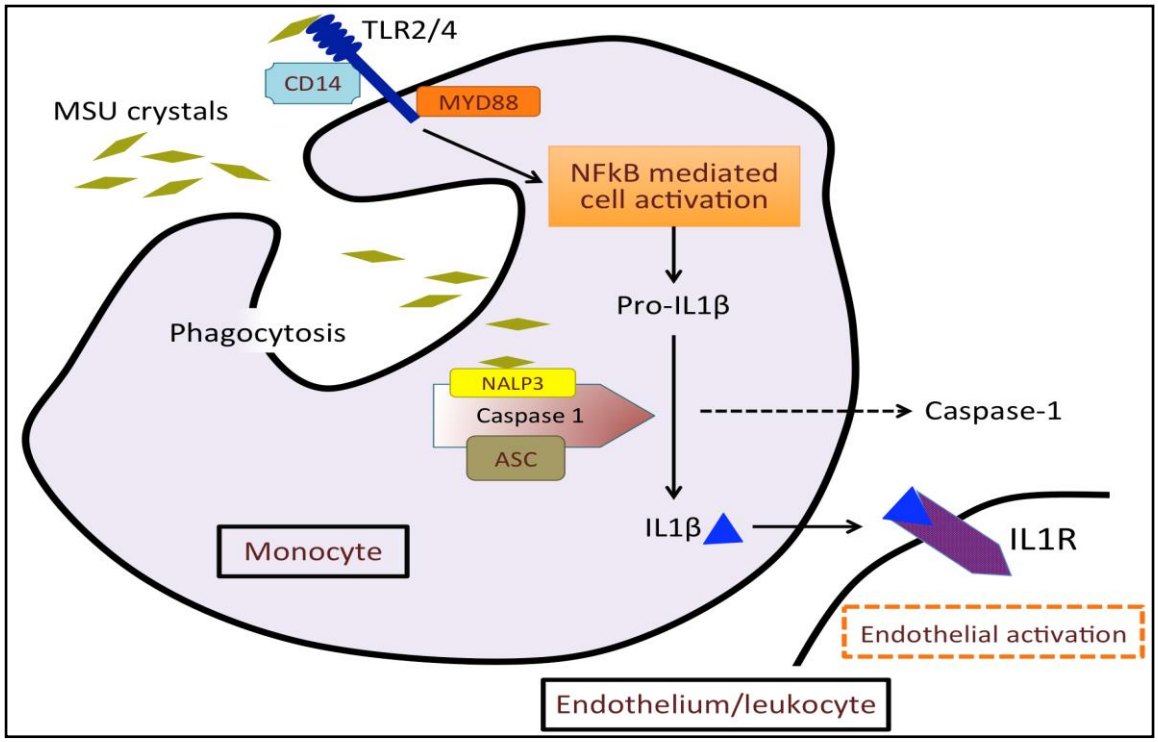


Figure 1.6: Immunological perspective of gout: MSU crystals activate monocytes via the Toll-like receptor (TLR) pathway and the inflammasome. Binding to TLR and CD14 promotes phagocytosis and cell activation through MYD88-dependent signalling mechanisms. In the cytosol, MSU crystals induce the formation of the NALP-3/NLRP3 (NACHT, LRR, and pyrin domain-containing-3) inflammasome and lead to caspase-1 processing of pro-IL-1β. Activation of the endothelium by IL-1β increases trafficking of neutrophils to the inflammatory site. ASC, apoptosis-associated speck-like protein containing a caspase-associated recruitment domain; IL, interleukin; NF-κB, nuclear factor-kappa-B. Modified and redrawn from So (2008).

1.1.4 Genetics of hyperuricaemia and gout

Hyperuricaemia is a complex trait. Variants in dozens of genes play a role in conferring risk to develop the phenotype. The interplay between the genetic risk variants and environmental risk factors is of crucial significance (Robinson and Horsburgh, 2014). Genome-wide association studies (GWAS), typically used to scan the genome to identify common genetic variants, mostly single nucleotide polymorphisms (SNPs), causally associated with a given phenotype/trait, is a methodology that has significantly improved our understanding of the genetic basis of complex traits including hyperuricaemia. At least 28 loci conferring risk to develop hyperuricaemia have been identified to date. Two major urate loci, namely *SLC2A9* and *ABCG2*, have been shown to collectively explain 3-4% of variance in serum urate concentration and thereby strongly influence serum urate levels in the body (Köttgen *et al.*, 2013).

Genome-wide association studies before 2013 identified a total of 11 urate loci, collectively explaining 5-6% variance in serum urate (Dehghan *et al.*, 2008a; Döring *et al.*, 2008; Kolz *et al.*, 2009; Yang *et al.*, 2010b). The loci were identified either in or near these genes; *PDZK1, GCKR, SLC2A9, ABCG2, RREB1, SLC17A1, SLC16A9, SLC22A11, NRXN2, INHBC.* The most recent serum urate GWAS data comes from Köttgen *et al.* (2013), who identified 18 novel urate related loci using the data from a population of >110,000 European individuals. The loci were identified in or near the *TRIM46, INHBB, SFMBT1, TMEM171, VEGFA, BAZ1B, PRKAG2, STC1, HNF4G, A1CF, ATXN2, UBE2Q2, IGF1R, NFAT5, MAF, HLF, ACVR1B/ACVRL1,* and *B3GNT4* genes and collectively explained ~7% variance in serum urate concentration (Köttgen *et al.*, 2013).

Given that hyperuricaemia is a prerequisite for the development of gout, it is highly likely that urate-associated genes are also associated with gout (Hollis-Moffatt *et al.*, 2012a; Hollis- Moffatt *et al.*, 2009; Köttgen *et al.*, 2013; Phipps-Green *et al.*, 2010). Recently 28 urate-associated loci, identified by Köttgen *et al.* (2013), were tested for their association with gout at the Merriman Laboratory (Phipps-Green *et al.*, 2016) using data from > 4,000 New Zealand European and Polynesian individuals. In addition to *SLC2A9, ABCG2, SLC17A1* and *GCKR*, significant associations with gout were identified, especially for four novel gout loci in or near *IGF1R, PDZK1, MAF* and *HLF* (Phipps-

Green *et al.*, 2016). These results also indicated a possible role of non-urate transporter genes in the development of gout.

1.1.5 Prevalence of hyperuricaemia and gout

As mentioned before, serum urate concentration of ≥ 0.41 mmol L^{-1} in men and ≥ 0.35 mmol L^{-1} in women is used to define hyperuricaemia (Dalbeth *et al.*, 2016). The fact that average recorded levels of serum urate may fall differently within this range for different populations (explained hereafter) means that the prevalence of hyperuricaemia varies between different regions of the world (Table 1.1). For example, the average serum urate concentrations have been recorded to be as low as 0.24 mmol L^{-1} and as high as 0.46 mmol L^{-1} in Brazilian and Taiwanese Aborigine males, respectively. For females, this range varies between 0.22 mmol L^{-1} to 0.37 mmol L^{-1} for Brazilian and Tibetan populations, respectively (Acheson and Florey, 1969; Chou and Lai, 1998; Gosling *et al.*, 2014). This average makes some populations to fall within the hyperuricaemic range e.g., Tibetans, Taiwanese Aborigine, Cook Island and New Zealand Māori (Brauer and Prior, 1978; Chou and Lai, 1998; Evans *et al.*, 1968). This, in turn, results in higher prevalence of hyperuricaemia in these populations (Chang *et al.*, 2001; Prior *et al.*, 1966).

Table 1.1: Country specific prevalence of hyperuricaemia

Country	Prevalence (%)	Country	Prevalence (%)
Brazil	13	Philippines	25
China	6 to 25	Russia	17
Indonesia	18	Samoa	33
Iran	8	Saudi Arabia	8
Italy	9 to 12	Seychelles	25
Japan	20 to 26	South Korea	5
Marshall Islands	85	Spain	5 to 11
Mexico	11	Sweden	10 to 16
Mongolia	5 to 18	Taiwan	10 to 52
New Zealand	8 (non-Māori)/17 to 19 (Māori)	Thailand	9 to 11
Nigeria	17	Turkey	12
Papua New Guinea	1	USA	21 to 22

Source: (Smith and March, 2015).

Although hyperuricaemia is essential for the development of gout, their prevalence may differ for the same population. The reason behind is evident – not all individuals with hyperuricaemia develop gout. For example, Taiwanese Aborigines have the highest prevalence of hyperuricaemia while the prevalence of gout has been recorded to be highest in Australasian populations (explained in the subsequent paragraph). The average level of urate is higher in males than females, and so is the prevalence of gout

(reviewed in Kuo *et al.* (2015)). An approximate range of the worldwide prevalence of gout was recorded to be 0.02% (for countries in South Asia: Afghanistan, Bhutan, India, Nepal and Pakistan) to 0.39% (Australasian countries; Australia and New Zealand) in a 2014 World Health Organisation (WHO) survey (Smith *et al.*, 2014), however this is likely to be an underestimate owing to the method of data capture. According to a recent study by Kuo *et al.* (2015), the overall prevalence of gout was recorded to be > 1% in most developed countries (especially in Europe and North America) (Figure 1.7). The overall prevalence of gout was recorded to be 3.9% in the adult US population (age ≥ 20 years) in the National Health and Nutritional Examination Survey (NHANES 2007-2008) (Zhu *et al.*, 2011). For Canada, an estimated general prevalence was suggested to be ~3% in adults (Badley and DesMeules, 2003; O'Donnell *et al.*, 2015). In Europe, Greece has been reported to have the highest prevalence of gout, at 4.75%, in the adult population (Anagnostopoulos *et al.*, 2010). Kuo *et al.* (2014b) provided the latest estimate of gout prevalence to be 3.22% in the adult (age ≥ 20 years) UK population, with an overall population estimate to be 2.49%. The overall estimates are similar to what have been reported for other European countries i.e., 3.3% for Spain (Sicras-Mainar *et al.*, 2013) and 3.7% for the Netherlands (Picavet and Hazes, 2003). On the other side of the picture, within Europe, the lowest prevalence of gout (0.3%) has been reported for the adult population of Portugal (Reis and de Queiroz, 2014) and Czech Republic (Hanova *et al.*, 2006).

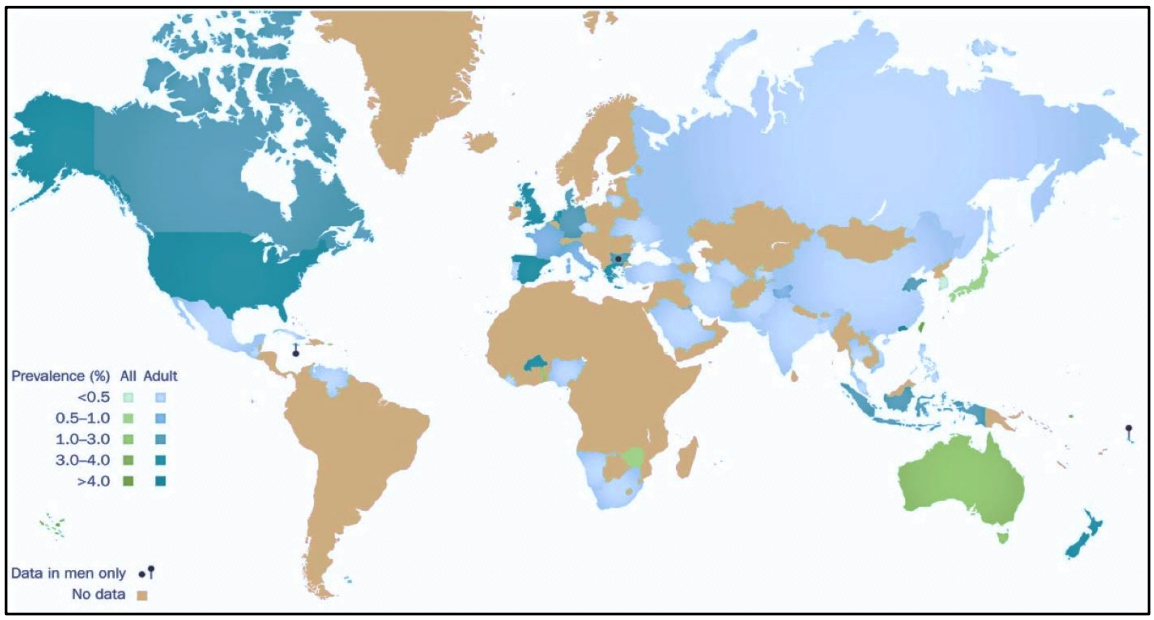


Figure 1.7: The estimated prevalence of gout across the world showing higher prevalence of gout in developed countries, especially in Oceania populations (Australia and New Zealand). Source: Kuo *et al.* (2015).

As illustrated in Figure 1.7 (Kuo *et al.*, 2015), the prevalence of gout is highest in Australasian countries i.e., Australia and New Zealand. Two equivalent studies indicated an overall prevalence of gout to be 1.5% for Australian (Robinson *et al.*, 2015) and 3.75% for New Zealand (Winnard *et al.*, 2012) adults. Interestingly, the prevalence and severity of gout differ strikingly among diverse ancestral groups in these regions, with Pacific Islanders and Māori having a higher prevalence of gout than people of European decent. About 3.8% of adults have been reported to have gout in an Australian aboriginal community in North Queensland, with 22% cases being identified as having severe gout (presence of subcutaneous tophi) (Minaur *et al.*, 2004). In contrast, an overall prevalence of 1.44% (all-age) was reported in Australian National Health Survey (ANHS) data (Skinner, 1997). For New Zealand, using data from the Aotearoa New Zealand Health Tracker (ANZHT), the prevalence of gout has been reported to be 3.2% in European individuals (Winnard *et al.*, 2012). In comparison, the prevalence of gout was as high as 7.6 and 6.1% for Pacific Island and Māori individuals, respectively (Winnard *et al.*, 2012). This indicated the rate of gout to be almost double in Pacific Island and Māori populations than Europeans. Over the passage of time, the prevalence of gout has progressively increased in the New Zealand Māori and European individuals (Figure 1.8), with a minimum rate of 2.7% and 0.3% recorded in 1958 to 6.06% and 3.24% in 2009 in both populations, respectively (Klemp *et al.*, 1997; Lennane *et al.*, 1960; Prior and Rose, 1966; Winnard *et al.*, 2012).

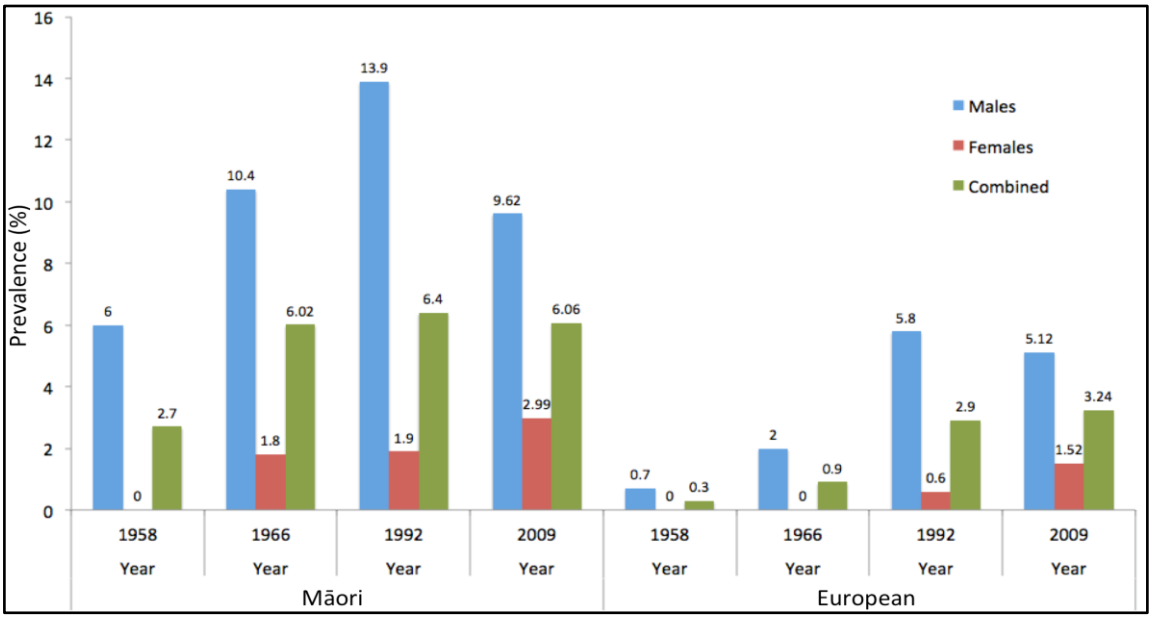


Figure 1.8: The estimated prevalence of gout in the New Zealand Māori and European population over time. Note: All values in the graph are presented for the year(s) when data were collected from the respective populations, which may differ from the year(s) it was published. A prevalence of 'zero' does not represent the absence of data.

1.1.6 Urate, gout and metabolic disease

1.1.6.1 Foreword

As mentioned earlier, gout has a high rate of co-occurrence with other metabolic conditions. These co-occurring conditions are generally referred to as 'comorbidities'. Considering hyperuricaemia and gout as index diseases under study, comorbidity is defined following Feinstein's definition (Feinstein, 1970) in my thesis. According to this definition:

"A comorbidity is any distinct additional entity that has existed or may occur during the clinical course of a patient who has the index disease under study."

Ample evidence is available linking gout to metabolic syndrome (or its components) (Puig and Martinez, 2008), renal disease (Kramer *et al.*, 2003; Kramer and Curhan, 2002; Yu *et al.*, 2012), hyperlipidaemia (reviewed in Kuo *et al.* (2014a)) and cardiovascular disease (CVD) (Choi and Curhan, 2007b; Krishnan *et al.*, 2006; Kuo *et al.*, 2013). However, evidence for its possible association with other, relatively less studied comorbidities, is still limited e.g., hypothyroidism (Durward, 1976; Kuzell *et al.*, 1955), cancer (Boffetta *et al.*, 2009; Kuo *et al.*, 2012) and anaemia (McAdams-DeMarco *et al.*, 2012).

The clinical guidelines and recommendations approved by American College of Rheumatology (ACR) (Nuki, 2014), American College of Physicians (ACP) (Shekelle *et al.*, 2017), European League against Rheumatism (EULAR) (Richette *et al.*, 2017) and the British Society for Rheumatology (BSR) (Roddy, 2016) have insisted and discussed more about the diseases that are pathophysiologically related to gout. Their guidelines also recommend urate-lowering therapy (ULT) only to the patients with either a high urate load or kidney-related comorbidities. Although it is understandable that a great focus was placed on comorbidities that are causally correlated with gout, multiple metabolic conditions may appear simply as co-occurring ailments rather than causes in many patients. Such metabolic conditions that may or may not indicate a direct (or causal) relationship with gout still merit provision of management guidelines.

Although the above-described observational studies have been useful in providing proof for a possible association between gout and other metabolic conditions, yet the information provided in these studies cannot ascribe the causality between cause and effect relationship. The main reason behind this uncertainty is that the data presented in such studies are purely observational based on the co-occurrence of the two conditions, which may in turn be prone to biases due to additional metabolic conditions as potential confounders of cause (exposure) and effect (disease) relationship (Bowden *et al.*, 2017). As complex traits (like gout) are affected by a number of environmental factors, getting false positive associations without accounting for confounders is highly likely (Hayden *et al.*, 2013; Shrank *et al.*, 2011). Unaccounted confounding is a reason why such studies may fail to replicate the reported associations or prove any causality when tested in randomised controlled trials (RCTs) (Kovesdy and Kalantar-Zadeh, 2012; Smith and Ebrahim, 2001, 2008). However, it may not always be possible to carry a large randomised controlled trial due to several social reasons. In such scenarios, Mendelian randomisation (MR) can be used as a successful approach, analogous to RCT, to infer a possible causal relationship between cause and effect (Iturrieta-Zuazo and Walter, 2015; Smith and Hemani, 2014). Mendelian randomisation makes use of genetic variants robustly associated with exposure (e.g., body iron levels) as instruments to identify their predictive effect on the outcome (e.g., serum urate/gout), adjusting simultaneously for potential confounders (Smith and Ebrahim, 2003). The MR exploits the basic principle of random assignment of alleles at conception that are not prone to confounding via environmental factors as well as reverse causation due to unidirectional flow of biological information i.e., information can be only translated from genome to protein and not in the reverse direction. Hence, exploring new information between gout and other metabolic conditions not only requires broadening the range of coexisting conditions being considered but also a more robust approach, like genetic association analysis and/or MR, to explain 'which is driving which' relationship between them.

1.1.6.2 Gout and metabolic disease in New Zealand

Aotearoa New Zealand has one of the highest documented prevalence of gout worldwide (as mentioned above) as well as its related comorbidities (Winnard *et al.*, 2013); worth mentioning is the co-prevalence of obesity, kidney disease, type 2 diabetes (T2DM) and CVDs, occurring individually or as components of the metabolic syndrome,

with hyperuricaemia or gout. Gout has generally been associated with all-cause mortality, mostly due to death from renal disease or CVD (Choi and Curhan, 2007b; Stamp and Chapman, 2013; Teng *et al.*, 2012). Despite their utmost clinical importance and subsequent increased demand for health care resources, studies exploring metabolic disease(s) as a comorbidity of gout have been seldom carried out in the New Zealand population. In the following paragraphs, using the data available in the literature, I have summarised the history of the prevalence of some well-known metabolic conditions. The summary is structured to specifically outline these comorbidities in the context of their coexistence with hyperuricaemia and gout, both globally and in New Zealand.

1.1.6.2.1 Obesity

Ranking first in the list of gout-associated comorbidities is obesity. Generally, a body mass index (BMI) > 30 kg/m^2 is referred as a cut off to specify obesity. Obesity, particularly visceral obesity, is also a well-recognised risk factor for other metabolic diseases e.g., T2DM and CVD (Després and Lemieux, 2006; Mokdad *et al.*, 2003). Being obese has also been reported to increase the xanthine oxidase (XO) catabolism of purines and, thereby, enhanced production of urate in adipose tissues (Tsushima *et al.*, 2013). High BMI measures and obesity have been described to be associated with increased urate levels in several human-based studies, mostly as components of metabolic syndrome (Billiet *et al.*, 2014; Johnson *et al.*, 2013b; Nejatinamini *et al.*, 2015). In addition, data from MR studies have also described a causal association between genetically determined adiposity measures and higher urate levels. These studies have described an elevation in urate to occur as a consequence of an increase in triglycerides, BMI, adiposity and its associated risk factors (Lyngdoh *et al.*, 2012; Palmer *et al.*, 2013; Rasheed *et al.*, 2014). Moreover, in a study including the data of 15,533 American individuals from a community-based cohort (Campaign Against Cancer and Heart Disease/CLUE II), McAdams-DeMarco *et al.* (2011) reported the onset of gout to be 3 years earlier on an average in obese (age > 21 years) compared to non-obese participants. An analysis of the US population data from NHANES 2007-2008 showed that 53% of the individuals with gout (age > 20 years) were obese (Zhu *et al.*, 2012).

An early survey of 115 Māori men revealed that individuals with hyperuricaemia and gout tend to be heavier than their comparative normouricaemic controls (Gibson *et*

al., 1984). In general, New Zealand is facing a continuous rise in the prevalence of obesity. Recent data from the New Zealand Ministry of Health reported an overall 31% of the individuals to be obese and 34% still overweight (www.health.govt.nz/). New Zealand has also been ranked as 3rd in the 2014 OECD (Organisation for Economic Co-operation and Development) report for the worldwide prevalence of obesity (Ng *et al.*, 2014). According to this survey, Polynesians largely contributed to the burden of obesity in New Zealand (www.oecd.org/). The New Zealand Ministry of Health further confirmed the prevalence of obesity for different ancestral groups in New Zealand, with Pacific Islanders being at the top (66.2%) followed by Māori (46.5%) and Europeans (and other ethnicities (29.2%) (Figure 1.9) (www.stats.govt.nz/ and www.health.govt.nz/).

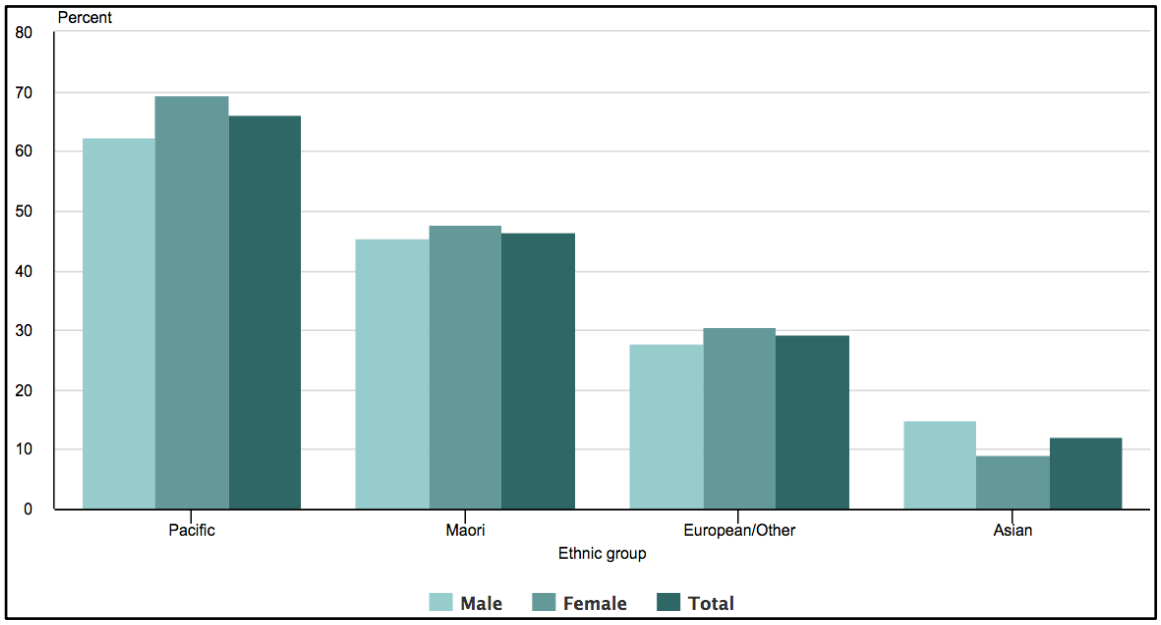


Figure 1.9: Graphical illustration of the proportion of the New Zealand obese population, stratified by ethnicity and gender. Source: www.stats.govt.nz/ and www.health.govt.nz/).

1.1.6.2.2 Renal impairment and chronic kidney disease

Renal dysfunction is common in patients with hyperuricaemia or gout. Increased serum creatinine and decreased glomerular filtration rate (GFR) are indicative of reduced kidney function. Higher serum creatinine levels have been classically associated with higher levels of serum urate in gout patients (Nishida, 1992). Contradictory to this, a Mendelian randomisation (MR) study found a causal association between genetically higher levels of serum urate and decreased serum creatinine (Hughes *et al.*, 2014). Evidence has also indicated reduced GFR to be associated with higher urate levels (Johnson *et al.*, 2013a; Mohandas and Johnson, 2008; Suliman *et al.*, 2006). A study

including data from NHANES 2009-2010 reported a 2-3 fold increase in the prevalence of gout for a decrease in GFR (data for every 30 mL/min/1.73m^2 decrease in GFR) (Krishnan, 2012). Several studies have reported an elevation in urate to be predictive for a progressive increase in pre-established chronic kidney disease (CKD) (Shi *et al.*, 2011) or the development of renal impairment in individuals with normal renal function (Ohno *et al.*, 2001; Shi *et al.*, 2011; Syrjänen *et al.*, 2000). Chronic hyperuricaemia has strongly been associated with CKD (Johnson *et al.*, 1999; Johnson *et al.*, 2013a). Higher prevalence of CKD has also been reported in gout patients. Data from the large US population-based study, NHANES 2007-2009, revealed about 71% of the gout patients to have CKD at stage 2 or higher (Zhu *et al.*, 2012). In a recent meta-analysis of data from 17 large studies including multiple populations, the global co-prevalence of gout with advanced stage CKD was found to be 24% (Roughley *et al.*, 2015). In contrast, two other recent studies reported a prevalence of gout in CKD patients to be 6.8% and 24.3% in the US and Dutch populations, respectively (Jing *et al.*, 2014; Wang *et al.*, 2015). In addition, a possible relationship of increased urate concentration has been marked as an independent predictor for the development of CKD in patients with T2DM (Altemtam *et al.*, 2011; Zoppini *et al.*, 2012). Having chronic kidney disease has also been regarded as a factor that increases the risk for the development of diabetes, CVD and other heart-related complications (Go *et al.*, 2004).

Likewise, a few studies have also reported the co-prevalence of gout with CKD in the New Zealand (NZ) population. A study reported a 3.5-fold higher rate of ESRD (end-stage renal disease) in Māori and Pacific Islanders than Europeans (Collins, 2010). Robinson *et al.* (2012) found that ~16-27% of gout patients who were admitted in any of the New Zealand hospitals between 1999 and 2009 had renal complications, with most patients being either Māori or Pacific Islanders. Although exact figures are still unknown, a 2014 national consensus statement from the New Zealand Ministry of Health (www.health.govt.nz/) suggested about 7% of the adult New Zealand population to be CKD sufferer (NZMH, 2014).

1.1.6.2.3 Type 2 diabetes

Type 2 Diabetes (T2DM) is another metabolic condition that has been consistently reported for its coexistence with hyperuricaemia and gout (Winnard *et al.*, 2013).

Elevated levels of serum urate have been shown to be strongly associated with an increased risk of T2DM (Sluijs *et al.*, 2015). However, the same study was unable to find a causal relationship between genetically high serum urate and T2DM using the MR approach (Sluijs *et al.*, 2015). Moreover, hyperuricaemia has been generally associated with a high risk of T2DM (Billiet *et al.*, 2014; Johnson *et al.*, 2009a). In contrast, the relationship between gout and diabetes is rather complex (Stamp and Chapman, 2013), with one study showing reduced risk of developing gout in patients with advanced T2DM (was ascribed to increased urination leading to increased urate excretion) (Choi and Ford, 2008). Additionally, a large study in the US population showed a higher risk of developing T2DM in male gout patients with higher risk of CVD (Choi *et al.*, 2008). Another study in the UK based population also found a similar relationship between gout and the risk of developing T2DM (Rho *et al.*, 2016). However, the risk ratio (RR) was higher in females (RR = 10.1) than males (RR = 9.5) (Rho *et al.*, 2016).

Consistent with the aforementioned study findings, a higher risk of developing T2DM has also been observed in the New Zealand individuals, with disproportionate adverse effects in the Polynesian populations (explained hereafter). According to a recent report from the New Zealand Ministry of Health (www.health.govt.nz/), being a Māori or Pacific Islander increases the risk of developing T2DM by 3-fold in comparison to other ancestral backgrounds. An overall prevalence of T2DM has been reported to be 6.6% in the New Zealand population (Winnard *et al.*, 2013), while other studies reported the prevalence to be highest in Pacific Islanders (Chan *et al.*, 2014). A recent study by Coppell *et al.* (2013), using data from the 2008/09 New Zealand Adult Nutrition Survey (NZANS), found an overall prevalence of prediabetes and diabetes to be 25.5% and 7.0%, respectively. The study also reported the prevalence of diabetes to be higher in males (8.3%) than females (5.8%). Winnard *et al.* (2013) reported a higher co-prevalence of gout and T2DM, with 25.6% of the gout patients (n = 119,234) in the ANZHT database also having T2DM.

1.1.6.2.4 Cardiovascular disease

A high prevalence of CVD and T2DM travels concomitantly with gout (Winnard *et al.*, 2013). Linking the above mentioned comorbidities, elevated levels of urate have been associated with an increased risk of cardiovascular mortality in patients with CKD

(Kanbay *et al.*, 2012; Madero *et al.*, 2009) and T2DM (Ito *et al.*, 2011). Hyperuricaemia has been described as a potential risk factor for a number of heart-related problems including hypertension (Grayson *et al.*, 2011), coronary heart disease (CHD)/ischemic heart disease (IHD) (Kim *et al.*, 2010), heart failure (Huang *et al.*, 2014), stroke (Kim *et al.*, 2009) and peripheral vascular disease (Baker *et al.*, 2007). More recent MR studies have reported a causal relationship between genetically predicted serum urate and adverse cardiovascular outcomes including sudden cardiac death (Kleber *et al.*, 2015). However, other studies using the same approach were unable to find such causal associations, between serum urate and CHD (Palmer *et al.*, 2013; White *et al.*, 2016) or biomarkers of CHD (Rasheed *et al.*, 2014). Despite these findings, hyperuricaemia has been associated with all-cause cardiovascular mortality in several populations (Chen *et al.*, 2015a; Li *et al.*, 2016; Moulin *et al.*, 2017). In line with these studies, gout has also been described to be strongly associated with cardiovascular mortality (Kuo *et al.*, 2009), especially CHD associated mortality (Choi and Curhan, 2007b; Clarson *et al.*, 2013) and myocardial infarction (Kuo *et al.*, 2013).

An overall prevalence of CVD in the New Zealand population has been reported to be 5.4%, with ~22.7% co-prevalence of gout and cardiovascular events (Winnard *et al.*, 2013). Heart failure and CVD were reported as potential comorbid conditions, respectively, in 27.6% and 39.1% of the gout patients who were admitted to the New Zealand hospitals between the years 1999-2009 (Robinson *et al.*, 2012). In the context of the New Zealand population, CHD has been declared as most common type of CVD. According to the 2014 data from the New Zealand Ministry of Health (www.health.govt.nz/), one in every eighteen adults has been diagnosed with CHD, with a death rate of one every 40 minutes. Again, the prevalence of CHD is higher in Māori (6.2%) and Pacific Islanders (5.04%) in comparison to the individuals of European descent (4.2%) (Thornley *et al.*, 2011). Bramley *et al.* (2004) showed that Māori had risk ratios of 1.9 and 5.7 for death by IHD and diabetes, respectively, when referenced against the non-indigenous New Zealand population. Data from The Auckland Region Coronary or Stroke Study (ARCOS) further supported this fact, in which Māori (Death rate = 68%) and Pacific Island (Death rate = 64%) individuals had higher death rates caused by a coronary event compared to Europeans (Death rate = 44%) (Bullen and Beaglehole, 1997). Another more recent analysis in the same database revealed that heart stroke

occurs at a comparatively younger age in Polynesians (61.9 years) than Europeans (74.6 years) (Feigin *et al.*, 2006).

1.1.7 Metabolic disease/comorbidities included in this study

The research in my thesis was broadly conducted and limited to the occurrence of (or changes in) the following three metabolic conditions/comorbidities of gout. These metabolic conditions have been abundantly reported for their possible coexistence with hyperuricaemia and/or gout as potential comorbidities.

1. Imbalanced iron homeostasis
2. Metabolic syndrome
3. Imbalanced lipid metabolism (dyslipidaemia)

To make this section precise for the readers and due to the discrete nature of selected comorbidities, extensive cumbersome text explanations are avoided here. Instead, section 1 of each chapter (Chapter 2 to Chapter 5) is dedicated to a detailed description of the history, background concept and biological mechanisms for each of the metabolic diseases plus rationale for its possible relationship with urate metabolism and gout, as described in the literature per se. The following paragraphs provide only a summary background for each of these comorbidities in a bigger picture before explaining the aims of this study.

1.1.7.1 Iron homeostasis and gout

Iron is an essential element of the human body and is vital for an array of metabolic functions, such as oxygen transport and oxidative phosphorylation. The human body is specialised to strictly maintain the normal blood levels of iron within a narrow range of 3 to 4 grams (3.8 g in males and 2.3 g in females) of the total body weight (Leong and Lonnerdal, 2012). However, the range may vary between 4 and 5 grams for well-nourished people (Gropper and Smith, 2013).

Food is the primary source of iron, providing heme and non-heme iron through animal and plant sources, respectively. In comparison to non-heme iron derived from plants, heme iron derived from animal sources (meat, fish and seafood) is more readily absorbable in the body. Dietary heme iron accounts for two-thirds of the average person’s

total iron stores (Bezwoda *et al.*, 1983; Carpenter and Mahoney, 1992; West and Oates, 2008) and due to its high bioavailability, the absorption of heme iron is 5 to 7 times higher than that of non-heme iron in the gastrointestinal tract (Björn-Rasmussen *et al.*, 1974; Reizenstein, 1979). The amount of iron gained from a non-heme source is totally dependent upon the individual's body iron status (Hurrell and Egli, 2010).

Maintaining iron in strict limits is crucial due to its biological toxicity and catalytic activity. Iron in its free floating form can primarily cause oxidative stress and, thereby, destruction and/or death of otherwise healthy cells due to the increased production of reactive oxygen species (ROS) (Bresgen and Eckl, 2015; Dixon and Stockwell, 2014). The human body is specialised to not only regulate its pre-existing endogenous iron pools but also the amount of iron from the diet, which is regulated at both the systemic and cellular levels (Ganz, 2013; Wang and Pantopoulos, 2011). The regulation of iron homeostasis involves a number of proteins, enzymes, and other cellular structures to properly distribute and store the metal. After reaching the stomach, iron in the food is absorbed by the duodenal enterocytes in a controlled fashion (Fuqua *et al.*, 2012). Once it enters the circulatory blood pool, the transport of iron in the blood is controlled by a glycoprotein named transferrin, which reversibly binds iron ions and carries them to the surrounding tissues (Rouault, 2003). The cellular uptake of iron is mainly governed by transferrin receptor mediated import from transferrin into the cells (Moos, 2002). As the human body does not possess any specialised excretory mechanism for the removal of iron, at each successive stage of its distribution, any extra amount of iron is required to be liganded and stored. This exceptionally important need to sequester iron in a suitable liganded form is fulfilled via 'ferritin'. Ferritin is a hollow globular protein, synthesised in the liver and specialised for the storage of excess iron in a non-toxic form and its release in a controlled fashion (Harrison *et al.*, 1986).

Failure to maintain normal body iron levels may end up with negligible iron imbalance to severe iron-related disorders. Two disorders of abnormal iron levels in the body are iron deficiency – too low iron, and iron overload – too high iron. In aggregate, iron deficiency and haemochromatosis (iron overload) have been reported to affect over 1 billion people around the globe (Crownover and Covey, 2013; Hentze *et al.*, 2004; WHO, 2015).

1.1.7.1.1 Body iron profile

Ferritin reflects cumulative iron stores in the body and is measured under regular laboratory practices as a surrogate marker to determine iron levels in the blood and iron-related disorders (WHO, 2011). In addition, body iron profile is reflected via the amount of total blood (serum) iron, serum transferrin, total iron binding capacity (TIBC) and transferrin saturation (TSAT). Total iron binding capacity (TIBC) is the capacity of the blood to bind iron with transferrin, that in turn indicates the maximum amount of iron that blood can carry and is an indirect measure of blood transferrin (Yamanishi *et al.*, 2003). Transferrin saturation (TSAT), on the other hand, is an indirect estimate of the amount of transferrin available for iron to bind. It is calculated as the ratio between total blood iron and TIBC and represents the percentage of transferrin's iron-binding sites occupied by iron i.e., a TSAT of 10% indicates that 10% of the iron-binding sites on the transferrin possess bound iron while 90% of the sites are still free for the iron to bind.

1.1.7.1.2 Iron homeostasis, urate and gout

Urate is a well known iron chelator i.e., it has the ability to bind to and protects the body tissues from metal (iron)-mediated free radical damage (Davies *et al.*, 1986; Ghio *et al.*, 1994). Iron, in turn, is known for its ability to modulate the activity of xanthine oxidase and subsequently the production of urate (Ghio *et al.*, 2002). Ferritin, TIBC and TSAT have been positively correlated with urate in the US National Health and Nutrition Examination Survey (NHANES III) (Ghio *et al.*, 2005; Mainous *et al.*, 2011) with elevation in serum urate suggested as an indicator of iron overload (Mainous *et al.*, 2011).

Supportive evidence for iron as a trigger for gout flares is provided by the ability of iron to form complexes with MSU crystals *in vitro*, their presence in the synovial fluid, stimulation of oxidative stress through the generation of ROS, granulocyte and complement activation and production of lymphocytes (Ghio *et al.*, 1994). Association of iron with a number of pro-inflammatory activities in animal models (Dabbagh *et al.*, 1992) and a decrease in gouty flares following phlebotomy to attain near iron-deficient levels in hyperuricaemic patients (Facchini, 2003) are also suggestive of its role in gouty arthropathy. Consistent with these observational studies, a recent study reported the association of a genetic variant in the transferrin receptor (*TFRC: rs1466085*) with gout.

Individuals carrying the risk allele of this variant also self-reported a higher likelihood of an iron-rich food as a trigger of flares (Merriman *et al.*, 2015).

Interestingly, purine-rich foods from animal-based diet have been associated with increased risk of recurrent gout attacks while purine-rich foods from plant-based diet did not show a strong correlation (Zhang *et al.*, 2012b). These observations, combined with the other observational and intervention data (Facchini, 2003; Ghio *et al.*, 2005; Mainous *et al.*, 2011), are consistent with the idea that iron in purine-rich foods (red meat for example) could be a causal factor of gout.

Despite the presence of evidence in literature suggesting a possible role of iron and its related disorders in predicting hyperuricaemia and gout, this important relationship has been excessively neglected in New Zealand population-based studies. In fact, apart from studies that are more observational and conventional, data providing information on a causal relationship between these two important metabolic components are not available in the literature. Therefore, imbalanced iron homeostasis was selected as one of the major metabolic complications of gout in my thesis. My research was conducted with the hope of shedding light on a possible causal relationship between these two metabolic conditions and to fill in the substantial gap in literature, especially in the context of the New Zealand population.

1.1.7.2 Metabolic syndrome and gout

Almost all of the metabolic diseases described in Section 1.1.6 are collectively considered as ‘metabolic syndrome’ or MetS. However, the debate of inclusion of hyperuricaemia per se in the definition is still controversial (Nejatinamini *et al.*, 2015; Wei *et al.*, 2015). The main reason for considering hyperuricaemia as one of the MetS components is the presence of plentiful data in a number of epidemiological studies reporting co-occurrence of and/or association between hyperuricaemia and MetS components (detailed in Section 1.1.6) (Billiet *et al.*, 2014; Chen *et al.*, 2007; Liu *et al.*, 2015; Yamasaki and Tomita, 2008).

Most of the epidemiological studies have suggested hyperuricaemia to be a condition occurring secondary to hyperinsulinaemia (Muscelli *et al.*, 1996; Soltani *et al.*, 2013) and thereby its positive association with insulin resistance and diabetes (Li *et al.*,

2013; Lippi *et al.*, 2008). Additionally, data for its positive association with other MetS components is abundant too e.g., obesity (Han *et al.*, 2014; Tang *et al.*, 2010), cardiovascular disorders and hypertension (Borghi *et al.*, 2014; Nakanishi *et al.*, 2003). A more recent study has indicated a positive correlation between elevated serum urate and a collection of risk factors that culminates as MetS (Nejatinamini *et al.*, 2015).

Components of MetS and the data for their possible relationship with hyperuricaemia and gout are already detailed in Section 1.1.6. However, despite the above-mentioned observational data (and data provided in Section 1.1.6), the number of studies providing a genetic association between gout and the metabolic syndrome are scarce. The research in my thesis was structured not only to fill in this substantial gap in the literature but also to provide a genetic basis to this observational relationship between the two metabolic conditions, especially with an emphasis given to the New Zealand-based ancestral population groups.

1.1.7.3 Lipid metabolism and gout

A number of lipid biomarkers and their transport proteins collectively make the blood lipid profile. However, most of this profile is characterised by two types of lipids, cholesterol and triglycerides (TG) and their carrier proteins called 'lipoproteins'. Lipoproteins are generally categorised by their density, the amounts of lipids and proteins in the molecule and the type of lipid biomarker they tend to transport through circulation. Whilst very low-density lipoproteins (VLDL) mainly transport TG, low-density lipoproteins (LDL) and high-density lipoproteins (HDL) are the major transporters of cholesterol (Voet *et al.*, 2006). An abnormal amount (hypo or hyper) of lipids in the blood is referred to as 'dyslipidaemia'.

As mentioned earlier, hyperuricaemia has been associated with dyslipidaemia as part of the metabolic syndrome (Mandal and Mount, 2015). A number of reports explained an association of hyperuricaemia and gout with all-cause and CVD mortality and myocardial infarction independent of confounders (Chen *et al.*, 2009a; Krishnan *et al.*, 2006; Kuo *et al.*, 2009). In addition, a recent MR study has provided evidence for a causal role of elevated triglyceride levels in raising serum urate in European individuals (Rasheed *et al.*, 2014). However, the biochemical basis for the causal relationship between serum urate and lipids is still poorly understood.

Genome-wide association studies have successfully identified hundreds of common variants with a significant association with the studied phenotype/disease. These variants, however, usually explain a relatively small proportion of disease heritability with a large proportion still 'missing' or 'hidden' (Goldstein *et al.*, 2013). For example, the most recent genome-wide association study from Köttgen *et al.* (2013), that combined data from > 110,000 European-ancestry individuals, provided evidence for association of 28 loci that collectively explain 7.0% of the variance in urate. Albeit considering all common variants included in the Köttgen *et al.* (2013) GWAS could collectively explain ~27-41% variance in urate, a significant portion of variance still remains unexplained. Non-urate loci (besides those involved in urate transport) may not only be helpful in exploring the 'missing' part of the urate variance but would also help to characterise the shared genetic basis of gout or urate with other metabolic conditions. Two lipid-related genes, *LRP2* (lipoprotein receptor-related protein 2/megalin) and *A1CF* (apolipoprotein B mRNA-editing enzyme 1 (APOBEC1) complementation factor) (explained below), are examples of such non-urate loci that have been selected to study in this thesis.

1.1.7.3.1 *LRP2* and *A1CF*

Lipoprotein receptor-related protein 2 (LRP2) or megalin is a protein which in humans is encoded by the *LRP2* gene (Fisher and Howie, 2006; Saito *et al.*, 2005). Lipoprotein receptor-related protein 2 gene is a non-urate transport locus that has been identified for its predominant function in lipid metabolism (Cabezas *et al.*, 2011; Christensen and Birn, 2002) and reabsorption and metabolism of glomerular-filtered substances (Hosaka *et al.*, 2009). The T allele of a common variant (*rs2544390*) within the *LRP2* gene has been reported to be associated with higher serum urate concentration in Japanese individuals (Hamajima *et al.*, 2012; Kamatani *et al.*, 2010). A more recent study (Rasheed *et al.*, 2013b) was able to identify the population-specific effects for the T allele of the *rs2544390* variant with an increased gout risk in New Zealand Māori and Pacific Island individuals but not Europeans. The study further indicated that alcohol intake in New Zealand Polynesian population over-rides the otherwise protective role of the C allele of this variant and thus causes increase in gout risk due to a non-additive gene (C allele of *LRP2* SNP *rs2544390*)-environment (alcohol intake) interaction (Rasheed *et al.*, 2013b). Consistently, Dong *et al.* (2015) also reported the T allele of *rs2544390* to be associated with increased susceptibility to gout in the Han Chinese population.

In addition to common variants, low frequency (or rare) variants within the *LRP2* gene have been reported for their association with several other disease phenotypes. A study identified disease mutation clusters of rare variants within *LRP2* to be associated with Autism Spectrum Disorders (ASD) in three different datasets (Ionita-Laza *et al.*, 2012). A more recent study that measured 13 urinary biomarkers in the Framingham Heart Study (FHS) Offspring Cohort (n = 2,640) reported a cluster of rare variants in the *LRP2* gene to be associated with urinary levels of Trefoil Factor-3 (*TFF3*; gene encoding the TFF3 protein, expressed in gastrointestinal mucosa and possibly involved in protecting, stabilizing and healing of mucus layer) along with multiple common variants of *LRP2* to be associated with TFF3 levels in urine and kidney injury molecule 1 (McMahon *et al.*, 2014).

Another emerging candidate in the non-urate transport gene list is apolipoprotein B mRNA-editing enzyme 1 (APOBEC1) complementation factor or *A1CF*, which plays a role in the production of two different protein isomers, apo B-48 and apo B-100 from one nuclear gene (Chen *et al.*, 1987; Powell *et al.*, 1987). In the recent GWAS, the *A1CF* variant (*rs10821905*) was associated with serum urate levels in Europeans (Köttgen *et al.*, 2013), along with significant association with gout risk in New Zealand Europeans although not in Māori and Pacific Islanders (Phipps-Green *et al.*, 2014). In terms of data from fine mapping and whole genome or exonic sequencing, no study has yet reported any disease-based association for other variants within *A1CF*.

Both, *LRP2* and *A1CF*, have also been reported as loci associated with kidney function in a recent GWAS (*LRP2* = 3.5E-08; *A1CF* = 1.07E-12) that included the data from > 130,000 European individuals (Pattaro *et al.*, 2016). Despite evidence for the involvement of *LRP2* and *A1CF* in a number of physiological processes, little is known about the effect of their possible causal variants on biological pathways, particularly those leading to gout. Based on the above described literature, it was hypothesised that other less common variants within *LRP2* and *A1CF* coding regions could contribute to gout risk in New Zealand Māori and Pacific Islanders and Europeans, which may prove helpful in explaining some part of the ‘unexplained heritability’ for gout.

1.1.8 Specific aims of the study

The research was carried out to add new data to the repository of 'gout and metabolic disease', especially by focusing on those coexisting metabolic conditions that have been regularly ignored and/or the data for which are scarce/absent in relation to hyperuricaemia and gout. While the broader aim of this study was to test the biochemical and genetic association of metabolic diseases with gout and hyperuricaemia in the European and New Zealand Polynesian populations, the specific aims were based on discrete hypotheses, outlined as follows:

1. To characterise iron profile biomarkers (total iron, transferrin, ferritin, transferrin saturation and total iron binding capacity) and test the hypothesis that changes in these markers are linearly correlated with a change in serum urate concentration and are predictable for hyperuricaemia and gout in the New Zealand European and Polynesian (Māori and Pacific Islanders) individuals.
2. To test the hypothesis that variants in the haemochromatosis gene (*HFE*) are independently involved in gout risk in the New Zealand Polynesian and European individuals.
3. To exploit the summary statistics data from recent genome-wide association studies (European population) and test the hypothesis that there is a causal relationship between the above-mentioned serum iron biomarkers and urate concentrations using the Mendelian randomisation approach.
4. To test the hypothesis that variants in metabolic syndrome related genes are independently involved in gout risk in Polynesian (Māori and Pacific Island) and European New Zealanders.
5. To identify and characterise other common variants obtained from exon sequencing of two lipid-related genes (*LRP2* and *A1CF*) and use a replication-based approach to test the hypothesis that common variants in these genes have an independent association with hyperuricaemia and gout in European and/or New Zealand Polynesian individuals.
6. To identify and characterise the rare and non-synonymous variants obtained from exon sequencing of two lipid-related genes (*LRP2* and *A1CF*) and test the hypothesis that these rare and/or non-synonymous variants collectively have an

influence on hyperuricaemia in European and New Zealand Polynesian individuals.

1.1.9 Thesis structure

This thesis is divided into six chapters. Each chapter, whilst describing experimental results/findings (Chapter 2 to Chapter 5), has its own comprehensive introduction and specified aims. This is primarily an attempt to bridge the distinct background of each comorbid condition as well as to provide a logical rationale in the context of hyperuricaemia and gout. Chapter 2 to Chapter 5 also provides relevant information available in literature, for the different ancestral groups in New Zealand.

The next two chapters (Chapter 2 and Chapter 3) provide the biochemical and genetic association analysis results of hyperuricaemia and gout with abnormal iron homeostasis as a coexisting potentially causal phenomenon. The findings explained in these two chapters mostly are 'first time ever' findings, especially within the New Zealand setting.

Chapter 2 was exclusively designed to provide information about the initial biochemical association analyses for blood iron biomarkers with hyperuricaemia and gout. In addition, the genetic association analysis for two well-known iron overload-related variants was carried out to assess their association with hyperuricaemia and/or gout within the NZ Polynesian and European populations. The positive observational outcomes from this chapter were the basic prompts to extend this work to Chapter 3.

Chapter 3 presents findings of the first-ever Mendelian randomisation analysis done using summary statistics from genome-wide association studies (GWAS) to find a possible causal relationship between serum iron biomarkers and urate concentrations.

Chapter 4 deals with the second comorbid condition, metabolic syndrome. This chapter provides outcomes of genetic association analyses to describe a relationship between several components of the metabolic syndrome, urate and gout. It also provides genetic evidence for an involvement of MetS components in the aetiology of gout for the first time in the New Zealand population.

Chapter 5 provides an extensive ‘rare variant analyses’ for exon sequencing data from the two well-recognised lipid-related genes (*LRP2* and *A1CF*). This study for the first time reports population-specific findings for common variants associated with hyperuricaemia and gout in the New Zealand population. The study also reports for the first time the population-specific rare variant burden of hyperuricaemia for *LRP2.*

Finally, Chapter 6 includes the overall discussion and conclusion of the thesis.

Anywhere in this thesis, unless specified in the footnotes, I performed all wet lab experiments and statistical analyses plus *in silico* data handling for Chapter 3 for MR software.

SECTION 1.2 POPULATION DATASETS USED IN THIS STUDY

Data of the participants collected from different regions of the world, including diverse populations, were used for demographic and clinical information and subsequent analyses. In order to avoid the unnecessary repetition of the text explaining the same details, a comprehensive summary of each dataset is provided here. In general, this summary provides the information about the recruitment criteria, gross total number, ancestral background and ethical approval for each study group. Hence these details are avoided while referring to these data sets in different chapters of this thesis.

1.2.1 Gout Case-Control Cohorts

1.2.1.1 New Zealand Gout Cohort

For the most part of this thesis, demographic and clinical data of subjects within the New Zealand Gout Cohort have been analysed. The New Zealand Gout Cohort is a case-control cohort that was developed for the evaluation of environmental and genetic risk factors for gout and hyperuricaemia in the New Zealand-resident European and Polynesian populations. Currently, the cohort contains data from > 1,800 European and > 2,800 Polynesian individuals.

For this cohort, the recruitment of gout cases was initiated back in 2001 at The University of Auckland, New Zealand, under the supervision of Dr. Lachy McLean. The project supervision was transferred to Professor Tony Merriman (University of Otago, Dunedin, New Zealand) in 2006. Under his overall supervision, the gout cases were initially recruited from two North Island cities, Auckland and Rotorua. The gout case recruitment was then extended to healthcare facilities within four major regions of New Zealand including Auckland, Wellington, Canterbury and Otago. The recruitment of the control group began in 2009, following which data of non-gout individuals were convenience sampled from above four regions of New Zealand. In 2010, the recruitment of a gout case-control sample set began in collaboration with Ngati Porou Hauora (NPH: operates in the *rohe* of Ngati Porou Hauora (tribal territory) located in the East Coast region (Te Tai Rāwhiti) of New Zealand) in collaboration with Ngati Porou Hauora Charitable Trust (NPHCT: www.nph.org.nz). The NPH cohort currently contains the data

of > 450 individuals. The purpose of the study was explained and a written informed consent was obtained from participants for the collection of their blood. In conjunction to this, every participant was asked to manually fill out a questionnaire about their socio-demographic status, family history of gout, grand-parental ancestry information, medical history and certain dietary habits. The American College of Rheumatology (ACR) gout criteria (Wallace *et al.*, 1977) were followed to ascertain the diagnosis of gout. Participants who fulfilled the ACR criteria were further asked to fill out a questionnaire to provide details of particular gout characteristics e.g., use of diuretics or other urate-lowering therapy (ULT), the frequency of acute attacks and presence of tophus/tophi (information about presence of tophi were obtained by physical examination). All gout participants were also asked to name and provide detailed information about the particular foods/drinks that trigger their gout flares. Following the questionnaires, blood and urine samples were collected from each participant and were sent to the Merriman Laboratory (University of Otago, Dunedin, New Zealand) for further processing. At the Merriman Laboratory, a portion of each serum and urine sample was sent to southern Community Laboratory (SCL), Dunedin (www.sclabs.co.nz) for biochemical analysis, and the remainder of the serum samples were aliquoted and were stored at -80°C for future use. Serum urate was measured using the uricase oxidation method (details in Section 2.2), while the DNA extraction was done using a standard chloroform/ethanol extraction protocol (Sambrook and Russell, 2006). All experimental procedures were followed in accordance with the ethical standards of the responsible committee on human experimentation and with the Helsinki Declaration of 1975, as revised in 1983.

The Lower South Ethics Committee (OTA/99/11/098) and the New Zealand Multi-Region Ethics Committee (MEC/05/10/130) granted ethical approval for the recruitment and subsequent study. The ethical approval for the Ngati Porou Hauora Cohort recruitment was granted by the Northern Y Region Committee (NTY/07/07/074) and University of Otago Human Ethics Committee (13/117). The formation and continued recruitment of the New Zealand Gout Cohort were funded by several funding bodies within New Zealand i.e., the New Zealand Health Research Council (www.hrc.govt.nz), University of Otago (www.otago.ac.nz), Arthritis New Zealand (www.arthritis.org.nz) and Lottery Heath Research New Zealand (www.communitymatters.govt.nz).

For the purpose of analysis at several instances in this thesis, the Polynesian data set in New Zealand Gout Cohort was divided into three ancestral sub-groups as described by Hollis-Moffatt *et al.* (2012a): Eastern Polynesian (EP: primarily included Cook Island and NZ Māori), Western Polynesian (WP; primarily included individuals from Samoa, Tonga, Tuvalu, Niue and Tokelau) and mixed Eastern and Western Polynesian (EPWP). Eastern Polynesian participants were further subdivided into two sub-groups: EPN (subjects with high EP ancestry) and EPZ (subjects with low EP ancestry).

1.2.1.2 Additional Gout Cohorts from Europe & Australia

Two additional, relatively smaller, European gout case-only cohorts were recruited from the Merriman Lab's collaborating institutes of Europe and Australia. The participants in these groups were recruited to increase the number of gout cases and also analyse them, treating non-gout samples from NZ Gout Cohort as their respective controls. The first group was recruited in collaboration with European Crystal Network (Eurogout Consortium) (Lioté *et al.*, 2013) and included the data of 827 individuals of European ancestry. The samples in the Eurogout Cohort were mainly recruited at the collection centres of these institutes; 1) University of Edinburgh, Keele University, Queen Elizabeth Hospital and City Hospital in the UK, 2) Universitario de Alicante, Hospital De Cruces, 3) Rijnstate Hospital, Radboud University Medical Centre, University Medical Centre Utrecht in Netherlands and 4) Universitätsklinikum Carl Gustav Carus an der Technischen Universität Dresden in Germany. Ethical review boards of the respective institutes granted ethical approval for the subject recruitment and subsequent analyses for the Eurogout Cohort. The details of ethical approvals are; Ethikkommission, Technische Universität Dresden (EK 8012012), South East Scotland Research Ethics Committee (04/S1102/41), Commission Cantonale (VD) D'éthique de la Recherche sur l'être Humain, Université de Lausanne, Commissie Mensgebonden Onderzoek regio Arnhem—Nijmegen and Partners Health Care System Institutional Review Board. The second group was recruited in collaboration with Arthritis Genomics Recruitment Initiative in Australia (AGRIA) and included data from 215 individuals of European ancestry. The samples in the AGRIA Cohort were mainly recruited at the collection centres of these institutes within Australia; 1) the University of Queensland and Diamantina Institute at the University of Queensland, 2) The University of New South Wales, St. Vincent's Hospital and 3) The University of Adelaide. The details of ethical approvals are:

Research and Ethics Committee, Repatriation General Hospital, South Australia (32/08); Research Ethics Committee, University of New South Wales. All subjects in these two cohorts had gout clinically ascertained using ACR criteria described above (Wallace *et al.*, 1977). A written informed consent was obtained from all subjects in Eurogout and AGRIA Cohorts.

1.2.1.3 UK Biobank Cohort

The UK Biobank is a large, non-commercial, repository that was established in the United Kingdom (UK) in 2007 (www.ukbiobank.ac.uk/). It was established (in Stockport, Greater Manchester, UK) to facilitate the investigation of the contribution of genetic predisposition and environmental exposure to the development of a range of diseases including heart disease, cancer, dementia, diabetes, osteoporosis, arthritis, eye disorders and forms of depression. The UK Biobank is a 25-year follow-up study that includes data from ~500,000 volunteers in the UK. All individuals in the study ranged from 40 to 69 years of age at the time of initial recruitment (2006-2009). The participants were invited to visit different assessment centres based in the UK and were interviewed about baseline demographic variables, medical history, lifestyle and nutritional habits.

For the purpose of this study, the gout case-control genotype and phenotype data were sourced and analysed from the UK Biobank Cohort under approval number 12611 (Ollier *et al.*, 2005).[1] The samples in the UK Biobank were genotyped using the Affymetrix Axiom array for 820,967 markers. The data of the participants were selected for analysis based on the European ancestry and availability of genotype information. Data of individuals that failed to either fulfil the genotype quality control assessment, had self-reported sex mismatch with genetic sex or had hospital diagnosed kidney disease (ICD10 I12, I13, N00-N05, N07, N11, N14, N17-N19, Q61, N25.0, Z49, Z94.0, Z99.2) were excluded from the study. Applying these selection criteria, genotyping data of ~73.3M SNPs were imputed using SHAPEIT3 and IMPUTE2 platforms and UK10K and 1000 Genomes as the combined reference panel.

[1] Professor Tony R Merriman (PhD supervisor) and Professor Nicola Dalbeth (Collaborator: The University of Auckland, New Zealand) supervised the UK Biobank project-associated work and a staff member of the Merriman Laboratory, Murray Cadzow (Assistant Research Fellow), determined the gout definitions and carried out the association analyses in the UK Biobank data.

For the UK Biobank Cohort, the diagnosis of gout was not ascertained following ACR criteria (described above). Therefore, gout status was defined following a combination of definitions (Cadzow *et al.*, 2016) provided in the previous epidemiological literature (Colhoun *et al.*, 2003; Dalbeth *et al.*, 2016; Köttgen *et al.*, 2013). The gout diagnosis criteria were broadly divided into four categories; self-reported, hospital diagnosed, use of ULT and Winnard-defined gout (Winnard *et al.*, 2012). Self-report of gout was defined as 'participant reporting as having gout at the time of interview'. Hospital diagnosed gout was defined as 'having a primary or secondary hospital discharge coding for gout (ICD10 M10 including subcodes). Use of ULT was defined as 'participant reporting as being on any ULT (allopurinol, febuxostat or sulphinpyrazone) and not being diagnosed as having leukemia or lymphoma (ICD10 C81-C96) by the hospital'. Winnard-defined gout was defined as 'having a hospital diagnosis of gout or being on gout-specific medication (any ULT or colchicine)' as described by Winnard *et al.* (2012).

1.2.2 Non-Gout (Control only) Cohorts

1.2.2.1 Jackson Heart Study Cohort

The Jackson Heart study is a population-based longitudinal study based in the City of Jackson, US (United States) state of Mississippi. The Jackson Heart Study was established in 1997 as a partnership among three major institutes in Jackson (Jackson State University, the University of Mississippi Medical Centre and Tougaloo College, Jackson Mississippi) and the National Institutes of Health's National Heart, Lung, and Blood Institute (NHLBI) and Office of Research on Minority Health, while the initial recruitment of the subjects was started in the year 2000. The JHS was designed to identify the risk factors for the cardiovascular disease and its associated manifestations including obesity, hypertension, diabetes, stroke, coronary heart disease and chronic kidney disease in more than 5000 African American individuals from Jackson, Mississippi (Taylor, 2005). The JHS represents the largest single-site, prospective study conducted in African Americans to investigate the inherited (genetic) factors that affect above defined metabolic conditions and to develop potential treatments that do more good and less harm than treatments that are available today. The JHS Cohort currently has data from ~5,300 male and female adults aged 35 to 84 years, from more than 400 families. The

information provided on the JHS official website (www.jacksonheartstudy.org/) indicates 3 phases (Exam 1 to 3) of the subject recruitment (information of year of recruitment was not provided). Publicly-accessible data includes demographic, clinical, dietary and social information of the recruited subjects in three directories (visit 1 to 3). Permission to access the data was granted to Professor Merriman under the dbGaP controlled access agreement (project name: "Genetic Basis of Gout"; project approval #384). The phenotype data for iron profile and baseline health information from Exam 1 phase 1 (visit 1) was used in this thesis for several analyses in Chapter 2 (Section 2.2).

1.2.2.2 Third National Health and Nutrition Examination Survey Cohort

The US Third National Health and Nutrition Examination Survey (NHANESIII) is a nationwide population-based study that was designed to investigate the prevalence and risk factors of multiple common diseases in the US population (www.cdc.gov/nchs/nhanes). The National Centre for Health Statistics (NCHS) and Centres and Disease Control and Prevention (CDC) are the supporting bodies for NHANESIII. The study contains data from 39,695 male and female individuals (at least 2 months of age or older), with the data being considered as illustrative for the general US population. The study started in 1988 and includes data from more than 85 regions within the US. The subjects in the NHANESIII were recruited in two phases. Phase 1 was conducted between the years 1988 to 1991 and collected data from 44 counties, while Phase 2 was conducted between the years 1991 to 1994 and collected the data from 45 different counties. Within the time period of six years, a total of 33,994 participants were interviewed at their home about their demographic status, dietary habits and medical histories followed by a visit for medical examination in a mobile examination centre (examination response rate was 78%). The phenotype data for NHANESIII is publicly available at their official website (www.cdc.gov/nchs/nhanes/nhanes3.htm) and can be downloaded without permission requirements. The phenotype data for iron profile and baseline health information from Phase 1 and 2 (European and African American individuals) were downloaded for various analyses in Chapter 2 (Section 2.2).

1.2.2.3 Atherosclerosis Risk in Community Study Cohort

The Atherosclerosis Risk in Community (ARIC) Study is a population-based, longitudinal study designed to investigate established and new risk factors for

atherosclerosis in the US population (www2.cscc.unc.edu/aric/; Database of Genotype and Phenotype/dbGaP; www.ncbi.nlm.nih.gov/gap accession # phs000280). The study includes data from 15,485 male and female adults mainly from four different communities in the US – Forsyth County, North Carolina; Northwest Minneapolis, Minnesota, Jackson, Mississippi and Washington County, Maryland. The subject recruitment for the study started in 1987. All individuals who agreed to participate in the study were interviewed in Exam 1 between 1987 and 1989 about the baseline demographic and clinical characteristics along with details of food intake. The average age of participants ranged between 45 and 64 years. The study then followed the clinical information of the participants on re-examination every 3-years along with annual detailed telephone interviews. The ARIC study currently has the data for a total of five clinical examinations completed – Exam 1: 1987-1989, Exam 2: 1990-1992, Exam 3: 1993-1995, Exam 4: 1996-1998 and Exam 5: 2011-2013. The study also had whole-genome genotyping performed as part of the GENEVA (Gene-Environment Association Studies) initiative (dbGaP accession #phs000090). The genotyping was performed for ~934,930 SNPs using the Affymetrix 6.0K Chip genotyping platform. Four major funding bodies - NHLBI: #N01-HC-55015, N01 HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, N01-HC-55022, R01HL087641, R01HL59367 and R01HL086694, National Human Genome Research Institute/NHGRI: #U01HG004402, National Institute of Health/NIH: #HHSN268200625226C and University of Carolina, Chapel Hill, collectively support the ARIC study. The permission to access the data was granted to Professor Merriman under the dbGaP controlled access agreement (project name: “Genetic Basis of Gout”; project approval #384). For the purpose of various analyses in this thesis (Chapter 4: Section 4.2), the phenotype and genotype data were obtained from visit 1 (1987-1989) only.

1.2.2.4 Framingham Heart Study Cohort (Offspring & Generation 3)

The Framingham Heart study (FHS), established in 1948, is a population-based longitudinal study aimed to identify and investigate common genetic and environmental risk factors that contribute to cardiovascular disease (www.framinghamheartstudy.org/; dbGaP accession # phs000007). The FHS cohort also includes data for other metabolic conditions including gout, hyperuricaemia, osteoporosis and diabetes. The current data in the FHS cohort were established over a period of time following a large number of

asymptomatic male and female individuals for the development of CVD or related symptoms. The original FHS cohort consisted of 5,209 individuals from Framingham, Massachusetts; all aged 28 to 62 years. Since the recruitment of the Original FHS cohort in 1948, two more cohorts were recruited – the Offspring and Generation 3 cohorts. The Offspring cohort was established in 1971, which essentially included 5,124 offspring of the participants in the 'Original' cohort and their spouses. The Generation 3 cohort was established in 2002 and included 4,148 grandchildren of the 'Original' cohort. Participants in both the Offspring (1971-1975) and Generation 3 (2002-2005) cohorts were asked to give written consent for the study. All interested participants were then interviewed about baseline demographic, social and clinical aspects. The individuals in the FHS cohort were also genotyped as part of the SNP Health Association Resource (SHARe) project (dbGaP accession # phs000342). The genotyping was performed for ~934,500 SNPs using the Affymetrix 500K mapping array. Two funding bodies - NHLBI and Boston University (www.bu.edu), collectively support the FHS. Permission to access the data was granted to Professor Merriman under the dbGaP controlled access agreement (project name: "Genetic Basis of Gout"; project approval #384). For the purpose of various analyses in this thesis (Chapter 4: Section 4.2), the phenotype and genotype data were obtained from examination 1 (Offspring: 1971-1975 and Generation 3: 2002-2005).

1.2.2.5 Cardiovascular Health Study Cohort

The Cardiovascular Health Study (CHS) is a prospective population-based observational study that was established to evaluate the risk factors, development and progression of cardiovascular disease (www.chs-nhlbi.org; dbGaP accession # phs000287). The CHS subject recruitment was started in 1989 with coverage of four major regions; California, Maryland, North Carolina and Pennsylvania. A total of 5,582 participants were interviewed and underwent extensive annual clinical examinations between the years 1989 and 1999. The cohort was divided into two age groups - Senior: aged between 65 and 79 years, and Aged: 80 years or older. The measurements included socio-demographic status, dietary habits and measurements of potential CVD risk factors such as hypertension, complete lipid profiles and the presence of subclinical disease (echocardiography, carotid ultrasound and cranial magnetic resonance imaging/MRI). During the follow-up, participants were contacted via phone calls to update their health status and hospitalisation. The main outcomes recorded in the follow-up were coronary

heart disease, heart failure, angina, heart stroke, transient ischemic attack and mortality. The CHS cohort was genotyped as part of SNP Typing for Association with Multiple Phenotypes from Existing Epidemiologic Data (STAMPEED; dbGaP accession # phs000226) using the Illumina HumanCNV370 duo bead chip as the genotyping platform. The National Institutes of Health's National Heart, Lung, and Blood Institute (NHLBI) is the main funding body for the CHS cohort. Permission to access the data was granted to Professor Merriman under the dbGaP controlled access agreement (project name: "Genetic Basis of Gout"; project approval #384). For the purpose of various analyses in this thesis (Chapter 4: Section 4.2), the phenotype and genotype data were obtained from visit 1 (1989-1991).

1.2.2.6 Coronary Artery Risk Development in Young Adults Study Cohort

The Coronary Artery Risk Development in Young Adults (CARDIA) Study was designed 'to examine the determinants and development of clinical and sub-clinical cardiovascular disease and its risk factors' (www.cardia.dopm.uab.edu; dbGaP accession # phs000285). The study began in 1985 with recruitment of 5,115 adult males and females aged 18-30 years. The participants were selected with equal numbers in sub-groups of race, ethnicity, gender, age and education. The recruitment was done in four centres - Birmingham, Alabama; Minneapolis, Minnesota; Chicago, Illinois and Oakland, California. After the initial interview about demographic and clinical history between 1985 and 1986, all participants were asked to participate in follow-up at year 2 (1987-1988), year 5 (1990-1991), year 7 (1992-1993), year 10 (1995-1996), year 15 (2000-2001) and year 20 (2005-2006). At a re-examination success rate of 72%, the current cohort includes the data completed for 3,622 individuals. The data collected were mainly for risk factors of cardiovascular disease e.g., blood pressure, glucose, cholesterol and other lipids. Baseline characteristics included the measurement of weight, skinfold fat, exercise pattern, dietary habits and alcohol consumption. Genotyping of the CARDIA cohort was also performed as a part of GENEVA initiative (dbGaP accession # phs000309) using the Affymetrix 6.0K genotype platform at Boston Massachusetts. In addition to NHLBI and NIH, the CARDIA study is funded by University of Alabama, Birmingham (www.uab.edu). Permission to access the data was granted to Professor Merriman under the dbGaP controlled access agreement (project name: "Genetic Basis of Gout"; project approval #384). For the purpose of various analyses in this thesis (Chapter

4: Section 4.2), the phenotype and genotype data were obtained from visit 2 of CARDIA cohort.

1.2.3 ReSequencing Cohort

In addition to a subset selected from NZ Gout Cohort for resequencing (details in Chapter 5), samples from two small European cohorts were also selected for resequencing. A description of these cohorts is provided below. The combined resequencing data set including all three subsets is referred to as 'ReSequencing Cohort' in the thesis text.

1.2.3.1 Nurse's Health Study Cohort

The Nurse's Health Study (NHS) is a female-based longitudinal study that was established in 1976 (www.channing.harvard.edu/nhs). The initial aim of the study was to investigate long-term health effects of oral contraceptives in female population of the USA. A total of 121,700 female nurses, aged between 30 and 55 years, were recruited from 11 different states of the USA. All participants were sent a questionnaire via email to record their demographic, medical and lifestyle details with success return rate of ~70%. Every two years, participants in the Nurse's Health Study are mailed a questionnaire to detail their medical history, hormonal medication usage, menopausal status and quality of life. The National Institute of Health (www.nih.gov), Brigham and Women's Hospital, and Harvard School of Public Health and Harvard Medical School (Boston, Massachusetts) are funding bodies for NHS.

1.2.3.2 Health Professionals Follow-up Study Cohort

The Health Professionals Follow-Up Study (HPFS) is a men's health based longitudinal study that began in 1986 (http://www.hsph.harvard.edu.hpfs). The aim of the study was to evaluate the potential link between nutritional factors and men's health via relating diet with incidence of serious illnesses e.g., heart disease, cancer, pulmonary and vascular disease. The HPFS all-male study was established to complement the all-female Nurse's Health Study (NHS; explained above in section 1.2.3.1), which inspects similar hypotheses in a population of females. More than 51,000 males, aged between 40 and 75 years, were recruited from different regions of United States of America (USA). Professions of the males included in the study were recognised as being pharmacists,

dentists, osteopaths, podiatrists, optometrists or veterinarians. All participants are mailed a questionnaire to provide their health, medication and exercise history every two years. The Harvard School of Public Health and the National Cancer Institute (www.cancer.gov) are financial sponsors for HPFS.

Chapter 2

Iron Metabolism & Gout

Biochemical and Genetic Association Study

SECTION 2.1 IRON HOMEOSTASIS, URATE AND GOUT: AN INTRODUCTION

2.1.1 Background

Iron is an essential element of the human body and is vital for an array of metabolic functions, such as oxygen transport and cellular respiration (oxidative phosphorylation). The importance of this metal lies in its ability to act as an electron donor in its ferrous (Fe^{2+}) state and acceptor in its ferric (Fe^{3+}) state. When present in abnormally excess amounts, iron can create potential hazards to the surrounding cells by catalysing the reaction of production of free radicals from hydrogen peroxide. These free radicals can end up in damaging a wide range of cellular structures, and ultimately killing the cells. While our body needs the right amount of iron for several metabolic functions, it also needs to protect its cells from the harm done by free floating iron. Being a trace element, iron is required only in small amounts in the body. This is why the human body is specialised to strictly maintain the normal blood levels of iron within a narrow range of 3 to 4 grams (3.8 g in males and 2.3 g in females) of the total body weight (Leong and Lonnerdal, 2012). For well-nourished people, this range may vary from 4 to 5 grams (Gropper and Smith, 2013).

2.1.2 Forms of dietary iron

Diet acts as the major source of iron for an adult human being. Food-based iron is typically categorised into heme and non-heme forms (Sharp, 2010). The readily absorbable heme is abundantly found in animal-based foods (meat, fish and seafood) as part of hemoproteins, myoglobin and hemoglobin. Studies estimate that dietary heme iron accounts for two-thirds of the average person's total iron stores (Bezwoda *et al.*, 1983; Carpenter and Mahoney, 1992; West and Oates, 2008). Due to its high bioavailability, the absorption of heme iron is 5 to 7 times higher in the gastrointestinal tract than non-heme iron (Björn-Rasmussen *et al.*, 1974; Reizenstein, 1979). The presence of proteolytic enzymes and low pH (which increases iron solubility) in the stomach and small intestine makes an ideal environment for the release of heme from hemoproteins (Collins and Anderson, 2012). Once entered into the enterocytes, iron is released from heme to be transported to the blood. In a meat-eating population, heme iron is estimated to contribute 10 to 15% of the total iron intake. However, it can contribute > 40% of the absorbed iron

due to its simpler form and higher absorption (Carpenter and Mahoney, 1992; Hunt, 2002). Non-heme iron, which comes from both plant and animal-based foods is much less well absorbed than heme iron. Heme iron being preferentially absorbed in the intestine, the amount of iron gained from a non-heme source is totally dependent upon the individual's iron status (Hurrell and Egli, 2010).

2.1.3 Mechanism of iron regulation

Iron homeostasis is a tightly regulated set of biochemical processes. A number of proteins, enzymes, and other cellular structures are involved in maintaining the normal body iron concentration. Figure 2.1 illustrates the series of mechanisms and their components involved in human iron homeostasis. In addition to the regulation of endogenous iron pools, the amount of iron from the diet is regulated in the body at two different levels namely systemic and cellular (Ganz, 2013; Wang and Pantopoulos, 2011).

2.1.3.1 Systemic regulation

The control of iron concentrations at the systemic level is accomplished by its precise uptake, distribution and storage. Duodenal enterocytes work in a controlled fashion to absorb a specific amount of the dietary iron via the villi/brush border (Fuqua *et al.*, 2012). The heme is directly imported inside the enterocytes via the enterocyte's cell membrane protein called divalent metal ion transporter 1/DMT1 (Collins and Anderson, 2012; Courville *et al.*, 2006). Conversely, the non-heme iron needs to be reduced to its ferrous form to act as DMT1 substrate. Two metalloreductases, duodenal cytochrome B/DCYTB and STEAP2 are known to perform this function (McKie, 2008), yet the exact mechanism involved in this reduction process is unknown (Zhang *et al.*, 2012a). In addition, a human cell culture study also presented some evidence of absorption of dietary ferritin in the enterocyte through endocytosis (San Martin *et al.*, 2008). Once the enterocyte uptake of heme or non-heme is completed, it is broken down to release iron (ferrous/Fe^{2+} form) by heme oxygenases (Fuqua *et al.*, 2012; Raffin *et al.*, 1974). Iron has two fates at this stage, either stored inside endogenous ferritin or exported out of the enterocytes into the circulation for its transport to other tissues of the body. A transmembrane protein at the basolateral surface of the enterocytes, ferroportin (FPN1), is the only known mammalian protein that makes this export possible (Anderson and Vulpe, 2009). Ferroportin exports iron in its ferrous form, but before release into the blood, it is

oxidised to its ferric form (only form that can bind the iron transporting protein) by ferroxidases (e.g., hephaestin) (De Domenico *et al.*, 2007). Once it reaches the circulatory blood pool, the transport of iron in the blood is controlled by a glycoprotein named transferrin/TF. Transferrin reversibly binds iron ions absorbed from the duodenum and carries them to the surrounding tissues (Rouault, 2003).

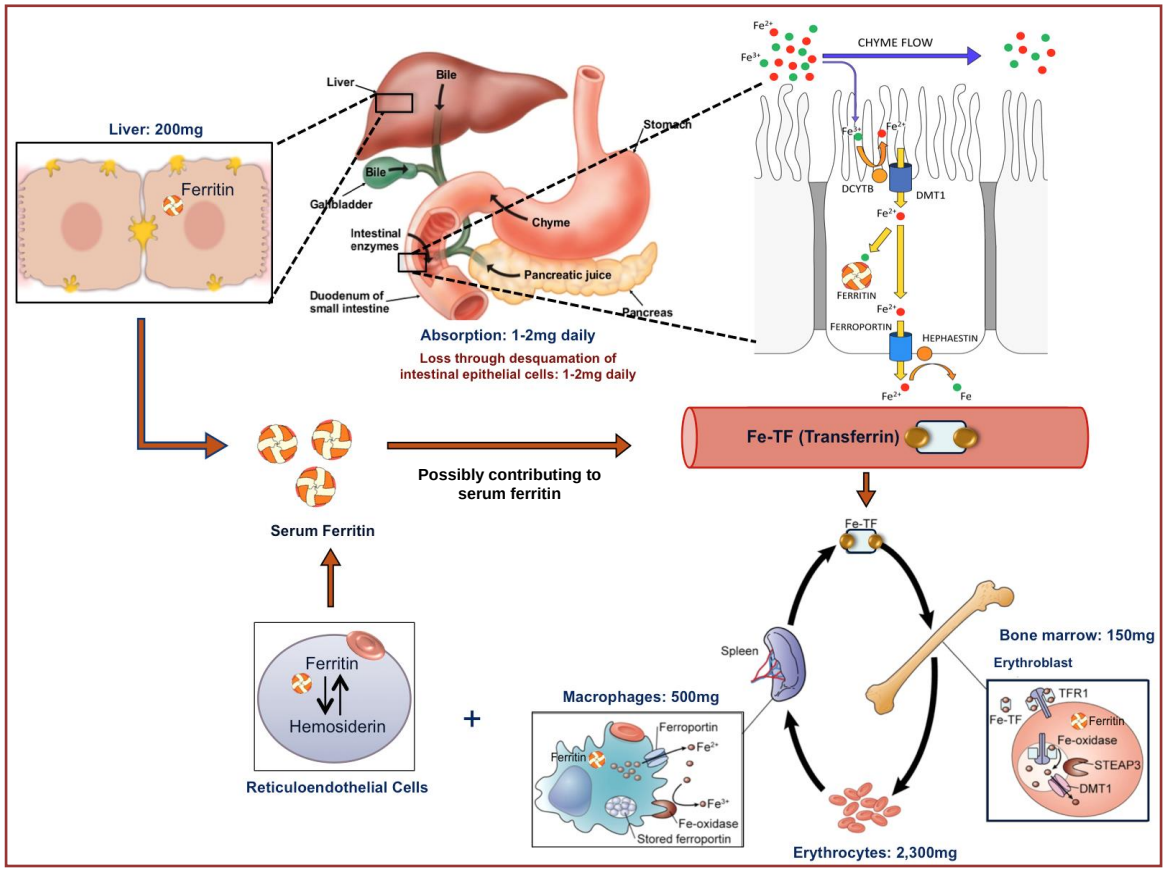


Figure 2.1: Steps involved in maintiaing iron balance in the human body.

The human body does not possess any specialised excretory mechanism for the removal of iron. Thus, it is essential to have tight regulation of iron absorption in order to match the body iron requirements. In conjunction to its absorption, two major mechanisms help to maintain the body iron levels within the normal range i.e., recycling and loss. The reticuloendothelial system, mainly comprised of monocytes and tissue macrophages, efficiently recycles iron through the breakdown of senescent (aged) red blood cells. In addition, a small but steady quantity of iron is lost through faeces, epithelial cell sloughing, sweating and menstrual bleeding in women (Knutson and Wessling-Resnick, 2003). According to a survey, the total amount of the average daily iron loss is 1 mg for men, 1.5 to 2 mg for pre- and 1 mg for postmenopausal women (Hunt *et al.*, 2009). A 25 amino acid peptide hormone ‘hepcidin’ is known for its ability

to inhibit the over-efflux of iron through intestinal enterocytes and reticuloendothelial cells (Knutson, 2010; Nemeth and Ganz, 2009). Hepcidin is a hepatic hormone, which freely circulates in the blood and can bind, internalise and degrade the iron exporter FPN1 to decrease blood iron release (Nemeth *et al.*, 2004).

2.1.3.2 Cellular regulation

Cellular iron levels are controlled differently by different cell types via the expression of particular iron regulatory and transport proteins. The cellular uptake of iron is mainly governed by transferrin receptor 1/TFR1 and transferrin receptor 2/TFR2 mediated import from transferrin into cells (Moos, 2002). In contrast to TFR2, TFR1 possesses 30 times higher affinity for transferrin-bound iron and is known to play the main role in iron uptake at the cellular level (Kawabata *et al.*, 2000; West *et al.*, 2000). Transferrin receptors can only recognise transferrin-bound iron, which means any free floating (unliganded) iron is not transported inside the cells via TFRs. This specific recognition is necessary to cause a conformational change on the cell surface to begin endocytosis of the iron and formation of endosome (Hentze *et al.*, 2010). The DMT1 and ZIP14 (Zrt-Irt-like protein) also allow iron to directly enter the cells via the plasma membrane (Lane *et al.*, 2015). As ferric iron can potentially introduce toxicity to cellular organelles, it needs to be reduced to its ferrous state before entering the cytoplasm. Two components, the DMT1 and STEAP family reductases, perform the functions of reduction and import, respectively (Hentze *et al.*, 2010). The ferrous form of iron imported into the cell makes up a labile iron pool where it stays in a soluble and chelateable state (Yehuda and Mostofsky, 2010).

Like enterocytes, iron is also exported out of other cells (neurons, erythrocytes and macrophages) that ultimately determine the systemic iron levels. The iron exporter FPN1 transports ferrous iron out of the cell (Ganz, 2005), simultaneously assisted with conversion of ferrous to ferric state before releasing it in the cytoplasm (Hentze *et al.*, 2010). Hepcidin performs its function of internalisation and degradation of FPN1 to regulate iron efflux at the cellular level. The exact mechanism being unknown, hepcidin may also downregulate TFR1 and DMT1 (Du *et al.*, 2011).

2.1.4 Iron storage in the body

Due to the high potential for biological toxicity and catalytic activity, iron ions should never stay 'free' or unliganded. This exceptionally important need to sequester iron in a suitable liganded form is fulfilled via 'ferritin'. Ferritin is a hollow globular protein, made of 24 subunits of heavy (FtH1) and light (FtL) chains (Arosio and Levi, 2010). Ferritin is the only protein specialised for the storage of excessive iron in a non-toxic or redox inactive form and its release in a controlled fashion (Harrison *et al.*, 1986). The ability of cells to regulate their iron uptake by modulating the conformation (endocytosis) and expression of the receptors (TFRs) on the cell surface and storage of excess iron as ferritin is one of the major features that tightly maintains iron homeostasis. Synthesis of ferritin mainly takes place in hepatocytes (liver cells), with minor amounts being synthesised by other cells (Anderson and Shah, 2013; Theil, 1987). It has been estimated that ferritin contains about 2 grams of the total iron (out of 3 to 4 grams) in the body, most of which is commonly found in blood regulatory tissues/organs i.e., bone marrow, spleen, liver, duodenum and skeletal muscle (Gropper and Smith, 2013; Saito, 2014). Iron stored in the heptaocytes makes up the primary physiologic iron reserve that can be mobilised to release iron according to the systemic metabolic demands (Anderson and Shah, 2013). Serum ferritin reflects cumulative iron stores in the body and it is measured under regular laboratory practices to determine the blood levels of iron (Saito *et al.*, 2013; Shoden *et al.*, 1953).

Hemosiderin is another iron storage complex, that works alongside ferritin to decrease and increase the iron supply to the cells under iron overload and deficiency, respectively. Mostly referred as 'inactive ferritin', hemosiderin is created by macrophages as an ill-defined complex of denatured ferritin and iron (Fischbach *et al.*, 1971). Any iron contained in hemosiderin has a minimal chance to be delivered to the body tissues. Ferritin has been shown to be actively converted to hemosiderin to protect cells from damage caused by iron overload (Saito and Hayashi, 2015). Also, as a protective mechanism towards iron deficiency, any drastic decrease in iron levels can instantly convert hemosiderin back to ferritin to normalise the total body iron concentration (Saito *et al.*, 2013).

2.1.5 Disorders of iron metabolism

Two disorders of abnormal body iron levels are iron deficiency – too low iron, and iron overload – too high iron. In aggregate, iron deficiency and overload have been reported to affect over 1 billion people around the globe (Hentze *et al.*, 2004). More recent studies indicated haemochromatosis (iron overload) to be more prevalent in men with an occurrence of 0.6% in European populations (Crownover and Covey, 2013). In contrast, a study based on 1995-2011 data collected from several populations (children and adult women only) estimated 800 million children (43% of total) and women (29% of total) to be affected with iron deficiency or anaemia globally (WHO, 2015).

2.1.5.1 Iron deficiency

Iron deficiency occurs when body iron concentrations drastically decrease to a level that could create potential hazards to health and life (CDC, 2006). Malnutrition is the most common cause of iron deficiency worldwide, especially in children and pre-menopausal women (Dlouhy and Outten, 2013; Robert and Xiaole, 2013). When dietary intake and systemic and cellular regulatory mechanisms are not able to compensate sufficiently for iron loss, the body develops a state of iron deprivation over time. The continuous deprivation could clinically manifest as iron deficiency anaemia (Njajou *et al.*, 2006) – a condition characterised by low levels of iron, reduced production of haemoglobin and oxygen supply to the tissues and microcytic erythrocytes. An iron-rich diet may be sufficient to treat mild iron deficiency, while anaemia may only be corrected following an appropriate therapy i.e., oral or parenteral iron intake (Camaschella, 2015; Lopez *et al.*, 2016).

2.1.5.2 Iron overload and haemochromatosis

Iron overload is characterised by accumulation of iron in the body due to genetic or environmental causes (Hider and Kong, 2013). Repeated blood transfusions and excessive intake of iron-rich supplements can lead to iron overload (Barton *et al.*, 2006; Robert and Xiaole, 2013). On the other hand, hereditary haemochromatosis (HHC) occurs due to one or two autosomal recessive mutations in the *HFE* (Human haemochromatosis) gene (Feder *et al.*, 1996) i.e., *rs1800562* (Cys282Tyr: replaces cysteine to tyrosine at amino acid 282) and *rs1799945* (His63Asp: replaces histidine with aspartate at amino

acid 63) (Merryweather-Clarke *et al.*, 1997). Haemochromatosis is considered the most common form of iron overload in Europeans (Merryweather-Clarke *et al.*, 1997; Powell *et al.*, 2016), however it accounts for less than 5% of known impaired iron metabolism conditions (Cherfane *et al.*, 2013). In addition, Pacific Islanders and Asians have been shown to have the highest geometric mean levels of ferritin and TSAT despite having lowest prevalence of C282Y homozygotes (Adams *et al.*, 2005). The presence of haemochromatosis mutations can disrupt a pathway of iron homeostasis and may manifest as reduced hepcidin production and increased intestinal iron absorption in the body (Powell *et al.*, 2016). Either through iron overload or haemochromatosis, too much iron in the blood could eventually overwhelm the storage capacity of the body. This can result in iron-mediated oxidative tissue damage, organ disease and complete organ failure (Kohgo *et al.*, 2008; Powell *et al.*, 2016).

2.1.6 Iron profile, urate and gout – evidence presenting possible correlation

Urate is well known for its physiological role as an antioxidant with an estimated 60% contribution to antioxidant activity in the human body (Nieto *et al.*, 2000). In addition, it is also known to protect body tissues from iron-mediated free radical damage via iron chelation (Davies *et al.*, 1986; Ghio *et al.*, 1994). Iron, in turn, can modulate the activity of xanthine oxidase and thereby the production of urate (Ghio *et al.*, 2002).

A number of studies have provided evidence for imbalanced iron homeostasis in renal and joint diseases e.g., chronic kidney (Macdougall *et al.*, 2016; Zumbrennen-Bullough and Babitt, 2014) and rheumatic disease (Baker and Ghio, 2009; Hachem and El-Zimaity, 2007). Ferritin, iron binding capacity and transferrin saturation have been positively associated with urate in European and African American individuals from the US National Health and Nutrition Examination Survey (NHANES) (Ghio *et al.*, 2005; Mainous *et al.*, 2011). One of these studies also suggested serum urate to be a potential indicator of iron overload (Mainous *et al.*, 2011).

There are a number of possible ways that iron could contribute to gouty inflammation. Iron is able to form complexes with MSU crystals *in vitro*, stimulate oxidative stress through the generation of reactive oxygen species, contribute to granulocyte and complement activation and production of lymphocytes (Ghio *et al.*,

1994). Association of iron with a number of pro-inflammatory activities in animal models (Dabbagh *et al.*, 1992) and a decrease in gouty flares following phlebotomy to attain near iron-deficient levels in hyperuricaemic patients (Facchini, 2003) are also suggestive of a role in gouty arthropathy. Consistent with these observational studies, data from an immune-focused GWAS over 450 NZ Europeans indicated a causal association of a variant in the transferrin receptor (*TFRC: rs1466085*) with an increased risk of gout (Merriman *et al.*, 2015). Individuals carrying the risk allele of this variant also self-reported a higher likelihood of an iron-rich food as a trigger of gout flares (Merriman *et al.*, 2015).

Food, as the basic source of iron for an adult human, provides heme (from animal-based food) and non-heme (from animal and plant-based foods) iron to the body, with dietary heme iron contributing two-thirds of a person's average iron stores (Bezwoda *et al.*, 1983; Carpenter and Mahoney, 1992; West and Oates, 2008). Diet is also a key source of purines. Interestingly, purine-rich foods from an animal-based diet have been associated with increased risk of recurrent gout attacks while purine-rich foods from a plant-based diet did not show a strong correlation (Zhang *et al.*, 2012b). These observations, combined with the other observational and intervention data (Facchini, 2003; Ghio *et al.*, 2005; Mainous *et al.*, 2011), are consistent with an alternative hypothesis that iron in purine-rich foods (red meat for example) could be a causal factor in gout.

This part of the thesis was based on an observational and genetic association study to test for any possible correlation of total iron, transferrin, ferritin, total iron binding capacity (TIBC) and transferrin saturation (TSAT) with serum urate and/or gout in individuals with different ancestries. The specific aims of this study were;

1. To evaluate any difference in the average levels of total iron, transferrin and ferritin between people with and without gout.
2. To replicate the association of serum ferritin with urate and to test for its association with hyperuricaemia in European and African American populations.
3. To test for an association of other blood iron profile markers (total iron, transferrin, TIBC and TSAT) with urate in NZ European and Polynesian datasets.
4. To test for association of serum ferritin with gout and flare frequency in European and NZ Polynesian individuals.

5. To analyse the genetic association of two haemochromatosis variants, *rs1800562* and *rs1799945*, within the *HFE* gene with urate and gout in NZ European and Polynesian populations.

SECTION 2.2 ASSOCIATION OF IRON BIOMARKERS AND HAEMOCHROMATOSIS VARIANTS WITH URATE AND GOUT

2.2.1 Background

As mentioned in Section 2.1, blood levels of total iron, transferrin and ferritin collectively make-up an adult body's iron profile and are measured in daily laboratory practices to assess a person's iron status (Sajeevan *et al.*, 2016). Ferritin directly reflects cumulative iron stores in the body and thereby acts as a reliable surrogate indicator of an iron-related disorder (WHO, 2011). In conjunction to these, the capacity of blood to bind iron with transferrin is calculated as total iron binding capacity (TIBC). Total iron binding capacity mirrors the maximum amount of iron that blood can carry and is an indirect measure of blood transferrin (Yamanishi *et al.*, 2003). Another important measure for iron status is transferrin saturation (TSAT), which indirectly estimates the amount of transferrin available for iron to bind. Calculated as a percentage, TSAT is a ratio between total iron and TIBC that represents the percentage of transferrin's iron-binding sites occupied by iron i.e., a TSAT of 20% means that, 80% sites on the transferrin are still free for iron to bind. Table 2.1 below shows the normal reference range for each of these parameters.

Table 2.1: Standardised normal levels/reference range for iron profile markers

Iron profile variable	Reference range	
	Male	**Female**
Serum iron	65-177 μg dL^{-1} (11.6-31.7 μmol L^{-1})	50-170 μg dL^{-1} (9.0-30.4 μmol L^{-1})
Serum ferritin	20-250 μg L^{-1} (ng mL^{-1})	15-150 μg L^{-1} (ng mL^{-1})
TIBC	250-370 μg dL^{-1} (45-66 μmol L^{-1})	
TSAT	20-50%	15-50%

Note: Reference range for each biomarker is presented as average for European population only. Source: (Fauci, 2008).

As described in section 2.1, elements of the iron profile have been positively associated with serum urate (SU) and present a possible relationship with gout and gout flares. The association of *TFRC: rs1466085* with gout in NZ Europeans somewhat provides the genetic basis to these observations (Merriman *et al.*, 2015). Despite such evidence, no study has yet provided a direct measure of iron markers in the context of their association with urate or gout in NZ populations. This part of my thesis is

specifically based on the hypothesis that body levels of iron contribute to gout risk, either by increasing serum urate or via another independent unknown metabolic route.

This section represents the biochemical and genetic association analyses carried out to investigate the association of iron with gout and urate in NZ European, Polynesian and African American populations. First a preliminary observational analysis was done, based on measuring serum levels of total iron, ferritin, transferrin, TSAT and TIBC and investigating their possible relationship with urate and gout. To support the previously reported findings of the intake of iron-rich food to be a possible factor in gout (Choi *et al.*, 2004b; Öztürk *et al.*, 2013; Williams, 2008), the observational analysis was further extended and food-based data collected from NZ individuals were analysed.

Given that ferritin has been reported as a marker of acute inflammation (Kell and Pretorius, 2014), C-reactive protein (CRP) was included as an adjustor in the regression models to rule out the possibility that any likely association between ferritin and risk of gout and flare frequency could be a consequence of inflammation. C-reactive protein is an acute-phase protein, which increases in inflammation and has previously been demonstrated to elevate in hyperuricaemia (Ruggiero *et al.*, 2006) and gout (Roseff *et al.*, 1987). Also, to rule out the possibility of getting biased results due to inflammation via liver damage, any participant with hepatic disease was excluded from the various analyses.

In addition, genetic association analysis of two haemochromatosis variants, *rs1800562* and *rs1799945*, was also done to test for any possible association with urate and gout in NZ European and Polynesian populations.

2.2.2 Methods

2.2.2.1 Biochemical analysis

2.2.2.1.1 Study participants

Four different datasets were used for this study, comprising two gout case-control sets from NZ and US (The United States of America) and two subsets of non-gout individuals selected from two publicly available cohorts, The Jackson Heat Study (JHS) and The US Third National Health and Nutrition Examination Survey (NHANES III)

(Chapter 1: Section 1.2). Table 2.2 and 2.3 report the demographic and clinical details for these study groups. The NZ data set included 320 European and Polynesian male individuals and was selected as a sub-set of New Zealand Gout Cohort (Chapter 1: Section 1.2). The US data set included 249 male individuals (gout and non-gout), recruited during 2014-2016 from community-based settings in the US[2]. The NZ sample set comprised male NZ European (100 cases and 60 controls) and Polynesian (100 cases and 60 controls) individuals. The data for consumption of iron-rich food (red meat, seafood and fish) were determined as previously reported (Flynn *et al.*, 2015). The US group comprised a mixture of Latino, African Americans and Europeans (189 cases and 60 controls). All US gout cases had crystal proven or clinically diagnosed gout, with no active acute gout at the time of sample draw (no NSAID or colchicine usage within 2 weeks before the sample draw). New Zealand participants without gout (controls) included in this study were convenience sampled from the Auckland, Otago and Canterbury regions of NZ. The US control group comprised sex- and ancestrally-matched volunteers. Individuals with a history of liver damage or disease were excluded from the NZ and US case groups and US control participants were included if they had never been diagnosed with gout and were not currently taking any non-steroidal anti-inflammatory drugs (NSAIDs) or colchicine. The Independent Ethics Committee E6 Good Clinical Practice granted ethical approval for the US samples. The collection was done in accordance with the Declaration of Helsinki (October 2008), and applicable local regulatory requirements (including Institutional Review Board approval). Written informed consent was obtained from all subjects for the collection of samples and subsequent analyses.

Publicly-available data from two larger cohorts, JHS and NHANES III, were used only for serum ferritin versus urate association analyses. These data sets have been detailed in Chapter 1 (Section 1.2). The JHS data included a total of 1,260 African American individuals, while the NHANES III data were comprised 4,355 African American and 5,112 European individuals. All individuals recruited from the JHS and

[2] Drs Jeffrey N Miner and Cory N Iverson (Biology, Ardea Biosciences, Inc., AstraZeneca Group, San Diego, US) kindly provided demographic and clinical information for the US data set.

NHANES III cohorts were at least 21 years of age. Subjects who self-reported as taking any diuretic or other urate-lowering medication, or had kidney disease or gout, or had first-degree relatives with gout were excluded from the serum/plasma urate association analyses. This categorisation was made to assess urate association only in non-gout subjects and to remove any chances of getting biased results due to gout or other potential factors affecting urate concentration and thus differed from the criteria applied to the previous study that used the NHANES III data (Ghio *et al.*, 2005). Additionally, none of the participants used here overlapped with Ghio *et al.* (2005), and the data were analysed separately in the NHANES III European and African American participants. In order to assess the association with hyperuricaemia, both cohorts were stratified on the basis of serum urate levels with a cut-off value of ≥ 0.41mmol L^{-1} between hyperuricaemic cases (serum urate ≥ 0.41 mmol L^{-1}) and normouricaemic controls (serum urate ≤ 0.37mmol L^{-1}).

The iron profile marker data provided for the US group included only ferritin, while the data sourced from JHS and NHANES III included total iron, ferritin, TIBC and TSAT.

Table 2.2: Demographic and clinical information for NZ, US, JHS and NHANES III data sets

Populations	NZ Europeans		NZ Polynesian		US		JHS	NHANES III (EUR)	NHANES III (AA)
Group	**Non-gout**	**Gout**	**Non-gout**	**Gout**	**Non-gout**	**Gout**	**Non-gout**	**Non-gout**	**Non-gout**
Baseline Information									
Total Participants (n)	60	100	60	100	60	193	1260	5112	4355
Males, n (%)	60 (100)	100 (100)	60 (100)	100 (100)	60 (100)	189 (97.92)	567 (45.00)	2460 (48.09)	1925 (44.41)
Age (years)^	53.41 ± 14.95	64.3 ± 11.73	41.98 ± 13.71	48.18 ± 13.42	52.1 ± 6.89	56.55 ± 10.76	47.59 ± 10.82	52.95 ± 19.87	41.90 ± 17.63
BMI (kg/m^2)^	26.30 ± 3.02	30.05 ± 4.46	31.09 ± 5.68	36.32 ± 7.75	-	-	31.16 ± 7.23	26.01 ± 5.05	27.65 ± 6.58
Serum Urate ($mmol\ L^{-1}$)^	0.33 ± 0.07	0.38 ± 0.10	0.32 ± 0.03	0.49 ± 0.10	0.35 ± 0.08	0.38 ± 0.11	0.31 ± 0.08	0.31 ± 0.07	0.32 ± 0.09
CRP ($mg\ dL^{-1}$)^	0.39 ± 0.25	0.59 ± 0.28	0.39 ± 0.26	0.60 ± 0.27	0.61 ± 0.81	1.03 ± 1.44	0.43 ± 0.71	0.41 ± 0.62	0.52 ± 0.82
Iron Profile Measurements									
Serum iron ($\mu g\ dL^{-1}$)^	105.74 ± 30.91	100.62 ± 33.76	90.01 ± 29.07	82.91 ± 27.77	-	-	81.38 ± 31.61	90.21 ± 36.39	80.18 ± 35.62
Serum ferritin ($ng\ mL^{-1}$)^	230.71 ± 182.91	268.42 ± 197.04	323.39 ± 173.22	462.42 ± 245.25	69.45 ± 63.41	206.63 ± 176.13	157.89 ± 156.95	130.62 ± 139.86	144.21 ± 163.74
Serum transferrin ($g\ L^{-1}$)^	2.52 ± 0.39	2.53 ± 0.34	2.62 ± 0.47	2.79 ± 0.57	-	-	-	-	-
TIBC ($\mu mol\ L^{-1}$)^	63.47 ± 9.99	63.65 ± 8.76	65.82 ± 12.01	70.18 ± 14.40	-	-	52.85 ± 8.91	62.70 ± 10.01	62.14 ± 10.54
TSAT (%)^	30.22 ± 9.12	28.8 ± 10.06	24.92 ± 8.05	21.56 ± 7.11	-	-	28.13 ± 11.05	26.37 ± 11.36	23.72 ± 10.81
Gout Characteristics									
No of gout flares/year^	-	3.72 ± 7.02	-	7.73 ± 23.71	-	2.02 ± 1.87	-	-	-
% Allopurinol treatment (% reported)	-	75 (100)	-	75.75 (99)	-	97.92 (84.56)	-	-	-
% Gout tophus (% reported)	-	27 (100)	-	66 (100)	-	NA	-	-	-

NZ: New Zealand, US: The United States of America, JHS: Jackson Heart Study, NHANES III: US Third National Health and Nutrition Examination Survey, EUR: White Caucasian/European, AA: African American, BMI: Body mass index, n (%): total number (percentage), % (% reported): total percentage[#] of individuals reported yes/no for the particular criteria (percentage who reported yes for the particular criteria out of #), CRP: C-reactive protein, TIBC: Total iron binding capacity, TSAT: Transferrin saturation. ^Data are shown as mean ± standard deviation.

Table 2.3: Detailed demographic and clinical information for non-gout individuals from the JHS and NHANES III study groups: stratified on the basis of gender

Populations	**JHS (African American)**		**NHANES III (White Caucasian/European)**		**NHANES III (African American)**	
Group	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**
Baseline Information						
Number (% of total)	567 (45.00)	693 (55.00)	2460 (48.09)	2652 (51.87)	1925 (44.41)	2430 (55.79)
Age (years)^	47.57 ± 10.63	47.61 ± 10.98	54.22 ± 19.54	51.77 ± 20.10	42.04 ± 17.67	41.79 ± 17.61
BMI (kg/m^2)^	29.57 ± 6.03	32.47 ± 7.85	26.34 ± 4.39	25.70 ± 5.57	26.32 ± 5.22	28.71 ± 7.31
Serum Urate ($mmol\ L^{-1}$)^	0.34 ± 0.07	0.26 ± 0.07	0.35 ± 0.07	0.27 ± 0.06	0.36 ± 0.08	0.28 ± 0.08
CRP ($mg\ dL^{-1}$)^	0.29 ± 0.49	0.55 ± 0.82	0.39 ± 0.63	0.42 ± 0.61	0.42 ± 0.67	0.60 ± 0.92
Iron Profile Measurements						
Serum iron ($\mu g\ dL^{-1}$)^	90.22 ± 30.23	74.14 ± 30.87	94.92 ± 35.14	85.83 ± 36.98	89.89 ± 35.11	72.48 ± 34.13
Serum ferritin ($ng\ mL^{-1}$)^	224.47 ± 180.17	103.38 ± 107.95	178.81 ± 152.05	85.97 ± 110.121	204.03 ± 175.96	96.77 ± 135.80
Serum transferrin ($g\ L^{-1}$)^	-	-	-	-	-	-
TIBC ($\mu mol\ L^{-1}$)^	51.02 ± 7.01	54.58 ± 9.32	60.91 ± 9.02	64.37 ± 10.58	60.03 ± 9.24	63.81 ± 11.19
TSAT (%)^	31.96 ± 10.52	24.99 ± 10.48	28.41 ± 11.17	24.49 ± 11.21	27.18 ± 10.57	20.98 ± 10.19

JHS: Jackson Heart Study, NHANES III: US Third National Health and Nutrition Examination Survey, EUR: White Caucasian/European, AA: African American, BMI: Body mass index, CRP: C-reactive protein, TIBC: Total iron binding capacity, TSAT: Transferrin saturation. ^Data are shown as mean ± standard deviation.

2.2.2.1.2 Biochemical measurements

The following paragraphs detail the methodologies adopted for the determination of biochemical markers (serum urate, total iron, transferrin, ferritin and CRP) for this study.

2.2.2.1.2.1 Plasma/serum urate measurements

Serum urate measurement in NZ subjects was carried out following uricase oxidation method. The endpoint measurement was done using fully automated Roche Cobas 8000 Modular P/D analyser and reagents provided by the manufacturers. The median coefficient of variation for this machine is 2.1% and the within-sample measurement correlation 99.6%. The test is based on an enzymatic colorimetric principle that eliminates interference intrinsic to chemical oxidation. Uricase was added to the serum samples as the specific reaction enzyme. This enzyme specifically catalyses the oxidation of urate to produce hydrogen peroxide, allantoin, and carbon dioxide. The amount of hydrogen peroxide produced in the uricase reaction is then used for the coupling of sulphonated dichlorophenol and 4-aminoantipyine to produce a colour. The difference in absorbance was measured at 515nm before and after addition of uricase, where the difference is directly proportional to the amount of urate present in the sample.

For US participants, urate was measured by Liquid chromatography-tandem mass spectrometry (LC-MS/MS) in a replicate plasma biomarker sample aliquot; hence urate is referred to as PU (plasma urate) in this data set. To measure PU levels, LC-MS/MS method was developed at Seventh Wave Laboratories (Missouri, USA). Separation was achieved on a Synergi Polar-RP 80A (4.6 x 50 mm, 4 µm) column with mobile phases of 0.1% formic acid in water and 0.1% formic acid in acetonitrile. Quantification was done using a multiple reaction-monitoring mode to monitor the precursor-to-product ion transitions of mass to charge ratio or m/z 167.0 to m/z 124.0 for urate and m/z 169 to m/z 125 for 1,3-(15) N urate in negative ionisation mode. The calibration curve was established over the range of 10-250 µg mL^{-1}, and the correlation coefficient was 0.999. The accuracy determined at eight concentrations ranged between 92.7 and 107%.

2.2.2.1.2.2 Measurement of iron profile markers

For NZ participants, serum iron profile markers were measured using Roche Cobas® systems (Roche Diagnostics GmbH D-68298 Mannheim, Germany). Standard laboratory protocols provided by Cobas® systems e602, c702 and c701/702 were followed to measure total iron, ferritin and transferrin, respectively[3].

The Roche ferritin assay module was based on an *in vitro* electrochemiluminescence immunoassay system. The assay consisted of two rounds of incubation followed by aspiration and calibration. A total of 10μL of each sample was incubated with monoclonal mouse antibodies – M-4.184 and M-3.170 to build up a sandwich complex. The plates were incubated for 18 minutes at room temperature. After addition of streptavidin-coated micro-particles followed by the second round of incubation, the complex became bound to the solid phase via interaction of biotin and streptavidin. The reaction mixture was aspirated into the measuring cell where the micro particles were magnetically captured onto the surface of the electrode. Unbound substances were then removed with ProCell/ProCell M. Chemiluminescent emission was then induced through voltage electrode and measured by a photomultiplier. Results were determined via a calibration curve (instrument specifically generated by 2-point calibration and a master curve provided via the reagent barcode).

In vitro tests for the quantitative determination of iron and transferrin in serum samples were done using a Roche Cobas c system (c701 and 702). The assay for determination of iron was based on colorimetric and photometric analysis and was performed by fully automated chemistry analysers (IRON2: ACN 8661). In the first step, serum detergent or acids were added to liberate Fe^{3+} from transferrin-Fe-complex, leaving apo-transferrin and Fe^{3+} as reaction products. Detergent was then used to further clarify any lipemic samples. The acid added was ascorbic acid, to provide ascorbate, which reduced the released Fe^{3+} ions to Fe^{2+} ions to react with FerroZine to form a coloured complex. The colour intensity was measured photometrically at 570nm using the Roche chemistry analyser. The colour intensity was directly proportional to the iron concentration in the serum sample. The measurement of transferrin was done using

[3] Vivienne Trethowen (Clinical Trials Assistant at Southern Community Laboratory, Dunedin Hospital, New Zealand) carried out the iron profile measurements for the New Zealand sample set.

module cobas c701 (TRSF2: ACN 8187). This test employed a 2-point end immnoturbidimetric assay i.e., human transferrin formed precipitates with a specific antiserum, which was then determined turbidimetrically at 700nm. The assay was performed in fully automated Roche/Hitachi chemistry analyser that calculates the concentration of the analyte in each sample on a pre-set protocol.

Total iron binding capacity (TIBC) and transferrin saturation (TSAT) were calculated using the following standardised formulas;

$$\text{TIBC } (\mu\text{mol L}^{-1}) = 25.1 \times \text{Transferrin } (\text{g L}^{-1})$$

$$\text{TSAT } (\%) = \text{Iron} / \text{TIBC} \times 100$$

Where '25.1' is the theoretical ratio of TIBC (in μmol L^{-1}) to transferrin (in g L^{-1})

2.2.2.1.2.3 Measurement of C-reactive protein

C-reactive protein levels in NZ subjects were measured by me in the Merriman Laboratory, using commercially available CRP human enzyme-linked immunosorbent assay (ELISA) kits from Abcam (ab99995 from R&D Systems, USA), according to manufacturer's instructions. This assay employs an antibody specific for human CRP coated on a 96-well plate. Standards and samples were pipetted into the wells to allow the binding of CRP to the immobilised antibody. The wells were then washed and biotinylated anti-Human CRP antibody was added. After washing away unbound biotinylated antibody, HRP-conjugated streptavidin was pipetted into the wells. The wells were again washed, followed by adding a 3,3',5,5'-Tetramethylbenzidine (TMB) substrate solution to each well. The intensity of colour developed after adding TMB was directly proportional to the bound CRP. Then a stop solution was added that changed the colour from blue to yellow. The intensity of the colour was measured at 450nm using a Thermo-Labsystems multiskan FC absorbance plate reader.

The Myriad Rules Based Medicine Human Multi-Analyte Profile system (Myriad RBM, Inc., Austin, Texas; https://myriadrbm.com/) was used to measure plasma ferritin (assession number P02794 and P02792) and CRP (assession number: P02714) in the US subjects. The samples were spun and transferred to a master microtiter plate. Using automated pipetting, an aliquot of each sample was added to individual microsphere

multiplexes of the selected Multi-Analyte Profile and blocker. This mixture was thoroughly mixed and incubated at room temperature for 1 hour. Multiplexed cocktails of biotinylated reporter antibodies were added robotically and after thorough mixing incubated for an additional hour at room temperature. Multiplexes were labelled using an excess of streptavidin-phycoerythrin solution, thoroughly mixed and incubated for one hour at room temperature. The volume of each multiplexed reaction was reduced by vacuum filtration and washed three times. After the final wash, the volume was increased by addition of buffer for analysis using a Luminex instrument and the resulting data interpreted using proprietary software developed by Myriad RBM. For each multiplex reaction, both calibrators and controls were included on each microtiter plate. Eight-point calibrators to form a standard curve were run in the first and last column of each plate and controls at three concentration levels were run in duplicate. The standard curve, control, and sample QC were performed to ensure proper assay performance. Study sample values for each of the analytes were determined using four and five parameter logistics, weighted and non-weighted curve fitting algorithms included in the data analysis package.

The methodologies followed for urate (Carpenter *et al.*, 2004), iron profile (Li *et al.*, 2015) and CRP (Fox *et al.*, 2008) measurements for the JHS are described elsewhere while Gunter (1996) provide details for the laboratory measurement protocols in the NHANES III cohort.

2.2.2.2 *HFE* Genetic association analysis

In order to validate the findings from the above observational analyses, two variants within haemochromatosis (*HFE*) gene, *rs1799945* and *rs1800562*, have been analysed for their association with urate and gout in New Zealand population.

2.2.2.2.1 Study participants

Participants from the NZ Gout Cohort (Chapter 1: Section 1.2), for whom genotype information was available, were included in these analyses. The NZ Gout Cohort was categorised into two major ancestral groups, NZ Polynesian (Māori and Pacific Islanders: n = 2,017; 941 cases and 1,076 controls) and NZ European (n = 1,421; 862 cases and 559 controls). In order to keep consistency with above analyses, the NZ Polynesian group was not divided further into any sub-groups. Data from the Ngati Porou

Hauora (NPH) individuals were also included as a separate Māori sample set (NPH: 144 cases and 65 controls). Gender-based stratification was done to analyse the associations separately in males and females. Data for non-gout or control subjects (from NZ Gout Cohort only) were used to test for an association of the two *HFE* variants (*rs1800562* and *rs1799945*) with serum urate. Of these, all subjects who self-reported as taking diuretic medication, or had renal failure, gout, or had first-degree relatives with gout were excluded from the various analyses. Table 2.4 provide demographic details of the study groups.

2.2.2.2.2 Genotyping

Genotype data for the two *HFE* gene variants, *rs1800562* and *rs1799945*, were sourced from the Illumina Chip CoreExome dataset. The genotyping in this dataset was done using the Illumina Infinium CoreExome (version 24) bead chip platform. Genotyping was performed at the University of Queensland (Centre for Clinical Genomics), followed by auto-clustering using GenomeStudio version 2011.1 software (Illumina, San Diego). The Illumina GenomeStudio best practice guidelines (Illumina, 2014) and quality control protocols of Guo *et al.* (2014) were followed to ensure that final genotype calls obtained from the auto-clustered genotypes were of highest possible quality.

2.2.2.2.3 Statistical analysis

Multiply-adjusted logistic and linear regression analyses were done using statistical software R version 3.3.2 (RCore, 2016) to test for an association of serum iron, transferrin, ferritin, TIBC and TSAT (explanatory variables) with gout and hyperuricaemia (binary response variables) and urate (continuous response variable), respectively. The regression model describes per unit change in response variable via per unit change in explanatory variable. Adjusted regression models were also used to assess the association of *rs1800562* and *rs1799945* (explanatory variables) with gout and serum urate, where again the regression model describes per unit change in response variables via addition of each affected allele (as per unit change in explanatory variables). The adjusted odds ratio (OR) and β-estimates were obtained by including age (in years), sex and BMI (wherever possible) as baseline covariates in the regression models. Beside these, two additional adjustors added in regression models were serum levels of CRP for

serum ferritin analysis, and number of self-reported Polynesian grandparents for the Polynesian sample set. Adherence to Hardy-Weinberg equilibrium (HWE) was calculated using the SHEsis package (Yong and Lin, 2005), with a significant deviation from HWE if $P_{HWE} \leq 0.016$ (0.05 divided by 3 – the number of data sets tested in Table 2.12). To increase the power of analysis, all ORs for logistic and all β-estimates for linear regression models were combined together in meta-analysis using meta package in R (version 4.3-2). A Q-statistic was calculated to measure the heterogeneity between the datasets. If the heterogeneity was significant ($P < 0.05$), the fixed-effect model was replaced with a random-effect model. Differences in the means for inter- and intragroup comparisons were calculated using an unpaired t-test in R version 3.3.2.

Table 2.4: Demographic and clinical information for study groups used for *HFE* variant analyses

Population	NZ Polynesian		NZ European		Ngati Porou Hauora	
Group	**Gout**	**Non-gout**	**Gout**	**Non-gout**	**Gout**	**Non-gout**
Baseline Information						
Total number (n)	941	1,076	862	559	144	65
Males, n (%)	804 (85.44)	555 (51.57)	729 (84.57)	416 (74.42)	122 (84.72)	34 (52.31)
Age (years)^	52.06 ± 13.14	43.15 ± 15.09	63.86 ± 12.83	54.32 ± 16.71	59.00 ± 11.46	46.30 ± 14.40
BMI (kg/m^2)^	36.07 ± 7.65	33.04 ± 7.59	30.29 ± 5.33	27.63 ± 5.61	36.54 ± 7.54	29.69 ± 5.69
Serum Urate ($mmol\ L^{-1}$)^	0.43 ± 0.11	0.37 ± 0.09	0.39 ± 0.11	0.34 ± 0.11	0.41 ± 0.10	0.36 ± 0.08
Co-morbidities						
Type 2 Diabetes, n (%)	198 (23.77)	138 (14.36)	123 (0.15)	31 (0.07)	42 (0.29)	5 (7.94)
Dyslipidaemia, n (%)	390 (46.81)	121 (12.59)	382 (0.47)	89 (0.19)	80 (0.56)	19 (30.16)
Heart problems, n (%)	244 (29.29)	102 (10.61)	300 (0.37)	69 (0.15)	46 (0.32)	6 (9.52)
Hypertension, n (%)	455 (54.62)	211 (21.95)	439 (0.54)	105 (0.23)	105 (0.73)	11 (17.46)
Liver problems, n (%)	6 (0.72)	3 (0.13)	17 (0.02)	7 (0.01)	2 (0.013)	0 (0.00)
Kidney problems, n (%)	173 (20.76)	33 (3.44)	180 (0.22)	23 (0.05)	15 (0.10)	2 (3.17)
Gout Characteristics						
Age at onset gout (years)^	38.06 ± 7.66	-	48.81 ± 16.87	-	39.23 ± 14.56	-
On Allopurinol n (%)	712 (85.47)	-	623 (0.769)	-	109 (0.762)	-
No of gout flares/year^	10.65 ± 33.95	-	7.86 ± 34.98	-	3.09 ± 4.67	-
Gout tophus, n (%)	338 (40.72)	-	280 (34.56)	-	126 (87.50)	-

NZ: New Zealand, BMI; Body mass index, n (%): total number (percentage), ULT: Urate lowering therapy. ^Data are shown as mean ± standard deviation. Data for the co-morbidities are self-reported.

2.2.3 Results

2.2.3.1 Biochemical analysis

Analyses of data from all four groups indicated only ferritin to have consistent significant associations with binary and continuous variables under study. The paragraphs below mostly provide association results for plasma/serum ferritin, while results for all other iron profile markers are provided in Appendix A, Table 2.1 through 2.4.

Serum ferritin concentration was positively associated with serum urate concentration in non-gout African American individuals from the JHS (Males: ß (mmol L^{-1}) = 0.04, $P_{ß}$ = 9.3E-03; Females: ß (mmol L^{-1}) = 0.08, $P_{ß}$ = 1.7E-03) and NHANES III (Males: ß (mmol L^{-1}) = 0.03, $P_{ß}$ = 5.1E-04; Females: ß (mmol L^{-1}) = 0.09, $P_{ß}$ = 1.76E-15) studies as well as Europeans (Males: ß (mmol L^{-1}) = 0.02, $P_{ß}$ = 6.6E-03; Females: ß (mmol L^{-1}) = 0.07, $P_{ß}$ = 3.4E-10) from the NHANES III study (Table 2.5). A positive association between ferritin and urate was also observed in NZ Polynesian (ß (mmol L^{-1}) = 0.09, $P_{ß}$ = 2.5E-04), but not NZ European (ß (mmol L^{-1}) = 0.05, $P_{ß}$ = 0.31) or US individuals (ß (mmol L^{-1}) = 0.24, $P_{ß}$ = 0.15) (Table 2.5). Combining all datasets in meta-analysis indicated a positive association between serum ferritin and urate (ß (mmol L^{-1}) = 0.05, P_{Het} = 0.27, $P_{ß}$ = 3.9E-32) (Table 2.5). Among other markers, serum iron, TIBC and TSAT were positively associated with urate in NHANES III Europeans (Iron: ß (mmol L^{-1}) = 0.16, $P_{ß}$ = 3.3E-10; TIBC: ß (mmol L^{-1}) = 0.31, $P_{ß}$ = 0.001; TSAT: ß (mmol L^{-1}) = 0.35, $P_{ß}$ = 1.1E-05). For NHANES III African Americans, positive associations were observed between serum iron, TSAT and urate (Iron: ß (mmol L^{-1}) = 0.16, $P_{ß}$ = 1.5E-06; TSAT: ß (mmol L^{-1}) = 0.54, $P_{ß}$ = 8.7E-04) (Appendix A, Table 2.1).

Analyses comparing normouricaemic with hyperuricaemic individuals also indicated association of ferritin with an increased risk of hyperuricaemia in African Americans from the JHS (Males: OR (95% CI) =1.012 (1.01 ; 1.02), P_{OR} = 1.4E-02; Females: OR (95% CI) = 1.040 (1.01 ; 1.07), P_{OR} = 3.3E-03) and NHANES III (Males: OR (95% CI) = 1.013 (1.01 ; 1.02), P_{OR} = 2.4E-06; Females: OR (95% CI) = 1.022 (1.01 ; 1.03), P_{OR} = 7.1E-06) studies (Table 2.6). Odds ratios represent change in serum urate per 10 ng mL^{-1} increase in concentration of serum ferritin. A similar association was observed when the normouricaemia and hyperuricaemia comparison was done for

NHANES III Europeans (Males: OR (95% CI) = 1.011 (1.01 ; 1.02), P_{OR} = 6.8E-04; Females: OR (95% CI) = 1.022 (1.01 ; 1.03), P_{OR} = 1.7E-08) (Table 2.6). Combining the JHS and NHANES III datasets together in meta-analysis showed a significant association of serum ferritin with an increased risk of hyperuricaemia (OR (95% CI) = 1.013 (1.01 ; 1.02), P_{Het} = 0.15, P_{OR} = 6.9E-23) (Table 2.6).

Average levels of ferritin were significantly elevated in both NZ Polynesian (P = 2.3E-04) and US (P = 2.4E-17) gout cases compared to controls. The values were not significantly different between NZ European gout case-controls (P = 0.21) (Appendix A, Table 2.4 and Figure 2.2). Serum ferritin was associated with an increased risk of developing gout in NZ Polynesian (OR (95% CI) = 1.032 (1.01 ; 1.05), P_{OR} = 1.8E-03) and US individuals (OR (95% CI) = 1.112 (1.06 ; 1.17), P_{OR} = 7.4E-06) (Table 2.7). An increase of 10 ng mL^{-1} of ferritin was associated with an increased risk of gout of 3% in NZ Polynesian and 11% in the US male individuals. Ferritin was not associated with an increased risk of gout in European male participants from NZ (OR (95% CI) = 0.99 (0.97 ; 1.02), P_{OR} = 0.83) that possibly introduced heterogeneity in meta-analysis. However, meta-analysis still indicated a direction of susceptibility to gout (OR (95% CI) = 1.039 (0.99 ; 1.08), P_{Het} = 0.003, P_{OR} = 0.09) (Table 2.7).

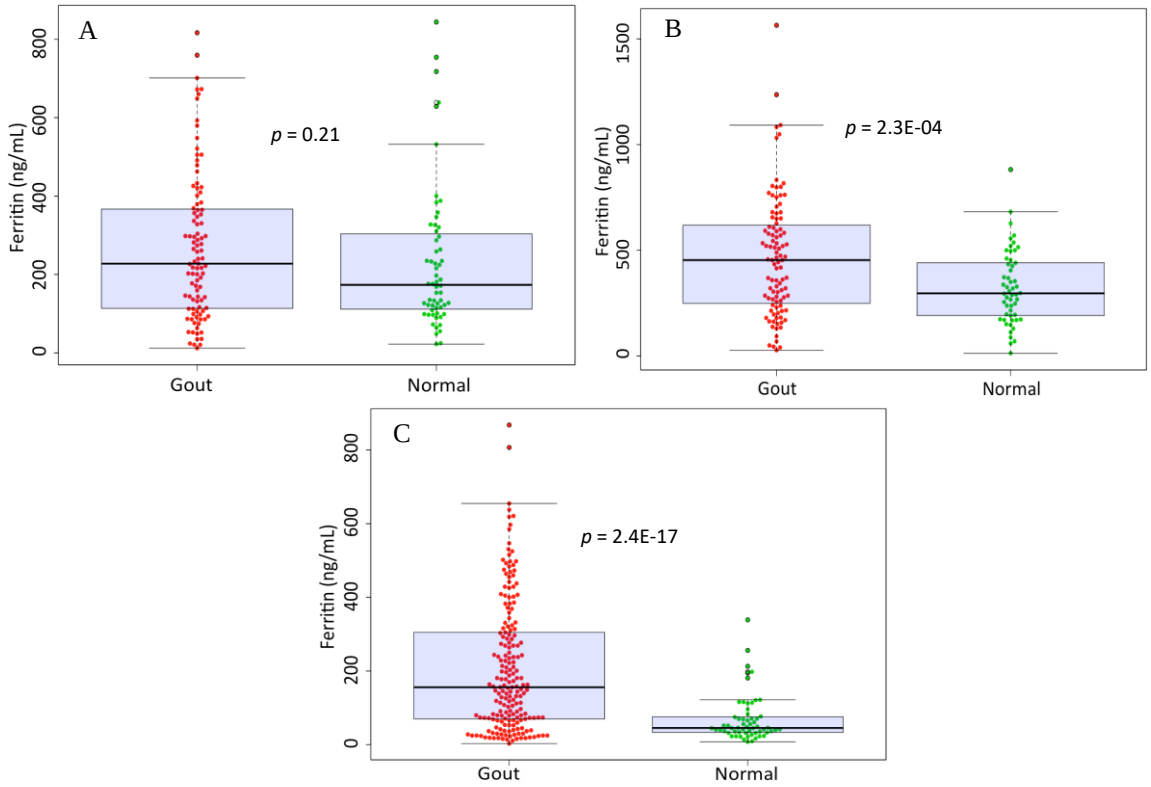


Figure 2.2: Distribution of serum ferritin and difference in means between gout case-control groups (males only) in (A) NZ European, (B) NZ Polynesian and (C) US datasets. *P*-values are given for a difference between the means.

Table 2.5: Association of ferritin (ng mL^{-1}) with serum urate (µmol L^{-1})

Population	n	ß [95% CI]	*P*	ß [95% CI]*	*P**
NZ European	60	0.062 [-0.041 ; 0.165]	0.24	0.055 [-0.054 ; 0.165]	0.31
NZ Polynesian	60	0.092 [0.045 ; 0.139]	1E-04	0.086 [0.040 ; 0.132]	2.5E-04
US	60	0.289 [-0.034 ; 0.612]	0.079	0.245 [-0.093 ; 0.584]	0.15
JHS (Males)	567	0.047 [0.011 ; 0.083]	9.5E-03	0.046 [0.011 ; 0.081]	9.3E-03
JHS (Females)	693	0.138 [0.087 ; 0.189]	1.6E-07	0.081 [0.030 ; 0.131]	1.7E-03
JHS (Combined)	1,260	0.163 [0.137 ; 0.196]	1.8E-27	0.059 [0.031 ; 0.087]	3.2E-05
NHANES III AA (Males)	1,925	0.061 [0.041 ; 0.082]	7.2E-09	0.035 [0.015 ; 0.056]	5.1E-04
NHANES III AA (Females)	2,430	0.167 [0.143 ; 0.191]	1.3E-42	0.096 [0.072 ; 0.121]	1.8E-15
NHANES III AA (Combined)	4,355	0.173 [0.158 ; 0.189]	1.9E-97	0.064 [0.049 ; 0.079]	2.1E-16
NHANES III EUR (Males)	2,460	0.036 [0.018 ; 0.055]	1.0E-04	0.024 [0.006 ; 0.042]	6.6E-03
NHANES III EUR (Females)	2,652	0.114 [0.091 ; 0.137]	1.3E-22	0.071 [0.048 ; 0.092]	3.4E-10
NHANES III EUR (Combined)	5,112	0.152 [0.136 ; 0.167]	6.9E-84	0.045 [0.032 ; 0.059]	5.1E-11
Meta-analysis (All)	10,907	0.149 [0.125 ; 0.173]	9.3E-34	0.056 [0.046 ; 0.065]	3.9E-32

NZ: New Zealand, US: The United States of America, JHS: Jackson Heart Study, NHANES III: US Third National Health and Nutrition Examination Survey, EUR: White Caucasian/European, AA: African American, n; number of non-gout individuals included in serum urate analysis. *Adjusted for age, sex, BMI, C-reactive protein and number of self-reported Polynesian grandparents for the NZ Polynesian group.

Table 2.6: Association of serum ferritin with hyperuricaemia

Population	OR [95% CI]	*P*	OR [95% CI]*	*P**
JHS (Males)	1.012 [1.010 ; 1.022]	1.7E-02	1.012 [1.010 ; 1.023]	1.4E-02
JHS (Females)	1.046 [1.021 ; 1.071]	1.7E-04	1.040 [1.011 ; 1.067]	3.3E-03
JHS (Combined)	1.031 [1.022 ; 1.041]	8.7E-11	1.021 [1.011 ; 1.026]	1.1E-03
NHANES III AA (Males)	1.022 [1.017 ; 1.027]	1.5E-17	1.013 [1.007 ; 1.018]	2.4E-06
NHANES III AA (Females)	1.042 [1.032 ; 1.051]	1.2E-20	1.022 [1.012 ; 1.032]	7.1E-06
NHANES III AA (Combined)	1.036 [1.032 ; 1.041]	1.3E-58	1.016 [1.011 ; 1.021]	9E-12
NHANES III EUR (Males)	1.013 [1.010 ; 1.017]	1.8E-13	1.011 [1.002 ; 1.012]	6.8E-04
NHANES III EUR (Females)	1.036 [1.029 ; 1.043]	3.1E-25	1.022 [1.012 ; 1.025]	1.7E-08
NHANES III EUR (Combined)	1.030 [1.026 ; 1.032]	1.2E-68	1.011 [1.007 ; 1.014]	4.2E-11
Meta-analysis (All)	1.032 [1.027 ; 1.037]	4.3E-35	1.013 [1.013 ; 1.016]	6.9E-23

JHS: Jackson Heart Study, NHANES III: US Third National Health and Nutrition Examination Survey, EUR: White Caucasian/European, AA: African American, n: number of non-gout individuals included in serum urate analysis, OR: Odds ratio, 95% CI: 95% confidence interval, *P*; *p*-values. *Adjusted for age, sex, BMI, C-reactive protein and number of self-reported Polynesian grandparents for the NZ Polynesian group. All values represent change in risk for every 10 ng mL^{-1} increase in serum ferritin.

Table 2.7: Association of serum ferritin with gout

Population	OR [95% CI]	*P*	OR [95% CI]*	*P**
NZ European	1.011 [0.993 ; 1.031]	0.22	0.997 [0.971 ; 1.023]	0.84
NZ Polynesian	1.026 [1.011 ; 1.045]	1.7E-03	1.032 [1.013 ; 1.055]	1.8E-03
US	1.115 [1.072 ; 1.171]	1.3E-06	1.112 [1.066 ; 1.170]	7.4E-06
Meta-analysis (All)	1.043 [1.014 ; 1.081]	0.02	1.039 [0.993 ; 1.088]	0.09

NZ: New Zealand, US: The United States of America, OR: Odds ratio, 95% CI: 95% confidence interval, *P*: *p*-values. *Adjusted for age, sex, BMI, C-reactive protein and number of self-reported Polynesian grandparents for the NZ Polynesian group. All values represent change in risk for every 10 ng mL^{-1} increase in serum ferritin.

As ferritin was positively associated with gout (in NZ Polynesian and US datasets), the analysis was extended to assess any possible association of serum ferritin, CRP and urate with self-reported data of frequency of gout flares (per year) in the NZ Polynesian, NZ European and the US sample sets. It was observed that each 10 ng mL^{-1} increase in ferritin was significantly associated with an increased frequency of gout flares in the US (ß (flares/year) = 0.02, $P_{ß}$ = 2.0E-03) and NZ European (ß (flares/year) = 0.09, $P_{ß}$ = 0.04) but not in NZ Polynesian (ß (flares/year) = -0.11, $P_{ß}$ = 0.14) individuals. C-reactive protein and urate concentrations were not significantly associated with for the self-reported number of gout flares in all three data sets (Table 2.8).

The urate-producing enzyme xanthine oxidase also releases iron from ferritin (Bolann and Ulvik, 1987). As allopurinol is widely used to treat gout as an inhibitor of this enzyme, the average levels of ferritin (and other iron measures) were also compared for NZ and US gout cases by stratifying the data sets according to the usage of allopurinol. Significantly increased serum ferritin levels were observed in the NZ Polynesian (P = 0.005) and the US (P = 0.02) participants who reported taking allopurinol for management of gout in comparison to those who didn't report allopurinol usage for gout treatment. The average levels of ferritin were not different in NZ European when the same stratification was done (P = 0.47) (Table 2.9). Assuming that allopurinol could be a possible confounder for an association of ferritin with gout and flare frequency, additional analyses were done excluding the individuals not taking allopurinol and adjusting for allopurinol usage for gout and flare frequency, respectively. The results obtained, however, did not indicate allopurinol as a confounder (Appendix A Table 2.7 and 2.8).

A significant difference (P = 0.04) between the mean levels of ferritin was found in NZ Polynesian gout participants when participants who self-reported iron-rich food as a trigger for gout attacks were compared to those who did not report any trigger. The results were however not significant in NZ Europeans (P = 0.78). Also, no significant difference was observed in average levels of total iron, transferrin, TIBC and TSAT when the same stratification criterion was applied (Table 2.10).

Table 2.8: Association of serum ferritin, CRP and urate with gout flares/year

Population	Ferritin (ng mL^{-1})†				C-reactive protein (mg dL^{-1})				Serum urate (mg dL^{-1})			
	ß [95% CI]	*P*	ß [95% CI]*	*P**	ß [95% CI]	*P*	ß [95% CI]*	*P**	ß [95% CI]	*P*	ß [95% CI]*	*P**
NZ European	0.08 [0.001 ; 0.16]	0.04	0.09 [0.003 ; 0.17]	0.042	1.55 [-3.86 ; 6.97]	0.57	1.75 [-3.85 ; 7.37]	0.54	-0.61 [-1.47 ; 0.23]	0.16	-0.61 [-1.48 ; 0.25]	0.16
NZ Polynesian	-0.11 [-0.24 to 0.03]	0.13	-0.11 [-0.24 ; 0.03]	0.14	-11.17 [-24.47 ; 2.11]	0.09	-9.95 [-23.75 ; 3.84]	0.16	-1.79 [-4.04 ; 0.44]	0.12	-1.59 [-3.88 ; 0.71]	0.17
US	0.02 [0.01 ; 0.041]	0.003	0.02 [0.01 ; 0.04]	0.002	0.07 [-0.17 ; 0.32]	0.58	0.01 [-0.01 ; 0.03]	0.39	0.07 [-0.09 ; 0.23]	0.39	0.13 [-0.03 ; 0.29]	0.12

NZ: New Zealand, US: The United States of America, ß: ß-estimates, 95% CI: 95% confidence interval, *P*: *p*-values. *Adjusted for age, sex, BMI and number of self-reported Polynesian grandparents in the NZ Polynesian group. †Values represent change in the annual frequency of flares for every 10ng mL^{-1} increase in serum ferritin. Levels of urate are as per recorded at the time of subject recruitment.

Table 2.9: Self-reported allopurinol use and iron profile comparison in gout patients

Population/Marker	n	Mean/n		SE	*P*	95%CI (for difference)
		NoAllop	Allop			
NZ European						
Total number	89	23	66	-	-	-
Serum Iron (μmol L^{-1})	-	100.44	98.64	6.35	0.77	[-10.66 ; 14.25]
Serum Transferrin (g L^{-1})	-	2.53	2.61	0.08	0.40	[-0.22 ; 0.09]
Serum Ferritin (ng mL^{-1})	-	253.71	284.31	43.46	0.47	[-115.78 ; 54.59]
TIBC (μmol L^{-1})	-	63.63	65.29	2.01	0.41	[-5.58 ; 2.26]
TSAT (%)	-	28.99	27.57	2.29	0.53	[-3.07 ; 5.91]
NZ Polynesian						
Total number	96	22	74	-	-	-
Serum Iron (μmol L^{-1})	-	91.21	85.08	5.46	0.26	[-4.59 ; 16.82]
Serum Transferrin (g L^{-1})	-	2.61	2.76	0.09	0.07	[-0.34 ; 0.01]
Serum Ferritin (ng mL^{-1})	-	373.81	494.32	42.83	0.005	[-204.46 ; -36.54]
TIBC (μmol L^{-1})	-	65.33	69.51	2.31	0.07	[-8.72 ; 0.35]
TSAT (%)	-	22.9	22.67	1.91	0.90	[-3.50 ; 3.96]
US						
Total number	189	29	160	-	-	-
Serum Ferritin (ng mL^{-1})	-	158.57	248.21	38.36	0.02	[-164.81 ; 14.45]

NZ: New Zealand, US: The United States of America, n: number of individuals, NoAllop: Not taking allopurinol, Allop: taking allopurinol, 95% CI: 95% confidence interval, *P*: *p*-values, TIBC: Total iron binding capacity, TSAT: Transferrin saturation.

Table 2.10: Comparison of average values of iron markers in iron-rich food trigger data for NZ European and Polynesian gout cases

Marker	Mean		SE	*P*	95% CI (for difference)
	Trigger	No trigger			
NZ European					
Serum Iron (μg dL^{-1})	92.81	103.29	6.079	0.08	[-1.439 ; 22.392]
Serum Transferrin (g L^{-1})	2.52	2.64	0.075	0.11	[-0.027 ; 0.268]
Serum Ferritin (ng mL^{-1})	283.45	270.87	39.241	0.78	[-121.21 ; 32.613]
TIBC (μmol L^{-1})	66.43	63.41	1.900	0.11	[-0.701 ; 6.748]
TSAT (%)	29.31	26.68	2.024	0.19	[-1.341 ; 6.593]
NZ Polynesian					
Serum Iron (μg dL^{-1})	81.82	90.41	6.446	0.18	[-4.04 ; 21.23]
Serum Transferrin (g L^{-1})	2.75	2.81	0.117	0.59	[-0.17 ; 0.29]
Serum Ferritin (ng mL^{-1})	528.26	402.27	62.153	0.04	[-247.81 ; -4.17]
TIBC (μmol L^{-1})	69.02	70.63	0.819	0.59	[4.31 ; 7.52]
TSAT (%)	21.99	23.31	1.607	0.41	[-1.83 ; 4.47]

NZ: New Zealand, US: The United States of America, n: number of individuals, 95% CI: 95% confidence interval, *P*: *p*-values, TIBC: Total iron binding capacity, TSAT: Transferrin saturation. Trigger: data from the gout subjects reporting any iron rich food (seafood, fish, red meat) as a trigger of gout flares, No trigger: data from gout subjects who didn't report any food as a trigger of gouty flares.

2.2.3.1.1 Analysis with log transformed ferritin

In contrast to the iron and transferrin, the distribution of the data for ferritin were not normal (linear) for all study groups (Appendix B Figure 2.1 and Figure 2.2). In order to validate the above significant findings with ferritin, the data were log transformed (Appendix B Figure 2.3 and Figure 2.4) to obtain a linear distribution and to test if the association for urate, hyperuricaemia, gout and frequency of gout flares still persisted. Results for the transformed (normalised) data are presented in Appendix A Table 2.5 through Table 2.8. Analyses of these data followed a similar pattern of associations as for the non-transformed data.

2.2.3.2 *HFE* Genetic association analysis

The two SNPs in the *HFE* gene were not in LD with each other in the European population (Figure 2.3). It was not possible to calculate LD in Chinese Han individuals, as the 1000 Genome database (www.browser.1000genomes.org) did not have any information available for *rs1800562* for this population. The minor allele (A allele) frequency for *rs1800562* was higher in Europeans (0.079) than Māori and Pacific Islanders (NZ Polynesian = 0.015, NPH = 0.021) (Table 2.12). Similarly, the minor allele (G allele) frequency for *rs1799945* was higher in NZ European (0.152) than Māori and Pacific Island (NZ Polynesian = 0.033, NPH = 0.048) individuals (Table 2.12).

When tested for an association, the minor allele (A) of *rs1800562* did not show any association with serum urate in NZ European (Males: ß (mmol L^{-1}) = 0.012, $P_{ß}$ = 0.40; Females: ß = -0.017, $P_{ß}$ = 0.31), NZ Polynesian (Males: ß = 0.004, $P_{ß}$ = 0.86; Females: ß (mmol L^{-1}) = 0.011, $P_{ß}$ = 0.69) or NPH (Males: ß (mmol L^{-1}) = 0.008, $P_{ß}$ = 0.91; Females: ß (mmol L^{-1}) = -0.075, $P_{ß}$ = 0.40) individuals (Table 2.11). Analysing males and female together also did not show any association of the A allele with serum urate in any population dataset (NZ European: ß (mmol L^{-1}) = 0.002, $P_{ß}$ = 0.83; NZ Polynesian: ß (mmol L^{-1}) = 0.002, $P_{ß}$ = 0.69; NPH: ß (mmol L^{-1}) = -0.019, $P_{ß}$ = 0.73) (Table 2.11). When combined in meta-analysis, the only difference was the direction of association between males and females, with males having positive (ß (mmol L^{-1}) = 0.010 mmol L^{-1}, $P_{ß}$ = 0.41, P_{Het} = 0.002) and females having negative (ß (mmol L^{-1}) = -0.012, $P_{ß}$ = 0.43, P_{Het} = 0.42) direction without any significant effect (Table 2.11). The meta-analysis also indicated significant heterogeneity (P_{Het} = 0.002) between male groups from

the NZ European, NZ Polynesian and NPH datasets (Table 2.11). Combining male and female groups together also did not indicate any association between the A allele of *rs1800562* and urate (ß (mmol L^{-1}) = 0.003, $P_{ß}$ = 0.75, P_{Het} = 0.89) (Table 2.11).

The allele A of *rs1800562* did not show any association with gout when tested in three data sets i.e., NZ European (Males: OR = 0.95, P_{OR} = 0.79; Females: OR = 1.06, P_{OR} = 0.87), NZ Polynesian (Males: OR = 1.09, P_{OR} = 0.80; Females: OR = 1.39, P_{OR} = 0.61) and NPH (Males: OR = 0.63, P_{OR} = 0.56; Females: OR = 1.64, P_{OR} = 0.82) (Table 2.13). The non-significant associations remained consistent for all population datasets when assessed for all individuals together (NZ European: OR = 1.01, P_{OR} = 0.98; NZ Polynesian: OR = 1.13, P_{OR} = 0.70; NPH: OR = 0.53, P_{OR} = 0.49) (Table 2.13). The A allele also did not show any association when sex-specific groups were combined separately (Males: OR = 0.96, P_{OR} = 0.82, P_{Het} = 0.81; Females: OR = 1.17, P_{OR} = 0.65, P_{Het} = 0.93) and together (Male: OR = 1.09, P_{OR} = 0.93, P_{Het} = 0.74) in meta-analysis (Table 2.13).

The G allele of *rs1799945* did not show any association with serum urate in NZ European males (ß (mmol L^{-1}) = 0.007, $P_{ß}$ = 0.51), while it indicated a nominal negative association in females (ß (mmol L^{-1}) = -0.028, $P_{ß}$ = 0.05) (Table 2.11). The G allele was also negatively associated with urate in NZ Polynesian males (ß (mmol L^{-1}) = -0.050, $P_{ß}$ = 0.02) with no association in females (ß (mmol L^{-1}) = 0.016, $P_{ß}$ = 0.97) (Table 2.11). However, for the NPH data set a positive association was observed between G allele and urate for males (ß (mmol L^{-1}) = 0.085, $P_{ß}$ = 0.02) but not females (ß (mmolL^{-1}) = 0.015, $P_{ß}$ = 0.82) (Table 2.11). The G allele, however, did not show any association with urate when all individuals were analysed together in different data sets (NZ European: ß (mmol L^{-1}) = -0.003, $P_{ß}$ = 0.75; NZ Polynesian: ß (mmol L^{-1}) = -0.021, $P_{ß}$ = 0.12; NPH: ß (mmol L^{-1}) = 0.044, $P_{ß}$ = 0.16) (Table 2.11). Combined meta-analysis did not show any significant association of the SNP *rs1799945* (G allele) with urate in both males (ß (mmol L^{-1}) = 0.008, $P_{ß}$ = 0.78, P_{Het} = 0.96) and females (ß (mmol L^{-1}) = -0.015, $P_{ß}$ = 0.16, P_{Het} = 0.52) and all individuals together (ß (mmol L^{-1}) = -0.005, $P_{ß}$ = 0.45, P_{Het} = 0.14) (Table 2.11).

The G allele of *rs1799945* indicated a positive association with the risk of gout in NZ European females, where addition of each allele (G) was associated with a 28% increase in the risk of developing gout (OR = 2.28, P_{OR} = 0.03) (Table 2.13). The

association was not significant for NZ European males (OR = 0.89, P_{OR} = 0.47), NZ Polynesian males (OR = 1.03, P_{OR} = 0.89) and females (OR = 1.24, P_{OR} = 0.61) and NPH males (OR = 0.62, P_{OR} = 0.62) (Table 2.13). The number of female individuals was not enough to test for an association of *rs1799945* with gout in the NPH data set (Table 2.13). However, the association did not remain when male and female individuals were combined together for different population data sets (NZ European: OR = 1.04, P_{OR} = 0.78; NZ Polynesian: OR = 1.07, P_{OR} = 0.74; NPH: OR = 0.75, P_{OR} = 0.70) (Table 2.13). Meta-analysis showed a trend towards susceptible association between the G allele and gout in females (OR = 1.71, P_{OR} = 0.06, P_{Het} = 0.29) but not males (OR = 0.96, P_{OR} = 0.82, P_{Het} = 0.81) (Table 2.13). However, combining sex-specific groups together in meta-analysis did not show any association between the G allele of *rs1799945* (OR = 1.03, P_{OR} = 0.73, P_{Het} = 0.90) (Table 2.13).

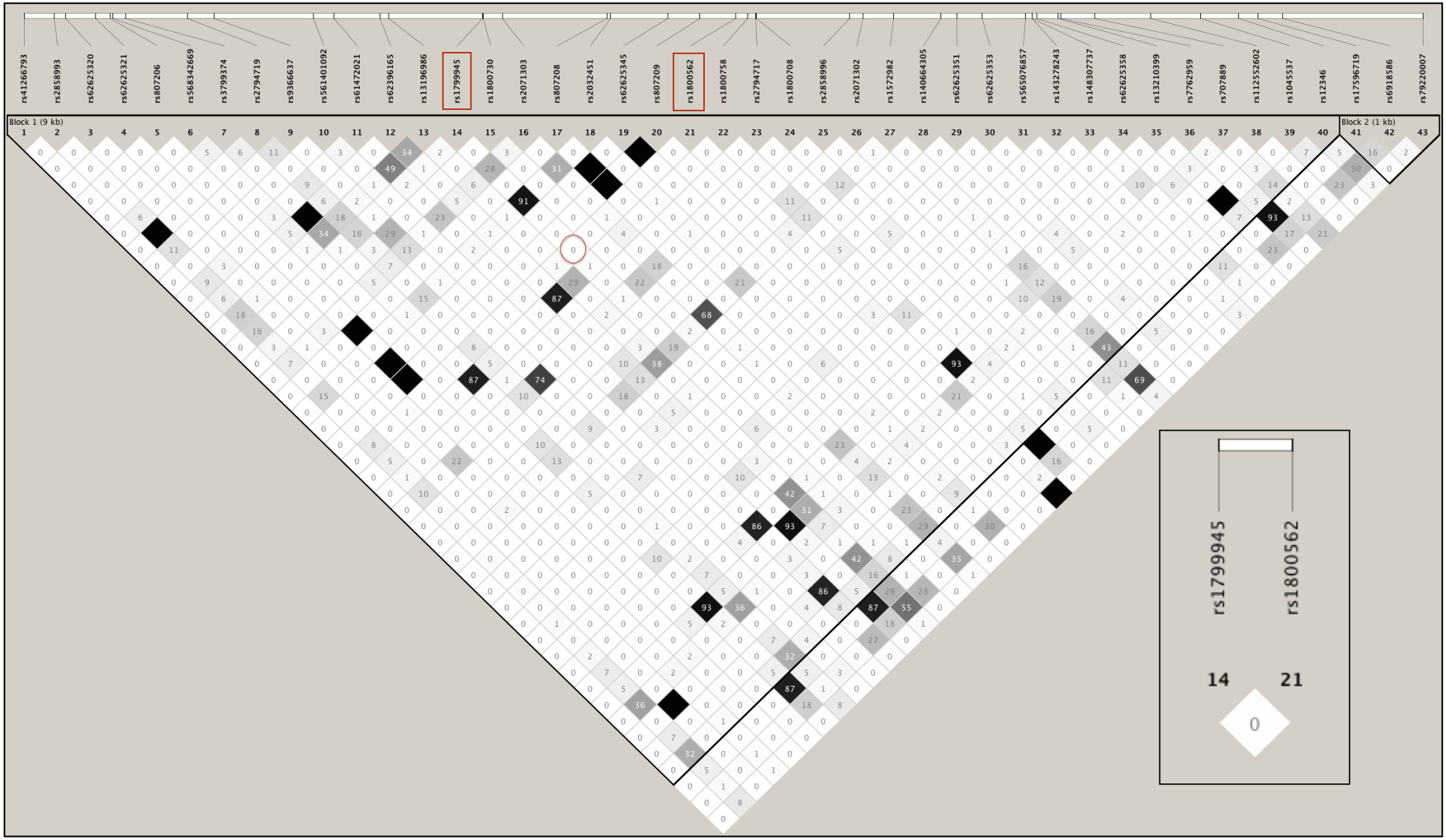


Figure 2.3: Linkage disequilibrium (LD) plot indicating 'R-squared/r^2' values between the *HFE* variants (*rs1800562* and *rs1799945* highlighted in the legend on the right side) in European population. Information for variant location, rs ID and LD values are from 1000 Genome database (http://browser.1000genomes.org/). The plot was generated using Haploview v4.2.

Table 2.11: Association analysis results for the *HFE* variants for serum urate (mmol L^{-1})

Population	ß [95% CI]†	*P*	ß [95% CI]*	*P**
		rs1800562		
NZ European (Males)	-0.008 [-0.041 ; 0.024]	0.59	0.012 [-0.016 ; 0.041]	0.40
NZ European (Females)	0.005 [-0.038 ; 0.050]	0.79	-0.017 [-0.051 ; 0.017]	0.31
NZ European (All)	0.001 [-0.024 ; 0.029]	0.95	0.002 [-0.019 ; 0.024]	0.83
NZ Polynesian (Males)	0.001 [-0.065 ; 0.064]	0.99	0.004 [-0.146 ; 0.164]	0.86
NZ Polynesian (Females)	0.027 [-0.041 ; 0.065]	0.65	0.011 [-0.043 ; 0.065]	0.69
NZ Polynesian (All)	0.007 [-0.036 ; 0.051]	0.74	0.007 [-0.013 ; 0.045]	0.69
NPH (Males)	0.012 [-0.153 ; 0.176]	0.88	0.008 [-0.146 ; 0.164]	0.91
NPH (Females)	-0.056 [-0.225 ; 0.113]	0.49	-0.075 [-0.266 ; 0.114]	0.40
NPH (All)	-0.023 [-0.142 ; 0.096]	0.70	-0.019 [-0.132 ; 0.094]	0.73
Meta-analysis (Males)	-	-	0.010 [-0.014 ; 0.034]	0.41
Meta-analysis (Females)	-	-	-0.012 [-0.037; 0.017]	0.43
Meta-analysis (All)	-	-	0.003 [-0.016 ; 0.022]	0.75
		rs1799945		
NZ European (Males)	0.013 [-0.012 ; 0.037]	0.30	0.007 [-0.014 ; 0.028]	0.51
NZ European (Females)	-0.015 [-0.048 ; 0.018]	0.36	-0.028 [-0.058 ; 0.001]	0.05
NZ European (All)	0.011 [-0.010 ; 0.032]	0.30	-0.003 [-0.019 ; 0.014]	0.75
NZ Polynesian (Males)	-0.059 [-0.107 ; -0.011]	0.01	-0.050 [-0.092 ; -0.008]	0.02
NZ Polynesian (Females)	0.004 [-0.033 ; 0.042]	0.83	0.016 [-0.131 ; 0.162]	0.97
NZ Polynesian (All)	-0.025 [-0.057 ; 0.005]	0.11	-0.021 [-0.047 ; 0.006]	0.12
NPH (Males)	0.019 [-0.049 ; 0.088]	0.56	0.085 [0.012 ; 0.158]	0.02
NPH (Females)	-0.010 [-0.133 ; 0.113]	0.87	0.015 [-0.131 ; 0.162]	0.82
NPH (All)	0.018 [-0.042 ; 0.080]	0.54	0.044 [-0.019 ; 0.108]	0.16
Meta-analysis (Males)	-	-	0.008 [-0.049 ; 0.065]	0.78
Meta-analysis (Females)	-	-	-0.015 [-0.037 ; 0.006]	0.16
Meta-analysis (All)	-	-	-0.005 [-0.019 ; 0.008]	0.45

NZ: New Zealand, NPH: Ngati Porou Hauora, ß: ß-estimates, 95% CI: 95% confidence interval, *P*: *p*-value. * All values are adjusted for age, BMI and grand-parental ancestry for Polynesian datasets.

Table 2.12: Genotype/allele frequencies of the *HFE* variants in gout case/control groups

Population	Group	Genotype/Allele frequency				P_{HWE}
		rs1800562				
		GG	AG	AA	A	
NZ European	Case	731 (0.848)	125 (0.145)	6 (0.070)	137 (0.079)	0.79
	Control	478 (0.855)	78 (0.139)	3 (0.054)	84 (0.075)	0.92
NZ Polynesian	Case	912 (0.969)	28 (0.029)	1 (0.001)	30 (0.015)	0.11
	Control	1047 (0.973)	29 (0.027)	0 (0.000)	29 (0.013)	0.65
NPH	Case	138 (0.958)	6 (0.042)	0 (0.000)	6 (0.021)	0.79
	Control	62 (0.954)	3 (0.046)	0 (0.000)	3 (0.023)	0.84
		rs1799945				
		CC	CG	GG	G	
NZ European	Case	613 (0.711)	235 (0.273)	14 (0.016)	263 (0.152)	0.11
	Control	406 (0.727)	142 (0.254)	10 (0.179)	162 (0.145)	0.54
NZ Polynesian	Case	879 (0.931)	59 (0.063)	2 (0.002)	63 (0.033)	0.34
	Control	1008 (0.936)	66 (0.061)	2 (0.002)	70 (0.032)	0.40
NPH	Case	131 (0.909)	12 (0.083)	1 (0.007)	14 (0.048)	0.23
	Control	60 (0.923)	4 (0.061)	1 (0.015)	6 (0.046)	0.05

NZ: New Zealand, NPH: Ngati Porou Hauora, P_{HWE}: *p*-value for Hardy Weinberg Equilibrium.

Table 2.13: Association analysis results for the *HFE* variants for gout

Population	Males				Females				Combined (Males and Females)			
	OR [95% CI]†	*P*†	OR [95% CI]*	*P**	OR [95% CI]†	*P*†	OR [95% CI]*	*P**	OR [95% CI]†	*P*†	OR [95% CI]*	*P**
HFE: rs1800562												
NZ European	0.89 [0.64 ; 1.25]	0.51	0.95 [0.65 ; 1.40]	0.79	1.69 [0.97 ; 2.99]	0.06	1.06 [0.04 ; 2.53]	0.87	1.06 [0.80 ; 1.41]	0.67	1.01 [0.72 ; 1.40]	0.98
NZ Polynesian	1.22 [0.64 ; 2.43]	0.53	1.09 [0.53 ; 2.31]	0.80	0.72 [0.17 ; 2.07]	0.58	1.39 [0.32 ; 4.65]	0.61	1.18 [0.71 ; 1.97]	0.52	1.13 [0.59 ; 2.13]	0.7
NPH	0.77 [0.16 ; 5.47]	0.76	0.63 [0.15 ; 3.61]	0.56	0.73 [0.03 ; 8.13]	0.81	1.64 [0.02 ; 1.06]	0.82	0.72 [0.23 ; 4.36]	0.88	0.53 [0.09 ; 3.74]	0.49
Meta-analysis	-	-	0.96 [0.69 ; 1.33]	0.82	-	-	1.17 [0.58 ; 2.35]	0.65	-	-	1.09 [0.75 ; 1.35]	0.93
HFE: rs1799945												
NZ European	0.92 [0.72 ; 1.18]	0.53	0.89 [0.67 ; 1.20]	0.47	1.36 [0.86 ; 2.16]	0.18	2.28 [1.04 ; 5.05]	0.03	1.11 [0.85 ; 1.32]	0.58	1.04 [0.79 ; 1.35]	0.78
NZ Polynesian	0.96 [0.62 ; 1.48]	0.84	1.03 [0.63 ; 1.71]	0.89	0.99 [0.47 ; 1.92]	0.99	1.24 [0.51 ; 2.47]	0.61	1.03 [0.73 ; 1.44]	0.86	1.07 [0.69 ; 1.64]	0.74
NPH	0.77 [0.17 ; 5.47]	0.76	0.62 [0.10 ; 5.34]	0.62	-	-	-	-	1.04 [0.44 ; 2.85]	0.92	0.75 [0.20 ; 3.73]	0.7
	-	-	0.92 [0.72 ; 1.18]	0.54	-	-	1.71 [0.97 ; 3.02]	0.06	-	-	1.03 [0.82 ; 1.29]	0.73

NZ: New Zealand, NPH: Ngati Porou Hauora, OR: Odds ratio, 95% CI: 95% confidence interval, *P*: *p*-values. †Unadjusted analysis, *All values are adjusted for age, BMI and grand-parental ancestry for Polynesian datasets. The combined male and female analysis is additionally adjusted for sex.

SECTION 2.3 DISCUSSION

The study successfully replicated the previously reported association of serum ferritin and TSAT with serum urate (Ghio *et al.*, 2005) in two NHANES III sample sets (Table 2.5). Additionally, including CRP as an adjustor ensured that the positive relationship with ferritin was not the consequence of the association of ferritin with inflammatory states (e.g., hyperuricaemia and gout). The association of serum ferritin with urate was extended from Europeans and African Americans to Polynesians (Table 2.5). For the first time, the study associated increased serum ferritin with the risk of gout and gout flares, although for both relationships association was observed in only two of the three data sets used (Table 2.7 and 2.8). Collectively the data associates ferritin with both serum urate levels and the risk of gout, including frequency of flares once gout is established.

The inconsistent association of ferritin with gout and gout flares between the NZ (ß (flares/year) = 0.09, P = 0.04) and US (ß (flares/year) = 0.02, P = 2E-03) European and NZ Polynesian (ß (flares/year) = -0.11, P = 0.14) data sets could be due to one of the several possible reasons. It is possible that the data represent a false positive association. Noting that the sample sets used here are relatively small, analysis of larger well-phenotyped sample sets to replicate the associations is required. Other factors contributing to inconsistent data could be genetic backgrounds of the populations, additive or non-additive influence from other environmental factors and population-specific dietary exposures. As far as usage of allopurinol is concerned, it was not found to be a potential confounder for positive association of ferritin with gout or gout flare frequency in my analyses (Appendix A Table 2.7 and 2.8). Urate was not found to be associated with frequency of gout flares in the literature that agrees with the findings in Table 2.8. An increase in ACTH in response to inflammation in gout and its consequent effect to increase urate excretion has been described as a possible mechanism behind this association (Bădulescu *et al.*, 2013).

Serum ferritin concentration has been demonstrated to correlate with body iron stores and reflect total metal accumulation in humans (Jacobs *et al.*, 1972; Lipschitz *et al.*, 1974) and reports of a possible direct relationship of iron and/or ferritin with urate and its related arthropathies are not recent (Green and Mazur, 1957; Mazur *et al.*, 1958; Muirden and Senator, 1968). It is known that metal ions can potentially cause oxidative stress

when bound to storage or transport proteins (Aust *et al.*, 1985) and urate is a well-known antioxidant in humans governing its antioxidant abilities by scavenging oxygen radicals, singlet oxygen and oxo-haem oxidants (Ames *et al.*, 1981; Howell and Wyngaarden, 1960; Kellogg and Fridovich, 1977). Thus it has been proposed that urate could reduce iron-catalysed oxidative stress by acting as a metal chelator by forming a highly stable 2:1 complex with Fe^{3+} ions, which doesn't support electron transport for iron oxidation (Davies *et al.*, 1986). Xanthine oxidase acts as the sole enzymatic source of urate in humans and exposure to increased iron may control the oxidation of hypoxanthine or xanthine by enhancing the activity of xanthine oxidase (Ghio *et al.*, 2002; Martelin *et al.*, 2002; Stonehuerner *et al.*, 1998). An increase in urate in both rodents (Muntane *et al.*, 1993; Ward *et al.*, 1993) and humans (Livrea *et al.*, 1996) in response to acute exposure to iron supports a direct link between levels of iron and urate, and elevated serum urate is sometimes used as a cue to screen for haemochromatosis (Mainous *et al.*, 2011).

Although hyperuricaemia is an essential requirement for gout, there are other factors involved in the formation and deposition of MSU crystals in the synovial fluid (Chhana *et al.*, 2015). There may be overlap between factors that reduce urate solubility (promoting MSU crystal formation) and those that are either associated with elevation of iron levels or iron-associated oxidative stress. One such factor is lower levels of albumin at a lower temperature (Kippen *et al.*, 1974), with one study demonstrating a negative relation between iron-induced acute oxidative stress and serum albumin level in a cohort of patients with chronic renal failure (Sezer *et al.*, 2007). The concentration of several cations (Füredi-Milhofer *et al.*, 1987; McNabb and McNabb, 1980) and a certain pH (7-8) also reduces urate solubility (Wilcox *et al.*, 1972); iron stays as a cation in its free toxic state and the presence of significant concentrations of this ionisable iron in the crystals from human tophus at physiological pH (Ghio *et al.*, 1994) suggests a connection between iron and MSU crystal formation.

Another important finding of the study is a positive correlation between the number of gout flares and ferritin in both NZ European and US gout subjects (Table 2.8). Iron deposits have been reported to be consistently present in the synovial membrane of people with rheumatoid arthritis (n = 23) but not those with other joint pathologies (Muirden and Senator, 1968). Some rodent-model studies have also demonstrated a remarkable improvement in joint-related inflammation following the removal of iron

from joints through chelation treatment(s) (Andrews *et al.*, 1987; Blake *et al.*, 1983). A 28-month follow-up study to maintain a near iron-deficiency by depleting the levels of metal via phlebotomy in hyperuricaemic patients (with gout) has been reported to induce either complete or marked reduction in incidence and severity of gout flares in humans (Facchini, 2003).

This study found an increased serum ferritin in NZ Polynesian people ($P = 0.04$) with gout who self-reported the consumption of an iron rich food (seafood, fish & red meat) as a trigger for gout (Table 2.10). Another recent study that provides self-reported food-trigger data for NZ Māori, Pacific Island and European individuals also indicated iron rich foods to be one of the top triggers for gout with 62.54% reporting either seafood or fish, and 35.18% reporting red meat as a trigger for gout flares (Flynn *et al.*, 2015). A positive correlation between the consumption of red meat and incident gout risk has been described previously in a 7.7-year follow-up study of 28,990, ostensibly healthy, men (Williams, 2008). Another 12-year prospective study in 47,150 men associated the consumption of meat and seafood but not purine-rich vegetables with an increased risk of gout (Choi *et al.*, 2004b). A Turkish retrospective study in people with gout indicated that higher consumption of total meat (including fish) acts as a precipitating factor for gout flares (Öztürk *et al.*, 2013). Also, purines from animal based, but not plant-based, food have been associated with increased risk of recurrent gout attacks (Zhang *et al.*, 2012b). In line with these findings, a large prospective study in NHANES III data, including 14,809 participants reported an association of increased consumption of red meat with hyperuricaemia (Choi *et al.*, 2005a). It is possible that additional to purines, the iron content of such animal-based foods are playing a role in determining the risk of gout and gout flares in several populations.

The study also indicated a significant elevation of serum ferritin levels in US European ($P = 0.02$) and NZ Polynesian ($P = 0.005$) gout individuals who were on allopurinol treatment (Table 2.9). This is of interest as xanthine oxidase is involved in the release of iron from ferritin and facilitating cellular stress through the production of hydroxyl radicals (Bolann and Ulvik, 1987). Use of allopurinol as a xanthine oxidase inhibitor has been attributed to an increased iron overload in rodent liver cells and elevated serum iron in patients with secondary gout (Powell and Emmerson, 1966). Although the debate of involvement of allopurinol in affecting iron metabolism in human

is controversial (Emmerson, 1966; Powell, 1970), my results suggest higher serum ferritin levels in gout patients on allopurinol therapy. Confirming the possible effect of allopurinol on iron metabolism in gout patients may be helpful in decision-making regarding treatment options for patients with a risk of iron overload.

The associations described above are based on observational data and are, therefore, prone to biases due to presence of possible confounding (Mann and Wood, 2012). Whilst its not possible to assign the causality based on my findings, the positive association of the G allele of *HFE*: *rs1799945* with gout in European females (OR = 2.28, *P* = 0.03; Table 2.12) and its association with serum urate in New Zealand population groups (Table 2.11) provided support to my observational data. Homozygosity of the same allele (G/G) has been classically associated with moderate risk of haemochromatosis (Hanson *et al.*, 2001). However, a single positive association in the NZ European female dataset could also be spurious and cannot be inferred unless assessed in the larger population dataset.

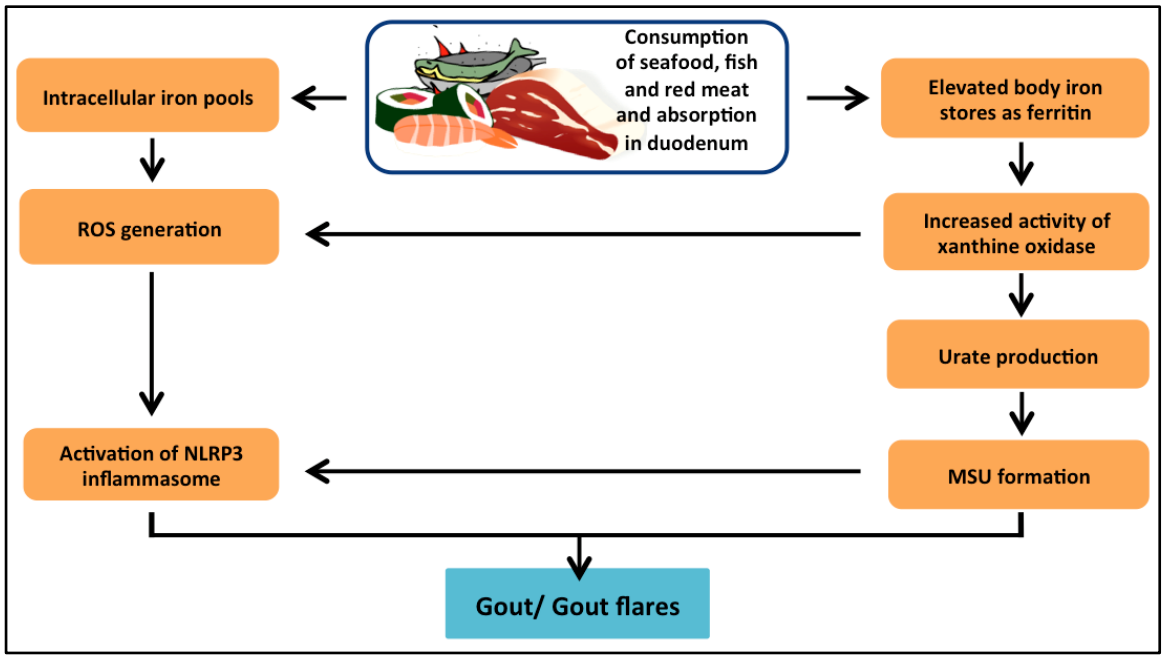


Figure 2.4: Proposed mechanism of involvement of iron in hyperuricaemia and gout pathophysiology.

Albeit it seems rather early to evaluate a possible mechanism underpinning the particular correlation of increased iron and/or consumption of red meat and seafood/fish with increased gout flares, the elevated urate production in the presence of ferritin still indicates a biological mechanism between iron availability and urate production in healthy humans. Based on my findings and previous data (Dabbagh *et al.*, 1992; Facchini,

2003; Ghio *et al.*, 1994), it is suggested that the production of reactive oxygen species directly through increased intracellular iron pools or indirectly through modulating the activity of XO via elevated ferritin, increased urate production, the formation of MSU crystals and hence activation of the NLRP3 inflammasome could lead to gout and gout flares (Figure 2.4).

2.3.1 Strengths and limitations of the study

Inclusion of acute-phase inflammatory protein CRP as an adjustor in various analyses represents the first strength of this study. While it is not possible to assign causality on the basis of the observational data in Section 2.2, the preliminary association of ferritin with urate independent of CRP suggests that the relationship is not due to the elevation of ferritin in inflammation (Kell and Pretorius, 2014). The consistency of the observational findings in this study with previously reported experimental and clinical intervention studies supports the argument for a relationship between serum ferritin and urate that predicts hyperuricaemia and gout.

Although serum urate associations were investigated in larger study cohorts, the study groups used for investigating the association of gout were comparatively small. A huge overlap between iron and purine rich foods could also be a difficult confounder to work with. However, using a larger cohort with richer information for food consumption (especially intake of iron rich low purine foods) and gout ascertained may provide a clearer picture of the observed relationship. Further investigation of a likelihood of intake of iron rich diet as a possible trigger for gout flares in larger cohorts may also help in advice for an appropriate diet with low iron content to avoid or decrease the severity of gout attacks in patients with gout.

2.3.2 Conclusion and future directions

This study replicates the association of ferritin with serum urate. Increased ferritin levels were also associated with gout and self-reported frequency of flares. As the associations reported here are derived from cross-sectional observational data, the causality of ferritin in control of urate concentration, the risk of gout and gouty flares cannot be inferred. Although the genetic association findings for *HFE* variant *rs1799945* provides some support to these observations, the NZ European female dataset was

relatively small and positive results warrant validation. This can either be done considering replication in a larger dataset, via randomised clinical trials or the use of genetic epidemiological approaches. One such approach is Mendelian randomisation that could shed light on any causal relationship (Robinson *et al.*, 2016). This approach is increasingly used as it can exploit existing data sets and is, therefore, a considerably cheaper alternative to a randomised clinical trial. The next chapter of my thesis is based on Mendelian randomisation analyses, which may help to provide a causal basis to the above observations.

Chapter 3

Mendelian Randomisation

Testing for a Causal Relationship between Iron and Urate

SECTION 3.1 MENDELIAN RANDOMISATION: AN INTRODUCTION

3.1.1 Background

Interventional and observational studies that apparently represent robust associations between environmental and physiological measures and disease risk are subject to a variety of biases or confounding (Hayden *et al.*, 2013; Shrank *et al.*, 2011). This means that statistical associations observed in the observational studies are not enough to draw a causal relationship between an exposure and an outcome. The biases in such studies sustain unless potential association confounders are appropriately adjusted for following their identification and perfect measurements (Bowden *et al.*, 2017). Unaccounted confounding along with misclassification and selection bias are the reasons why such studies may fail to replicate the reported associations or prove any causality when tested in randomised controlled trials (RCTs) (Kovesdy and Kalantar-Zadeh, 2012; Smith and Ebrahim, 2001, 2008). Detecting association in an observational study could be the result of influence of a confounder rather than the direct effect of the biomarker on a disease itself. As it is hard to separate the causal associations from confounding and/or reverse causation (Lipsitch *et al.*, 2016), incorrect causal inferences have been reported in such studies even after careful statistical adjustments for possible measured confounders and careful study design (Fewell *et al.*, 2007; Smith and Ebrahim, 2001). In addition, effect estimates from interventional studies may be prone to errors e.g., regression dilution bias and/or incorrect measurement of biomarkers for technical (or biological) reasons (Smith and Ebrahim, 2005). After the identification of these problematic aspects of epidemiological investigation, a number of methods, aimed at improving causal inference, were developed (Lipsitch *et al.*, 2012; Smith, 2008). One such successful approach, analogous to RCT (Iturrieta-Zuazo and Walter, 2015; Smith and Hemani, 2014), is Mendelian randomisation or MR (Smith and Ebrahim, 2003).

3.1.2 Basic principle of Mendelian randomisation

Presented for the first time by Smith and Ebrahim (2003), the Mendelian randomisation approach is based on ‘Mendel’s law of independent assortment’ that states that ‘during gamete-formation, individual hereditary factors (alleles) are assorted independent of each other and thus their associated traits get equal opportunity to occur

by chance (or together)'. Also known as nature's randomised trial, this statistical approach exploits random assignment of alleles at conception to disentangle cause and effect in the presence of confounding. Genetic variants robustly associated with exposure phenotype (usually a physiological biomarker also known as the risk factor) can be used to identify the predictive (causal or non-causal) effect of the variant on the outcome (usually a disease) (Didelez and Sheehan, 2007). The promising applicability of MR is made possible via two unique characteristics of the genotype. First, the random allocation of alleles at conception independent of the environmental exposures lowers the chances of potential confounding to occur in genetic association studies in the same way as randomised treatment allocation in RCT (Hingorani and Humphries, 2005). Second, the invariable nature of DNA sequence and unidirectional flow of the information from DNA to complex phenotype or disease make reverse causation impossible (Swerdlow *et al.*, 2016). Thus MR can also provide valid cause-effect information that is not prone to variability via environmental influences.

3.1.3 Mendelian randomisation model and instrumental variable

The first step of any MR study is to select a valid instrumental variable (IV), which is usually a genetic variant (G) robustly associated with the exposure/risk factor (X) that can be used as a proxy indicator for 'X'. The number of instrumental variables is solely dependent on their validity, which is in turn determined by three major assumptions (Smith and Ebrahim, 2003). Illustrated in the MR model in Figure 3.1, these assumptions are detailed below;

1. IV1: The instrument (G) is associated with the exposure/risk factor (X).
2. IV2: The instrument (G) is not associated with any known and unknown confounders (U) of the exposure (X)-outcome (Y) association.
3. IV3: The instrument (G) is independent of the outcome (Y) conditional on the exposure (X) and confounders (U) i.e., no horizontal pleiotropy (Burgess *et al.*, 2013; Hemani *et al.*, 2016).

In addition to finding causality, the selection of a valid IV can provide a clear picture of two attributes of the relationship between exposure and outcome i.e., the 'direction of association' and 'horizontal pleiotropy' (Smith and Hemani, 2014). As the whole point of MR is to indicate a cause and effect relationship between 'X' and 'Y'

independent of 'U', provided the direction of association (either positive or negative) between 'G' and 'X' is known, the direction of causal association between 'X' and 'Y' can be inferred. Also, interpretation of results from MR analyses must be done carefully if the genetic variant (G) chosen as the IV exhibits horizontal pleiotropy i.e., 'G' is associated with 'X' but also influences 'Y' through a pathway that does not include 'X'. If any horizontal pleiotropy is found, the IV3 assumption in the MR model is violated (Haycock *et al.*, 2016; Smith and Ebrahim, 2003). Unlike the IV1 assumption in the MR model, the chances of selecting an instrument that completely fulfils IV2 and IV3 are low (Bowden *et al.*, 2016), the IV3 assumption being reported to be nearly impossible to prove (Didelez and Sheehan, 2007).

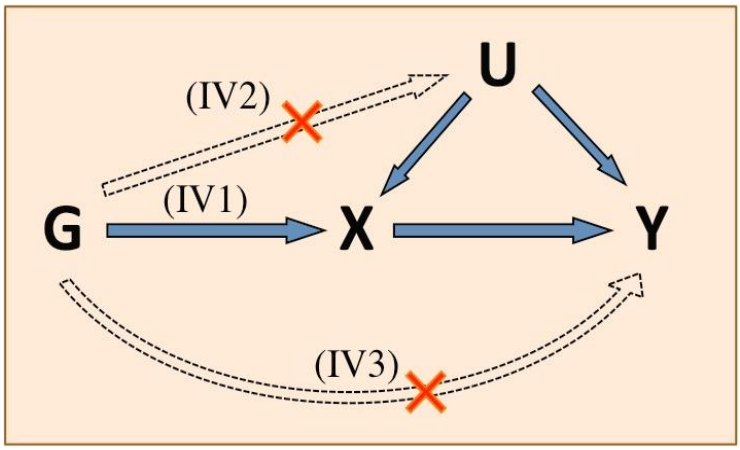


Figure 3.1: The Mendelian randomisation model detailing selection of exposure (X) instrument variable (IV or G) and the assumptions (solid blue arrows) it should fulfil with regard to its association with potential confounders (U) and the outcome (Y). Dotted arrows indicate potential violation of the MR model. Adapted from Burgess *et al.* (2013) and Hemani *et al.* (2016).

3.1.4 Extensions to basic MR in the wake of time

Given that the use of MR as a genetic epidemiological approach could shed light on a causal relationship (Robinson *et al.*, 2016), there have been a number of developments and extensions to the basic MR study design since its first proposal by Smith and Ebrahim (2003). Like conventional MR, each extension has its own advantages and limitations. More broadly and in the milieu of this chapter, MR study designs can be categorised on the basis of the number of variants and samples (datasets) used.

3.1.4.1 Multiple-instrument MR

Mendelian randomisation is conventionally performed using a single genetic variant as IV, whose biological aspects for an association with exposure are well understood. If a single variant is used as a proxy indicator (or IV) for an exposure phenotype, it should have a strong association with the exposure i.e., it must explain sufficiently large variance in the exposure (trait) in the said population and must be exclusively associated with the exposure independent of other traits (Smith and Ebrahim, 2008; Taylor *et al.*, 2014). Using as an instrument a variant that is known for its association(s) with traits other than exposure would increase the chances of pleiotropy. This means that the same genetic variant may be a proxy for more than one environmentally modifiable exposure or may be in linkage disequilibrium (LD) with another variant that is known to be affecting the outcome (Smith and Ebrahim, 2008; Solovieff *et al.*, 2013). Both of these scenarios would make the selected variant a weak instrument and could increase the degree of biases in MR analyses.

The limitation to the MR approach arising from usage of a single weak IV can be partially alleviated by increasing the number of selected instruments (genetic variants). For example, using multiple instruments instead of one would increase the proportion of explained variance for an exposure and thereby improve the accuracy in two-stage least squares (2SLS) regression (Smith and Hemani, 2014). In this situation, the variance explained by all of the selected instruments is combined into a weighted allele score as an optimal approach (Burgess and Thompson, 2013). When using a multiple-instrument MR approach, the selection of variants should be done carefully as the presence of a complete LD between the selected instruments can cause drastic biases in the MR outcomes. Similar predicted causal effects of low-density lipoprotein cholesterol (LDL-C) on coronary heart disease (CHD) through 9 genetic variants from 6 genes provides a convincing example of pleiotropy via intra-instrument LD (Ference *et al.*, 2012).

3.1.4.2 Two-sample MR

Another advantageous extension to the basic MR model is to utilise data from two independent samples rather than one, an approach known as two-sample MR (Inoue and Solon, 2010; Pierce and Burgess, 2013). The idea behind two-sample MR is to obtain the estimates of instrument-exposure association and instrument-outcome association from

two different sets of participants. The phenomenon of underestimation of true causal effects or having 'winners curse' in one-sample MR (Taylor *et al.*, 2014) is unlikely to happen when the two-sample approach is applied. Also, unlike one-sample MR using multiple instruments, the impact of weak instrument bias (biased due to confounding in multivariable regression) is towards the null in the two-sample model (Smith and Hemani, 2014). This approach can also successfully reduce the finite (small) sample biases, especially in the cases where the instruments are weak and individual-level data are tested through 2SLS regression (Burgess *et al.*, 2016).

Two-sample MR holds best when no information about the intermediate phenotype (individual-level data) is available and the data to be tested only provide values of estimates and effect sizes e.g., the summary data from GWA studies' or large consortia (Smith and Hemani, 2014). Use of multiple instruments from two large studies not only increase the statistical power (Burgess and Thompson, 2015) but also provide the opportunity to use proxy variants for the instruments whose information is not available in the outcome study data (Gao *et al.*, 2016). While considering large samples, however, a huge overlap of the participants between the two studies should be avoided as it could lead to potential biases, especially if the outcome is a continuous variable (Burgess *et al.*, 2016).

3.1.5 Use of summary-level data in MR and possible statistical approach

As the field of biological research develops, more and more data becomes publicly-available on a larger number of participants e.g., GWA studies that often provide genetic association results over tens of thousands of samples (Benyamin *et al.*, 2014; Köttgen *et al.*, 2013). For most GWAS, however, it may not be possible for the authors to provide public access to the individual-level data due to ethical reasons. The availability of limited data, in turn, makes it challenging to carry out MR analysis especially via using the 2SLS regression approach in the absence of intermediate phenotype information. A number of statistical models (explained further in the next paragraph) have been introduced that could exploit the summarised data from GWAS in MR analyses, each having its own advantages and limitations.

In order to obtain MR estimates for each variant in summary data, the Wald ratio method (Lawlor *et al.*, 2008) can be applied. This approach employs the ratio method to calculate the estimate (Wald estimate: equivalent to regression coefficient) wherein the regression coefficient of 'Y' on 'G' is divided by the regression coefficient of 'X' on 'G'. The estimates from individual variants can be combined using the inverse-variance weighted (IVW) method using summary data. The IVW method averages the estimates (ß) of causality of 'X' on 'Y' across the sample(s) and does not require information of intermediate phenotype (Burgess *et al.*, 2013). Another method that can combine estimates across the sample is the likelihood-based method, constructed assuming a linear relationship between 'X' and 'Y' and normal distribution of 'ß' (Burgess *et al.*, 2013).

Although, the IVW and likelihood methods provide equivalent statistics to 2SLS, yet both methods are unable to cope with possible horizontal pleiotropy that usually arises in the presence of multiple weak variants (violation of IV3 assumption). A more robust approach that can address this problem is sensitivity analysis e.g., Egger-regression (termed as MR-Egger for MR analysis) (Bowden *et al.*, 2015). A recent study indicating a causal association of adiposity-related traits with breast cancer provides a convincing example of the applied use of IVW and MR-Egger on the summary data from GWA studies (Gao *et al.*, 2016).

3.1.6 Mendelian randomisation studies for iron biomarkers

In the context of rheumatology, the MR approach has predominantly been applied to explain the cause-effect relationship between urate levels and cardio-metabolic traits e.g., hypertension (Kleber *et al.*, 2015), coronary heart disease/CHD (Palmer *et al.*, 2013; White *et al.*, 2016), diabetes (Pfister *et al.*, 2011; Sluijs *et al.*, 2015), BMI (Palmer *et al.*, 2013; White *et al.*, 2016), obesity (Chen *et al.*, 2017; Lyngdoh *et al.*, 2012) and renal function (Hughes *et al.*, 2014). However, there are only a couple of studies using MR analyses to explain a causal relationship between iron metabolism and other traits, including an MR study that provided evidence supporting a causal association between blood iron and decreased risk of Parkinson disease using iron-related variants as instruments from GWAS data (from European and Australian individuals) (Pichler *et al.*, 2013), and another more recent study that indicated a protective effect of iron and ferritin on kidney function via MR analysis (del Greco *et al.*, 2016). The latter is the only MR

study available until now (to my knowledge) that has reported that exposure to iron is causal in decreasing the risk of kidney disease.

Since it is impossible to exclude pleiotropy due to the presence of potential confounders in observational studies, an RCT should be done to determine if the association is causal. Whilst it was not possible to carry out an RCT, I chose Mendelian randomisation (MR) as an alternative approach to investigate the causal relationship of the associations between iron profile markers and urate found in the biochemical analysis in Chapter 2. The specific aim of my study was ‘to use Mendelian randomisation as an alternative to RCT to test for a causal role of total serum iron, serum ferritin, serum transferrin and TSAT (the exposures) in raising or lowering urate levels (the outcome) using multiple variants as instruments by exploiting summarised data from recent GWAS. In addition, the causal role of urate in increasing/decreasing the levels of total serum iron, serum ferritin, serum transferrin and TSAT was also assessed by selecting urate-related variants as instruments for iron profile biomarkers as outcome using summary statistics from same GWAS datasets.

The results are expected to provide additional information to address the broader question of whether or not iron homeostasis is causally involved in affecting urate concentrations and its associated aetiologies.

SECTION 3.2 MENDELIAN RANDOMISATION ANALYSES USING SUMMARY DATA FROM GENOME-WIDE ASSOCIATION STUDIES

3.2.1 Background

As described in Section 3.1, whilst searching for a causal effect of a risk factor on a disease outcome, it is mandatory to carefully adjust error due to confounding to increase the reliability of results in observational studies. This section was aimed to make use of large summary-level data from GWA studies to run MR analysis through a valid statistical approach that is equivalent or even more robust than 2SLS (used when individual-level data are available).

3.2.2 Methods

The two-sample bidirectional (iron as a cause for urate as an effect and urate as a cause for iron as an effect) MR approach was adopted in this study using multiple instruments. This approach allows the selection of variant-exposure associations (instruments) from one GWAS, and variant-outcome associations from a different GWAS (Burgess *et al.,* 2015a). Two-sample MR increases the sample size, which in turn provides MR analysis with adequate statistical power to identify small-sized causal effects of common variants (Gage *et al.,* 2016). The MR model was built on the basis of a hypothesis that states that "an increase in exposure (iron profile biomarkers or urate) is causal for an increase in the outcome (urate or iron profile biomarkers)".

3.2.2.1 Selection of GWAS datasets

Mendelian randomisation analysis in this study was based on publicly-available summary-level data from two large previously reported GWAS for iron (Benyamin *et al.,* 2014) and urate (Köttgen *et al.,* 2013).

3.2.2.1.1 Iron as exposure for urate as outcome

3.2.2.1.1.1 Exposure instrument data

Summary statistics from a recent GWAS of iron-related phenotypes (Benyamin *et al.*, 2014) were used to select the exposure-associated instruments. Overall, this GWA study reported genetic association data for a sample set of over 48,900 European male and female individuals from 18 study cohorts (Discovery and Replication cohorts; see Appendix A Table 3.1) within 'Genetics of Iron Status Consortium' (GIS). Publicly-available meta-analysis effect estimates (in standard deviation/SD units) for serum iron (2,096,457 SNPs), serum ferritin (2,036,124 SNPs), serum transferrin (2,104,242 SNPs) and transferrin saturation/TSAT (2,102,226 SNPs) were sourced from 'GeneEpi platform' (Genetic Epidemiology, Psychiatric Genetics, Asthma Genetics and Statistical Genetics Laboratories Brisbane, Australia; Link: www.genepi.qimr.edu.au).

As it was not possible to calculate *F*-statistics (*F*-statistics is calculated as a measure of strength of association of IV to exposure in MR studies) for variants based on the summary information in the available data, the threshold of $p < 5 \times 10^{-8}$ was set for a SNP to be a valid instrument for exposure. The GWAS-level ($p < 5 \times 10^{-8}$) threshold is known to correspond to an *F*-statistic of 30 (Burgess *et al.*, 2013), and as most variants were well below the above threshold, their *F*-statistics would be even greater than 30. Following this criteria, all SNPs that were significantly associated ($p < 5 \times 10^{-8}$) with serum iron, transferrin, ferritin or TSAT, were selected as MR instruments (Table 3.1) for the study. A number of these SNPs were in high LD with each other. For each iron-related trait, where $r^2 > 0.6$, one SNP was randomly designated as a potential instrument for the MR analysis and all other highly correlated SNPs were omitted. Any variant that was in LD with a known urate-related variant or was in high eQTL (i.e., coexpressed with another gene) with any known urate loci, was also excluded. Two different gene-expression platforms, GTEx portal (www.gtexportal.org) and HaploReg v4.1 (www.archive.broadinstitute.org) were used to determine the degree of coexpression of the instrument SNPs with other genes to check if any of the instrument carries a pleiotropic effect by up/down-regulating expression of other genes. Using these criteria resulted in a list of three instruments for serum iron (*rs1525892, rs1800562, rs855791*), eight for serum transferrin (*rs1495741, rs174577, rs1800562, rs744653, rs7646473,*

rs8177240, rs9268633, rs9990333), five for serum ferritin (*rs12693541, rs1800562, rs2413450, rs651007, rs411988*) and three for TSAT (*rs1800562, rs8177272, rs855791*). Table 3.2 provides complete detail for all selected exposure-instruments. Any instrument showing possible pleiotropy due to high LD with another instrument (*rs1525892, rs1495741, rs411988*), high co-expression with urate-related genes (*rs1799945*), and possible pleiotropic effect reported in the literature (*rs7385804*) has been removed from the various MR analyses.

3.2.2.1.1.2 Outcome data

Summary statistics data from the largest urate GWAS (Köttgen *et al.*, 2013) were used for outcome-instrument association estimates in the MR model. This GWA study provided the meta-analysis data for urate-associated loci in > 110,000 European male and female individuals from 48 study cohorts (Appendix A Table 3.1) within Global Urate Genetics Consortium (GUGC). Publicly-available effect estimates (for urate) of 2,450,547 SNPs were used as outcome-association variables (www.metabolomics.helmholtz-muenchen.de/gugc/).

3.2.2.1.2 Urate as exposure for iron as outcome

3.2.2.1.2.1 Exposure instrument and outcome data

To assess the exposure-outcome relationship between urate and iron, a reverse MR analysis was also carried out via instrumenting the urate-related variants from Köttgen *et al.* (2013) GWAS summary statistic data while summary GWAS statistics from Benyamin *et al.* (2014) were used as outcome data.

Using the above-described criteria for selection of IVs, four variants from known urate transporter genes were selected as potential urate instruments from the exposure GWAS (Köttgen *et al.*, 2013); *SLC2A9: rs12498742, SLAC16A9: rs1171614, SLC22A11: rs2078267* and *SLC22A12: rs478607* (Table 3.1). Although the effect estimate at *SLC2A9* locus was strong enough (ß (mg dL^{-1}) = 0.373, P = 1E-700) to be selected as a single IV in the MR model, additional loci were required to be selected to fulfil the IV selection requirements of the statistical approach applied in this study (explained in Section 3.2.2.2 below).

3.2.2.1.3 Assessment of GWAS sample overlap

As mentioned earlier, the use of two datasets in the two-sample MR model can lead to false positive results and can create bias due to large sample overlap between the exposure and outcome datasets (Burgess *et al.*, 2016). The GWA studies used in this chapter provide the meta-analysis results from several study cohorts of European ancestry. Due to the same ancestral background of the participants within two selected GWA studies (Benyamin *et al.*, 2014; Köttgen *et al.*, 2013), it was highly likely to expect a large sample overlap between them making the two-sample MR prone to biases. Before running the MR analysis, the percentage of the sample overlap was also measured between the two studies.

Table 3.1: Details for the exposure-associated instrument variants selected for the study

Expo/Outc phenotype	rs ID	Allele		Chr	Gene/Nearest gene	Exp-GWAS_ß	SE	Exp-GWAS_*p*	OutC-GWAS_*p*	r^2 (in LD with)	eQTL-gene (*p*)/ pleiotropy/Status
		Eff	Ref								
Iron profile-related instruments											
Iron/Urate	*rs1525892*	A	G	3	*TF*	0.074	0.0104	1.65E-12	0.096	-	-
	rs1800562	A	G	6	*HFE*	0.372	0.02	3.96E-77	0.001	-	-
	rs855791	A	G	22	*TMPRSS6*	-0.186	0.0101	4.31E-77	0.979	-	-
	rs1799945	C	G	6	*HFE*	-0.189	0.01	1.10E-81	2.94E-06	-	*SLC17A3* (8.5E-06)
	rs7385804	A	C	7	*TFR2*	0.064	0.007	1.36E-18	0.396	-	Pleiotropic
	rs8177240	T	G	3	*TF*	-0.089	0.011	2.37E-12	0.091	1 (*rs1525892*)	-
Trf/Urate	*rs1495741*	A	G	8	*NAT2*	0.082	0.0122	1.57E-11	0.309	-	-
	rs174577	A	C	11	*FADS1*	0.068	0.0107	1.90E-10	0.002	-	-
	rs1800562	A	G	6	*HFE*	-0.549	0.0208	1.26E-153	0.001	-	-
	rs744653	T	C	2	*SLC40A1*	0.092	0.0144	2.00E-10	0.558	-	-
	rs7646473	A	G	3	*SLC20A1*	-0.062	0.0107	6.84E-09	0.457	-	-
	rs8177240	T	G	3	*TF*	-0.423	0.0107	3.82E-340	0.091	-	-
	rs9268633	A	G	6	*HLA-DRA*	-0.072	0.0128	2.31E-08	0.522	-	-
	rs9990333	T	C	3	*TFRC*	-0.067	0.0101	3.01E-11	0.381	-	-
	rs4921915	A	G	8	*NAT2*	0.082	0.012	1.74E-11	0.289	1 (*rs1495741*)	-
	rs1799945	C	G	6	*HFE*	0.114	0.01	9.36E-30	2.94E-06	-	*SLC17A3* (8.5E-06)
Ferritin/Urate	*rs12693541*	T	C	2	*SLC40A1*	-0.106	0.014	4.18E-14	0.375	-	-
	rs1800562	A	G	6	*HFE*	0.211	0.0187	1.43E-29	0.001	-	-
	rs2413450	T	C	22	*TMPRSS6*	-0.056	0.0095	3.57E-09	0.913	-	-
	rs368243	T	C	17	*TEX14*	0.051	0.0093	3.80E-08	0.620	0.9 (*rs411988*)	-
	rs651007	T	C	9	*ABO*	-0.05	0.009	1.31E-08	0.021	-	-
	rs411988	A	G	17	*TEX14*	-0.044	0.007	1.59E-10	0.692	-	-
	rs1799945	C	G	6	*HFE*	-0.065	0.01	1.36E-18	2.94E-06		*SLC17A3* (8.5E-06)
TSAT/Urate	*rs1800562*	A	G	6	*HFE*	0.577	0.0203	1.52E-178	0.001	-	-
	rs8177272	A	G	3	*TF*	-0.097	0.0106	5.52E-20	0.107	-	-
	rs855791	A	G	22	*TMPRSS6*	-0.192	0.0101	3.50E-80	0.979	-	-
	rs1799945	C	G	6	*HFE*	-0.231	0.01	5.13E-109	2.94E-06	-	*SLC17A3* (8.5E-06)

Urate-related instruments											
Urate/Iron	*rs12498742*	A	G	4	*SLC2A9*	0.373	0.006	1E-700	0.312	-	-
	rs1171614	T	C	10	*SLC16A9*	-0.079	0.007	2.3E-28	0.940	-	-
	rs2078267	T	C	11	*SLC22A11*	-0.073	0.006	9.4E-38	-	-	Not in outcome data
	rs478607	A	G	11	*SLC22A12*	-0.047	0.007	4.4E-11	0.871	-	-
Urate/Trf	*rs12498742*	A	G	4	*SLC2A9*	0.373	0.006	1E-700	0.929	-	-
	rs1171614	T	C	10	*SLC16A9*	-0.079	0.007	2.3E-28	0.333	-	-
	rs2078267	T	C	11	*SLC22A11*	-0.073	0.006	9.4E-38	-	-	Not in outcome data
	rs478607	A	G	11	*SLC22A12*	-0.047	0.007	4.4E-11	0.546	-	-
Urate/Ferritin	*rs12498742*	A	G	4	*SLC2A9*	0.373	0.006	1E-700	0.460	-	-
	rs1171614	T	C	10	*SLC16A9*	-0.079	0.007	2.3E-28	0.551	-	-
	rs2078267	T	C	11	*SLC22A11*	-0.073	0.006	9.4E-38	-	-	Not in outcome data
	rs478607	A	G	11	*SLC22A12*	-0.047	0.007	4.4E-11	0.059	-	-
Urate/TSAT	*rs12498742*	A	G	4	*SLC2A9*	0.373	0.006	1E-700	0.108	-	-
	rs1171614	T	C	10	*SLC16A9*	-0.079	0.007	2.3E-28	0.857	-	-
	rs2078267	T	C	11	*SLC22A11*	-0.073	0.006	9.4E-38	-	-	Not in outcome data
	rs478607	A	G	11	*SLC22A12*	-0.047	0.007	4.4E-11	0.523	-	-

Trf: Transferrin, Ferritin: Log ferritin, TSAT: Transferrin saturation, Exp/Outc: Exposure and outcome traits, rs ID: SNP reference number, Eff: effect allele, Ref: reference allele, Chr: chromosome, SE: standard error, Exp-GWAS_ß: Effect estimate for association with exposure trait in Exposure GWAS, Exp-GWAS_*p*: *p*-value from exposure association GWAS, OutC-GWAS_*p*: *p*-value from outcome association GWAS, r^2: R-squared value for the LD (linkage disequilibrium) with another SNP in the list, eQTL-gene: gene with which the said SNP is coexpressed, eQTL: expression quantitative trait loci, Status: Absence of selected IV in the outcome data.

3.2.2.2 Statistical analysis

Two-sample MR analysis was performed using the MR-Base (Mendelian Randomisation-Base) platform (www.mrbase.org) (Hemani *et al.*, 2016). The files for exposure instruments were manually uploaded in the MR-Base. In order to assure no LD between the instruments, the 'LD clumping' command was used that excluded one out of any two instruments that were in LD with each other at the threshold of 60% or above (r^2 > 0.6) for each exposure trait. For any instrument (SNP) that was absent in the outcome data, a proxy instrument was used (if it was available in the outcome data) at LD 60% or above (r^2 > 0.6) to assure that the power of using multiple instruments was not lost (Table 3.1). The proxy variants were obtained within MR-Base from outcome data in the European population using the 1000 Genomes database (www.internationalgenome.org), with a limit of +/- 250kb or +/- 1000 SNPs for each variant. In order to make sure that the effect estimate of exposure-instrument corresponds to the same allele as their effects on the outcome, the 'allele harmonisation' command was used that attempts to align strands for palindromic sequences (at MAF > 0.3) before the MR model is run.

The following four statistical approaches were used to assure robust analysis and to attain reliable causal outcomes in the MR model. All of these methodologies are considered valid while dealing with summarised data from GWA studies in presence of a continuous exposure and a continuous outcome (Section 3.1.5).

1. Wald ratio method (single SNP MR estimate analysis)
2. Inverse-variance weighted (IVW) method (combined MR estimate analysis)
3. Leave-one out method (sensitivity analysis for an outlier IV)
4. Egger-regression method (sensitivity analysis for horizontal pleiotropy)

3.2.2.2.1 Wald ratio method

First, the MR estimates (Wald estimate) were calculated and obtained within MR-Base using the ratio of coefficients or Wald ratio method (Wald, 1940). The Wald ratio is the simplest method that estimates the causal effect of the exposure (X) on the outcome (Y) using each IV as a separate entity. If the IV is indexed as '*k*', the Wald estimate represents the ratio between estimates of two linear regressions i.e., a regression of outcome/*Y* on '*k*' and the regression of exposure/*X* on '*k*' (Burgess *et al.*, 2015b; Lawlor

et al., 2008). For the continuous outcome under study i.e., urate, the Wald ratio for each variant '*k*' was calculated as;

$$\text{Wald ratio} = \frac{Y_k}{X_k}$$

Where, Y_k is the coefficient of '*k*' (IV) in the regression of 'Y' on '*k*' and X_k is the coefficient of '*k*' in the regression of 'X' on '*k*'.

3.2.2.2.2 Inverse-variance weighted (IVW) method

As the Wald ratio method is more efficient for single IV MR, and the data employed in this study included multiple-variants, the combined cause-effect relationship was analysed via the inverse-variance weighted (IVW) method that requires at least two IVs to provide accurate MR measurements (Burgess *et al.*, 2013; Burgess *et al.*, 2015b). Under the IVW model, the combined ratio estimate (β_{IVW}) of the effect of the exposure '*X*' on the outcome '*Y*' using genetic variants $k = 1, \ldots, K$ was calculated as;

$$\beta_{IVW} = \frac{\sum_k X_k Y_k \sigma_{Yk}^{-2}}{\sum_k X_k^2 \sigma_{Yk}^{-2}}$$

Where X_k is the per-allele addition effect of SNP *k* (IV) with exposure (e.g., iron biomarkers in SD units), Y_k is the per-allele addition change in the outcome (e.g., urate in mg dL^{-1} units), σ_{Yk}^{-2} indicates inverse-variance of the gene-outcome associations. The IVW analysis assumes all genetic variants used in the MR model to be valid instruments.

Further to these, two sensitivity analyses were done to provide robustness to the MR outputs. The details as to why these methods were adopted are described in the paragraphs below.

3.2.2.2.3 Leave-one-out method

Leave-one-out analysis (Corbin *et al.*, 2016) was applied as an extension to the IVW analyses above. The basic calculations in the leave-one-out method are same as IVW, except that it provides the MR estimates between exposure and outcome by excluding each IV, one at a time, from the IVW model. The test checks if one particular

SNP (IV) disproportionately influences the summary causal estimates or is an outlier. Briefly, the method sequentially removes each SNP from the MR model as 'leave-one-out permutation analysis'.

3.2.2.2.4 Egger-regression method

Given that MR analysis is prone to horizontal pleiotropy that violates IV3 in the MR model (Section 3.1), the MR-Egger test (Egger-regression) was applied as a sensitivity analysis to detect directional (horizontal pleiotropy) within MR-Base (Bowden *et al.*, 2015). The MR-Egger test is a regression of the Y_k (gene-outcome) on the X_k (gene-exposure) associations where the intercept is not constrained to pass through zero and requires at least three IVs to provide accurate assessment of the pleiotropy. In detail, the MR-Egger intercept allows the information on the directional pleiotropy to be detected i.e., if the X_k is zero, the Y_k should also be zero and represent average pleiotropic effect across the genetic variants. Any contrary result indicates the presence of a 'direct effect', that is a pathway between the gene (G or IV or *k*) and the outcome (Y) independent of the exposure (X). In practice, the bias detected via MR-Egger is equivalent to the detection of finite-study or weak instrument bias (Burgess *et al.*, 2015b). In contrast to IVW, MR-Egger assumes all instruments to be violating the IV assumptions in the MR-model.

Both IVW and MR-Egger are efficient enough to calculate small causal effects or information on the gene-outcome relationship independent of the exposure under variable number (at least two for IVW and three for MR-Egger) and strength of instruments (Bowden *et al.*, 2017; Bowden *et al.*, 2015; Burgess *et al.*, 2015b). Heterogeneity of estimates across the variants was also calculated using the Cochran Q test via IVW and MR-Egger methods, with a threshold of $p \leq 0.05$ indicating the presence of heterogeneity.

3.2.3 Results

3.2.3.1 Selection of iron exposure instruments

Of the variants associated with iron-related phenotypes in the combined discovery and replication data from the GIS consortium (Benyamin *et al.*, 2014), the *HFE* variant *rs1799945* was excluded from the list because of the possible pleiotropic effect with a known urate locus. The SNP *rs1799945* has been reported to be in a *cis*-eQTL in the

adrenal gland and is associated with the expression of solute carrier family 17 member 3 gene (*SLC17A3*: *P* = 8.5E-06) that encodes a urate transporter (www.gtexportal.org). Also, the SNP was observed to be significantly associated with urate (*P* = 2.9E-06) in the outcome data (Köttgen *et al.*, 2013), which might be indicative of this SNP to be a urate-association signal independent of iron and its related serum biomarkers. Another iron-associated SNP within the *TFR2* gene, *rs7385804*, has been reported as a *cis*-eQTL in whole blood (Westra *et al.*, 2013) and is associated with the expression of solute carrier family 12 member 9 gene (*SLC12A9*: *P* = 1.29E-05) and Ephrin type-B receptor 4 gene (*EPHB4*: *P* = 2.7E-34) (www.archive.broadinstitute.org). The *SLC12A9* gene encodes a protein that acts as the inhibitor of member 2 gene within same family (*SLC12A2*) (Caron *et al.*, 2000). Expressed ubiquitously in the body, *SLC12A2* is a Na^+-K^+-Cl^- cotransporter, which is inhibited by diuretics (Markadieu and Delpire, 2014) and its increased activity is associated with increased blood pressure in animals (Orlov *et al.*, 2015). As blood pressure could influence glomerular filteration rate independent of iron metabolism, a pleiotropic effect could not be ruled out. The *EPHB4* gene is known to upregulate podocyte activity in glomerulonephritis (Wnuk *et al.*, 2012), and thereby suggested to have an additional pleiotropic effect via the iron-independent influence on renal function. Two SNPs, *NAT2: rs4921915* and *TEX14: rs368243* were excluded from the list of potential instruments due to their high LD with *NAT2: rs1495741* (r^2 = 1.0) and *TEX14: rs411988* (r^2 = 0.93), respectively. For other selected instruments, no evidence of a pleiotropic effect was found based on literature and gene-expression search on bioinformatics platforms.

3.2.3.2 Selection of urate exposure instruments

The variants selected as instruments for urate exposure were exclusively selected from urate-transporter genes and have not been reported in the literature to possess pleiotropic effect for iron metabolism (to my knowledge). The *SLC2A9* (solute carrier family 2 member 9) is a member of the SLC2A facilitative glucose transporter family that encodes for a protein GLUT9/human glucose transporter 9 (www.genecards.org). Members of the SLC2A family play a significant role in maintaining glucose homeostasis with *SLC2A9* specifically being recognised as major urate transporter gene (Vitart *et al.*, 2008). The *SLC2A9: rs12498742* was the top urate-associated locus in Köttgen *et al.* (2013) GWAS. No *cis*-eQTL has been reported for the SNP *rs12498742* in two gene-

expression platforms used in this study (HaploRegv4: www.archive.broadinstitute.org and GTEx portal: www.gtexportal.org) except *SLC2A9* in the whole blood (*P* = 6.42E-09). A number of other variants have been reported to be in complete LD with *rs12498742*, all within or near the *SLC2A9* locus (www.archive.broadinstitute.org and www.gtexportal.org). The second IV (*rs1171614*) was selected from a urate-transport gene called solute carrier family 16 member 9 (*SLC16A9*), which encodes for MCT9 (monocarboxylate transporter 9) protein. The SNP *rs1171614* has been reported to be mainly associated with expression of only *SLC16A9* in several tissues in *cis*-eQTL data (Westra *et al.*, 2013) i.e., artery aorta (*P* = 1.92E-06), esophagous mucosa: (*P* = 2.31E-08) and thyroid gland (*P* = 1.77E-06) (www.archive.broadinstitute.org and www.gtexportal.org). The variant *rs1171614* was not reported to alter the expression of another gene that could have suggested its possible pleiotropic effect on iron metabolism. The third IV (*rs478607*) was selected from the gene *SLC22A12/URAT1* (solute carrier family 22 member 12). The protein encoded by this gene is a member of the organic anion transporter (OAT) family, and acts as a urate transporter to regulate urate levels in the blood (www.genecards.org). The SNP *rs478607* has been reported to be a *cis*-eQTL for expression of *SLC22A12* in several tissues and expression of splicing factor 1 (*SF1*) gene in the whole blood (*P* = 3.15E-08) (www.archive.broadinstitute.org and www.gtexportal.org), which is not suggestive for possible pleiotropy. No data were found for the *SLC22A11: rs2078267* variant in the outcome GWAS (Benyamin *et al.*, 2014), which is why this variant was not included as an IV in the MR model.

3.2.3.3 Sample overlap in selected GWAS datasets

The two GWA studies selected for two-sample MR (Benyamin *et al.*, 2014; Köttgen *et al.*, 2013) were found to have only 14% participant overlap (Appendix B Figure 3.1) when their corresponding European cohorts (Appendix A Table 3.1) were observed. As the overlap was sufficiently low, it was considered to not drastically affect the MR model with sample biases.

3.2.3.4 Iron biomarkers as exposure for urate as outcome

3.2.3.4.1 Iron as exposure for urate as outcome

Single variant two-sample MR using the Wald ratio method indicated a positive causal association between iron and urate via one instrument only i.e., *rs1800562* (ß = 0.107, *P* = 0.0008) (Table 3.2, Figure 3.2: A). No causal iron-urate relationship was observed in primary combined MR analysis using IVW (ß = 0.05, *P* = 0.16) (Table 3.2, Figure 3.2: A). Heterogeneity between the effects across the variants was not significant for IVW (*Q-P* = 0.11), but significant when MR-Egger (*Q-P* = 0.02) was done (Table 3.2, Figure 3.3: A). However, a sensitivity MR-Egger indicated little evidence of pleiotropy (intercept: ß = -0.006, OR = 0.99, *P* = 0.63) for an association between iron and urate using all 3 SNPs (Table 3.3, Figure 3.3: B). The MR-Egger combined causal estimate was also not significant and similar to the primary (IVW) analysis (ß = 0.083, *P* = 0.37) (Table 3.2, Figure 3.2: A). Leave-one-out IVW meta-analysis showed a significant causal iron-urate relationship (ß = 0.110, *P* = 1.96E-04) after removing *rs855791* from the iron-instrument list (Table 3.4). The result also indicated that the non-causal association in the IVW meta-analysis was driven by *rs855791* (Table 3.4, Figure 3.2: B). All MR ß-estimates are presented as an effect of an SD unit increase in iron on urate (mg dL^{-1}). This analysis provided evidence that iron plays a causal role in increasing urate concentrations.

3.2.3.4.2 Transferrin as exposure for urate as outcome

Two independent instruments showed a causal transferrin-urate association when assessed via the Wald ratio method i.e., *rs174577* (ß = 0.263, *P* = 0.002) and *rs1800562* (ß = -0.073, *P* = 0.0008) (Table 3.2, Figure 3.4: A). Transferrin was not causally associated with urate in combined IVW meta-analysis (ß = 0.001, *P* = 0.97) (Table 3.2, Figure 3.4: A). Significant heterogeneity (IVW: *Q-P* = 0.002, MR-Egger: *Q-P* = 0.001) was found for effect estimates across the instruments (Table 3.2, Figure 3.5: A). A sensitivity MR-Egger, however, did not indicate evidence of pleiotropy for an association between transferrin and urate using the full list of 8 SNPs (intercept: ß = 0.005, OR = 1.00, *P* = 0.45) (Table 3.3, Figure 3.5: B). The combined causal estimate via MR-Egger did not show a significant association (ß = -0.016, *P* = 0.61) (Table 3.2, Figure 3.4: A). Leave-one-out permutation analysis also did not indicate any variant to be unproportionately altering the non-significant association in IVW (Table 3.4, Figure 3.4:

B). All MR ß-estimates are presented as an effect of an SD unit increase in transferrin on urate (mg dL^{-1}). Overall, this analysis did not provide any evidence of a causal role of serum transferrin in changing serum urate concentrations.

3.2.3.4.3 Ferritin as exposure for urate as outcome

The Wald ratio method showed two instruments indicating a positive causal association between ferritin and urate i.e., *rs1800562* (ß = 0.190, *P* = 0.0002) and *rs651007* (ß = 0.320, *P* = 0.022) (Table 3.2, Figure 3.6: A). However, when combined in standard IVW analysis, no significant ferritin-urate association was observed (ß = 0.089, *P* = 0.17) (Table 3.2, Figure 3.6: A). Heterogeneity between the effects across the variants was not significant for IVW (*Q-P* = 0.06), but significant when assessed through MR-Egger regression (*Q-P* = 0.01) (Table 3.2, Figure 3.7: A). A sensitivity MR-Egger indicated very little pleiotropy (intercept: ß = -0.007, OR = 0.99, *P* = 0.57) in association between ferritin and urate using the full list of 5 SNPs (Table 3.3, Figure 3.7: B). The combined causal estimate obtained using MR-Egger was not significant similar to standard MR analysis (ß = 0.160, *P* = 0.32) (Table 3.2, Figure 3.6: A). The secondary sensitivity approach using leave-one-out analysis did not indicate any variant to be considerably changing the non-significant association in IVW, except *rs12693541*, removal of which showed a positive causal association between ferritin and urate (ß = 0.14, *P* = 0.02) (Table 3.4, Figure 3.6: B). All MR ß-estimates are presented as an effect of an SD unit increase in ferritin (log_{10}) on urate (mg dL^{-1}). The analysis provided evidence of a causal relationship of ferritin with urate.

3.2.3.4.4 Transferrin saturation as exposure for urate as outcome

None of the 3 instruments indicated a causal TSAT-urate association when assessed via the Wald ratio method, except *rs1800562* (ß = 0.069, *P* = 0.001) (Table 3.2, Figure 3.8: A). Transferrin saturation was also not causal for urate in combined IVW meta-analysis (ß = 0.035, *P* = 0.31) (Table 3.2, Figure 3.8: A). No significant heterogeneity (IVW: *Q-P* = 0.106, MR-Egger: *Q-P* = 0.924) was found for effect estimates across the instruments (Table 3.2, Figure 3.9: A). A sensitivity MR-Egger, however, did not indicate evidence of pleiotropy for an association between TSAT and urate using full list of 3 SNPs (intercept: ß = -0.019, OR = 0.98, *P* = 0.20) (Table 3.3, Figure 3.9: B). The combined causal estimate via MR-Egger did not show a significant

association (ß = 0.102, P = 0.17) (Table 3.2, Figure 3.8: A). Leave-one-out permutation analysis also did not indicate any variant to be unproportionately altering the non-significant association in IVW (Table 3.4, Figure 3.8: B). All MR ß-estimates are presented as an effect of an SD unit increase in TSAT on urate (mg dL^{-1}). This analysis did not provide any evidence of a causal role of transferrin saturation in changing serum urate concentrations.

Table 3.2 Association between iron-related traits and urate using two-sample Mendelian randomisation

MR-analysis Method	Phenotype		Gene/locus	Instrument variant	ß-estimate	SE	[95% CI]	*p*-causal	*Q-p*
	Exposure	Outcome							
Wald ratio	Iron	Urate	*TF*	*rs1525892*	0.127	0.076	[-0.02 ; 0.28]	0.093	-
-	Iron	Urate	*HFE*	*rs1800562*	0.107	0.032	[0.04 ; 0.17]	0.0008	-
-	Iron	Urate	*TMPRSS6*	*rs855791*	-0.001	0.031	[-0.06 ; 0.06]	0.986	-
All - IVW	Iron	Urate	-	All	0.056	0.039	[-0.02 ; 0.13]	0.151	0.179
All - MR Egger	Iron	Urate	-	All	0.064	0.109	[-0.15 ; 0.28]	0.661	0.011
Wald ratio	Transferrin	Urate	*NAT2*	*rs1495741*	-0.079	0.077	[-0.23 ; 0.07]	0.310	-
-	Transferrin	Urate	*FADS1*	*rs174577*	0.263	0.083	[0.10 ; 0.43]	0.001	-
-	Transferrin	Urate	*HFE*	*rs1800562*	-0.073	0.022	[-0.12 ; -0.03]	0.0008	-
-	Transferrin	Urate	*SLC40A1*	*rs744653*	0.049	0.084	[-0.12 ; 0.21]	0.559	-
-	Transferrin	Urate	*SLC20A1*	*rs7646473*	0.071	0.101	[-0.13 ; 0.27]	0.485	-
-	Transferrin	Urate	*TF*	*rs8177240*	0.023	0.013	[0.01 ; 0.05]	0.092	-
-	Transferrin	Urate	*HLA-DRA*	*rs9268633*	0.064	0.100	[-0.13 ; 0.26]	0.523	-
-	Transferrin	Urate	*TFRC*	*rs9990333*	-0.070	0.081	[-0.23 ; 0.09]	0.384	-
All - IVW	Transferrin	Urate	-	All	0.001	0.021	[-0.04 ; 0.04]	0.970	0.002
All - MR Egger	Transferrin	Urate	-	All	-0.016	0.031	[-0.08 ; 0.04]	0.611	0.0004
Wald ratio	Log Ferritin	Urate	*SLC40A1*	*rs12693541*	-0.068	0.076	[-0.22 ; 0.08]	0.374	-
-	Log Ferritin	Urate	*HFE*	*rs1800562*	0.190	0.057	[0.08 ; 0.30]	0.0002	-
-	Log Ferritin	Urate	*TMPRSS6*	*rs2413450*	0.011	0.106	[-0.20 ; 0.22]	0.919	-
-	Log Ferritin	Urate	*TEX14*	*rs411988*	-0.048	0.123	[-0.29 ; 0.19]	0.697	-
-	Log Ferritin	Urate	*ABO*	*rs651007*	0.320	0.140	[0.05 ; 0.59]	0.022	-
All - IVW	Log Ferritin	Urate	-	All	0.089	0.066	[-0.04 ; 0.22]	0.176	0.063
All - MR Egger	Log Ferritin	Urate	-	All	0.160	0.134	[-0.10 ; 0.42]	0.320	0.014
Wald ratio	TSAT	Urate	*HFE*	*rs1800562*	0.069	0.021	[0.03 ; 0.11]	0.0001	-
-	TSAT	Urate	*TF*	*rs8177272*	-0.094	0.059	[-0.21 ; 0.02]	0.110	-
-	TSAT	Urate	*TMPRSS6*	*rs855791*	-0.001	0.030	[-0.06 ; 0.06]	0.986	-
All - IVW	TSAT	Urate	-	All	0.035	0.034	[-0.03 ; 0.10]	0.306	0.106
All - MR Egger	TSAT	Urate	-	All	0.102	0.027	[0.05 ; 0.16]	0.168	0.924

TSAT: Transferrin saturation, All - IVW: Meta-analysis using inverse-variance method, All - MR Egger: Mendelian randomisation using Egger regression, ß: beta estimates, SE: standard error, 95% CI: 95% confidence interval, *p*-causal: *p*-value using MR analysis, *Q-p*: Cochran's heterogeneity test *p*-value for heterogeneity.

Table 3.3: Results of horizontal pleiotropy for two-sample MR-Egger test

Phenotype		Horizontal pleiotropy-MR Egger		
Exposure	**Outcome**	**Egger intercept-ß (OR)**	**SE**	***p*-value**
Iron	Urate	-0.0016 (0.99)	0.019	0.946
Transferrin	Urate	0.0052 (1.00)	0.006	0.451
Log Ferritin	Urate	-0.0067 (0.99)	0.011	0.577
TSAT	Urate	-0.0193 (0.98)	0.006	0.205

TSAT: Transferrin saturation, MR Egger: Mendelian randomisation using Egger regression, ß (OR): beta estimates (odds ratio for ß), SE: standard error, *p*-value: *p*-value for directional pleiotropy.

Table 3.4: Results of leave-one-out sensitivity analysis for association between iron-related traits and urate using two-sample Mendelian randomisation

Phenotype		Instrument variant excluded from IVW analysis	ß-estimate	[95% CI]	*p*-causal
Exposure	**Outcome**				
Iron	Urate	*rs1525892*	0.050	[-0.06 ; 0.16]	0.348
Iron	Urate	*rs1800562*	0.017	[-0.07 ; 0.10]	0.697
Iron	Urate	*rs855791*	0.110	[0.05 ; 0.17]	1.96E-04
Iron	Urate	All	0.050	[-0.02 ; 0.13]	0.151
Transferrin	Urate	*rs1495741*	0.002	[-0.04 ; 0.05]	0.917
Transferrin	Urate	*rs174577*	-0.004	[-0.04 ; 0.03]	0.838
Transferrin	Urate	*rs1800562*	0.025	[-0.01 ; 0.06	0.153
Transferrin	Urate	*rs744653*	0.0001	[-0.05 ; 0.05]	0.999
Transferrin	Urate	*rs7646473*	0.0001	[-0.05 ; 0.05]	0.999
Transferrin	Urate	*rs8177240*	-0.041	[-0.11 ; 0.02]	0.220
Transferrin	Urate	*rs9268633*	0.0001	[-0.05 ; 0.05]	0.999
Transferrin	Urate	*rs9990333*	0.002	[-0.04 ; 0.05]	0.927
Transferrin	Urate	All	0.001	[-0.04 ; 0.04]	0.970
Log Ferritin	Urate	*rs12693541*	0.141	[0.02 ; 0.27]	0.027
Log Ferritin	Urate	*rs1800562*	0.007	[-0.14 ; 0.15]	0.929
Log Ferritin	Urate	*rs2413450*	0.101	[-0.05 ; 0.26]	0.204
Log Ferritin	Urate	*rs411988*	0.103	[-0.04 ; 0.25]	0.168
Log Ferritin	Urate	*rs651007*	0.070	[-0.06 ; 0.20]	0.303
Log Ferritin	Urate	All	0.089	[-0.04 ; 0.22]	0.176
TSAT	Urate	*rs1800562*	-0.019	[-0.09 ; 0.05]	0.604
TSAT	Urate	*rs8177272*	0.046	[-0.02 ; 0.11]	0.158
TSAT	Urate	*rs855791*	0.051	[-0.05 ; 0.15]	0.319
TSAT	Urate	All	0.035	[-0.03 ; 0.10]	0.306

IVW: meta-analysis using inverse-variance method, ß: beta estimate, 95% CI: 95% confidence interval, *p*-causal: *p*-value using IVW meta-analysis.

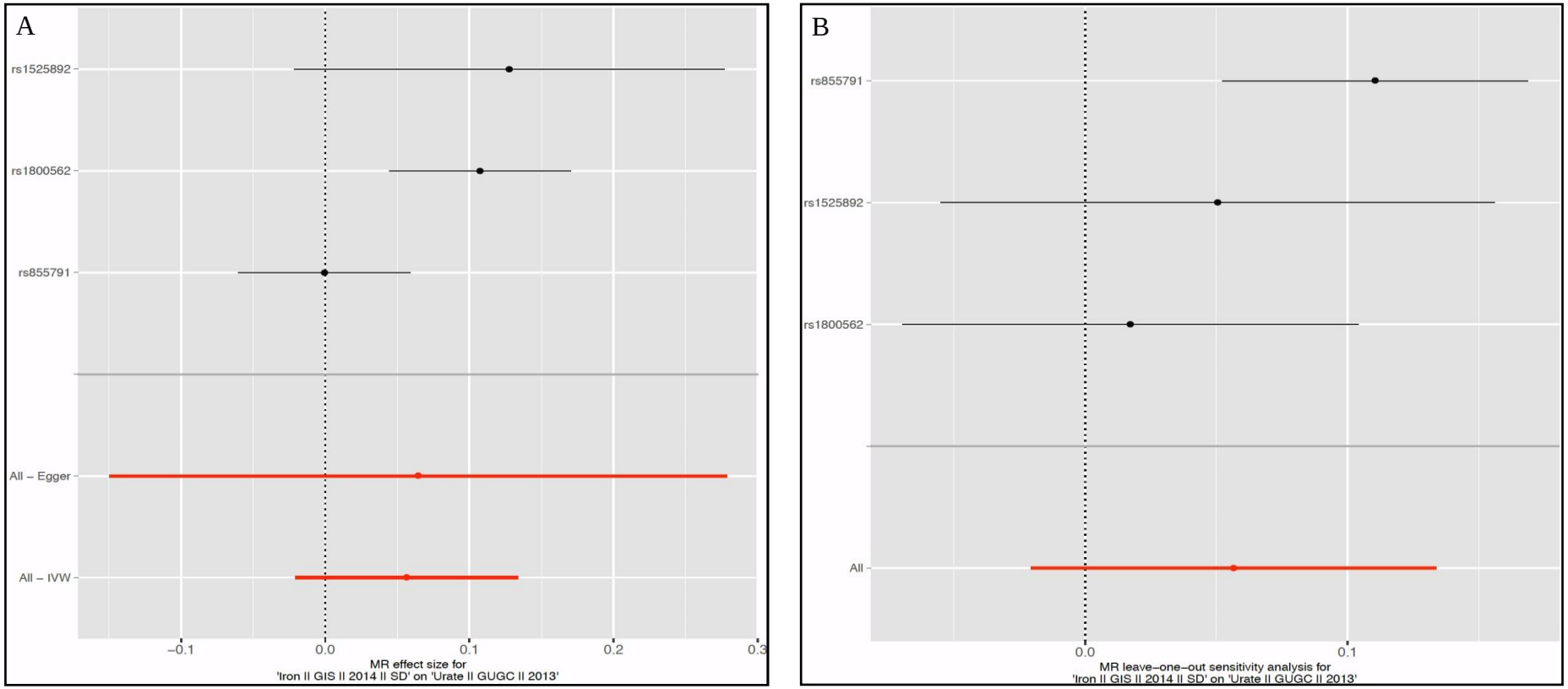


Figure 3.2: Forest graphs showing results of MR-analysis for three variants (SNPs) instrumented for serum iron as exposure (and serum urate as outcome) with their effect sizes on the x-axis and IDs on the y-axis (A) Graph showing results from the Wald ratio, 'Inverse-variance weighted (IVW)' and 'MR-Egger' methods; black dots indicate Wald ratio estimate for each SNP with black horizontal lines across the dots representing 95% confidence intervals (CI) corresponding to each estimate. 'All-IVW' red dot at the bottom indicates regression estimate from IVW meta-analysis with the red line across the dot representing 95% CI, while the 'All-Egger' red dot indicates regression estimate from MR-Egger sensitivity meta-analysis with the red line across the dot representing the 95% CI (B) Graph showing results from 'Leave-one-out' sensitivity analysis; each black dot indicates IVW estimate for all instrument SNPs excluding the one indicated on y-axis, while the black horizontal lines across the dots represents the 95% CI corresponding to each estimate. The 'All' red dot at the bottom indicates estimate from IVW analysis including all SNPs that are the same as 'All-IVW' in graph (A) for a comparison. The vertical dashed line in (A) and (B) represents effect size = zero (null).

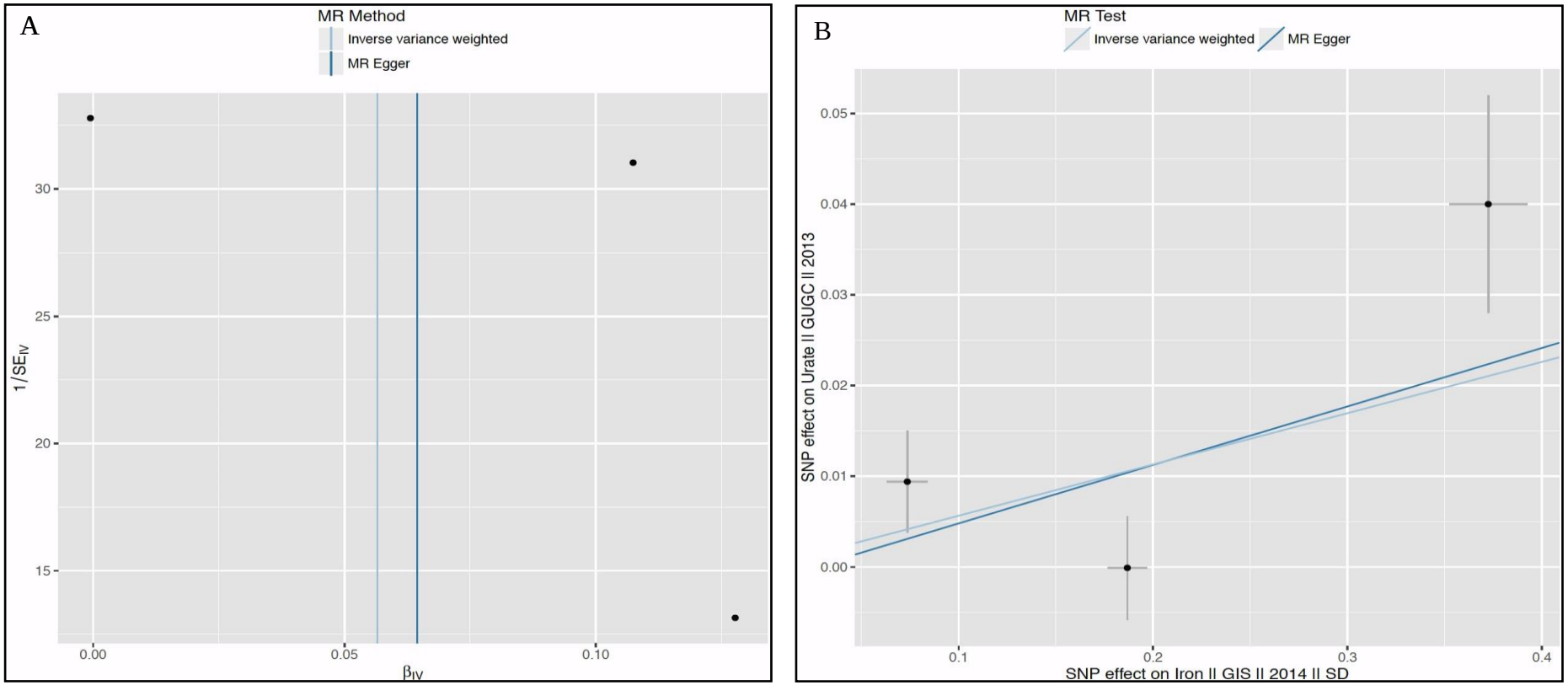


Figure 3.3: Graphs showing results of MR-analysis for three variants (SNPs) instrumented for serum iron as exposure on serum urate as an outcome via 'Inverse-variance weighted (IVW)' (indicated with light blue lines) and 'MR-Egger' (indicated with dark blue lines) methods (A) Funnel graph plotted to compare heterogeneity assessed via IVW and MR-Egger methods. Values on the x and y-axis represent effect estimates ($ß_{IV}$) and reciprocal standard errors ($1/SE_{IV}$) for the instrument variables (IVs indicated by black dots). Less precise estimates (lower values on y-axis) 'funnel' in as they increase in precision, while asymmetry in the funnel plot indicates directional heterogeneity between the estimates (B) Scatter plot to compare IVW and Egger-regression MR results: SNP effects (black dots with their 95% CI as vertical grey lines through the dots) on the outcome [urate (mg dL^{-1}) on y-axis] are plotted against SNP effects (black dots with their 95% CI as horizontal grey lines through the dots) on the exposure [iron (SD) on x-axis]. All SNPs with negative effects on the exposure are shown to be positive, with the sign of the effect on the outcome flipped. The slope of the line represents the causal association.

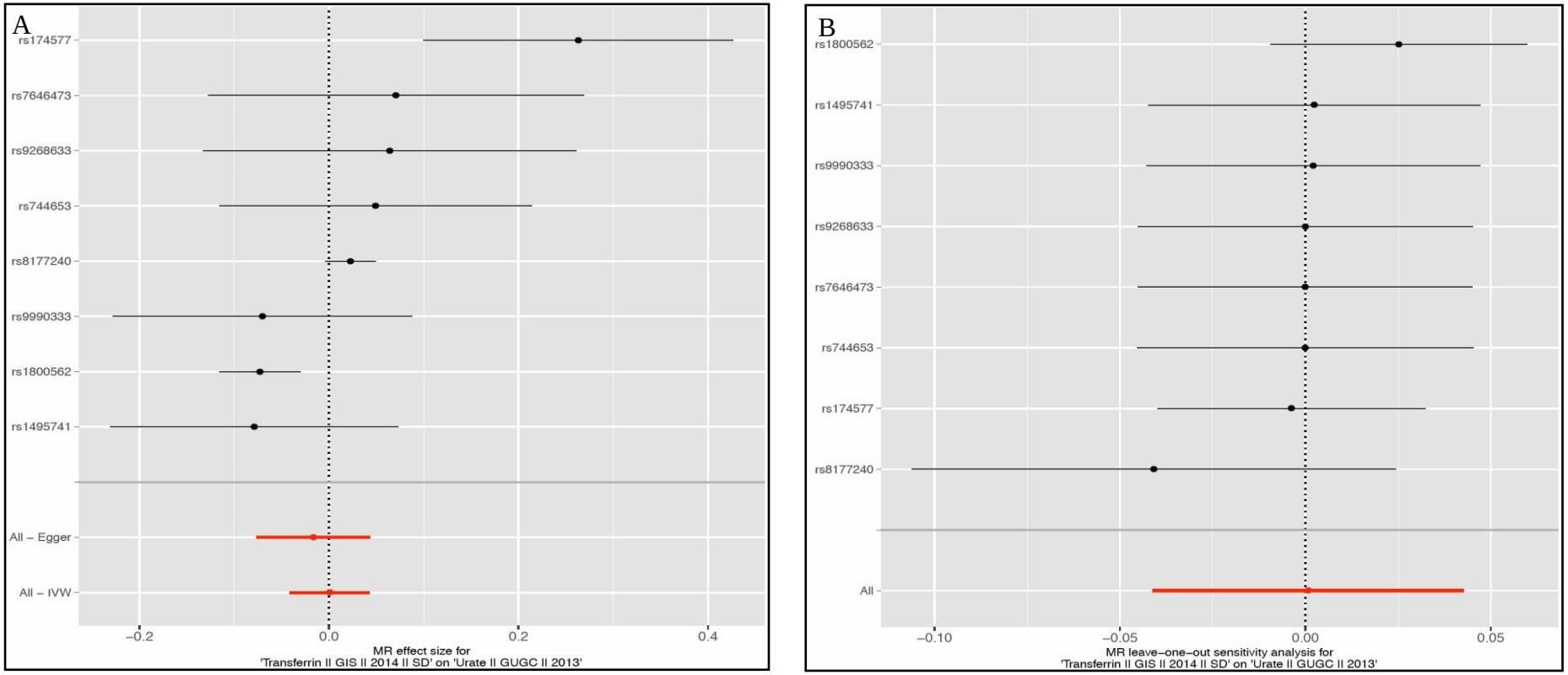


Figure 3.4: Forest graphs showing results of MR-analysis for eight variants (SNPs) instrumented for serum transferrin as exposure (and serum urate as outcome) with their effect sizes on the x-axis and IDs on the y-axis (A) Graph showing results from the Wald ratio, 'Inverse-variance weighted (IVW)' and 'MR-Egger' methods; black dots indicate Wald ratio estimate for each SNP with black horizontal lines across the dots representing 95% confidence intervals (CI) corresponding to each estimate. 'All-IVW' red dot at the bottom indicates regression estimate from IVW meta-analysis with the red line across the dot representing 95% CI, while the 'All-Egger' red dot indicates regression estimate from MR-Egger sensitivity meta-analysis with the red line across the dot representing the 95% CI (B) Graph showing results from 'Leave-one-out' sensitivity analysis; each black dot indicates IVW estimate for all instrument SNPs excluding the one indicated on y-axis, while the black horizontal lines across the dots represents the 95% CI corresponding to each estimate. The 'All' red dot at the bottom indicates estimate from IVW analysis including all SNPs that are the same as 'All-IVW' in graph (A) for a comparison. The vertical dashed line in (A) and (B) represents effect size = zero (null).

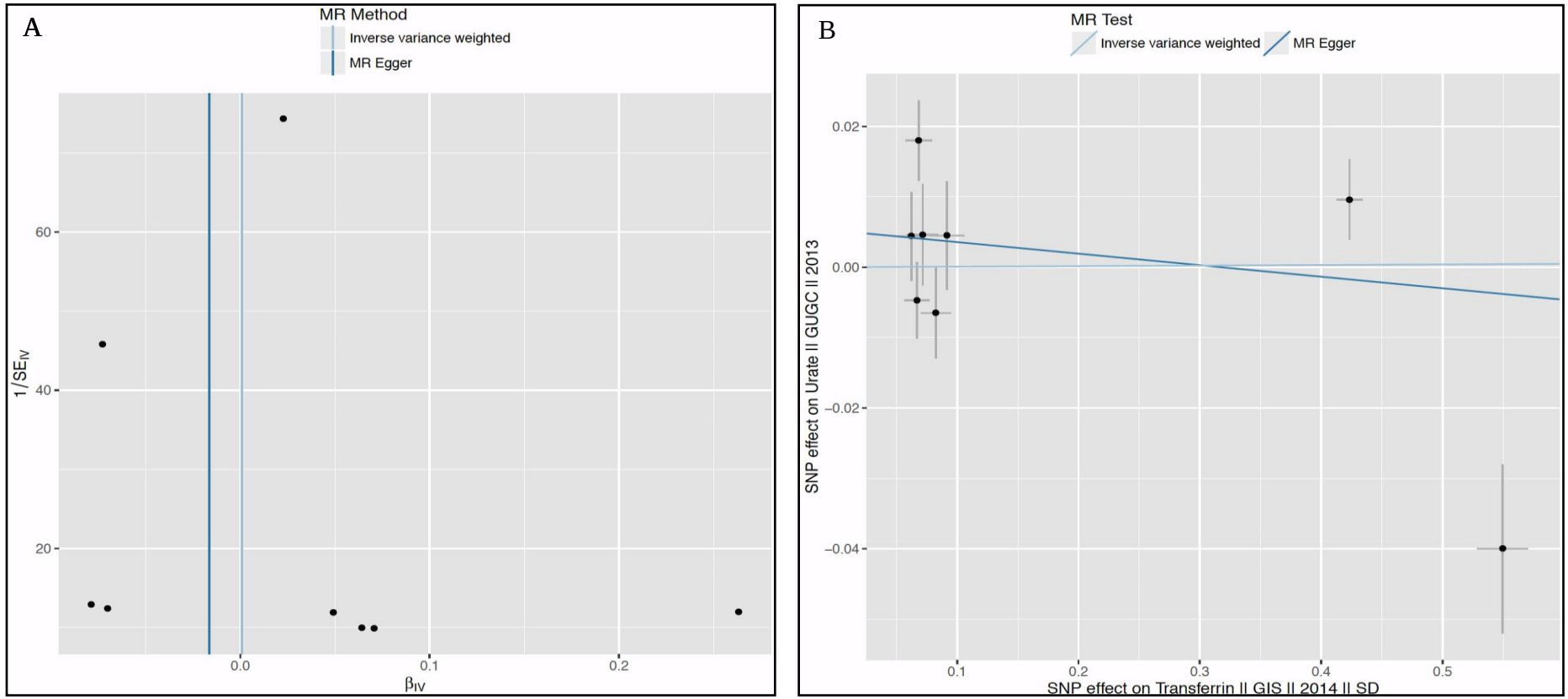


Figure 3.5: Graphs showing results of MR-analysis for eight variants (SNPs) instrumented for serum transferrin as exposure on serum urate as an outcome via 'Inverse-variance weighted (IVW)' (indicated with light blue lines) and 'MR-Egger' (indicated with dark blue lines) methods (A) Funnel graph plotted to compare heterogeneity assessed via IVW and MR-Egger methods. Values on the x and y-axis represent effect estimates (ßIV) and reciprocal standard errors (1/SEIV) for the instrument variables (IVs indicated by black dots). Less precise estimates (lower values on y-axis) 'funnel' in as they increase in precision, while asymmetry in the funnel plot indicates directional heterogeneity between the estimates (B) Scatter plot to compare IVW and Egger-regression MR results: SNP effects (black dots with their 95% CI as vertical grey lines through the dots) on the outcome [urate (mg dL^{-1}) on y-axis] are plotted against SNP effects (black dots with their 95% CI as horizontal grey lines through the dots) on the exposure [transferrin (SD) on x-axis]. All SNPs with negative effects on the exposure are shown to be positive, with the sign of the effect on the outcome flipped. The slope of the line represents the causal association.

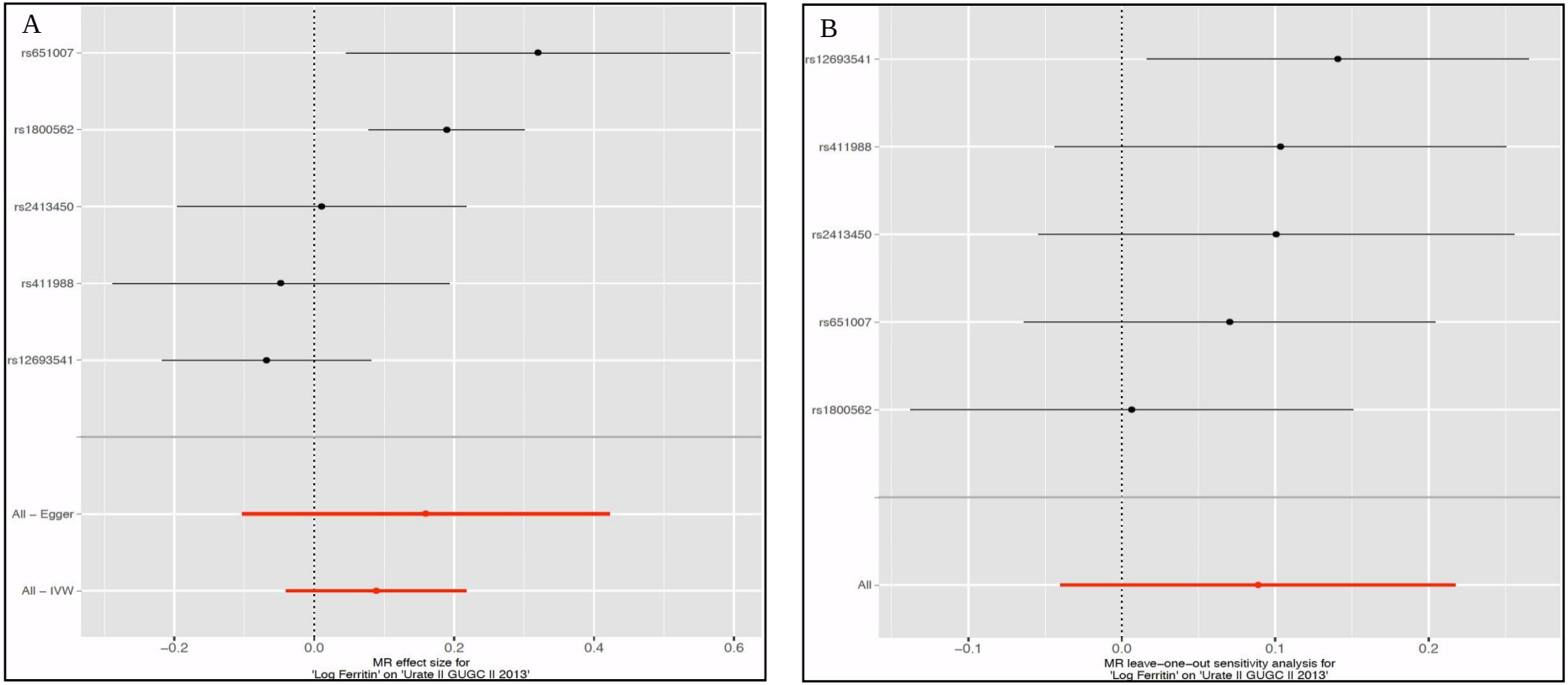


Figure 3.6: Forest graphs showing results of MR-analysis for five variants (SNPs) instrumented for serum ferritin (log) as exposure (and serum urate as outcome) with their effect sizes on the x-axis and IDs on the y-axis (A) Graph showing results from the Wald ratio, 'Inverse-variance weighted (IVW)' and 'MR-Egger' methods; black dots indicate Wald ratio estimate for each SNP with black horizontal lines across the dots representing 95% confidence intervals (CI) corresponding to each estimate. 'All-IVW' red dot at the bottom indicates regression estimate from IVW meta-analysis with the red line across the dot representing 95% CI, while the 'All-Egger' red dot indicates regression estimate from MR-Egger sensitivity meta-analysis with the red line across the dot representing the 95% CI (B) Graph showing results from 'Leave-one-out' sensitivity analysis; each black dot indicates IVW estimate for all instrument SNPs excluding the one indicated on y-axis, while the black horizontal lines across the dots represents the 95% CI corresponding to each estimate. The 'All' red dot at the bottom indicates estimate from IVW analysis including all SNPs that are the same as 'All-IVW' in graph (A) for a comparison. The vertical dashed line in (A) and (B) represents effect size = zero (null).

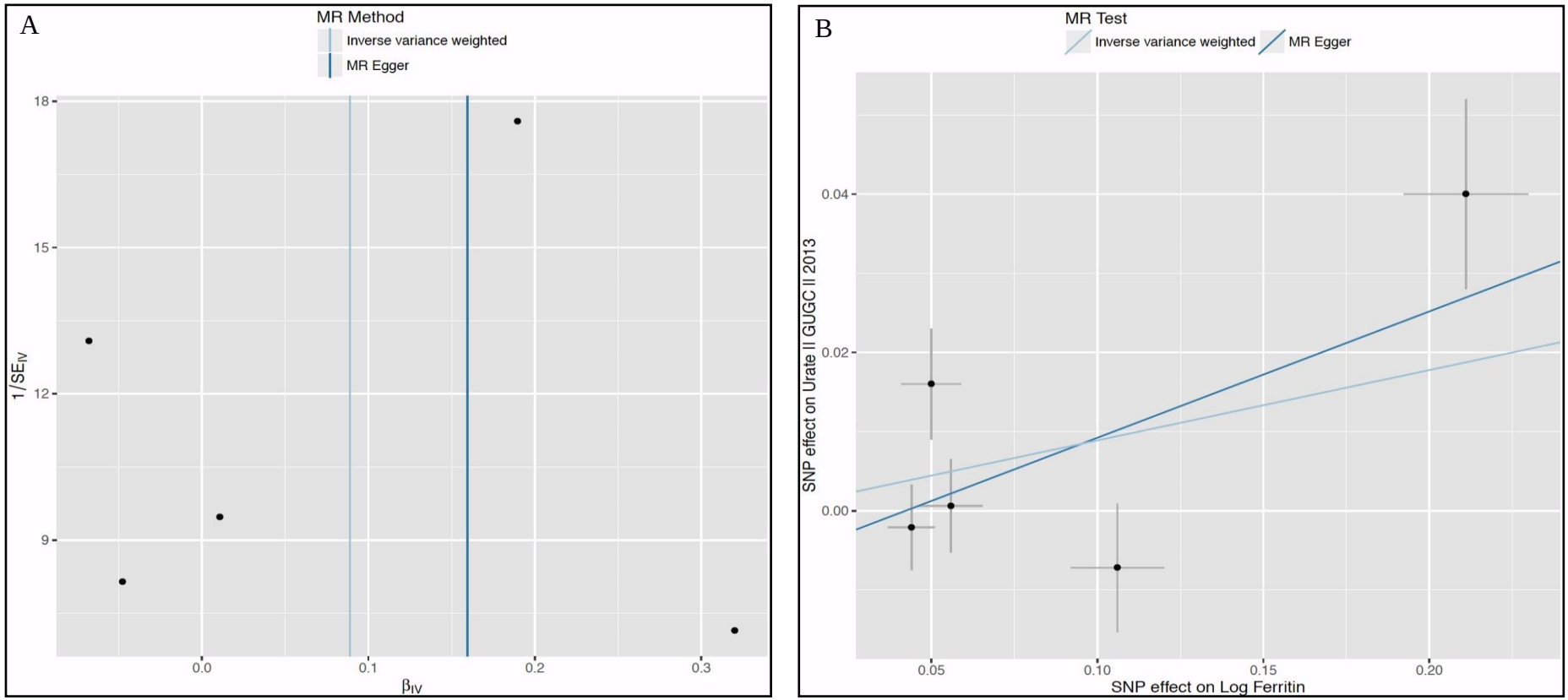


Figure 3.7: Graphs showing results of MR-analysis for five variants (SNPs) instrumented for serum ferritin (log) as exposure on serum urate as an outcome via 'Inverse-variance weighted (IVW)' (indicated with light blue lines) and 'MR-Egger' (indicated with dark blue lines) methods (A) Funnel graph plotted to compare heterogeneity assessed via IVW and MR-Egger methods. Values on the x and y-axis represent effect estimates (ßIV) and reciprocal standard errors (1/SEIV) for the instrument variables (IVs indicated by black dots). Less precise estimates (lower values on y-axis) 'funnel' in as they increase in precision, while asymmetry in the funnel plot indicates directional heterogeneity between the estimates (B) Scatter plot to compare IVW and Egger-regression MR results: SNP effects (black dots with their 95% CI as vertical grey lines through the dots) on the outcome [urate (mg dL^{-1}) on y-axis] are plotted against SNP effects (black dots with their 95% CI as horizontal grey lines through the dots) on the exposure [ferritin (log) (SD) on x-axis]. All SNPs with negative effects on the exposure are shown to be positive, with the sign of the effect on the outcome flipped. The slope of the line represents the causal association.

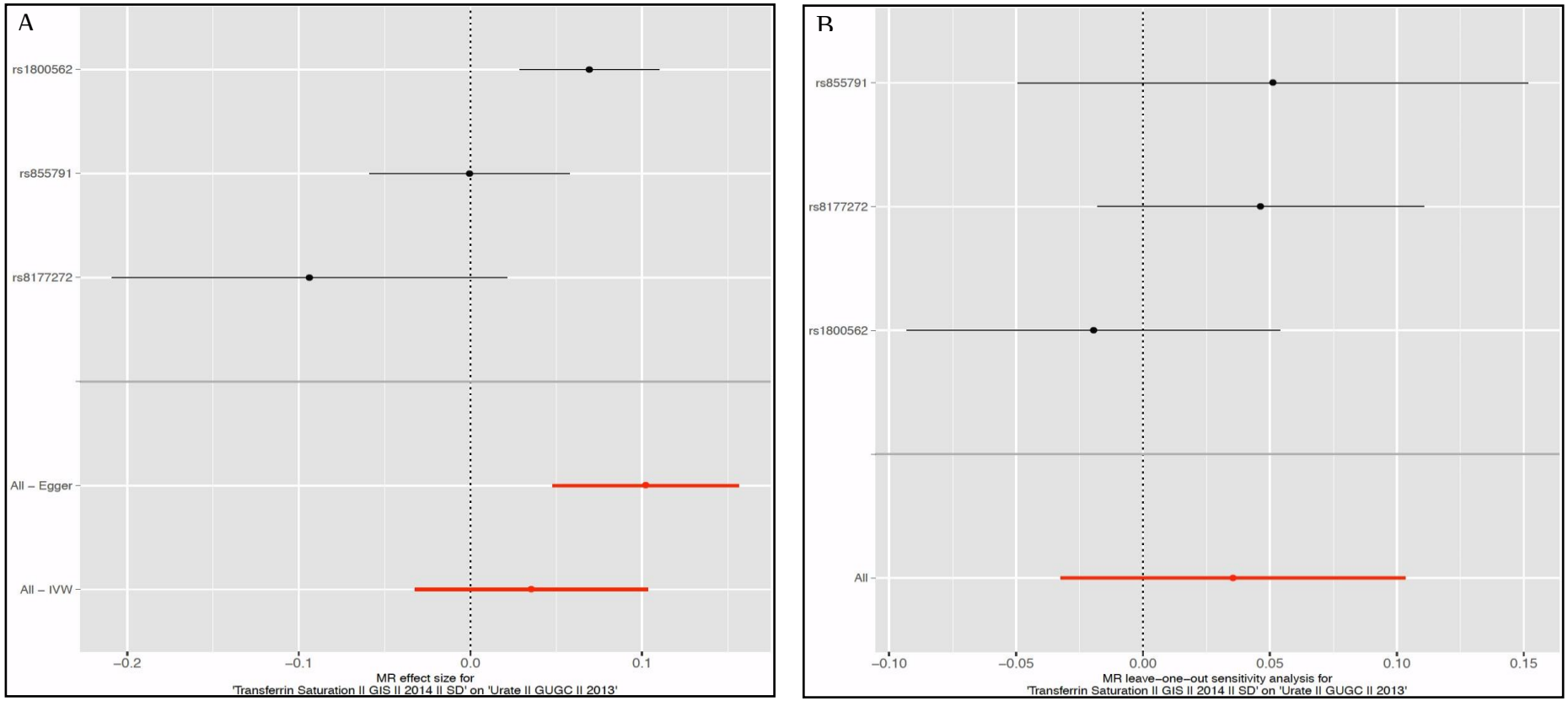


Figure 3.8: Forest graphs showing results of MR-analysis for three variants (SNPs) instrumented for serum transferrin saturation (TSAT) as exposure (and serum urate as outcome) with their effect sizes on the x-axis and IDs on the y-axis (A) Graph showing results from the Wald ratio, 'Inverse-variance weighted (IVW)' and 'MR-Egger' methods; black dots indicate Wald ratio estimate for each SNP with black horizontal lines across the dots representing 95% confidence intervals (CI) corresponding to each estimate. 'All-IVW' red dot at the bottom indicates regression estimate from IVW meta-analysis with the red line across the dot representing 95% CI, while the 'All-Egger' red dot indicates regression estimate from MR-Egger sensitivity meta-analysis with the red line across the dot representing the 95% CI (B) Graph showing results from 'Leave-one-out' sensitivity analysis; each black dot indicates IVW estimate for all instrument SNPs excluding the one indicated on y-axis, while the black horizontal lines across the dots represents the 95% CI corresponding to each estimate. The 'All' red dot at the bottom indicates estimate from IVW analysis including all SNPs that are the same as 'All-IVW' in graph (A) for a comparison. The vertical dashed line in (A) and (B) represents effect size = zero (null).

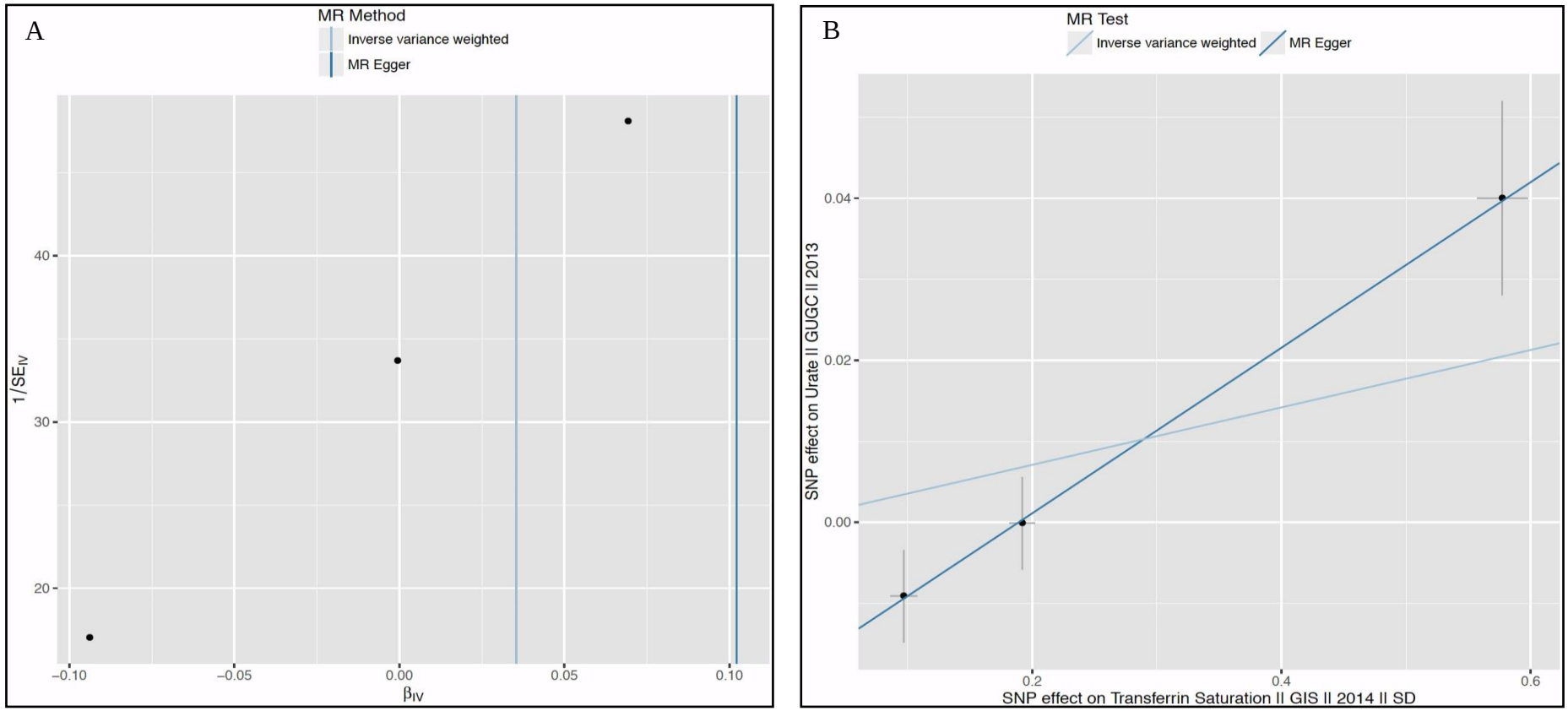


Figure 3.9: Graphs showing results of MR-analysis for three variants (SNPs) instrumented for serum transferrin saturation/TSAT as exposure on serum urate as an outcome via 'Inverse-variance weighted (IVW)' (indicated with light blue lines) and 'MR-Egger' (indicated with dark blue lines) methods (A) Funnel graph plotted to compare heterogeneity assessed via IVW and MR-Egger methods. Values on the x and y-axis represent effect estimates (ßIV) and reciprocal standard errors (1/SEIV) for the instrument variables (IVs indicated by black dots). Less precise estimates (lower values on y-axis) 'funnel' in as they increase in precision, while asymmetry in the funnel plot indicates directional heterogeneity between the estimates (B) Scatter plot to compare IVW and Egger-regression MR results: SNP effects (black dots with their 95% CI as vertical grey lines through the dots) on the outcome [urate (mg dL^{-1}) on y-axis] are plotted against SNP effects (black dots with their 95% CI as horizontal grey lines through the dots) on the exposure [TSAT (SD) on x-axis]. All SNPs with negative effects on the exposure are shown to be positive, with the sign of the effect on the outcome flipped. The slope of the line represents the causal association.

3.2.3.5 Urate as exposure for iron profile biomarkers as outcome

3.2.3.5.1 Urate as exposure for iron as outcome

Single variant two-sample MR using the Wald ratio method did not indicate a causal association between urate and iron for any instrument (*rs12498742*: ß = -0.032, *P* = 0.31; *rs1171614*: ß = 0.011, *P* = 0.943; *rs478607*: ß = 0.047, *P* = 0.872) (Table 3.5, Figure 3.10: A). No causal urate-iron relationship was observed in primary combined MR analysis using IVW (ß = -0.029, *P* = 0.33) or MR-Egger (ß = -0.043, *P* = 0.53) (Table 3.5, Figure 3.10: A). Heterogeneity between the effects across the variants was not significant for IVW (*Q-P* = 0.96) or MR-Egger (*Q-P* = 0.99) (Table 3.5, Figure 3.11: A). However, a sensitivity MR-Egger indicated an evidence of little pleiotropy (intercept: ß = 0.004, OR = 1.01, *P* = 0.77) for an association between urate and iron using all 3 SNPs (Table 3.6, Figure 3.11: B). Removing any variant in leave-one-out IVW meta-analysis also did not indicate any significant causal urate-iron relationship (Table 3.7, Figure 3.10: B). All MR ß-estimates are presented as an effect of a mg dL^{-1} increase in urate on iron (per SD unit). This analysis did not provide any evidence of a causal role of urate in changing blood iron levels.

3.2.3.5.2 Urate as exposure for transferrin as outcome

None of the 3 instruments indicated a causal urate-transferrin association when assessed via the Wald ratio method, except *rs478607* that showed only nominal significance (ß = -0.519, *P* = 0.058) (Table 3.5, Figure 3.12: A). Urate was also not causal for transferrin in combined IVW meta-analysis (ß = 0.106, *P* = 0.74) (Table 3.5, Figure 3.12: A). No significant heterogeneity (IVW: *Q-P* = 0.743, MR-Egger: *Q-P* = 0.803) was found for effect estimates across the instruments (Table 3.5, Figure 3.13: A). A sensitivity MR-Egger only indicated an evidence of little pleiotropy for an association between urate and transferrin using full list of 3 SNPs (intercept: ß = 0.0124, OR = 1.02, *P* = 0.48) (Table 3.6, Figure 3.13: B). The combined causal estimate via MR-Egger did not show a significant urate-transferrin association (ß = -0.030, *P* = 0.65) (Table 3.5, Figure 3.12: A). Leave-one-out permutation analysis also did not indicate any variant to be unproportionately altering the non-significant association in IVW meta-analysis (Table 3.7, Figure 3.12: B). All MR ß-estimates are presented as an effect of a mg dL^{-1} increase

in urate on transferrin (per SD unit). This analysis did not provide any evidence of a causal role of urate in changing blood transferrin levels.

3.2.3.5.3 Urate as exposure for ferritin as outcome

The Wald ratio method did not indicate a positive urate-ferritin relationship via any of the 3 instruments (*rs12498742*: ß = 0.003, *P* = 0.97; *rs1171614*: ß = 0.156, *P* = 0.33; *rs478607*: ß = 0.183, *P* = 0.302) (Table 3.5, Figure 3.14: A). The primary combined MR analysis using IVW (ß = 0.018, *P* = 0.65) or MR-Egger (ß = 0.056, *P* = 0.59) regression also did not indicate any causal urate-ferritin relationship (Table 3.5, Figure 3.14: A). Heterogeneity between the effects across the variants was not significant for IVW (*Q*-*P* = 0.361) or MR-Egger regression (*Q*-*P* = 0.089) (Table 3.5, Figure 3.15: A). A sensitivity MR-Egger indicated very little pleiotropy (intercept: ß = -0.007, OR = 0.99, *P* = 0.57) in association between urate and ferritin using the full list of 3 SNPs (Table 3.6, Figure 3.15: B). The sensitivity analysis using leave-one-out approach did not indicate any variant to be considerably changing the non-significant association in IVW meta-analysis (Table 3.7, Figure 3.14: B). All MR ß-estimates are presented as an effect of a mg dL^{-1} increase in urate on ferritin (log_{10}) (per SD unit). This analysis did not provide any evidence of a causal role of urate in changing ferritin levels.

3.2.3.5.4 Urate as exposure for transferrin saturation as outcome

No instrument showed a causal urate-TSAT association when assessed individually via the Wald ratio method (*rs12498742*: ß = -0.0501, *P* = 0.11; *rs1171614*: ß = 0.027, *P* = 0.86; *rs478607*: ß = -0.187, *P* = 0.52) (Table 3.5, Figure 3.16: A). Urate was also not causally associated with TSAT in combined IVW meta-analysis (ß = -0.051, *P* = 0.096) or MR-Egger sensitivity regression analysis (ß = -0.042, *P* = 0.542) (Table 3.5, Figure 3.16: A). No heterogeneity (IVW: *Q*-*P* = 0.94, MR-Egger: *Q*-*P* = 0.66) was found for effect estimates across the instruments (Table 3.5, Figure 3.17: A). A sensitivity MR-Egger, however, did indicate an evidence of little pleiotropy for an association between urate and TSAT using the full list of 3 SNPs (intercept: ß = -0.005, OR = 0.99, *P* = 0.85) (Table 3.6, Figure 3.17: B). Leave-one-out permutation analysis also did not indicate any variant to be unproportionately altering the non-significant association in IVW meta-analysis (Table 3.7, Figure 3.16: B). All MR ß-estimates are presented as an effect of a

mg dL^{-1} increase in urate on TSAT (per SD unit). This analysis did not provide any evidence of a causal role of urate in changing transferrin saturation levels.

Table 3.5 Association between urate and iron-related traits using two-sample Mendelian randomisation

MR-analysis Method	Phenotype		Gene/locus	Instrument variant	ß-estimate	SE	[95% CI]	*p*-causal	*Q-p*
	Exposure	Outcome							
Wald ratio	Urate	Iron	*SLC2A9*	*rs12498742*	-0.032	0.031	[-0.09 ; 0.03]	0.309	-
-	Urate	Iron	*SLC16A9*	*rs1171614*	0.011	0.158	[-0.30 ; 0.32]	0.943	-
-	Urate	Iron	*SLC22A12*	*rs478607*	0.046	0.291	[-0.52 ; 0.62]	0.872	-
All - IVW	Urate	Iron	-	All	-0.029	0.031	[-0.09 ; 0.03]	0.335	0.965
All - MR Egger	Urate	Iron	-	All	-0.043	0.047	[-0.14 ; 0.05]	0.534	0.996
Wald ratio	Urate	Transferrin	*SLC2A9*	*rs12498742*	0.022	0.029	[-0.04 ; 0.08]	0.461	-
-	Urate	Transferrin	*SLC16A9*	*rs1171614*	0.091	0.152	[-0.21 ; 0.39]	0.548	-
-	Urate	Transferrin	*SLC22A12*	*rs478607*	-0.519	0.274	[-1.06 ; 0.02]	0.058	-
All - IVW	Urate	Transferrin	-	All	0.011	0.031	[-0.05 ; 0.07]	0.736	0.743
All - MR Egger	Urate	Transferrin	-	All	-0.030	0.049	[-0.13 ; 0.07]	0.653	0.803
Wald ratio	Urate	Log Ferritin	*SLC2A9*	*rs12498742*	0.003	0.032	[-0.06 ; 0.07]	0.926	-
-	Urate	Log Ferritin	*SLC16A9*	*rs1171614*	0.156	0.163	[-0.16 ; 0.48]	0.336	-
-	Urate	Log Ferritin	*SLC22A12*	*rs478607*	0.183	0.302	[-0.41 ; 0.77]	0.544	-
All - IVW	Urate	Log Ferritin	-	All	0.018	0.041	[-0.06 ; 0.10]	0.656	0.361
All - MR Egger	Urate	Log Ferritin	-	All	0.057	0.077	[-0.09 ; 0.21]	0.595	0.089
Wald ratio	Urate	TSAT	*SLC2A9*	*rs12498742*	-0.050	0.031	[-0.11 ; 0.01]	0.109	-
-	Urate	TSAT	*SLC16A9*	*rs1171614*	-0.028	0.158	[-0.34 ; 0.28]	0.860	-
-	Urate	TSAT	*SLC22A12*	*rs478607*	-0.187	0.293	[-0.76 ; 0.39]	0.523	-
All - IVW	Urate	TSAT	-	All	-0.051	0.031	[-0.11 ; 0.01]	0.096	0.942
All - MR Egger	Urate	TSAT	-	All	-0.042	0.048	[-0.14 ; 0.05]	0.542	0.668

TSAT: Transferrin saturation, All - IVW: Meta-analysis using inverse-variance method, All - MR Egger: Mendelian randomisation using Egger regression, ß: Beta estimates, SE: Standard error, 95% CI: 95% confidence interval, *p*-causal: *p*-value using MR analysis, *Q-p*: Cochran's heterogeneity test *p*-value for heterogeneity.

Table 3.6: Results of horizontal pleiotropy for two-sample MR-Egger test

Phenotype		Horizontal pleiotropy-MR Egger		
Exposure	**Outcome**	**Egger intercept-ß (OR)**	**SE**	***p*-value**
Urate	Iron	0.0042 (1.004)	0.0114	0.772
Urate	Transferrin	0.0124 (1.012)	0.0117	0.481
Urate	Log Ferritin	-0.0118 (0.988)	0.0183	0.634
Urate	TSAT	-0.0026 (0.997)	0.0114	0.855

TSAT: Transferrin saturation, MR Egger: Mendelian randomisation using Egger regression, ß (OR): beta estimates (odds ratio for ß), SE: standard error, *p*-value: *p*-value for directional pleiotropy.

Table 3.7: Results of leave-one-out sensitivity analysis for association between urate and iron-related traits using two-sample Mendelian randomisation

Phenotype		Instrument variant excluded from IVW analysis	ß-estimate	[95% CI]	*p*-causal
Exposure	**Outcome**				
Urate	Iron	*rs12498742*	0.094	[-0.25 ; 0.29]	0.888
Urate	Iron	*rs1171614*	-0.031	[-0.09 ; 0.03]	0.319
Urate	Iron	*rs478607*	-0.030	[-0.09 ; 0.03]	0.324
Urate	Iron	All	-0.029	[-0.09 ; 0.03]	0.335
Urate	Transferrin	*rs12498742*	0.162	[-0.12 ; 0.44]	0.256
Urate	Transferrin	*rs1171614*	0.004	[-0.06 ; 0.07]	0.876
Urate	Transferrin	*rs478607*	0.008	[-0.05 ; 0.07]	0.782
Urate	Transferrin	All	0.010	[-0.05 ; 0.07]	0.735
Urate	Log Ferritin	*rs12498742*	-0.052	[-0.56 ; 0.45]	0.841
Urate	Log Ferritin	*rs1171614*	0.015	[-0.10 ; 0.13]	0.786
Urate	Log Ferritin	*rs478607*	0.024	[-0.03 ; 0.08]	0.402
Urate	Log Ferritin	All	0.018	[-0.06 ; 0.10]	0.656
Urate	TSAT	*rs12498742*	-0.064	[-0.34 ; 0.21]	0.647
Urate	TSAT	*rs1171614*	-0.052	[-0.11 ; 0.01]	0.097
Urate	TSAT	*rs478607*	-0.049	[-0.11 ; 0.01]	0.109
Urate	TSAT	All	-0.051	[-0.11 ; 0.01]	0.096

TSAT: Transferrin saturation, IVW: meta-analysis using inverse-variance method, ß: beta estimate, 95% CI: 95% confidence interval, *p*-causal: *p*-value using IVW meta-analysis.

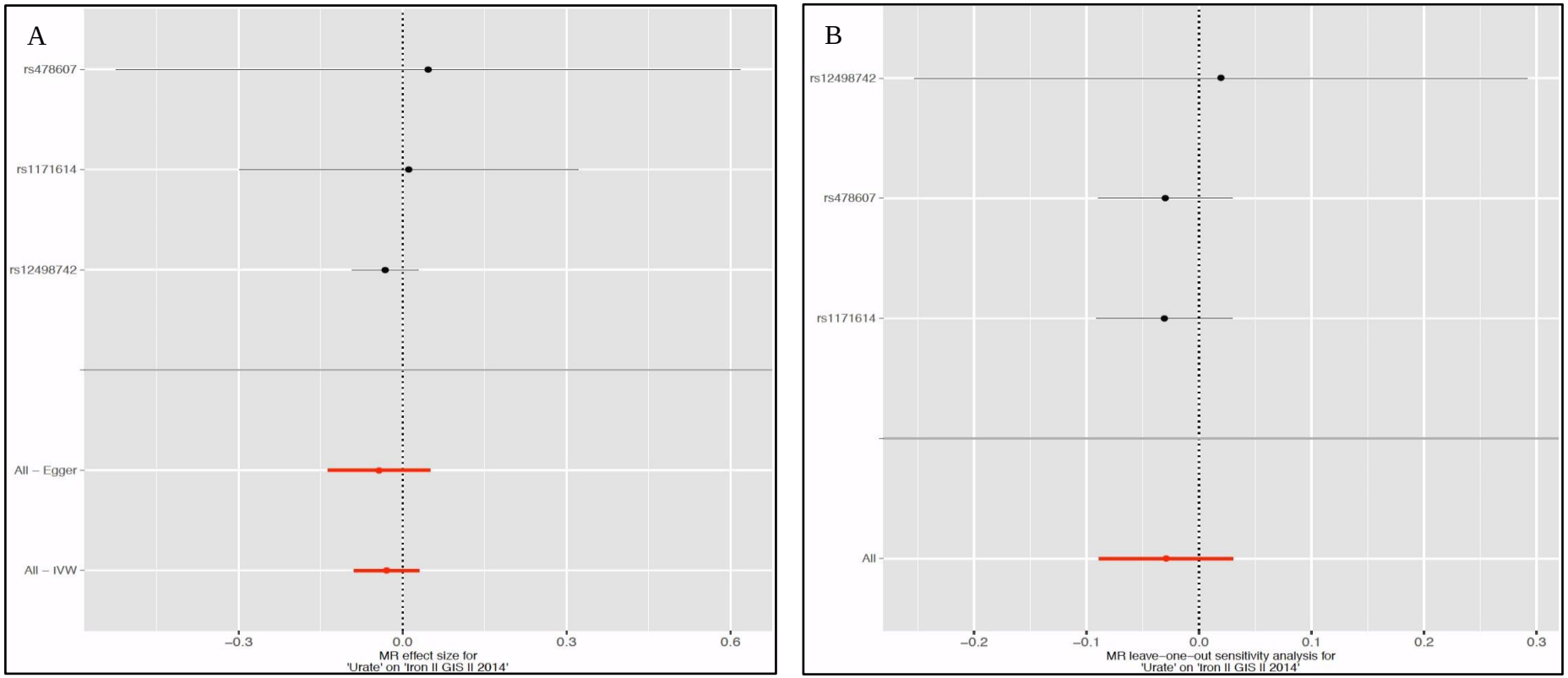


Figure 3.10: Forest graphs showing results of MR-analysis for three variants (SNPs) instrumented for serum urate as exposure (and serum iron as outcome) with their effect sizes on the x-axis and IDs on the y-axis (A) Graph showing results from the Wald ratio, 'Inverse-variance weighted (IVW)' and 'MR-Egger' methods; black dots indicate Wald ratio estimate for each SNP with black horizontal lines across the dots representing 95% confidence intervals (CI) corresponding to each estimate. 'All-IVW' red dot at the bottom indicates regression estimate from IVW meta-analysis with the red line across the dot representing 95% CI, while the 'All-Egger' red dot indicates regression estimate from MR-Egger sensitivity meta-analysis with the red line across the dot representing the 95% CI (B) Graph showing results from 'Leave-one-out' sensitivity analysis; each black dot indicates IVW estimate for all instrument SNPs excluding the one indicated on y-axis, while the black horizontal lines across the dots represents the 95% CI corresponding to each estimate. The 'All' red dot at the bottom indicates estimate from IVW analysis including all SNPs that are the same as 'All-IVW' in graph (A) for a comparison. The vertical dashed line in (A) and (B) represents effect size = zero (null).

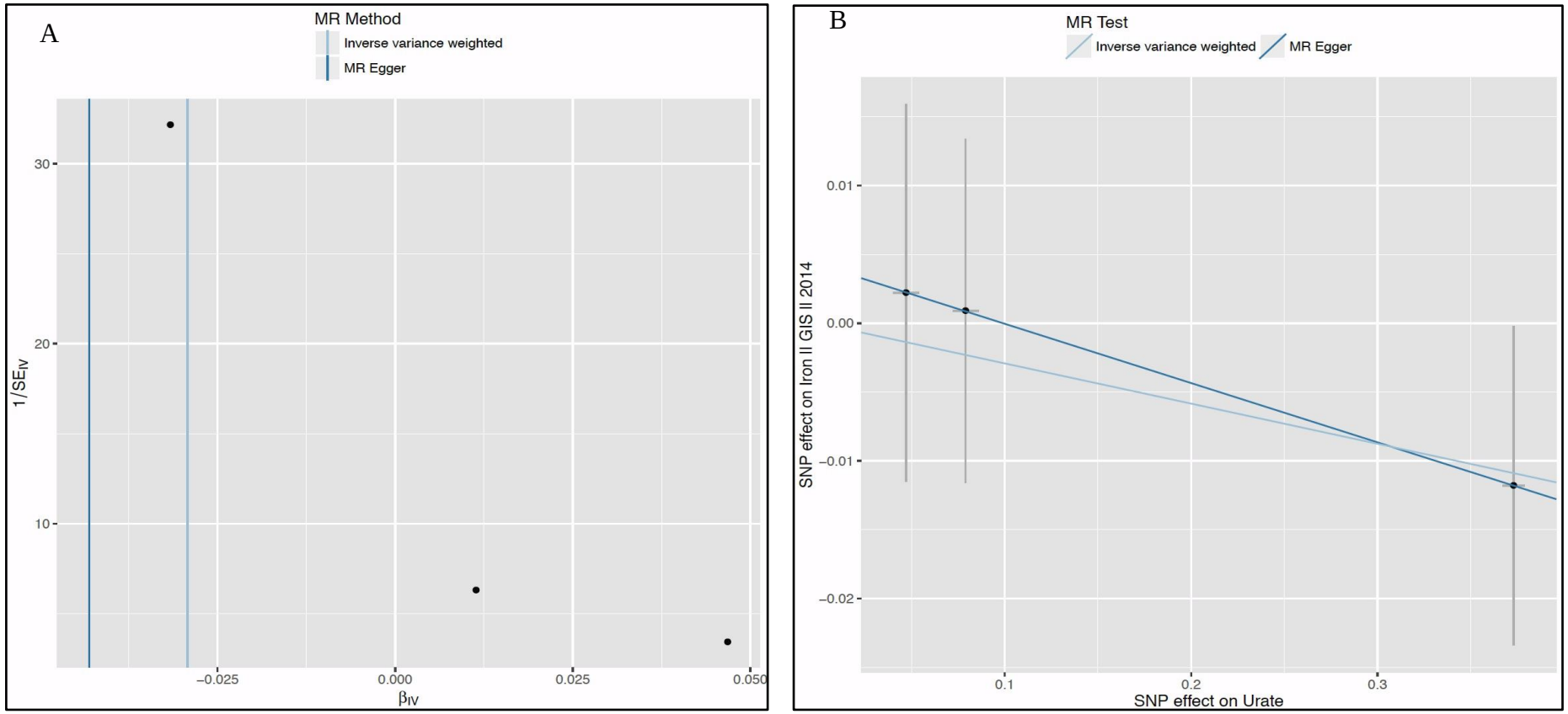


Figure 3.11: Graphs showing results of MR-analysis for three variants (SNPs) instrumented for serum urate as exposure on serum iron as an outcome via 'Inverse-variance weighted (IVW)' (indicated with light blue lines) and 'MR-Egger' (indicated with dark blue lines) methods (A) Funnel graph plotted to compare heterogeneity assessed via the IVW and MR-Egger methods. Values on the x and the y-axis represent effect estimates (ßIV) and reciprocal standard errors (1/SEIV) for the instrument variables (IVs indicated by black dots). Less precise estimates (lower values on y-axis) 'funnel' in as they increase in precision, while asymmetry in the funnel plot indicates directional heterogeneity between the estimates (B) Scatter plot to compare the IVW and Egger-regression MR results: SNP effects (black dots with their 95% CI as vertical grey lines through the dots) on the outcome [iron (SD) on the y-axis] are plotted against SNP effects (black dots with their 95% CI as horizontal grey lines through the dots) on the exposure [urate (mg dL^{-1}) on the x-axis]. All SNPs with negative effects on the exposure are shown to be positive, with the sign of the effect on the outcome flipped. The slope of the line represents the causal association.

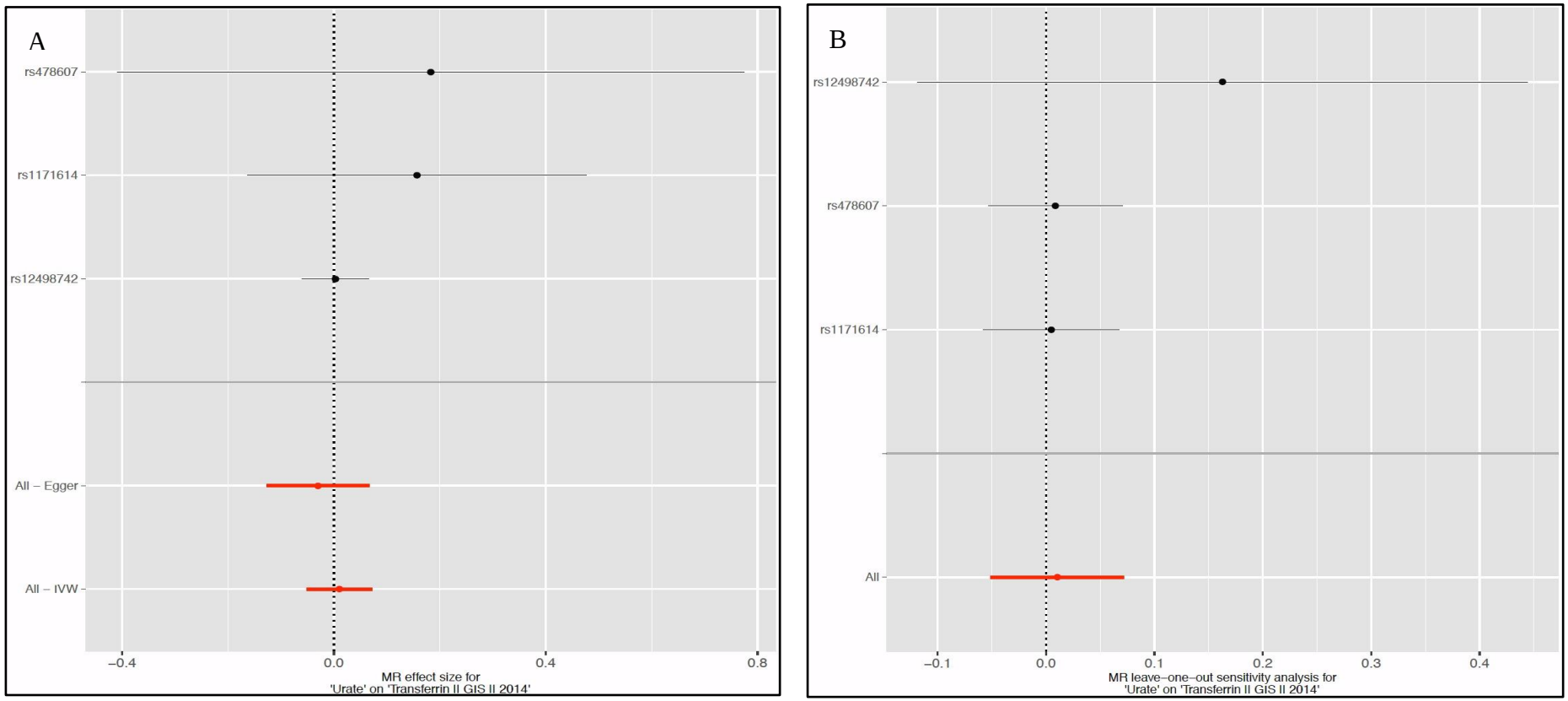


Figure 3.12: Forest graphs showing results of MR-analysis for three variants (SNPs) instrumented for serum urate as exposure (and serum transferrin as outcome) with their effect sizes on the x-axis and IDs on the y-axis (A) Graph showing results from the Wald ratio, 'Inverse-variance weighted (IVW)' and 'MR-Egger' methods; black dots indicate Wald ratio estimate for each SNP with black horizontal lines across the dots representing 95% confidence intervals (CI) corresponding to each estimate. 'All-IVW' red dot at the bottom indicates regression estimate from IVW meta-analysis with the red line across the dot representing 95% CI, while the 'All-Egger' red dot indicates regression estimate from MR-Egger sensitivity meta-analysis with the red line across the dot representing the 95% CI (B) Graph showing results from 'Leave-one-out' sensitivity analysis; each black dot indicates IVW estimate for all instrument SNPs excluding the one indicated on y-axis, while the black horizontal lines across the dots represents the 95% CI corresponding to each estimate. The 'All' red dot at the bottom indicates estimate from IVW analysis including all SNPs that are the same as 'All-IVW' in graph (A) for a comparison. The vertical dashed line in (A) and (B) represents effect size = zero (null).

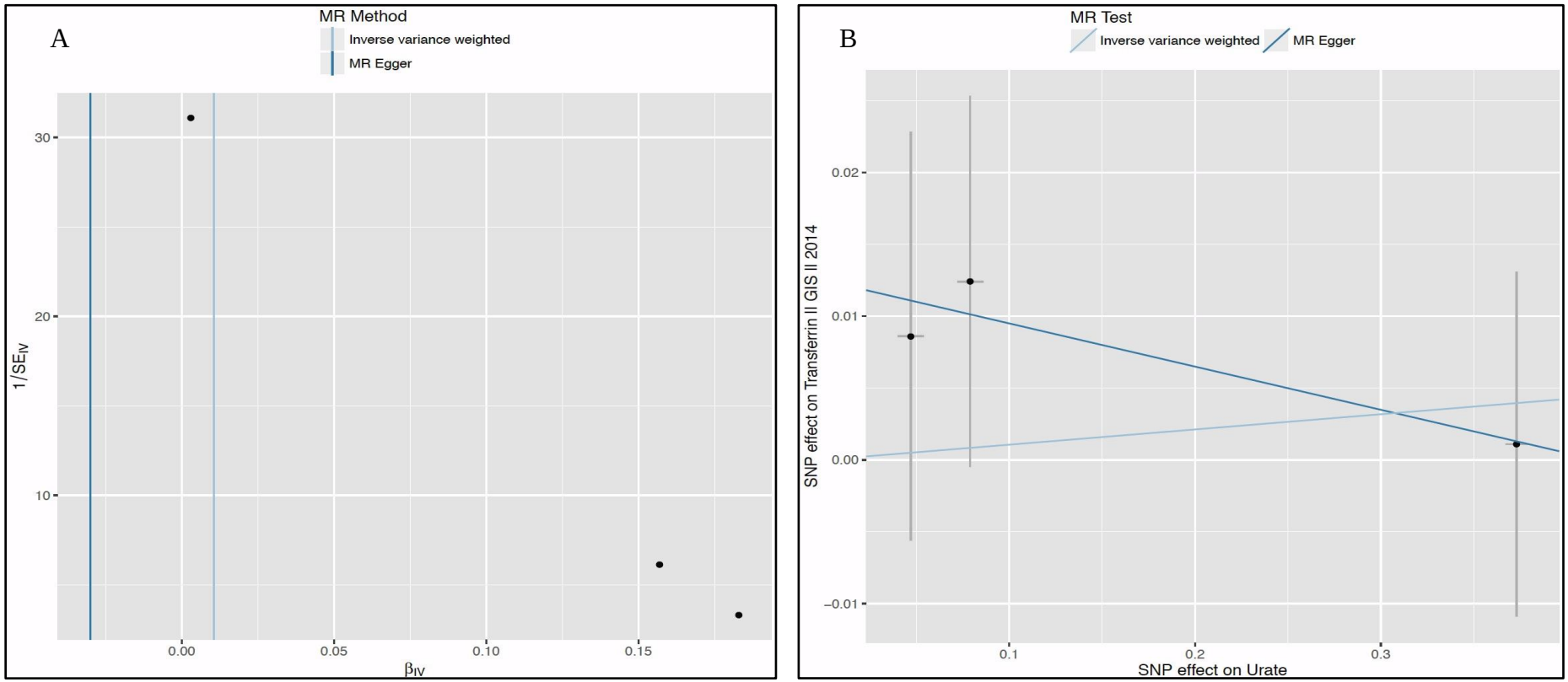


Figure 3.13: Graphs showing results of MR-analysis for three variants (SNPs) instrumented for serum urate as exposure on serum transferrin as an outcome via 'Inverse-variance weighted (IVW)' (indicated with light blue lines) and 'MR-Egger' (indicated with dark blue lines) methods (A) Funnel graph plotted to compare heterogeneity assessed via IVW and MR-Egger methods. Values on the x and the y-axis represent effect estimates (ßIV) and reciprocal standard errors (1/SEIV) for the instrument variables (IVs indicated by black dots). Less precise estimates (lower values on y-axis) 'funnel' in as they increase in precision, while asymmetry in the funnel plot indicates directional heterogeneity between the estimates (B) Scatter plot to compare IVW and Egger-regression MR results: SNP effects (black dots with their 95% CI as vertical grey lines through the dots) on the outcome [transferrin (SD) on the y-axis] are plotted against SNP effects (black dots with their 95% CI as horizontal grey lines through the dots) on the exposure [urate (mg dL^{-1}) on the x-axis]. All SNPs with negative effects on the exposure are shown to be positive, with the sign of the effect on the outcome flipped. The slope of the line represents the causal association.

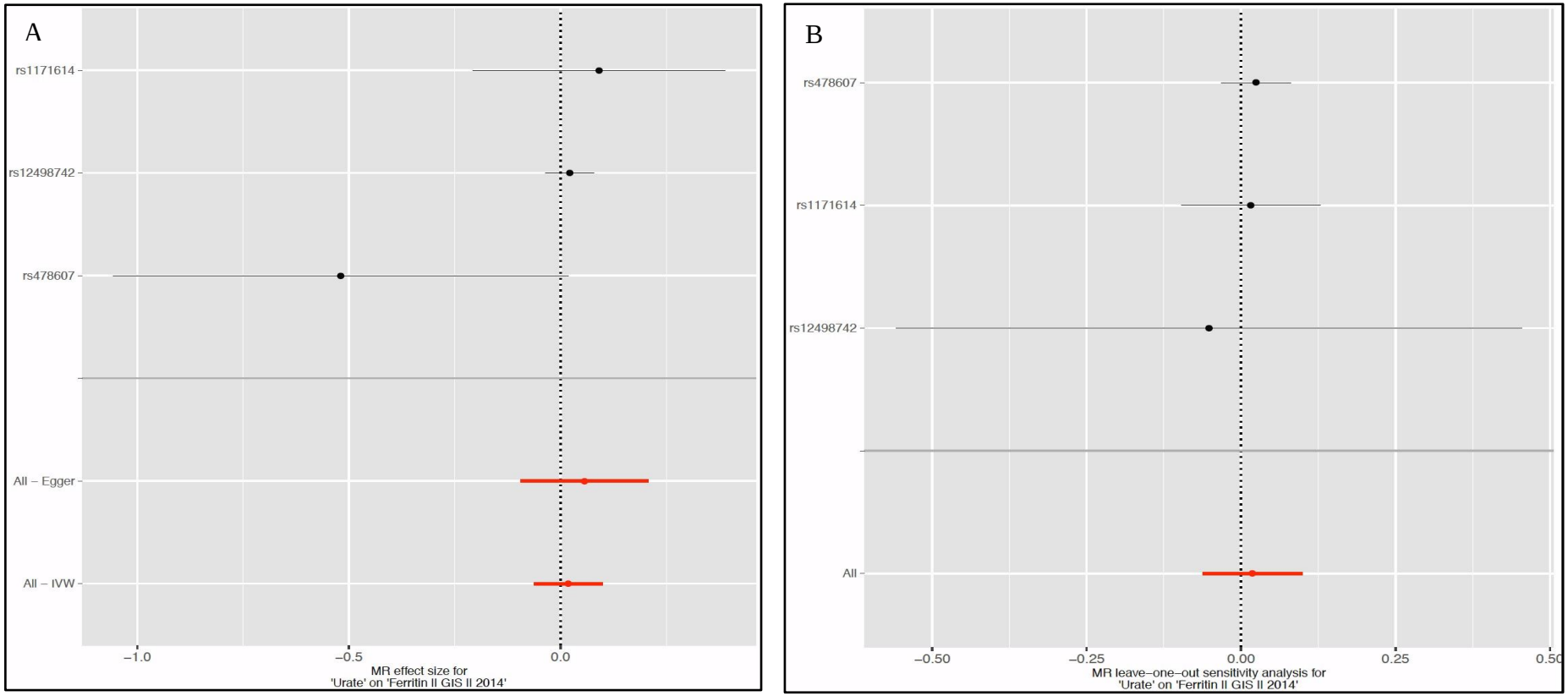


Figure 3.14: Forest graphs showing results of MR-analysis for three variants (SNPs) instrumented for serum urate as exposure (and serum ferritin (log) as outcome) with their effect sizes on the x-axis and IDs on the y-axis (A) Graph showing results from the Wald ratio, 'Inverse-variance weighted (IVW)' and 'MR-Egger' methods; black dots indicate Wald ratio estimate for each SNP with black horizontal lines across the dots representing 95% confidence intervals (CI) corresponding to each estimate. 'All-IVW' red dot at the bottom indicates regression estimate from IVW meta-analysis with the red line across the dot representing 95% CI, while the 'All-Egger' red dot indicates regression estimate from MR-Egger sensitivity meta-analysis with the red line across the dot representing the 95% CI (B) Graph showing results from 'Leave-one-out' sensitivity analysis; each black dot indicates IVW estimate for all instrument SNPs excluding the one indicated on y-axis, while the black horizontal lines across the dots represents the 95% CI corresponding to each estimate. The 'All' red dot at the bottom indicates estimate from IVW analysis including all SNPs that are the same as 'All-IVW' in graph (A) for a comparison. The vertical dashed line in (A) and (B) represents effect size = zero (null).

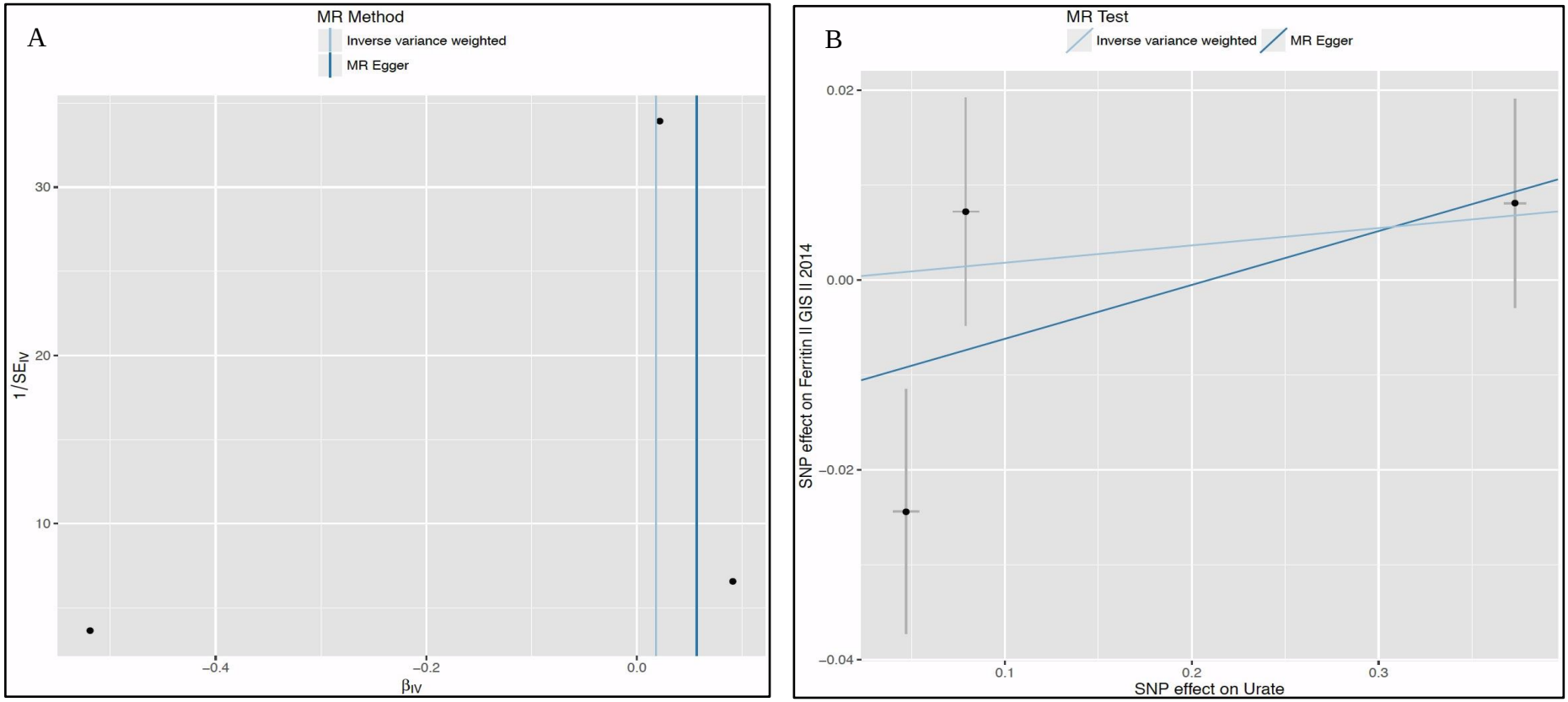


Figure 3.15: Graphs showing results of MR-analysis for three variants (SNPs) instrumented for serum urate as exposure on serum ferritin (log) as an outcome via 'Inverse-variance weighted (IVW)' (indicated with light blue lines) and 'MR-Egger' (indicated with dark blue lines) methods (A) Funnel graph plotted to compare heterogeneity assessed via IVW and MR-Egger methods. Values on the x and the y-axis represent effect estimates (ßIV) and reciprocal standard errors (1/SEIV) for the instrument variables (IVs indicated by black dots). Less precise estimates (lower values on y-axis) 'funnel' in as they increase in precision, while asymmetry in the funnel plot indicates directional heterogeneity between the estimates (B) Scatter plot to compare IVW and Egger-regression MR results: SNP effects (black dots with their 95% CI as vertical grey lines through the dots) on the outcome [Ferritin (log) (SD) on the y-axis] are plotted against SNP effects (black dots with their 95% CI as horizontal grey lines through the dots) on the exposure [urate (mg dL^{-1}) on the x-axis]. All SNPs with negative effects on the exposure are shown to be positive, with the sign of the effect on the outcome flipped. The slope of the line represents the causal association.

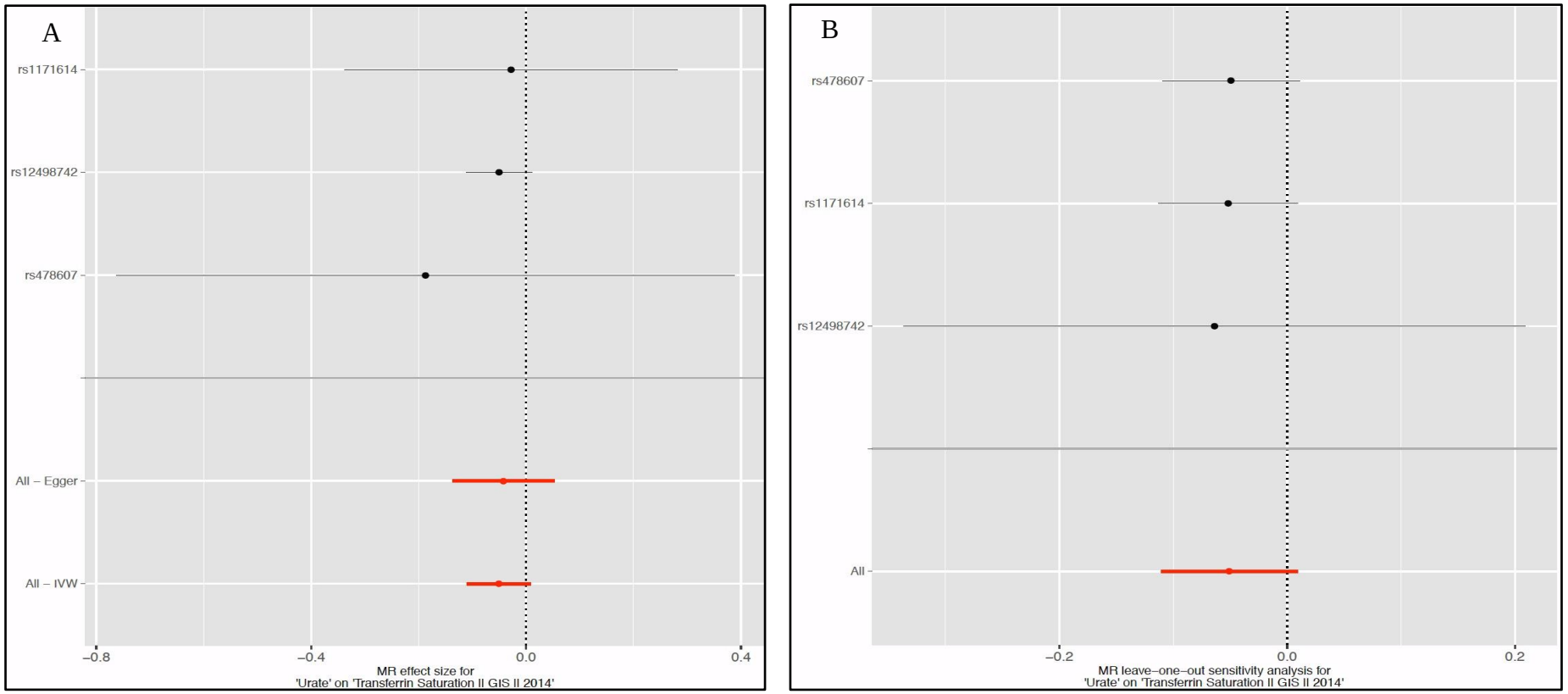


Figure 3.16: Forest graphs showing results of MR-analysis for three variants (SNPs) instrumented for serum urate as exposure (and serum transferrin saturation (TSAT) as outcome) with their effect sizes on the x-axis and IDs on the y-axis (A) Graph showing results from the Wald ratio, 'Inverse-variance weighted (IVW)' and 'MR-Egger' methods; black dots indicate Wald ratio estimate for each SNP with black horizontal lines across the dots representing 95% confidence intervals (CI) corresponding to each estimate. 'All-IVW' red dot at the bottom indicates regression estimate from IVW meta-analysis with the red line across the dot representing 95% CI, while the 'All-Egger' red dot indicates regression estimate from MR-Egger sensitivity meta-analysis with the red line across the dot representing the 95% CI (B) Graph showing results from 'Leave-one-out' sensitivity analysis; each black dot indicates IVW estimate for all instrument SNPs excluding the one indicated on y-axis, while the black horizontal lines across the dots represents the 95% CI corresponding to each estimate. The 'All' red dot at the bottom indicates estimate from IVW analysis including all SNPs that are the same as 'All-IVW' in graph (A) for a comparison. The vertical dashed line in (A) and (B) represents effect size = zero (null).

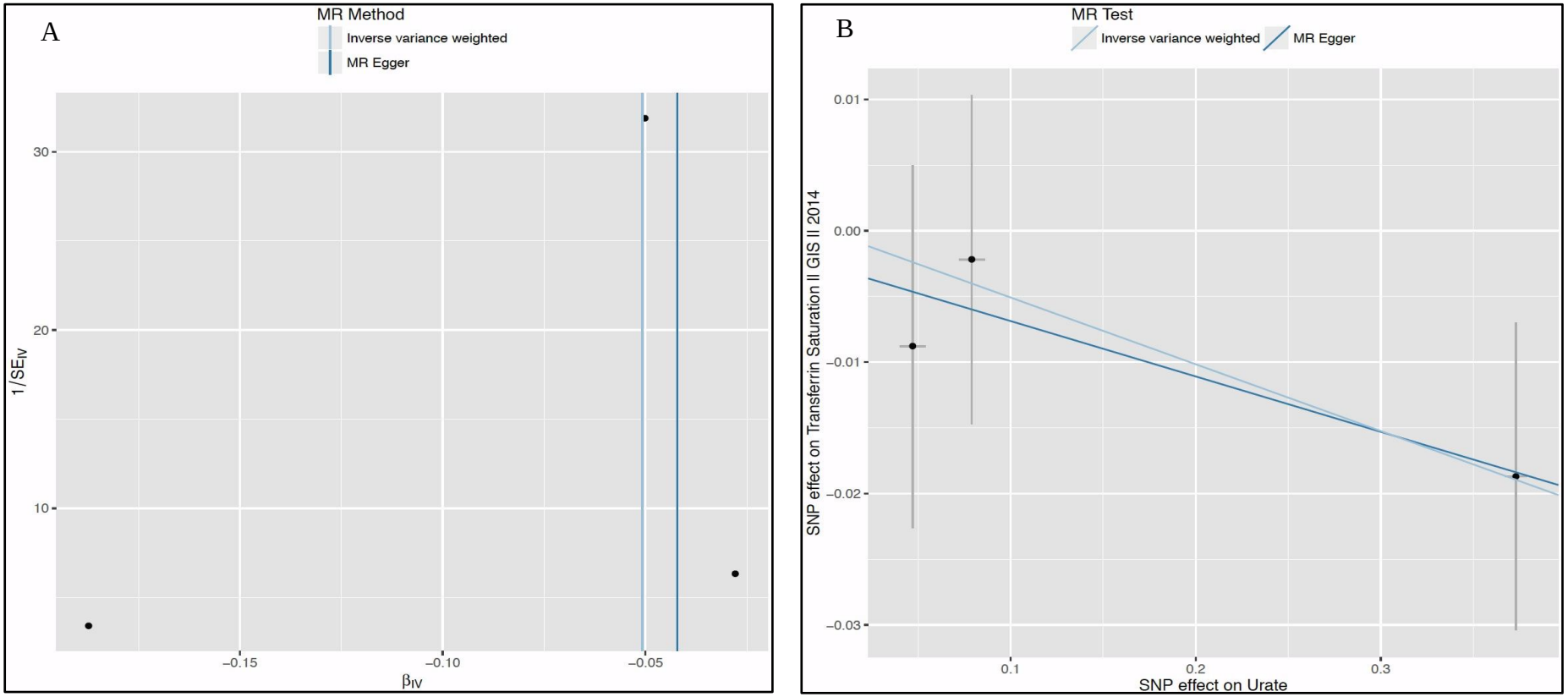


Figure 3.17: Graphs showing results of MR-analysis for three variants (SNPs) instrumented for serum urate as exposure on serum transferrin saturation as an outcome via 'Inverse-variance weighted (IVW)' (indicated with light blue lines) and 'MR-Egger' (indicated with dark blue lines) methods (A) Funnel graph plotted to compare heterogeneity assessed via IVW and MR-Egger methods. Values on the x and the y-axis represent effect estimates (ßIV) and reciprocal standard errors (1/SEIV) for the instrument variables (IVs indicated by black dots). Less precise estimates (lower values on y-axis) 'funnel' in as they increase in precision, while asymmetry in the funnel plot indicates directional heterogeneity between the estimates (B) Scatter plot to compare IVW and Egger-regression MR results: SNP effects (black dots with their 95% CI as vertical grey lines through the dots) on the outcome [transferrin saturation (SD) on the y-axis] are plotted against SNP effects (black dots with their 95% CI as horizontal grey lines through the dots) on the exposure [urate (mg dL^{-1}) on the x-axis]. All SNPs with negative effects on the exposure are shown to be positive, with the sign of the effect on the outcome flipped. The slope of the line represents the causal association.

SECTION 3.3 DISCUSSION

Two-sample single-instrument Mendelian randomisation using GWAS summary-level data indicated a causal effect of exposure to iron or its biomarkers in increasing urate concentration. Combined meta-analyses, using the selective set of iron-related instruments, also indicated a causal association of an exposure of iron and ferritin in increasing serum urate concentrations.

3.3.1 Single-instrument MR and pleiotropy

Instrumenting iron biomarkers as exposure for urate as outcome in single-instrument MR via the Wald ratio method showed a significant causal relationship between all four iron profile biomarkers (iron, transferrin, ferritin and TSAT) and urate using a missense SNP, *HFE*: *rs1800562*, as the instrument (Table 3.2). This variant has been described as a *cis*-eQTL for tripartite motif-containing protein 8 (*TRIM38*), histone gene cluster 1 (*HIST1H-2AC, 2BD, 4A* and *4H*) and mostly for lincRNA of unknown function (*U91328.19* and *22*) (www.archive.broadinstitute.org and www.gtexportal.org). On the basis of evidence available in the literature, *TRIM38* was found to be involved in production of E3 ubiquitin-protein, which negatively regulates the immune response via toll-like receptors (TLRs) (Zhao *et al.*, 2012), although the protein doesn't seem to be involved in urate metabolism (Jeong *et al.*, 2009). Overexpression of histone cluster 1 family genes has been reported in human meningioma (Pérez-Magán *et al.*, 2010), while no evidence has been found for a direct involvement in any renal function or urate metabolism. The variant *rs1800562* is present within the *HFE* gene on chromosome 6 in close proximity to known urate loci *SLC17A1-3* and therefore its causality may be questionable. Although, the variant *rs1800562* was shown to have a low LD ($r^2 < 0.6$) with only one variant within the *SLC17A1* gene in Köttgen *et al.* (2013) data (Appendix B Figure 3.2), the causal signal at *rs1800562* in my study was still not considered to be ambiguous.

In addition, two non-coding variants showed a significant causality of transferrin for urate via *rs174577* and ferritin for urate via *rs651007* (Table 3.2). Since these variants are found in intronic regions, no functionality has been described in the literature (to my knowledge). The SNP *rs174577* is located in the intronic region of the fatty acid

desaturase2 (*FADS2*) gene and has been reported to be co-expressed with multiple loci at gene-expression (eQTL: expression quantitative trait loci) platforms (www.archive.broadinstitute.org and www.gtexportal.org), while no association with urate has been reported for *rs174577* so far except Köttgen *et al.* (2013) that reported a positive association between the A allele of *rs174577* and serum urate concentration (ß = 0.018, *P* = 0.002) in European individuals. However, no pleiotropic effect has been reported yet for this variant in the context of iron metabolism. The second SNP, *rs651007*, is located in the non-coding region of the *ABO* gene and is known to be associated with a number of traits including red blood cell count (Van Der Harst *et al.*, 2012). The *ABO* gene encodes the glycosyltransferase responsible for the A-B-O blood groups. This variant has been reported to have high coexpression with only two other genes, Surfeit 6 (*SURF6*) and Globoside alpha-1,3-N-acetyl-galactossaminyl-transferase (*GBGT1*) in the *cis*-eQTL database (www.gtexportal.org and www.archive.broadinstitute.org). Based on the evidence available in the literature, *SURF6* is a member of the family of genes that codes for RNA-binding proteins in the nucleolus (Magoulas and Fried, 2000), while *GBGT1* not only encodes for the ABO-related glycosyltransferase but also participates in the biosynthesis of a glycolipid heterophil protein called 'the Forssman antigen' (Haslam and Baenziger, 1996). These findings are suggestive for 'no pleiotropy' between the exposure and outcome and provide the evidence of involvement of a change in iron homeostasis (as exposure) to be causal for a change in urate metabolism (as outcome) only if certain variants are selected as potential IVs in the Mendelian randomisation model.

Instrumenting urate as exposure for iron (and its biomarkers) as outcome did not indicate any urate-iron positive causal relationship for any of the variant (Table 3.5). As detailed in Section 3.2.3.2, none of the 3 variants selected as instruments for urate exposure suggested a possible pleiotropy with iron metabolism. Leave-one-out sensitivity analysis also did not provide any evidence for any of these three variants to be outliers in combined meta-analysis (Table 3.7). Therefore, my study does not indicate that an increase in urate concentration can increase blood levels of iron or its related biomarkers.

3.3.2 Combined multiple-instrument MR

The findings from the combined multiple-instrument MR did not show any causality in standard IVW MR analyses for iron biomarkers on urate (Table 3.2). Although, these results do not support the observational findings presented in Chapter 2, they do not oppose them either. A possible explanation that can address these differences is the presence of unmeasured confounding factors in the observational results that were responsible for overexpression of large significant values. On the other side of the picture, there seems to exist a high degree of pleiotropy between iron and urate metabolism that makes it difficult to separate these two mechanisms on the basis of available summarised GWAS data. For example, *rs1799945* has been reported as one of the top iron-related SNPs within the *HFE* gene in European and African American individuals (Benyamin *et al.*, 2014; Li *et al.*, 2015). Not only was this variant found to have high co-expression (*P* = 8.5E-06) with *SLC17A3* (a known urate-locus), but also was significantly associated with urate in Köttgen *et al.* (2013) data (*P* = 2.94E-06), and having an LD (r^2 < 0.6) with variants within *SLC17A1*, *SLC17A2* and *SLC17A3* (Appendix B Figure 5.2). This pleiotropy made it impossible to include this strong exposure-associated SNP as an IV in MR analysis, especially for serum iron and TSAT. As described earlier, the second variant within the *HFE* gene, *rs1800562*, provided nearly equally significant causal signals for all four iron-related traits (Table 3.2). Although, there is no proof available to support pleiotropy for *rs1800562* with urate and its risk factors, the highly significant MR findings still require validation and may be considered less reliable than causal signals from *rs174577* and *rs651007* (Table 3.2).

Besides, when a sensitivity analysis was carried out using the leave-one-out approach, a significant causal effect of iron and ferritin was observed in increasing urate concentration via combined IVW meta-analysis (Table 3.4). For iron, keeping *TMPRSS6: rs855791* out of the meta-analysis showed a highly significant cause-effect relationship between iron and urate (ß = 0.110, *P* = 1.96E-04), which in turn indicated *rs855791* to be the variant driving non-significant association in the standard IVW analysis (Table 3.4). Excluding another SNP, *SLC40A1: rs12693541*, out of the meta-analysis also showed significant casual association of ferritin exposure in increasing urate concentration (ß = 0.141, *P* = 0.027) (Table 3.4). The gene *TMPRSS6* (transmembrane protease serine 6) encodes a transmembrane enzyme called ‘serine 6’, and is reported to be involved in

extracellular matrix remodelling within the liver (www.ncbi.nlm.nih.gov), sensing iron deficiency and promoting its absorption via blocking the gene encoding for hepcidin (Du *et al.*, 2008). The SNP *rs855791* is a missense variant that increases the enzyme efficiency for inhibiting hepcidin (Nai *et al.*, 2011) and has been reported to be protective against iron deficiency anaemia in aged women (Pei *et al.*, 2014). The SNP *rs12693541* is a non-coding variant within *SLC40A1* (solute carrier family 40 member 1) gene. The *SLC40A1* gene (also known as ferroportin gene) encodes 'ferroportin' protein. As described in Chapter 2, ferroportin is so far the only known mammalian iron exporter in duodenal and other epithelial cells. Defects in the *SLC40A1* gene can cause haemochromatosis type 1 and 4 (Altès *et al.*, 2009; Camaschella, 2006) and reduced iron export (Moreno- Carralero *et al.*, 2014). The SNP *rs12693541* has been reported as a *cis*-eQTL for two other genes 'asparagine synthetase domain containing 1 (*ASNSD1*) and ORMDL sphingolipid biosynthesis regulator 3 (*ORMDL3*). These genes encode for proteins involved in the synthesis of aspartate and sphingolipids, respectively (Siow *et al.*, 2015). Although, excluding the variants within these genes gave significant causal associations for iron and ferritin in MR analysis (Table 3.4), the roles of these two genes in iron homeostasis cannot be denied. These facts may not only explain how the non-significant associations in MR analysis were driven by these two variants (*rs855791* and *rs12693541*) alone, but are also suggestive of an underpinning unmeasured correlation between pathways involved in iron and urate metabolism. In a nutshell, excluding *rs855791* and *rs12693541* from meta-analysis provides evidence for a causal relationship between iron exposure and a change in urate concentration (Table 3.4). Considering together that urate is a natural chelator for the metal 'iron', studies explaining the upregulation and increased production of urate following exposure of iron (explained in Section 2.1) and causal signals in my study via *rs174577* and *rs651007* (detailed in Section 3.3.1), the presence of a possible iron to urate pathway linking these two metabolites cannot be ruled out. An MR analysis from del Greco *et al.* (2016) explained the protective causality between iron and ferritin as exposure and eGFR (estimated glomerular filtration rate) as outcome. Given that a decrease in eGFR has been associated with an increase in urate concentration (Johnson *et al.*, 2013a; Mohandas and Johnson, 2008; Suliman *et al.*, 2006), the positive causal association in my study can be compared with del Greco *et al.* (2016) results as the set of genetic loci explaining renal function and serum urate concentration have 20% overlap (Köttgen *et al.*, 2013; Pattaro *et al.*, 2016). On the basis these findings, a possible role of iron in increasing urate concentration via

renal dysfunction can be speculated. However, there is no literature explaining a genetic correlation between iron and urate till date.

3.3.3 Strengths and limitations of the study

This study provides the first-ever two-sample MR analysis using summary data from large GWAS datasets to investigate a cause-effect relationship between iron and urate metabolism. The prime strength of my study is the use of large-sized datasets from two GWAS with minimal sample overlap, which provided strength to the MR analysis by excluding finite sample study bias and biases created by sample overlap. The removal of possible pleiotropy on the basis of an *a priori* search of literature and gene coexpression information available on the bioinformatics platforms (GTEx, HaploReg 4.1, ENCODE) and intra-instrument LD made the MR analysis more valid. Another attribute of the study is the use of a robust and updated statistical approach available to analyse the summarised data in the context of two-sample MR and *a posteriori* adjustment for horizontal pleiotropy in various analyses after detection of intra-instrument heterogeneity. While the results presented from various approaches used in the MR model may seem inconsistent, it should be considered that each approach has its own limitations and differ in terms of how the validity of an instrument and possible pleiotropy is determined in each model. Single MR clearly provides causal signals for iron and ferritin, specifically for the top loci in the iron GWAS. In this study, the causality of iron and ferritin for a change in urate has been described via multiple-instrument analysis for a selected set of instruments using the leave-one-out approach. The results from IVW and MR-Egger approaches are considered to be robust; however, carrying out a sensitivity analysis may totally end up changing the status of causality. A recent study has depicted this post-test to be valid enough to describe a causal effect while using summary data in any MR analysis (Burgess *et al.*, 2017).

A possible limitation of the study could be the presence of unmeasured pleiotropic confounders that were impossible to remove, especially those other than intra-instrument heterogeneity e.g., gene-environment interaction (Smith and Ebrahim, 2003). Also, the effect estimates described by single variant or sensitivity analyses were of limited magnitude. A possible solution to this limitation is the availability of larger GWAS data

encompassing wide ranges of iron biomarkers and urate and carrying out similar MR analyses using such datasets.

3.3.4 Conclusion and future directions

In conclusion, this study provides causal association signals for iron-urate correlation using single or selective instruments only. This is the first study explaining a cause-effect relationship between iron biomarkers and urate using a two-sample MR approach while incorporating GWAS summary data. A follow-up of this study using *in vitro* functional model experiments can provide evidence of any intermediate pathway/phenotype involved in controlling both iron and urate metabolism simultaneously. Several established human/animal cell lines, primary tissues or disease model animals (mice, zebrafish) are now available for such in vitro analyses. Programmable nucleases can be used to successfully introduce mutations in these cell lines followed by determining changes in gene regulatory function(s) and their possible effect on other gene expression and functionality. The gene expression study could be an appropriate follow-up to MR studies where any change in expression of exposure associated genes and their possible effect on the functionality of outcome-associated gene/phenotype can be recorded. This would be useful especially since urate is a natural chelator of iron and has been proposed to be elevated in response to the metal exposure. Such an experimental approach would also be helpful to explain the cause and effect relationship between iron and urate elevation on the basis of the 'which is driving which' principle. The significant causal effects from single variant and leave-one-out sensitivity analyses should be replicated in large-sized datasets where individual-level data are available. Due to the involvement of iron in increasing oxidative stress and enhancing inflammation, similar MR analysis using GWAS data for gout (no such data was available to carry out similar MR analysis in my thesis) will be helpful in analysing if there exists any stand-alone causality for gout aetiology via inflammation independent of urate metabolism.

Chapter 4

Gout & Metabolic Syndrome

Genetic Association Analyses

SECTION 4.1 URATE, GOUT AND METABOLIC SYNDROME: AN INTRODUCTION

4.1.1 Background

Metabolic syndrome (MetS) is a combination of physiological and anthropometric anomalies that act as risk factors for the development of cardiovascular disease and diabetes (Sookoian and Pirola, 2011; Wilson *et al.*, 2005). The components that constitute MetS typically include abdominal obesity, hypertension, dyslipidaemia and insulin resistance (Alberti *et al.*, 2005; Grundy *et al.*, 2004). These risk factors are highly interrelated and, when clustered together, can increase the risk of developing atherosclerosis and type-2 diabetes mellitus (T2DM) by 3 and 5-fold, respectively (AHA, 2005; Grundy *et al.*, 2005; Sookoian and Pirola, 2011). Metabolic syndrome has also been associated with a 1.6-fold increase in the rate of mortality worldwide (Harris, 2013). Whilst MetS has classically been categorised as a combination of various pathophysiological conditions, several studies have described MetS as a binary factor including some GWA studies (Kraja *et al.*, 2011; Kristiansson *et al.*, 2012; Zabaneh and Balding, 2010). More recent studies in comparative physiology indicate that the syndrome and many of its associated factors can arise simply as a consequence of excessive fat storage (Johnson *et al.*, 2013c). Since not all of these factors sufficiently explain all cardiovascular events, several other factors have been considered to be included in the definition of MetS. These factors broadly include increased inflammatory biomarkers, microalbuminuria and coagulation-related anomalies (Fu *et al.*, 2009; Guo *et al.*, 2012; Meigs *et al.*, 2000; Ramakrishna and Jailkhani, 2008). In addition, elevated urate concentration (Kanbay *et al.*, 2016) and gout (Billiet *et al.*, 2014) have been described to be associated with single or multiple components of the syndrome.

4.1.2 Prevalence of metabolic syndrome

The worldwide prevalence of MetS is increasing, specifically in accordance with age and increased BMI (Ervin, 2009). Despite the ambiguity in accurately defining the term ‘MetS’, a large number of studies (refer to subsequent paragraphs) have been undertaken to determine its prevalence in different parts of the world. Subject to attributes such as ethnic background, age and gender, the recent worldwide prevalence of MetS ranges between 10 to 84% (Kaur, 2014). A population-based study by Pal and Ellis

(2010) suggests that MetS affects nearly 20% of adults in the Western world. Table 4.1 indicates some of the recent studies and their reported MetS prevalence averaged between males and females.

Table 4.1 Prevalence of metabolic syndrome in the different geographical regions

Region	MetS prevalence (%)	Reference
Asia Pacific	11.9 to 37.1	(Ranasinghe *et al.*, 2017)
South Asia	8.6 to 46.1	(Aryal and Wasti, 2016)
Europe	11.6 to 26.3	(van Vliet-Ostaptchouk *et al.*, 2014)
USA (overall)	22.9 to 34.2	(Beltrán-Sánchez *et al.*, 2013; Mozumdar and Liguori, 2011)
USA (Central)	23.0 to 35.1	(Wong-McClure *et al.*, 2015)
USA (South)	18.8 to 43.3	(Márquez-Sandoval *et al.*, 2011)
Africa	12.5 to 62.5	(Okafor, 2012)
Middle East	13.6 to 36.3	(Sliem *et al.*, 2012)

MetS: metabolic syndrome, USA: United States of America.

A comparative study using data from NHANES III (6,423 US individuals) and NHANES 1999-2006 (6,962 US individuals) showed an increase in the prevalence of MetS in the US population between the two survey periods (Mozumdar and Liguori, 2011). The study also reported nearly 68 million adult US individuals (32.4 million males and 35.3 million females) as having MetS, with an overall prevalence ranging between 29.2 to 34.2%. However, another more recent study using data from NHANES 2009-2010 (> 1,800 US individuals) showed a decrease in MetS prevalence with an average of 22.9% (Beltrán-Sánchez *et al.*, 2013). A meta-analysis including data of adult South American individuals from 12 different studies also revealed the prevalence of MetS to be 23.2% in men and 25.3% in women (Márquez-Sandoval *et al.*, 2011). A study in the Australian population reported 34.4% males and 27.4% females as having MetS, with males having greater prevalence than females with a progression of age (Cameron *et al.*, 2007). Two studies, including data of European individuals from Ireland (Waterhouse *et al.*, 2009) and Denmark (Jeppesen *et al.*, 2007), showed an overall prevalence of MetS to be 21.4 and 23.4%, respectively. A more recent study combining data from 51 countries in the Asia-Pacific region reported an overall prevalence of MetS in between the range of 11.9 to 37.1% (Ranasinghe *et al.*, 2017).

Like other regions, a higher prevalence of MetS, or some of its components, have been reported among multicultural populations residing within New Zealand (Simmons

and Thompson, 2004). An initial survey in the Auckland region reported an average MetS prevalence of 32% for Māori, 39% for Pacific Island and 16% for European individuals (Gentles *et al.*, 2007). Moreover, a number of studies showed a higher prevalence of MetS components in the New Zealand population, especially in Polynesian individuals. Some of the examples also include reports on increased obesity (Ng *et al.*, 2014), insulin resistance (IR) (Simmons and Thompson, 2004), diabetes (Chan *et al.*, 2014), cardiovascular disease (CVD) (Linhart *et al.*, 2016) and CHD (Winnard *et al.*, 2013). Having the highest prevalence of gout in the world (explained in Chapter 1), the New Zealand Polynesian population also has a considerably higher rate on occurrence of T2DM and CVD as comorbidities of gout (Winnard *et al.*, 2013; Winnard *et al.*, 2012).

4.1.3 Urate and components of metabolic syndrome

While a number of studies have argued that hyperuricaemia is one of the components of MetS, the debate of inclusion of increased urate concentrations in the definition of the metabolic syndrome is still controversial (Nejatinamini *et al.*, 2015; Wei *et al.*, 2015). The simultaneous occurrence of hyperuricaemia and conditions contributing to MetS has been excessively observed (Billiet *et al.*, 2014), with a number of epidemiological studies confirming the association between the two (Chen *et al.*, 2007; Liu *et al.*, 2015; Yamasaki and Tomita, 2008). Although a number of studies have suggested hyperuricaemia to be a condition occurring secondary to hyperinsulinaemia (Muscelli *et al.*, 1996; Soltani *et al.*, 2013), data suggesting its association with insulin resistance, diabetes (Li *et al.*, 2013; Lippi *et al.*, 2008), obesity (Han *et al.*, 2014; Tang *et al.*, 2010), cardiovascular disorders, hypertension (Borghi *et al.*, 2014; Nakanishi *et al.*, 2003) and interlinking complex correlation of all of these conditions (Desai *et al.*, 2010; Soltani *et al.*, 2013) are still growing.

A recent study by Johnson *et al.* (2013c) redefined MetS and several of its associated conditions as consequences of excessive fat storage. The study also indicated that in most mammals and birds, excess fat is not only stored in adipose tissues but also in the liver and serum (as triglycerides/TG). This storage, in turn, has been shown to correlate with the development of decreased insulin sensitivity and hypertension in the form of increased blood pressure (Johnson *et al.*, 2013c). In addition to other biological aspects, nucleic acid metabolism plays an important role in controlling the process of fat

storage in our bodies. This metabolism regulates fat storage and insulin resistance via the stimulation of adenosine monophosphate (AMP) deaminase, fat degradation and reduced gluconeogenesis through activation of AMP-activated protein kinase (Cicerchi *et al.*, 2014; Lanaspa *et al.*, 2012a; Lanaspa *et al.*, 2015). As urate is a key product of AMP deaminase (details in Chapter 1), a high likeliness of its major role in promoting fat storage have been proposed (Choi *et al.*, 2014; Cicerchi *et al.*, 2014; Lanaspa *et al.*, 2012a; Lanaspa *et al.*, 2012b). Another recent study by Nejatinamini *et al.* (2015) reported an association of urate with factors involved in developing MetS. The major components of MetS included in this study are detailed in the subsequent paragraphs, while Figure 4.1 below illustrates the possible link between these components and increased urate concentration.

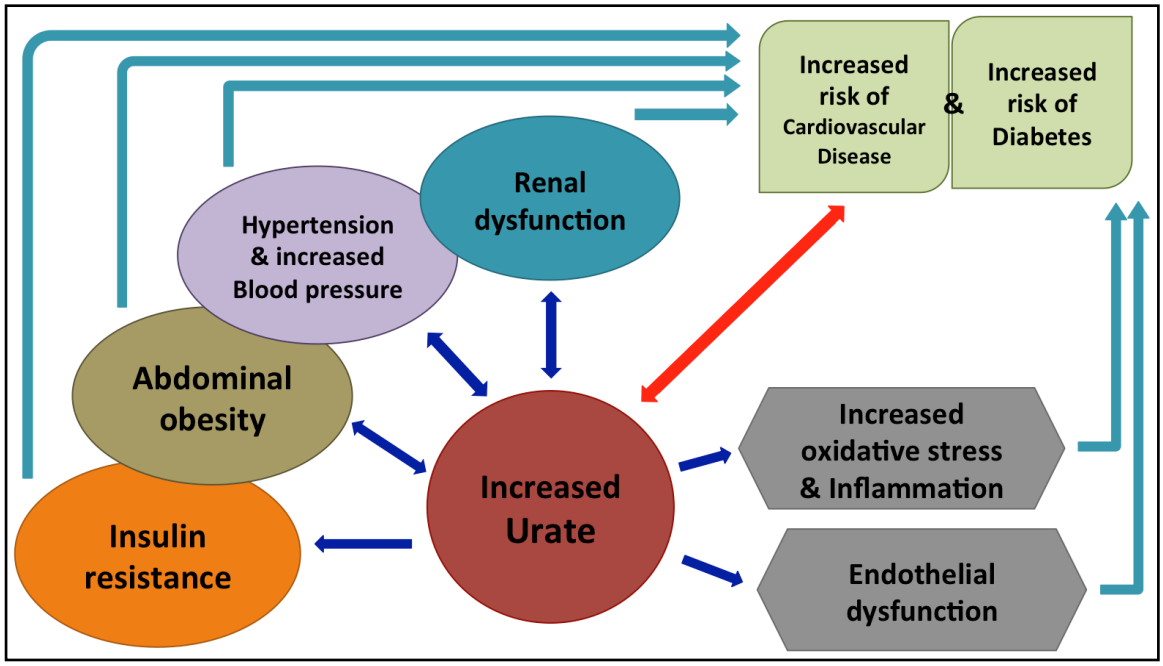


Figure 4.1: Schematic diagram illustrating the interaction of urate with components of the metabolic syndrome (MetS), the risk of developing diabetes and cardiovascular disease. Double-headed dark-blue arrows indicate a two-sided interaction between urate and MetS components. The light blue arrows indicate all possible routes that could contribute to the development of either cardiovascular disease or diabetes, while the red arrow indicates interrelationship of these two anomalies with increased urate.

4.1.3.1 Urate, insulin resistance and diabetes

Without question, insulin resistance (IR) contributes to the development of diabetes. The involvement of IR in the epidemiology of MetS was first proposed in 1988 (Reaven, 1988). The study, for the first time, attributed insensitivity to insulin or IR as a major cause of MetS. Determined as being the major factor linking to gout and its associated co-morbidities, hyperuricaemia was first described as a contributory factor to

diabetes in the late 1800s (Duckworth, 1889). Johnson *et al.* (2013b) provided a detailed review of an overwhelming epidemiological literature explaining the coexistence and role of hyperuricaemia in the development of IR and T2DM. A study by Meshkani *et al.* (2011) found a higher prevalence of IR, hyperinsulinaemia, hypertension, dyslipidaemia and obesity in subjects with higher urate levels.

Historically, hyperuricaemia has been considered as a condition secondary to IR (Facchini *et al.*, 1991; Muscelli *et al.*, 1996). However, more recent evidence brings to light a plausible causal role of hyperuricaemia in the development of IR (Johnson *et al.*, 2013b), especially since a decrease in insulin sensitivity is preceded by an increase in urate concentration (Krishnan *et al.*, 2012). In fact, a decrease in urate concentration has been proposed to increase insulin sensitivity in MetS murine models (Baldwin *et al.*, 2011). A 15-year follow-up study by Krishnan *et al.* (2012) investigated an increase in baseline urate levels in 5,012 US young adults and found a significant hazard of onset of both, diabetes (HR = 1.87) and IR (HR = 1.36). However, no association was reported between the elevated baseline urate and plasma insulin concentrations, suggesting urate as being a risk factor for the development of IR and, thus, diabetes. A 10-year population-based follow-up study in the Netherland also reported urate as an independent risk factor for diabetes (Dehghan *et al.*, 2008b). In a rat-model experiment, Scott *et al.* (1981) found a 26% decrease in serum insulin and 38% increase in serum glucose 4 weeks following inhibition of uricase. A study on sugar-induced diabetes in rats showed that elevated levels of urate increased oxidative stress and up-regulated the activity of urate transporters in pancreatic islet cells (Roncal-Jimenez *et al.*, 2011). Although not many examples are available to explain the effect of reduced urate level on IR in humans, two small randomised controlled trials reported a reduction in IR following the administration of the urate lowering drugs allopurinol (Facchini *et al.*, 2016) and benzbromarone (Ogino *et al.*, 2009). In contrast to these observational studies, findings from the MR studies failed to report a causal association between serum urate and risk of T2DM (Keenan *et al.*, 2016; Pfister *et al.*, 2011; Sluijs *et al.*, 2015). Sluijs *et al.* (2015) further suggested that use of urate lowering therapy may not be beneficial in reducing diabetes risk. These MR findings make a possible role of urate levels in increasing the risk of diabetes controversial (Johnson *et al.*, 2015). However, as almost all of these MR studies used selected variants from urate transport genes (especially *SLC2A9*) as MR instruments, it may point towards the role of these genes in improved kidney functions as described by

Hughes *et al.* (2014). Hughes *et al.* (2014) found that the individual genetic variants of the genetic risk score with the strongest effect on serum urate did not have the strongest beneficial effect on renal function and suggested the possibility of pathways distinct from serum urate levels to be involved in enhanced renal function. The authors further described that an increase in urate levels could possibly influence tubule biochemistry by exchanging uric acid for other metabolites and cofactors. For example, two well-known urate transporters, SLC2A9 and ABCG2 play a vital role in exchanging hexose sugars for uric acid and an adenosine triphosphate-dependent secretion of uric acid, respectively (reviewed in Anzai and Endou (2011)). While most of the observational studies showed a positive correlation between high urate levels and increased risk of developing diabetes and insulin resistance, these detrimental effects have not yet been proved via the role of best choice urate transport genes in MR studies.

4.1.3.2 Urate and obesity

Obesity, especially abdominal obesity, is considered as a major manifestation of MetS (Després and Lemieux, 2006). Urate has been described as a potential predictor of fatty liver (Sirota *et al.*, 2013) and obesity (Masuo *et al.*, 2003). The simultaneous occurrence of hyperuricaemia and obesity has been reported to co-exist with other components of the syndrome e.g., diabetes and IR (Johnson *et al.*, 2013b). An *in vitro* experiment using cultured liver cells, showed that an increase in urate concentration increased triglyceride levels in cells (Lanaspa *et al.*, 2012b). Another *in vitro* study model demonstrated that urate induced fat accumulation in hepatocytes (Choi *et al.*, 2014). Further to these, hyperuricaemia was also found to increase the triglyceride levels in hepatic cells using animal model (rat) *in vitro* experiments (Tapia *et al.*, 2013). The fact that an increase in hepatic triglyceride levels is a known risk factor for increased adiposity irrespective of alcohol consumption (Fabbrini *et al.*, 2010; Jung and Choi, 2014), abnormally high urate concentrations could, therefore, be directly related to inducing this increase. Contrary to these, reports of genetically determined higher urate levels being causal for increased triglycerides via the MR studies provide different findings, with studies describing elevated serum urate as a consequence rather than a cause of an increased BMI, adiposity and/or its risk factors (Lyngdoh *et al.*, 2012; Palmer *et al.*, 2013; Rasheed *et al.*, 2014). The possible factor playing a role in contradictory

findings between observational and MR studies has been described in the preceding paragraph.

When it comes to weight gain and adiposity, without a doubt, diet plays a central role. Most of the purine-rich foods that increase body concentrations of urate have been shown to act as risk factors for the development of MetS or its components (Choi *et al.*, 2004a). Historically, foods high in fat are also known to increase serum urate levels (Ogryzlo, 1965). A fructose-rich diet plays an interlinking role between elevated serum urate and increased risk of obesity and other components of MetS (Nakagawa *et al.*, 2006), especially the amount of fructose in 'added sugars' (Johnson *et al.*, 2007). In fact, the use of sugar-sweetened beverages has been found to have a link with the epidemic of obesity as well as MetS (Basu *et al.*, 2013; Malik *et al.*, 2013; Malik *et al.*, 2010). Additionally, several lines of experimental evidence suggest that the primary fructose-related mechanism involved in increasing urate levels and inducing hyperuricaemia is in fact one of the master pathways that can lead to the development of obesity, insulin resistance and cardiovascular disorders (collectively called MetS) (Baldwin *et al.*, 2011; Lanaspa *et al.*, 2012b; Lanaspa *et al.*, 2012c; Sánchez-Lozada *et al.*, 2008; Tapia *et al.*, 2013).

A number of studies have reported that considerable weight loss can reduce serum urate levels in a clinically significant way (Dalbeth *et al.*, 2014). In a dataset of 4,047 individuals from the Swedish Obese Subjects Study, bariatric surgery was found to decrease urate levels by 14% at 2 years and 8% in 10 years in 1,845 and 641 Swedish individuals in comparison to their controls (Sjöström *et al.*, 2004). Decreasing urate concentrations have also been shown to reduce the fructose sensitised fatty liver development in human hepatocytes. In addition, reports on reduction in liver fat following a urate-lowering treatment are available in both animal models with components of MetS (Lanaspa *et al.*, 2012c) and in those with alcohol-induced fatty liver (Kono *et al.*, 2000).

4.1.3.3 Urate and cardiovascular disease

In addition to an association with IR, diabetes and obesity, hyperuricaemia has been associated with CVD for decades (Gertler *et al.*, 1951). Since the first suggestion of a urate-CVD relationship back in the 19th century (Haig, 1889), a number of studies

reported urate to be a risk factor for cardiovascular events (Bos *et al.*, 2006; Moriarity *et al.*, 2000; Okura *et al.*, 2009). Albeit the direct independent role of urate in causing heart disease in the general population is still controversial (Soltani *et al.*, 2013), many examples report its indirect role in occurrence of CVD via IR, hypertension, and renal disease or a combination of these conditions (Johnson *et al.*, 1999; Johnson and Tuttle, 2000). A number of studies have also shown an association of hyperuricaemia with surrogate markers of arterial plaque build-up (atherosclerosis) e.g., coronary artery calcification (Krishnan *et al.*, 2011), carotid artery thickness (Zhang *et al.*, 2012c) and brachial pulse velocity (Ishizaka *et al.*, 2007).

In a large study including participants with hypertension and T2DM, a higher serum urate concentration (> 7 mg dL^{-1}) was found to be associated with increased cardiovascular mortality (Chen *et al.*, 2009b). Another similar study suggested increased mortality in patients with coronary syndrome to be associated with a serum urate concentration of 7.5 mg dL^{-1} or higher (Ndrepepa *et al.*, 2012). The occurrence of hyperuricaemia in patients with congestive heart failure was also reported to be associated with a higher rate of mortality (Kim *et al.*, 2010). A large meta-analysis study including data from 15 prospective cohorts suggested a link between high serum urate and greater incidence of heart stroke and associated mortality (Li *et al.*, 2014). More recent MR studies have reported a causal relationship between the genetically predicted serum urate and adverse cardiovascular outcomes including sudden cardiac death (Kleber *et al.*, 2015), especially cardiometabolic disease in T2DM patients (Yan *et al.*, 2016). Other similar studies, however, were unable to find any causal association of genetically determined elevated urate concentration with an increased risk of cardiometabolic anomalies including CHD, hypertension, heart failure and/or ischemic stroke (Keenan *et al.*, 2016; Palmer *et al.*, 2013).

Since the direct causal effect of hyperuricaemia in the context of heart disease is still unclear, the possible relationship between them has been a topic of interest for a number of clinical randomised trial studies. One such study showed that the administration of allopurinol to reduce urate levels improved hypertension and carotid intimal thickness in CHD patients with and without renal disease (Higgins *et al.*, 2014). Other such examples include studies reporting benefits of urate-lowering therapy in

improving arterial stiffness (Grimaldi-Bensouda *et al.*, 2014), angina (Noman *et al.*, 2010) and ventricular hypertrophy (Kao *et al.*, 2011).

4.1.4 Gout and components of metabolic syndrome

The five components of MetS described above (insulin resistance, T2DM, obesity, hypertension and CVD) are well-known comorbidities for gout and have long been speculated for a link with gout (Chen *et al.*, 2012; Choi *et al.*, 2005b; Fam, 2002). Given that hyperuricaemia is the main risk factor for gout aetiology and MetS being potentially associated with urate, it is highly unlikely that MetS would not simultaneously co-exist with gout. Gout, being a complex disease, has been reported to be associated with all-cause mortality in a number of studies. However, most of the studies have reported the rate of mortality in gout patients to be higher for two major factors of MetS i.e., cardiovascular or renal disease (Choi and Curhan, 2007a; Stamp and Chapman, 2013).

A quantitative study by Rho *et al.* (2005) was the first report that combined the prevalence of MetS and gout in the Korean population. Their data showed that gout cases were more likely to develop components of MetS (43.6%) in comparison to healthy controls (5.2%). The data also found increased BMI and HDL to be the variables most significantly associated with the development of MetS in individuals with gout when compared to otherwise healthy controls in the same population (Rho *et al.*, 2005). A similar study including US population data from NHANES III also reported the prevalence of MetS in gout cases to be as high as 62.8% in comparison to the individuals without gout (25.4%) (Choi *et al.*, 2007). The study further reported that 3.5 million US adults who had ever suffered from gout also simultaneously suffered from metabolic conditions included in the definition of MetS. Another more recent example comes from the study by Kuo *et al.* (2014a) that assessed data from more than 75,000 European individuals (from the United Kingdom) for gout and its associated comorbidities prior to and following the diagnosis. Kuo *et al.* (2014a) found that participants in the gout cohort were more likely to have at least one or more MetS components unlike the participants in the non-gout cohort. The gout individuals were also on the higher hazard of developing cardiovascular disease, obesity, diabetes and other endocrine disorders within an average timeframe of ~3.5 years unlike their ancestrally-matched controls who had an average time period of ~9.2 years to develop any of the aforementioned comorbidities (Kuo *et al.*,

2014a). Given the higher prevalence of the coexistence of MetS (or its individual components) and gout in several populations, the components should be appropriately recognised and taken into account while suggesting any long-term treatment to individuals with gout.

4.1.5 Genetics of metabolic syndrome

Running a large GWA study in order to identify the genetic components of MetS is particularly complicated owing to its complexity at both genetic and clinical levels. This is the reason as to why, until now, most of the genetic association studies have been done on individual components of MetS or a combination of a few rather than the syndrome per se as a binary phenotype (O'Neill and O'Driscoll, 2015).

Family and twin studies have provided a large amount of evidence with regard to the heritability of MetS. A genetic familial study by Lin *et al.* (2005) analysed the data of 803 participants from 89 Hispanic families in Northern Manhattan Family Study (NMFS) and reported an overall heritability of 24% for MetS. The heritability for individual components of MetS was shown to be 46%, 24%, 47% and 60% for waist circumference, higher glucose, TG and HDL cholesterol levels, respectively (Lin *et al.*, 2005). The heritability of MetS was reported to be 29.9% when the data of 293 Italian individuals (51 families) within the Linosa Study were analysed (Bellia *et al.*, 2009). Moreover, the results from twin studies showed a concordance in clustering of hypertension, diabetes and obesity to be higher (31.6%) in monozygotic twins than dizygotic (6.3%) twins (Carmelli *et al.*, 1994).

A number of research groups also have performed MetS-related GWAS. The classic example comes from an intronic SNP within the fat mass obesity-associated gene (*FTO*) that was found to be associated with T2DM in a UK-based cohort (Frayling *et al.*, 2007). The study further explained that the *FTO* gene increases the risk for developing T2DM via its effect on BMI (Frayling *et al.*, 2007). Similar results were also reported in other populations including individuals of European (Loos *et al.*, 2008; Polašek *et al.*, 2009; Speliotes *et al.*, 2010; Willer *et al.*, 2009) and Asian decent (Wen *et al.*, 2012). These GWA studies also reported some other genes to be associated with increased BMI and risk of obesity i.e., genes coding for transmembrane protein 13 (*TMEM13*),

potassium channel tetramerisation domain containing 15 (*KCTD15*) and melanocortin receptor 4 (*MC4R*).

To date, only three GWA studies have reported the data for MetS as a binary trait. A genetic association study in an Indian-Asian male population (n = 2,300) reported a number of loci to be associated with discrete MetS components, but was unable to find any loci associated with MetS as a binary trait (Zabaneh and Balding, 2010). Another GWA study conducted in European population of ~22,000 male and female individuals reported five SNPs within three loci (*LPL*: Lipoprotein lipase, *APOA5*: Apolipoprotein A5 and *CETP*: Cholesteryl ester transfer protein) to be associated with MetS at a genome-wide threshold (Kraja *et al.*, 2011). This and another study showed a total of 17 loci to be associated with either MetS as a whole or a combination of some of its components (Kristiansson *et al.*, 2012).

While the genetic association data for MetS are limited, a bulk of literature presents the GWA results for its individual components reporting dozens of loci, with most loci including genes implicated in IR, obesity or lipid metabolism (O'Neill and O'Driscoll, 2015). However, the data describing a possible genetic association of components of MetS with urate and gout are still scarce. This chapter was structured to investigate this substantial gap in the literature by studying selected variants that have been reported to be robustly associated with major MetS components and analysing them for their association with urate and gout within the NZ Polynesian and European populations. While the particular details of each MetS-related genetic variant selected for this study from the *ADRB3, MC3R, MC4R* and *ADTRP* genes and their possible association with urate and gout are provided in the next section (Section 4.2), the specific aims of the study were;

1. To genotype and test the IR and BMI associated variant, *ADRB3: rs4994*, for its association with urate and gout in the European and NZ Polynesian populations.
2. To genotype and test the obesity and BMI associated variants, *MC3R: rs3827103*, *MC4R: rs17700633* and *MC4R: rs17782313* for their association with urate and gout in the European and NZ Polynesian populations.
3. To genotype and test the CHD associated variant, *ADTRP: rs6903956*, for its association with urate and gout in the European and NZ Polynesian populations.

SECTION 4.2 GENETIC ASSOCIATION OF COMPONENTS OF METABOLIC SYNDROME WITH URATE AND GOUT

4.2.1 Background

Hyperuricaemia and gout are closely related to most of the components of MetS, especially, dyslipidaemia, obesity, hypertension and glucose intolerance (Choi and Ford, 2007; Choi *et al.*, 2007). These, and other similar reports (Facchini *et al.*, 1991) has led to construction of the basic hypothesis of this chapter. According to this hypothesis 'if urate is associated with the components of MetS in observational studies, the genetic variants contributing to these components should also contribute to hyperuricaemia and gout'. This section was designed to do genetic association analyses of the selected MetS-related variants with serum urate and gout in European and NZ Māori and Pacific Island (Polynesian) individuals. Although the basic hypothesis of this study is the same as Mendelian randomisation, a different approach (explained in the next section) was used to assess the causal association due to the unavailability of genetic association data for selected genes/variants for the New Zealand population.

4.2.2 Methods

4.2.2.1 Selection of variants

Five MetS-related variants (SNPs) were selected based on their association with urate and gout in the literature. Table 4.2 provides a list of these variants, while a summary of each of these variants is provided in the subsequent paragraphs.

Table 4.2: List of the MetS-associated variants selected in this study

Variant (SNP)	Chromosome	Gene/nearest gene	Consequence	Gene-related trait
rs4994	8	*ADRB3*	Missense	BMI, IR
rs3827103	20	*MC3R*	Missense	LM, Obesity, Anti-inflam
rs17700633	18	*MC4R*	Intergenic	BMI, LM, Obesity
rs17782313	18	*MC4R*	Intergenic	BMI, LM, T2DM, IR, Obesity
rs6903956	6	*ADTRP*	Intronic	CHD

SNP: Single nucleotide polymorphism, BMI: body mass index, LM: lipid metabolism, Anti-inflam: anti-inflammatory response, T2DM: Type 2 diabetes mellitus, CHD: Coronary heart disease.

4.2.2.1.1 Insulin resistance & BMI associated variant

The first variant that was selected for this study is a missense variant, *rs4994*, within the beta-3 adrenergic receptor gene (*ADRB3*) on chromosome 8 (Table 4.1). The SNP causes a substitution of tryptophan (Trp) to arginine (Arg) at codon 64 (Trp64Arg) within the first transmembrane domain of the ADRB3 protein. The *ADRB3* gene in humans encodes the beta-3 adrenergic receptor or β3 adrenoreceptor protein (Blocker, 2013). This protein is a part of the adrenergic nervous system, which releases adrenaline and/or norepinephrine as neurotransmitters. The adrenergic system acts as one of the main neurohormonal regulatory centres to maintain normal smooth muscle tone within the heart and overall cardiac action-response. The *ADRB3* gene is classically known for its role in the regulation of lipid metabolism and glucose homeostasis through its expression in adipose tissues (Krief *et al.*, 1993). However, more recently, *ADRB3* was reported to be abundantly expressed in the acetylcholine-releasing nerve fibres in the bladder, which was suggestive of its role in the regulation of metabolic functions in the bladder (Coelho *et al.*, 2017). The presence of the Trp64Arg polymorphism has been extensively reported for its association with early onset diabetes (Gjesing *et al.*, 2008; Nagase *et al.*, 1997), obesity, BMI (Gjesing *et al.*, 2008; Valve *et al.*, 1998), and insulin resistance (Allison *et al.*, 1998; De Luis *et al.*, 2009; Widén *et al.*, 1995). In a meta-analysis of 97 cohorts (n = 44,833 individuals), Kurokawa *et al.* (2008) found Arg64 to be significantly associated with BMI having a stronger effect in East Asian sample sets. Consistent reports of association of the Arg64 allele with increased adiposity measures, high blood pressure and elevated serum urate was provided in a longitudinal study that analysed data of elderly male individuals in Olivetti Prospective Heart Study (OPHS) (Strazzullo *et al.*, 2001).

The variant *rs4994* was selected based on the above hypothesis (reported association with urate and gout) and the rationality of the MetS-related evidence in previous literature. A combination of increased BMI and the presence of Arg64 was reported to be associated with a 4-fold increase in the risk of developing hyperuricaemia in a postprandial diabetic group from Chinese population (Wang *et al.*, 2002). Similar association of the Arg64 allele and hyperuricaemia were also reported in data sets of Korean (Rho *et al.*, 2007) and Chinese (Huang *et al.*, 2013) male individuals. The Arg64 allele was also indicated to be significantly associated with high risk of hyperuricaemia in

a 6-year follow-up study including male and female individuals from Southern Spain (Morcillo *et al.*, 2010). In addition to hyperuricaemia, Arg64 (*rs4994*: G allele) also has been reported to be associated with gout susceptibility in Chinese Han population (Wang *et al.*, 2011a).

4.2.2.1.2 Obesity & BMI associated variants

Three variants, associated with obesity and BMI, were selected from the genes within the melanocortin receptor (MCR) family. The selected variants included a missense variant, *rs3827103*, within the *MC3R* gene on chromosome 20 and two intronic variants, *rs17700633* and *rs17782313*, within the *MC4R* gene on chromosome 18 (Table 4.1).

The MCR family includes five members of the melanocortin system, expressed in the central nervous system (CNS) and several peripheral tissues, each holding distinctive specification for different melanocortins (Dores, 2009; Hadley and Dorr, 2006; Voisey *et al.*, 2003). The melanocortins are members of a class of naturally occurring peptide hormones derived from a larger precursor molecule named pro-opiomelanocortin (POMC) (Hadley and Haskell, 1999; Raffin-Sanson *et al.*, 2003). During post-translational modification, the POMC molecule cleaves to form smaller melanocortins that, in turn, act as potential substrates for MCRs. Two melanocortin molecules, adrenocorticotropic hormone (ACTH) and melanocyte-stimulating hormones (MSH; α- and β-forms), are known to be involved in the regulation of anti-inflammatory response and hunger cycle via binding to MC3R and MC4R, respectively. The gene encoding MC3R has also been observed to be expressed in a range of peripheral cells including placenta, heart, gut, pancreas and macrophages (found within gout inflamed knee joint) (Getting *et al.*, 2002). Getting *et al.* (2002) further explained that ACTH and other smaller fragments of α- and β-MSH can inhibit monosodium urate crystal-induced neutrophil migration and release of proinflammatory cytokines and chemokines in gouty arthritis. Development of severe IR and higher levels of adipose tissue inflammation have also been reported in experiments using *MC4R* and *MC3R* knockout mice, respectively (Trevaskis *et al.*, 2007).

The *MC3R* gene has been classically known for its role in weight regulation, energy metabolism, and regulation of the cardiovascular system (Getting *et al.*, 2002;

Rediger *et al.*, 2012; Tao and Segaloff, 2004). Polymorphisms in this gene can cause the potential loss of receptor function and expression or defective receptor activation (Lee *et al.*, 2007; Savastano *et al.*, 2009). The selected SNP, *rs3827103,* is a missense variant within the *MC3R* gene that causes a substitution of valine (Val) to isoleucine (Ile) at codon 81, denoted as Val81Ile. Ile81 allele homozygosity has been reported to be associated with an increased risk of obesity and higher IR in children (Feng *et al.*, 2005; Savastano *et al.*, 2009). A familial study including data from T2DM patients also showed a marginal association between the Ile81 allele and increased IR in adult French Caucasians (Hani *et al.*, 2001). More recent studies have also indicated the association of the Ile81 allele with significantly higher risk of developing MetS (Suazo *et al.*, 2013). Alsmadi *et al.* (2014) also suggested a possible role of the Ile81 allele in elevated blood pressure and thus increased hypertension.

The *MC4R* gene is known for its role in regulating appetite, food intake and its associated behaviours and energy metabolism via signalling for α-MSH and agouti-related peptide (AgRP). Common variants within/near the *MC4R* gene (including the *MC4R* variants selected for this study) have been associated with increased adiposity, reduced insulin sensitivity and higher BMI in GWA studies (Chambers *et al.*, 2008; Loos *et al.*, 2008). The variant *rs17700633* is located 188 kb and *rs17782313* is located 109 kb downstream of the *MC4R* gene. The literature has reported these variants to be involved in increasing the risk of obesity and its related phenotypes (explained hereafter). Large case-control studies have confirmed the association of the A allele of *rs17700633* and the C allele of *rs17782313* with obesity and increased BMI in several populations (Beckers *et al.*, 2011; Srivastava *et al.*, 2014; Zobel *et al.*, 2009). Additionally, the same alleles have been reported to be significantly associated with higher intakes of total energy and dietary fat (Kring *et al.*, 2010; Qi *et al.*, 2008) and increased risk of diabetes in adults (Marcadenti *et al.*, 2013; Mutombo *et al.*, 2014). As both of these variants are located within non-coding regions, their direct influence on gene function is unclear. However, the presence of these variants is considered to alter the function of the gene as the pattern of phenotypic associations provided in the literature for these variants are similar to those mediated via altering the function of the *MC4R* gene (Zobel *et al.*, 2009).

Although, the selected variants from the *MCR* genes have not been reported for their direct potential association with urate or gout, higher adiposity, weight gain and

diabetes are still strong risk factors/common comorbidities for gout (detailed in Chapter 1). The pieces of evidence described above are supportive for strong role of *MC4R* and *MC3R* polymorphisms in the onset of obesity, decreased inflammatory response and increased insulin resistance. These facts, if incorporated into the hypothesis of this study, support the idea of a possible role of MetS-related variants in the aetiology of gout.

4.2.2.1.3 Coronary heart disease associated variant

A CHD-associated variant, *rs6903956*, was also selected for this study. The SNP *rs6903956* is an intronic variant located on chromosome 6, at position 24.1, within the *ADTRP* gene encoding the androgen-dependent tissue factor pathway inhibitor (TFPI) regulating protein (ADTRP). The *ADTRP: rs6903956* variant is a comparatively nascent locus known to be associated with increased CHD susceptibility at GWAS level (Nikpay *et al.*, 2015; Wang *et al.*, 2011c). The ADTRP protein functions to up-regulate TFPI expression and maintain the anticoagulant protection of the endothelium to avoid endothelial dysfunction in response to androgen stimulus (Lupu *et al.*, 2011). More recently, *ADTRP* gene expression has been shown to regulate the expression of other genes involved in the sustenance of cell cycle progression via proliferation and apoptosis (Luo *et al.*, 2016). A study in the Han Chinese population reported that the A allele of *rs6903956* can increase the risk of developing asymptomatic hyperuricaemia (aHU) in adults (Meng *et al.*, 2015).

4.2.2.2 Study participants

Four different data sets were used to assess the association of selected variants with gout. The first dataset was a sub-set of the New Zealand Gout Cohort (Section 1.2), including 1,872 European and 2,464 Polynesian case-control individuals from New Zealand. The NZ Gout Cohort was stratified into four ancestral groups as detailed in Section 1.2.1: NZ European (NZ EUR; 910 cases and 962 controls), Eastern Polynesian (EP; 510 cases and 698 controls), Western Polynesian (WP; 365 cases and 320 controls) and mixed Eastern and Western Polynesian (EPWP; 33 cases and 76 controls). The Māori dataset from NPH consisted of 270 gout cases and 192 controls. Data for the EP group were separately analysed in EPN (334 cases and 392 controls) and EPZ (157 cases and 311 controls) sub-groups (details provided in Section 1.2.1). Two additional datasets of European gout case individuals were selected from the European Crystal Network Cohort

(EUROGOUT; n = 818) and the Arthritis Genomics Recruitment Initiative in Australia Cohort (AGRIA; n = 198) (details in Section 1.2). The gout cases from EUROGOUT and AGRIA Cohorts were combined with the NZ European dataset to increase the power of the study, hereafter referred to as 'EUR (Combined European)' dataset. The fourth dataset for gout association analysis was selected from the UK Biobank Gout Cohort (please refer to details in Section 1.2). This dataset included 2,432 gout cases and 102,989 controls. Table 4.3 represents a summary of demographic details for these datasets.

Data for non-gout (control) individuals were sourced from four publicly-available datasets (detailed in Section 1.2.6 through 1.2.9). The information for individuals in these cohorts was only used for serum urate association analyses. The data were sourced for a total of 5,367 individuals in ARIC, 5,109 in FHS (Offspring and Generation 3), 1,432 in CARDIA and 2,421 individuals in CHS Cohort. Table 4.4 provides demographic and clinical details of these datasets. Analyses for serum urate association were also performed in NZ European and Polynesian non-gout individuals. For this purpose, any individuals who self-reported as taking diuretic medication or were on any other urate lowering therapy (ULT), had renal failure, gout or had first-degree relatives with gout were excluded from the various analyses.

Table 4.3: Demographic details of the datasets included in the study for serum urate (non-gout individuals only) and gout association analyses

Population	NZ Polynesian								European					
Sub-population	**EP**		**WP**		**EPWP**		**NPH**		**NZ EUR**		**UK Biobank**		**EUROGOUT**	**AGRIA**
Group	Gout	Non-gout	Gout	Non-gout	Gout	Non-gout	Gout	Non-gout	Gout	Non-gout	Gout	Non-gout	Gout	Gout
Baseline Information														
Total Participants	510	698	365	320	33	76	270	192	910	962	2432	102989	818	198
Male, n (%)	449 (88.03)	329 (38.75)	400 (88.49)	201 (56.94)	34 (85.00)	43 (51.80)	226 (83.70)	107 (50.95)	789 (82.87)	523 (54.53)	2201 (90.50)	90414 (87.78)	703 (85.94)	176 (88.88)
Age (years)^	56.21 ± 12.97	45.21 ± 15.71	48.88 ± 12.65	40.44 ± 14.08	41.67 ± 13.21	36.86 ± 16.18	57.93 ± 12.96	43.07 ± 14.96	63.64 ± 13.01	49.77 ± 18.21	60.59 ± 6.56	60.46 ± 6.63	62.25 ± 12.91	59.80 ± 11.41
BMI (kg/m2)^	35.38 ± 8.34	32.33 ± 7.84	63.26 ± 7.99	34.45 ± 6.68	36.84 ± 7.89	33.64 ± 6.84	36.14 ± 8.10	30.62 ± 5.88	30.45 ± 6.72	27.71 ± 5.83	30.93 ± 4.73	27.64 ± 4.41	30.18 ± 7.95	29.69 ± 4.24
Serum Urate (mmol L^{-1})^	0.41 ± 0.11	0.36 ± 0.08	0.44 ± 0.11	0.39 ± 0.09	0.46 ± 0.09	0.38 ± 0.11	0.44 ± 0.11	0.36 ± 0.08	0.39 ± 0.11	0.34 ± 0.10	-	-	0.42 ± 0.14	0.33 ± 0.12
Comorbidities														
Type II Diabetes†	27.59 (83.30)	13.91 (85.03)	22.40 (93.92)	16.15 (88.64)	22.5 (100)	15.18 (95.18)	34.03 (95.97)	6.34 (97.61)	85.12 (98.84)	5.84 (94.08)	-	-	23.17 (55.09)	64.70 (25.37)
Dyslipidaemia†	48.97 (77.16)	14.73 (80.27)	49.64 (91.32)	15.45 (85.67)	39.47 (95.00)	11.84 (91.56)	50 (95.47)	18.81 (96.19)	48.59 (97.05)	17.32 (89.83)	-	-	56.17 (51.07)	53.84 (19.40)
Heart problems†	40.11 (84.40)	13.11 (85.84)	20.36 (95.87)	5.93 (91.08)	22.5 (100)	8.64 (97.59)	36.31 (95.47)	11.11(98.57)	36.62 (98.94)	11.11 (98.96)	-	-	10.77 (51.93)	72.72 (32.83)
Hypertension†	61.62 (83.30)	27.22 (83.52)	48.96 (93.92)	17.11 (90.00)	45.00 (100)	20.77 (92.77)	64.58 (96.58)	17.47 (98.09)	53.83 (98.52)	21.13 (89.83)	-	-	72.14 (69.01)	81.25 (47.76)
Kidney problems†	24.14 (82.83)	3.77 (76.79)	18.73 (92.62)	3.52 (84.32)	21.05 (95.00)	4.34 (83.13)	11.76 (93.96)	2.23 (85.23)	22.88 (98.21)	5.51 (46.99)	-	-	13.18 (50.07)	70.00 (29.85)
Gout Characteristics														
On Diuretics/ULT†	43.66 (57.16)	-	26.25 (82.86)	-	47.28 (87.5)	-	18.51 (33.33)	-	43.22 (61.13)	-	-	-	35.26 (61.83)	-
Gout attacks per year^	10.24 ± 36.37	-	12.60 ± 38.07	-	7.47 ± 9.86	-	3.01 ± 5.01	-	7.62 ± 33.39	-	-	-	3.42 ± 4.76	-
Gout tophus†	35.55 (87.85)	-	44.31 (91.54)	-	40 (100)	-	8.29 (96.98)	-	33.33 (99.89)	-	-	-	33.45 (77.61)	43.13 (76.11)

NZ; New Zealand, EP; East Polynesian, WP; West Polynesian, EPWP; Mixture of East and West Polynesian, NPH; Māori cohort from Ngati Porou Hauora, NZ EUR; NZ European, UK; United Kingdom, EUROGOUT; European Crystal Network Gout Cohort, AGRIA; Arthritis Genomics Recruitment Initiative in Australia Gout Cohort, BMI; Body mass index, n (%); total number (percentage). ^Data are shown as mean ± standard deviation. †Data are shown as percentage reported with comorbidities (total percentage who answered the question of having comorbidities as 'yes' or 'no'), Data for the comorbidities are self-reported. Data presented as an average for five MetS-related variants genotyped through TaqMan PCR. The genotype success rate was 99.37% for *rs4994*, 98.99% for *rs3827103*, 98.89% for r*s17700633*, 98.93% for *rs17782313* and 99.01% for *rs6903956*.

Table 4.4: Demographic details of publicly-available datasets used only for serum urate analysis in non-gout individuals

Population	ARIC	FHS	CARDIA	CHS
Baseline Information				
Total Participants	5367	5109	1432	2421
Male, n (%)	2459 (45.81)	2360 (46.19)	671 (46.85)	987 (40.76)
Age (years)^	53.48 ± 5.58	36.99 ± 9.41	25.54 ± 3.33	72.13 ± 5.35
BMI (kg/m2)^	25.98 ± 4.32	25.81 ± 4.91	23.65 ± 3.94	25.87 ± 4.18
Serum Urate (mmol L^{-1})^	0.33 ± 0.07	0.30 ± 0.08	0.31 ± 0.07	0.31 ± 0.07
Comorbidities				
Type II Diabetes†	2.92 (99.96)	0.72 (94.78)	0.49 (99.44)	6.21(99.71)
Dyslipidaemia†	-	-	2.73 (97.28)	-
Heart problems†	1.76 (99.96)	0.48 (66.14)	5.83 (99.25)	7.98 (98.31)
Hypertension†	6.12 (99.49)	0.99 (93.87)	7.41 (99.19)	17.71 (99.09)
Kidney problems†	-	-	3.09 (100)	0 (99.46)

ARIC; Atherosclerosis Risk in Community Study Cohort, FHS; Framingham Heart Study Cohort (Offspring and Generation 3), CARDIA; Coronary Artery Risk Development in (Young) Adults Study Cohort, CHS; Cardiovascular Health Study Cohort, BMI; Body mass index, n (%); total number (percentage). ^Data are shown as mean ± standard deviation. †Data are shown as percentage reported with comorbidities (total percentage who answered the question of having comorbidities as 'yes' or 'no'), Data for the comorbidities are self-reported.

4.2.2.3 Genotyping

All five variants, *rs4994, rs3827103, rs17700633, rs17782313* and *rs6903956* were genotyped over all European (NZ European, EUROGOUT and AGRIA) and NZ Polynesian serum samples. Genotyping was performed by an allelic discrimination assay i.e., TaqMan® SNP Genotyping assay (C_2215549_20; Applied Biosystems, Foster City, USA) following the instruction provided in the manufacturer's protocol. The TaqMan SNP genotyping were auto-called via reporter dye signal plots on Lightcycler® 480 Real-Time Polymerase Chain Reaction (RT-PCR) System (Roche Applied Science, Indianapolis, IN, USA). The details about the sources for genotype data for UK Biobank, ARIC, FHS, CARDIA and CHS cohorts are already described in Section 1.2.

4.2.2.3.1 Genotype data imputation

Publicly-available European data sets used for serum urate analyses were missing the genotype data for a number of variants. In such groups, genotype data were imputed using 1000 Genomes haplotype data [phase 1; 2013, NCBI (National Center for Biotechnology Information) build 37] as a reference panel for all populations. A platform 'IMPUTE2' was used for the purpose of imputation as described by Howie and Marchini (2011). Table 4.4 below provides the details of variants and populations for which the genotype data were imputed.

Table 4.5: Details of variants and study cohorts for which the genotype data were imputed

Variant	Study Cohort	Reference panel	Imputation platform
rs4994	ARIC, FHS, CHS, CARDIA	1000 Genomes haplotype data (phase 1; 2013)	IMPUTE2
rs3827103	-	1000 Genomes haplotype data (phase 1; 2013)	IMPUTE2
rs17700633	CHS	1000 Genomes haplotype data (phase 1; 2013)	IMPUTE2
rs17782313	CHS	1000 Genomes haplotype data (phase 1; 2013)	IMPUTE2
rs6903956	CHS	1000 Genomes haplotype data (phase 1; 2013)	IMPUTE2

ARIC; Atherosclerosis Risk in Community Study Cohort, FHS; Framingham Heart Study Cohort (Offspring and Generation 3), CARDIA; Coronary Artery Risk Development in (Young) Adults Study Cohort, CHS; Cardiovascular Health Study Cohort.

4.2.2.4 Statistical analysis

A multiply-adjusted regression analysis approach was used to measure the association of the five MetS-related variants (explanatory variables) with gout (binary response variables) and serum urate (continuous response variable). All logistic and linear regression analyses were done using statistical software R version 3.3.2 (RCore, 2016). Any participant with missing data for any variable was excluded from the various analyses. Allelic ORs and β-estimates were calculated for each variant including age, sex and BMI as primary adjustors in the regression model. For Polynesian individuals, self-reported grandparental ancestry was included as an additional adjustor in the various analyses. To increase the power of analysis, all NZ Polynesian subgroups were combined separately and with the European data sets in meta-analysis using the Meta package within R (http://CRAN.R-project.org/package=meta, 2014) using a fixed-effect model. For a meta-analysis showing heterogeneity ($P_{Het} < 0.05$), the fixed-effect model was replaced with a random-effect model. A $P \leq 0.05$ was used to indicate the threshold for nominal statistical significance between response and explanatory variables in regression and meta models.

Linkage disequilibrium was calculated between two *MC4R* variants using the information from 1000 Genome database (http://browser.1000genomes.org/) for both European and Chinese datasets. Haploview v4.2 was used to generate LD plots where $r^2 \geq 60$ was set as the threshold for significant LD. Power to detect a $P_{OR} < 0.05$ was calculated in the NZ Polynesian, European and UK Biobank data sets following Johnson *et al.* (2001) methodology. Additionally, power to detect an effect size (ß) of 0.02 was calculated for non-gout datasets using the methodology described by Cohen (1988) and Selya *et al.* (2012).

4.2.3 Results

The two selected variants from the *MC4R* gene, *rs17700633* and *rs17782313*, were not in LD with each other, both in the European and Chinese sample sets in the 1000 Genomes database (Figure 4.2).

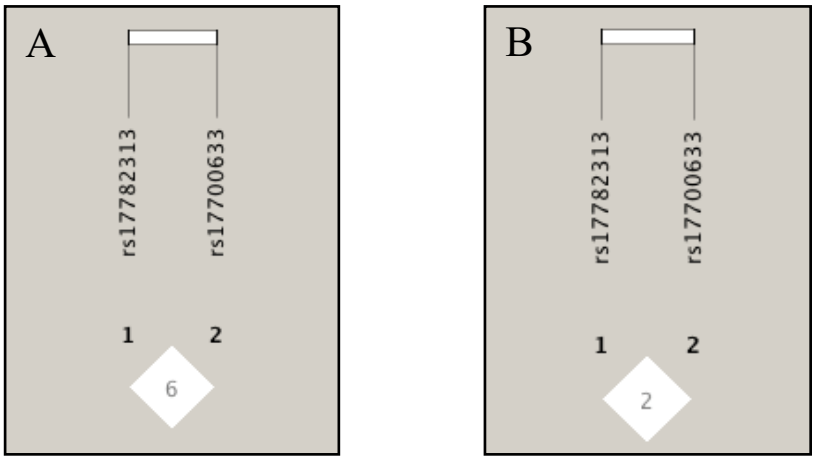


Figure 4.2: Linkage disequilibrium (LD) plot indicating 'R-squared/r^2' values between the 2 *MC4R* variants in (A) European and (B) Chinese populations. An $r^2 \geq 60$ in the above plot indicates LD between the particular variants. Information for variant location, rs ID and LD values are from 1000 Genome database (http://browser.1000genomes.org/). The plots were generated using Haploview v4.2.

Both NZ Polynesian and combined European data sets were highly powered (> 90%) to detect a moderate effect (OR = 1.5) at an altered allele frequency > 0.1 (Figure 4.3). The UK Biobank data set was highly powered (100%) to detect a weaker effect (OR = 1.2) at an altered allele frequency > 0.1 (Figure 4.4).

All non-gout European sample sets used for serum urate association analysis were adequately powered (> 80%) to detect an effect size (ß) of > 0.02 (Figure 4.5).

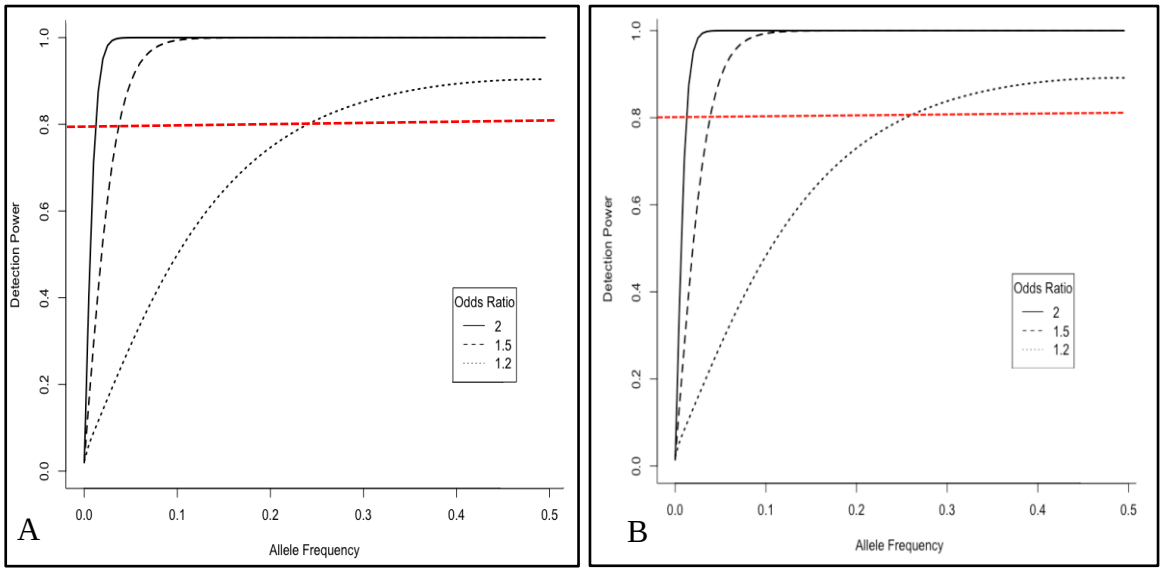


Figure 4.3: Association detection power in New Zealand European (A) and Polynesian (B) sample sets across a range of odds ratio (effect sizes) and altered allele frequencies. The broken red line indicates an adequate detection power ≥ 80%.

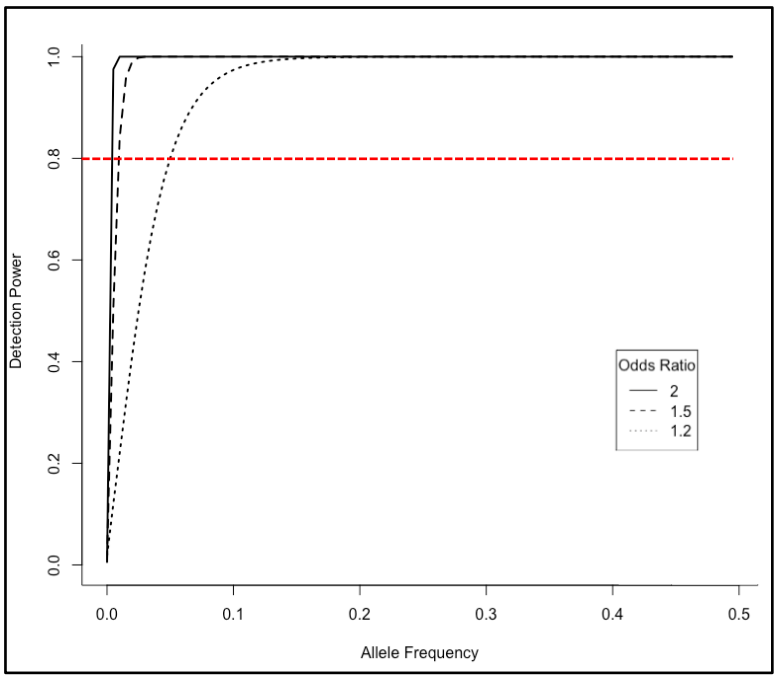


Figure 4.4: Association detection power in the UK Biobank dataset across a range of odds ratio (effect sizes) and altered allele frequencies. The broken red line indicates an adequate detection power ≥ 80%.

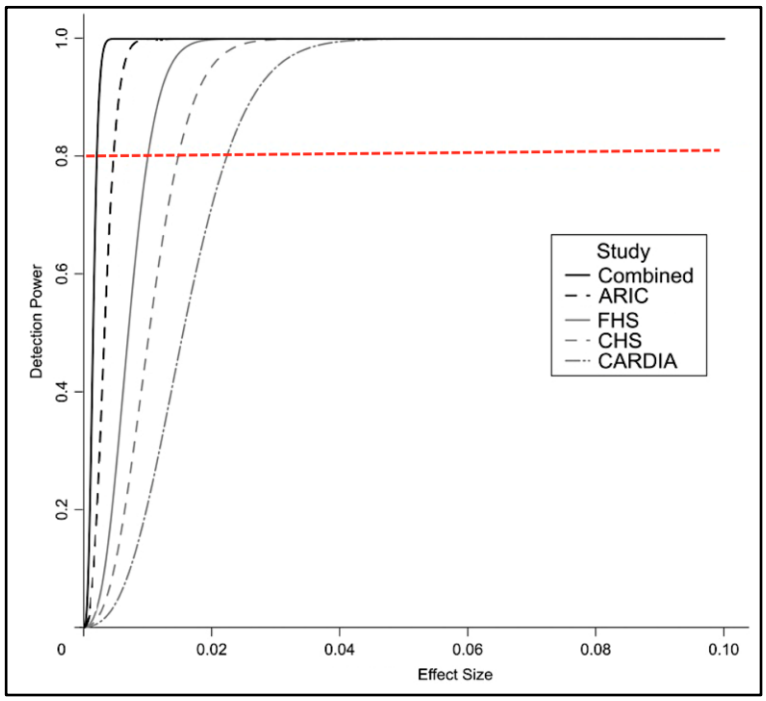


Figure 4.5: Detection power in the European datasets used for serum urate association analyses across a range of effect sizes. The broken red line indicates an adequate detection power ≥ 80%.

4.2.3.1 Association analysis results for the *ADRB3* variant

The G allele (Arg64 allele) of *rs4994* was significantly associated with reduced risk of gout in the UK Biobank data (OR = 0.88, P = 0.04) after including age, sex and BMI as potential confounders. For the NZ Polynesian datasets, similar association was observed only in the WP sample set (OR = 0.62, $P_{\text{Unadjusted}}$ = 0.04). However, the association became insignificant when adjusted for potential confounders (OR = 0.61, P = 0.08) (Table 4.6). The G allele did not show any association with gout in any other Polynesian and European data set (Table 4.6). The protective effect of the G allele was still observed in WP sample set when adjusted for T2DM (OR = 0.57, P = 0.05), hypertension (OR = 0.56, P = 0.05) and renal dysfunction (OR = 0.53, P = 0.03) (Table 4.7). Including renal dysfunction as a potential adjustor also showed that the addition of each G allele was significantly associated with increasing susceptibility to gout by 86% in the EPZ data set (OR = 1.86, P = 0.05) (Table 4.7). In order to increase the power of analysis, all European and NZ Polynesian case-control groups were combined together and separately with the UK Biobank group. No significant association of the G allele of *rs4994* with gout was observed for the combined NZ Polynesian data set (OR = 0.98, P = 0.88, P_{Het} = 0.10) (Table 4.9, Appendix B Figure 4.1). However, on combining the European and UK Biobank data sets together, the G allele showed a significant protective association with gout (OR = 0.89, P = 0.03, P_{Het} = 0.22) (Table 4.9, Appendix B Figure 4.1). The protective association was sustained when all Polynesian and European data sets were combined together in single meta-analysis (OR = 0.91, P = 0.04, P_{Het} = 0.14) (Table 4.9, Appendix B Figure 4.1).

When tested for an association with serum urate in non-gout sample sets, the Arg64 (G) allele of *rs4994* showed a strong positive association with serum urate only in WP individuals (β (mmol L^{-1}) = 0.036, P = 0.004) (Table 4.8). The G allele did not show any association with serum urate in any other sample set (Table 4.8). Combining all NZ Polynesian data sets together by a meta-analysis did not show any association of the G allele of *rs4994* with urate (β (mmol L^{-1}) = 0.008, P = 0.16, P_{Het} = 0.08) (Table 4.9, Appendix B Figure 4.2). Also, no association of the G allele of *rs4994* with urate was found when all European sample sets were combined together (β (mmol L^{-1}) = 0.0003, P = 0.84, P_{Het} = 0.62) (Table 4.9, Appendix B Figure 4.2) and with Polynesian sample sets (β (mmol L^{-1}) = 0.0007, P = 0.61, P_{Het} = 0.19) (Table 4.9, Appendix B Figure 4.2).

4.2.3.2 Association analysis results for the *MC3/4R* variants

4.2.3.2.1 *MC3R: rs3827103*

The A allele (Ile81 allele) of *rs3827103* was significantly associated with a decreased risk of gout in the EPWP data set (OR = 0.43, $P_{\text{Unadjusted}}$ = 0.02) (Table 4.6). However, including age, sex, BMI and grand-parental ancestry estimates as potential adjustors weakened the association (OR = 0.44, *P* = 0.07) (Table 4.6). A similar but weaker trend of protective association between the Ile81 allele and gout risk was observed in the EPZ data set (OR = 0.64, *P* = 0.06) in multiply-adjusted regression (Table 4.6). No such association was observed in any other NZ Polynesian and/or European data sets (Table 4.6). Including various comorbidities in the analyses showed a nominal protective association between the A allele of *rs3827103* and gout in the EPZ, EPWP and EUR sample sets when adjusted for hypertension (OR = 0.60, *P* = 0.05), renal dysfunction (OR = 0.41, *P* = 0.05) and dyslipidaemia (OR = 0.75, *P* = 0.05), respectively (Table 4.7). The NZ Polynesian sample showed a trend towards protective association of the Ile81 allele with gout when combined in meta-analysis (OR = 0.85, *P* = 0.07, P_{Het} = 0.22) (Table 4.9, Appendix B Figure 4.3). Increasing the sample size in meta-analysis (combining all NZ Polynesian and European datasets) indicated a positive, yet no difference in effect size, association between the A allele and gout (OR = 0.92, *P* = 0.03, P_{Het} = 0.27) (Table 4.9, Appendix B Figure 4.3) in comparison to when all European samples were combined together (OR = 0.92, *P* = 0.14, P_{Het} = 0.27) (Table 4.9, Appendix B Figure 4.3).

Linear regression analysis was performed to test for association of the Ile81 allele of *rs3827103* with urate in non-gout data sets. A weak trend of both negative and positive association of the A allele with urate was observed in the EPN (β (mmol L^{-1}) = -0.01, *P* = 0.06) and EPZ (β (mmol L^{-1}) = 0.02, *P* = 0.06) sample sets, which did not retain after adjustment for potential confounders (EPN: β (mmol L^{-1}) = -0.001, *P* = 0.10; EPZ: (β (mmol L^{-1}) = 0.004, *P* = 0.66) (Table 4.8). Meta-analysing NZ Polynesian (β (mmol L^{-1}) = -0.001, *P* = 0.88, P_{Het} = 0.25) (Table 4.9, Appendix B Figure 4.4) and European (β (mmol L^{-1}) = 0.0001, *P* = 0.95, P_{Het} = 0.62) (Table 4.9, Appendix B Figure 4.4) populations separately and together (β (mmol L^{-1}) = 0.0001, *P* = 0.99, P_{Het} = 0.53) also

did not show any association of the A allele with serum urate (Table 4.9, Appendix B Figure 4.4).

4.2.3.2.2 *MC4R: rs17700633*

The A allele of *rs17700633* did not show any association with gout in any of the Polynesian and European data sets (Table 4.6). There was only a weak trend of protective association of the A allele with gout in the WP sample set in adjusted analysis (OR = 0.19, P = 0.08) (Table 4.6). The association followed a similar pattern when adjusted for comorbid conditions, T2DM (OR = 0.73, P = 0.06) and renal dysfunction (OR = 0.76, P = 0.09) (Table 4.7). Combining the Polynesian and European study groups together (OR = 1.03, P = 0.31, P_{Het} = 0.48) (Table 4.9, Appendix B Figure 4.5) and separately (Polynesian: OR = 0.97, P = 0.76, P_{Het} = 0.29; European: OR = 1.03, P = 0.24, P_{Het} = 0.76) (Table 4.9, Appendix B Figure 4.5) in meta-analyses did not show any association of the A allele with gout (Table 4.9).

No association of the A allele of *rs17700633* with serum urate was found in both unadjusted and adjusted regression analyses in any of the study groups (Table 4.8). However, combining the Polynesian and European data sets together in meta-analysis indicated a positive but weak association of the A allele with serum urate (β (mmol L^{-1}) = 0.001, P = 0.07, P_{Het} = 0.96) (Table 4.9, Appendix B Figure 4.6). No such trend was observed when the Polynesian (β (mmol L^{-1}) = 0.005, P = 0.18, P_{Het} = 0.97) and European (β (mmol L^{-1}) = 0.001, P = 0.12, P_{Het} = 0.79) data sets were combined separately (Table 4.9, Appendix B Figure 4.6).

4.2.3.2.3 *MC4R: rs17782313*

The C allele of *rs17782313* was significantly associated with an increased risk of gout in the EPZ sample set (OR = 1.51, $P_{Unadjusted}$ = 0.03). The association was, however, not significant when potential confounders were added as adjustors in the analysis (OR = 1.56, P = 0.11) (Table 4.6). In the EUR group, each additional C allele was found to be associated with a 23% increase in gout risk (OR = 1.23, P = 0.02) (Table 4.6). No such association was observed for any other population group (Table 4.6). The association followed a consistent, but stronger, pattern after including renal dysfunction as an adjustor in the analysis (OR = 1.23, P = 0.01) (Table 4.7). Combining European datasets

in meta-analysis showed a 7% increase in the risk of developing gout with each additional C allele (OR = 1.07, P = 0.02, P_{Het} = 0.11) (Table 4.9, Appendix B Figure 4.7). The association retained a similar pattern in Polynesian and European combined meta-analysis (OR = 1.06, P = 0.03, P_{Het} = 0.23) that was not observed when only Polynesian data sets were combined (OR = 0.96, P = 0.77, P_{Het} = 0.23) (Table 4.9, Appendix B Figure 4.7).

The C allele of *rs17782313* was found to be positively associated with serum urate concentrations in non-gout individuals from the FHS data set (β (mmol L^{-1}) = 0.004, P = 0.01) (Table 4.8). A strong association of the C allele was also observed in the CHS data set (β (mmol L^{-1}) = 0.006, $P_{Unadjusted}$ = 0.008), which became weaker when adjusted for covariates (β (mmol L^{-1}) = 0.004, P = 0.07) (Table 4.8). Meta-analysis indicated a positive association of the C allele of *rs17782313* with serum urate in non-gout European individuals (β (mmol L^{-1}) = 0.002, P = 0.007, P_{Het} = 0.47) (Table 4.9, Appendix B Figure 4.8). No such association was observed in the combined Polynesian data set (β (mmol L^{-1}) = 0.004, P = 0.52, P_{Het} = 0.57) (Table 4.9, Appendix B Figure 4.8). However, when the Polynesian and European data sets were combined together, the C allele showed a positive association with serum urate (β (mmol L^{-1}) = 0.002, P = 0.006, P_{Het} = 0.69) (Table 4.9, Appendix B Figure 4.8).

4.2.3.3 Association analysis results for the *ADTRP* variant

No association was observed with the A allele of *rs6903956* and gout in any of the NZ Polynesian or European data sets (Table 4.6). The association remained insignificant even after including comorbidities, T2DM, hypertension, renal dysfunction and dyslipidaemia as potential confounders (Table 4.7). Moreover, the A allele was not associated with gout when Polynesian and European data sets were combined together (OR = 0.99, P = 0.71, P_{Het} = 0.74) and separately (Polynesian: OR = 1.02, P = 0.83, P_{Het} = 0.49; European: OR = 0.98, P = 0.66, P_{Het} = 0.95) in meta-analyses (Table 4.9, Appendix B Figure 4.9).

The A allele was found to be negatively associated with serum urate in the NPH data set (β (mmol L^{-1}) = -0.035, P = 0.004) (Table 4.8). However, the same allele indicated a positive association with serum urate concentrations in the ARIC European data set (β (mmol L^{-1}) = 0.002, P = 0.05) (Table 4.8). Combining all NZ Polynesian data sets together in a meta-analysis showed significant negative association of the A allele

with serum urate (β (mmol L^{-1}) = -0.011, P = 0.02, P_{Het} = 0.08) (Table 4.9, Appendix B Figure 4.10). No such association was observed when all European datasets were combined (β (mmol L^{-1}) = 0.001, P = 0.11, P_{Het} = 0.61) and when European and Polynesian data sets were combined together in meta-analyses (β (mmol L^{-1}) = 0.001, P = 0.22, P_{Het} = 0.05) (Table 4.9, Appendix B Figure 4.10).

Table 4.6: Genotype/allele frequencies and association analysis of MetS-related variants with gout

Population	Group	Genotype/Altered Allele Frequency				Unadjusted		Adjusted*		HWE
						OR [95% CI]	*P*	OR [95% CI]	*P*	
ADRB3: rs4994										
		AA	**AG**	**GG**	**G**					
EPN	Case	266 (0.796)	66 (0.197)	2 (0.006)	70 (0.10)	0.89 [0.64 ; 1.25]	0.53	1.08 [0.69 ; 1.68]	0.72	0.33
	Control	306 (0.780)	82 (0.209)	4 (0.011)	90 (0.114)					0.56
EPZ	Case	122 (0.777)	33 (0.210)	2 (0.012)	37 (0.117)	1.49 [0.95 ; 2.35]	0.07	1.66 [0.91 ; 3.08]	0.09	0.89
	Control	262 (0.842)	47 (0.151)	2 (0.006)	51 (0.081)					0.94
WP	Case	336 (0.915)	31 (0.084)	0 (0.000)	31 (0.042)	0.62 [0.38 ; 0.99]	0.04	0.61 [0.34 ; 1.06]	0.08	0.39
	Control	272 (0.877)	34 (0.109)	4 (0.012)	42 (0.067)					0.02
EPWP	Case	27 (0.931)	2 (0.069)	0 (0.000)	2 (0.034)	0.31 [0.04 ; 1.21]	0.13	0.31 [0.04 ; 1.44]	0.17	0.84
	Control	57 (0.814)	13 (0.185)	0 (0.000)	13 (0.092)					0.39
NPH	Case	213 (0.803)	52 (0.196)	0 (0.000)	52 (0.098)	1.07 [0.66 ; 1.73]	0.77	0.95 [0.47 ; 1.90]	0.89	0.07
	Control	160 (0.833)	28 (0.144)	4 (0.020)	36 (0.093)					0.05
EUR	Case	1397 (0.838)	262 (0.157)	7 (0.004)	276 (0.083)	1.00 [0.76 ; 1.30]	0.99	1.11 [0.77 ; 1.61]	0.57	0.15
	Control	801 (0.834)	151 (0.157)	4 (0.004)	159 (0.083)					0.27
UK Biobank	Case	2099 (0.863)	321 (0.132)	12 (0.004)	345 (0.071)	0.85 [0.78 ; 0.97]	0.02	0.88 [0.79 ; 0.99]	0.04	0.94
	Control	87102 (0.845)	15252 (0.148)	635 (0.006)	16522 (0.080)					0.24
MC3R: rs3827103										
		GG	**AG**	**AA**	**A**					
EPN	Case	128 (0.389)	147 (0.446)	54 (0.164)	255 (0.387)	0.95 [0.77 ; 1.18]	0.69	1.04 [0.78 ; 1.38]	0.77	0.28
	Control	133 (0.341)	204 (0.523)	53 (0.135)	310 (0.397)					0.06
EPZ	Case	88 (0.560)	61 (0.388)	8 (0.051)	77 (0.245)	1.05 [0.76 ; 1.45]	0.73	0.64 [0.39 ; 1.02]	0.06	0.53
	Control	179 (0.581)	113 (0.366)	16 (0.051)	145 (0.235)					0.73
WP	Case	239 (0.654)	108 (0.295)	18 (0.049)	144 (0.197)	0.97 [0.74 ; 1.26]	0.82	0.79 [0.57 ; 1.09]	0.15	0.21
	Control	203 (0.636)	103 (0.322)	13 (0.041)	129 (0.202)					0.98
EPWP	Case	20 (0.606)	13 (0.393)	0 (0.000)	13 (0.196)	0.43 [0.21 ; 0.88]	0.02	0.44 [0.17 ; 1.04]	0.07	0.15
	Control	32 (0.421)	36 (0.473)	8 (0.105)	52 (0.342)					0.64
NPH	Case	84 (0.381)	104 (0.472)	32 (0.145)	168 (0.381)	0.90 [0.67 ; 1.21]	0.49	0.91 [0.61 ; 1.39]	0.68	0.98

EUR	Control	58 (0.308)	108 (0.574)	22 (0.117)	152 (0.404)	0.89 [0.72 ; 1.09]	0.27	0.81 [0.61 ; 1.06]	0.12	0.01
	Case	1427 (0.857)	229 (0.137)	9 (0.005)	247 (0.074)					0.95
	Control	812 (0.844)	141 (0.146)	9 (0.009)	159 (0.082)					0.31
UK Biobank	Case	2057 (0.845)	366 (0.150)	9 (0.004)	384 (0.078)	0.94 [0.85 ; 1.05]	0.34	0.96 [0.86 ; 1.06]	0.46	0.08
	Control	86619 (0.841)	15689 (0.152)	681 (0.006)	17051 (0.082)					0.31
MC4R: rs17700633										
		GG	**AG**	**AA**	**A**					
EPN	Case	128 (0.455)	125 (0.444)	28 (0.099)	181 (0.322)	1.14 [0.89 ; 1.45]	0.28	1.14 [0.85 ; 1.54]	0.36	0.75
	Control	188 (0.484)	171 (0.441)	29 (0.074)	229 (0.295)					0.24
EPZ	Case	67 (0.527)	50 (0.393)	10 (0.078)	70 (0.275)	1.10 [0.79 ; 1.52]	0.56	1.21 [0.78 ; 1.91]	0.38	0.87
	Control	169 (0.555)	114 (0.375)	21 (0.069)	156 (0.256)					0.76
WP	Case	246 (0.677)	102 (0.281)	15 (0.041)	132 (0.181)	0.84 [0.65 ; 1.09]	0.19	0.75 [0.55 ; 1.03]	0.08	0.28
	Control	205 (0.641)	95 (0.296)	20 (0.062)	135 (0.211)					0.05
EPWP	Case	19 (0.575)	14 (0.424)	0 (0.000)	14 (0.212)	0.68 [0.34 ; 1.29]	0.26	0.77 [0.35 ; 1.64]	0.52	0.12
	Control	42 (0.552)	24 (0.315)	10 (0.131)	44 (0.289)					0.04
NPH	Case	114 (0.518)	94 (0.427)	12 (0.545)	118 (0.268)	0.91 [0.66 ; 1.26]	0.6	0.95 [0.61 ; 1.51]	0.84	0.18
	Control	92 (0.473)	81 (0.437)	12 (0.064)	105 (0.283)					0.29
EUR	Case	770 (0.464)	726 (0.437)	162 (0.097)	1050 (0.316)	1.07 [0.95 ; 1.21]	0.23	1.01 [0.85 ; 1.19]	0.91	0.63
	Control	475 (0.494)	394 (0.410)	92 (0.095)	578 (0.300)					0.43
UK Biobank	Case	1150 (0.482)	1016 (0.426)	217 (0.091)	1450 (0.304)	1.03 [0.97 ; 1.11]	0.23	1.04 [0.97 ; 1.11]	0.27	0.72
	Control	49892 (0.495)	42032 (0.417)	8829 (0.087)	59690 (0.296)					0.86
MC4R: rs17782313										
		TT	**CT**	**CC**	**C**					
EPN	Case	241 (0.899)	21 (0.078)	6 (0.022)	33 (0.061)	1.11 [0.71 ; 1.71]	0.65	1.12 [0.64 ; 1.96]	0.68	1.40E-07
	Control	347 (0.892)	41 (0.105)	1 (0.002)	43 (0.055)					0.85
EPZ	Case	78 (0.634)	40 (0.325)	5 (0.040)	50 (0.203)	1.51 [1.02 ; 2.22]	0.03	1.56 [0.91 ; 2.76]	0.11	0.96
	Control	222 (0.730)	76 (0.250)	6 (0.019)	88 (0.144)					0.86
WP	Case	220 (0.700)	79 (0.251)	15 (0.047)	109 (0.173)	0.88 [0.67 ; 1.16]	0.39	0.83 [0.60 ; 1.16]	0.29	0.02
	Control	209 (0.655)	97 (0.304)	13 (0.041)	123 (0.192)					0.68
EPWP	Case	21 (0.700)	9 (0.300)	0 (0.000)	9 (0.15)	0.66 [0.28 ; 1.38]	0.29	0.61 [0.22 ; 1.44]	0.27	0.33
	Control	49 (0.644)	21 (0.276)	6 (0.078)	33 (0.217)					0.11
NPH	Case	192 (0.868)	27 (0.1222)	2 (0.009)	31 (0.070)	1.18 [0.69 ; 2.07]	0.53	0.86 [0.41 ; 1.84]	0.70	0.34

EUR	Control	167 (0.893)	18 (0.096)	2 (0.012)	22 (0.058)	1.12 [0.99 ; 1.25]	0.06	1.23 [1.03 ; 1.49]	0.02	0.07
	Case	925 (0.554)	629 (0.376)	115 (0.068)	859 (0.257)					0.56
	Control	566 (0.589)	339 (0.352)	56 (0.058)	451 (0.234)					0.57
UK Biobank	Case	1394 (0.573)	893 (0.367)	145 (0.059)	1183 (0.243)	1.05 [0.98 ; 1.13]	0.12	1.05 [0.98 ; 1.12]	0.13	0.91
	Control	60408 (0.586)	37009 (0.359)	5572 (0.054)	48153 (0.233)					0.32

ADTRP: rs6903956

		GG	**AG**	**AA**	**A**					
EPN	Case	254 (0.765)	73 (0.219)	5 (0.015)	83 (0.125)	0.83 [0.61 ; 1.14]	0.26	0.95 [0.64 ; 1.44]	0.83	0.92
	Control	282 (0.723)	103 (0.264)	5 (0.012)	113 (0.144)					0.19
EPZ	Case	88 (0.567)	58 (0.374)	9 (0.058)	76 (0.245)	0.89 [0.64 ; 1.22]	0.48	1.04 [0.66 ; 1.63]	0.85	0.89
	Control	164 (0.539)	118 (0.388)	22 (0.072)	162 (0.266)					0.90
WP	Case	313 (0.778)	85 (0.211)	4 (0.010)	93 (0.115)	0.83 [0.61 ; 1.15]	0.27	0.83 [0.57 ; 1.24]	0.37	0.51
	Control	239 (0.749)	74 (0.232)	6 (0.018)	86 (0.134)					0.92
EPWP	Case	22 (0.666)	11 (0.333)	0 (0.000)	11 (0.166)	1.33 [0.59 ; 2.90]	0.47	1.81 [0.65 ; 4.95]	0.24	0.25
	Control	58 (0.783)	13 (0.175)	3 (0.040)	19 (0.128)					0.06
NPH	Case	182 (0.674)	81 (0.300)	7 (0.025)	95 (0.176)	1.22 [0.85 ; 1.77]	0.27	1.33 [0.80 ; 2.24]	0.26	0.56
	Control	138 (0.722)	49 (0.256)	4 (0.021)	57 (0.149)					0.88
EUR	Case	604 (0.372)	791 (0.487)	227 (0.140)	1345 (0.414)	0.99 [0.88 ; 1.12]	0.91	0.99 [0.84 ; 1.17]	0.93	0.21
	Control	361 (0.377)	453 (0.473)	142 (0.148)	737 (0.385)					0.99
UK Biobank	Case	940 (0.386)	1137 (0.467)	355 (0.145)	1847 (0.379)	0.98 [0.93 ; 1.05]	0.66	0.99 [0.93 ; 1.04]	0.67	0.71
	Control	39294 (0.381)	48551 (0.471)	15144 (147)	78839 (0.383)					0.46

*****All values are adjusted for age, sex and body mass index, plus, for grand-parental ancestry estimates for Polynesian datasets. EPN: East Polynesian subjects with high EP ancestry, EPZ: East Polynesian subjects with low EP ancestry, WP: West Polynesian, EPWP; Mixture of East and West Polynesian, NPH; Māori cohort from Ngati Porou Hauora, EUR; NZ European, EUROGT (European Crystal Network Gout Cohort) and AGRIA (Arthritis Genomics Recruitment Initiative in Australia Gout Cohort) combined, UK Biobank; United Kingdom Biobank Cohort. OR [95% CI]; Odds ratio [95% confidence interval], *P = p-value* for ORs, HWE; Values for Hardy Weinberg Equilibrium.

Table 4.7: Association analysis of MetS-related variants with gout adjusted for comorbidities

Population	Baseline adjustments		Comorbidities							
			T2DM*		Hypertension*		Renal dysfunction*		Dyslipidaemia*	
	OR [95% CI]	*P*	OR [95% CI]	*P*	OR [95% CI]	*P*	OR [95% CI]	*P*	OR [95% CI]	*P*
ADRB3: rs4994										
EPN	1.08 [0.69 ; 1.68]	0.72	1.03 [0.66 ; 1.62]	0.86	1.08 [0.69 ; 1.70]	0.72	0.93 [0.58 ; 1.49]	0.77	0.92 [0.14 ; 1.10]	0.72
EPZ	1.66 [0.91 ; 3.08]	0.09	1.78 [0.96 ; 3.34]	0.06	1.56 [0.80 ; 3.05]	0.18	1.86 [0.98 ; 3.60]	0.05	1.72 [0.90 ; 3.13]	0.09
WP	0.61 [0.34 ; 1.06]	0.08	0.57 [0.32 ; 1.01]	0.05	0.56 [0.31 ; 1.01]	0.05	0.53 [0.29 ; 0.95]	0.03	0.62 [0.33 ; 1.13]	0.12
EPWP	0.31 [0.04 ; 1.44]	0.17	0.31 [0.04 ; 1.45]	0.17	0.35 [0.04 ; 1.76]	0.24	0.41 [0.05 ; 2.02]	0.31	0.61 [0.07 ; 3.30]	0.58
NPH	0.95 [0.47 ; 1.90]	0.89	0.93 [0.46 ; 1.89]	0.85	0.92 [0.45 ; 1.90]	0.83	0.98 [0.48 ; 1.98]	0.96	0.96 [0.47 ; 1.94]	0.91
EUR	1.11 [0.77 ; 1.61]	0.57	1.12 [0.77 ; 1.63]	0.55	1.12 [0.76 ; 1.65]	0.56	1.11 [0.76 ; 1.63]	0.58	1.13 [0.11 ; 0.63]	0.52
MC3R: rs3827103										
EPN	1.04 [0.78 ; 1.38]	0.77	1.03 [0.77 ; 1.37]	0.81	1.06 [0.79 ; 1.43]	0.65	1.05 [0.78 ; 1.42]	0.66	1.01 [0.73 ; 1.37]	0.96
EPZ	0.64 [0.39 ; 1.02]	0.06	0.65 [0.40 ; 1.04]	0.08	0.60 [0.36 ; 1.00]	0.05	0.68 [0.41 ; 1.11]	0.12	0.74 [0.45 ; 1.21]	0.24
WP	0.79 [0.57 ; 1.09]	0.15	0.77 [0.56 ; 1.07]	0.13	0.79 [0.57 ; 1.10]	0.16	0.78 [0.56 ; 1.08]	0.14	0.87 [0.62 ; 1.22]	0.42
EPWP	0.44 [0.17 ; 1.04]	0.07	0.43 [0.17 ; 1.01]	0.06	0.44 [0.16 ; 1.08]	0.08	0.41 [0.14 ; 0.98]	0.05	0.56 [0.19 ; 1.45]	0.25
NPH	0.91 [0.61 ; 1.39]	0.68	0.87 [0.56 ; 1.34]	0.53	0.84 [0.55 ; 1.29]	0.45	0.90 [0.58 ; 1.39]	0.64	0.91 [0.59 ; 1.38]	0.65
EUR	0.81 [0.61 ; 1.06]	0.12	0.77 [0.58 ; 1.03]	0.08	0.77 [0.58 ; 1.04]	0.08	0.81 [0.61 ; 1.08]	0.15	0.75 [0.56 ; 1.01]	0.05
MC4R: rs17700633										
EPN	1.14 [0.85 ; 1.54]	0.36	1.18 [0.87 ; 1.59]	0.27	1.17 [0.86 ; 1.59]	0.29	1.16 [0.86 ; 1.58]	0.31	1.27 [0.91 ; 1.78]	0.14
EPZ	1.21 [0.78 ; 1.91]	0.38	1.24 [0.79 ; 1.96]	0.33	1.21 [0.75 ; 1.95]	0.42	1.27 [0.79 ; 2.04]	0.32	1.28 [0.81 ; 2.06]	0.28
WP	0.75 [0.55 ; 1.03]	0.08	0.73 [0.53 ; 1.01]	0.06	0.79 [0.57 ; 1.09]	0.15	0.76 [0.55 ; 1.04]	0.09	0.77 [0.55 ; 1.07]	0.12

EPWP	0.77 [0.35 ; 1.64]	0.52	0.76 [0.33 ; 1.64]	0.50	0.82 [0.36 ; 1.81]	0.63	0.92 [0.41 ; 2.06]	0.85	0.85 [0.36 ; 1.94]	0.71
NPH	0.95 [0.61 ; 1.51]	0.84	0.98 [0.61 ; 1.57]	0.95	0.97 [0.62 ; 1.55]	0.92	0.98 [0.61 ; 1.58]	0.95	1.02 [0.63 ; 1.63]	0.94
EUR	1.01 [0.85 ; 1.19]	0.91	1.01 [0.84 ; 1.21]	0.91	1.02 [0.81 ; 1.15]	0.70	1.00 [0.84 ; 1.20]	0.94	1.01 [0.83 ; 1.21]	0.97
MC4R: rs17782313										
EPN	1.12 [0.64 ; 1.96]	0.68	1.11 [0.63 ; 1.99]	0.7	0.13 [0.65 ; 2.01]	0.66	1.24 [0.71 ; 2.22]	0.45	0.85 [0.45 ; 1.59]	0.61
EPZ	1.56 [0.91 ; 2.76]	0.11	1.55 [0.89 ; 2.74]	0.12	1.50 [0.84 ; 2.69]	0.16	1.71 [0.95 ; 3.11]	0.07	1.60 [0.91 ; 2.91]	0.11
WP	0.83 [0.60 ; 1.16]	0.29	0.84 [0.61 ; 1.17]	0.32	0.92 [0.65 ; 1.31]	0.65	0.86 [0.62 ; 1.21]	0.41	0.86 [0.61 ; 1.22]	0.41
EPWP	0.61 [0.22 ; 1.44]	0.27	0.59 [0.22 ; 1.42]	0.27	0.54 [0.18 ; 1.38]	0.22	0.57 [0.19 ; 1.44]	0.26	0.49 [0.16 ; 1.29]	0.17
NPH	0.86 [0.41 ; 1.84]	0.70	0.83 [0.38 ; 1.81]	0.64	0.83 [0.38 ; 1.81]	0.63	0.85 [0.39 ; 1.89]	0.69	0.91 [0.42 ; 2.01]	0.81
EUR	1.23 [1.03 ; 1.49]	0.02	1.19 [0.98 ; 1.44]	0.06	1.17 [0.97 ; 1.43]	0.09	1.23 [1.02 ; 1.51]	0.01	1.19 [0.97 ; 1.45]	0.08
ADTRP: rs6903956										
EPN	0.95 [0.63 ; 1.44]	0.83	0.95 [0.62 ; 1.44]	0.81	0.93 [0.61 ; 1.44]	0.75	0.84 [0.54 ; 1.29]	0.43	0.81 [0.51 ; 1.28]	0.37
EPZ	1.04 [0.66 ; 1.63]	0.85	1.03 [0.65 ; 1.62]	0.88	1.15 [0.72 ; 1.85]	0.55	1.08 [0.67 ; 1.74]	0.72	1.09 [0.68 ; 1.75]	0.71
WP	0.83 [0.56 ; 2.24]	0.37	0.78 [0.52 ; 1.16]	0.22	0.85 [0.56 ; 1.27]	0.42	0.84 [0.57 ; 1.26]	0.40	0.94 [0.62 ; 1.44]	0.79
EPWP	1.81 [0.65 ; 4.95]	0.24	1.81 [0.65 ; 4.92]	0.24	1.82 [0.62 ; 5.17]	0.25	1.81 [0.61 ; 5.21]	0.27	1.36 [0.41 ; 4.24]	0.59
NPH	1.33 [0.81 ; 2.24]	0.26	1.24 [0.75 ; 2.11]	0.39	1.21 [0.72 ; 2.08]	0.47	1.22 [0.73 ; 2.08]	0.43	0.98 [0.82 ; 1.17]	0.86
EUR	0.99 [0.84 ; 1.17]	0.93	0.97 [0.82 ; 1.17]	0.77	0.98 [0.83 ; 1.17]	0.9	0.98 [0.82 ; 1.17]	0.86	1.25 [0.75 ; 2.11]	0.39

*****All values are adjusted for age, sex and body mass index, plus, for grand-parental ancestry estimates for Polynesian datasets. EPN: East Polynesian subjects with high EP ancestry, EPZ: East Polynesian subjects with low EP ancestry, WP: West Polynesian, EPWP; Mixture of East and West Polynesian, NPH; Māori cohort from Ngati Porou Hauora, EUR; NZ European, EUROGT (European Crystal Network Gout Cohort) and AGRIA (Arthritis Genomics Recruitment Initiative in Australia Gout Cohort) combined. OR [95% CI]; Odds ratio [95% confidence interval], *P = p-value* for ORs.

Table 4.8: Association analysis of MetS-related variants with serum urate

Gene: rs ID	Population	Unadjusted		Adjusted *	
		ß [95% CI]	*P*	ß [95% CI]	*P*
ADRB3: rs4994	EPN	0.004 [-0.016 ; 0.026]	0.66	0.003 [-0.015 ; 0.023]	0.7
	EPZ	0.017 [-0.018 ; 0.054]	0.33	0.011 [-0.018 ; 0.040]	0.45
	WP	0.036 [0.009 ; 0.063]	0.01	0.036 [0.011 ; 0.062]	0.004
	EPWP	-0.021 [-0.088 ; 0.044]	0.51	-0.018 [-0.076 ; 0.039]	0.53
	NPH	-0.006 [-0.041 ; 0.026]	0.68	-0.008 [-0.051 ; 0.008]	0.16
	NZ EUR	-0.025 [-0.054 ; 0.004]	0.09	-0.012 [-0.037 ; 0.012]	0.29
	ARIC EUR	0.001 [-0.003 ; 0.006]	0.56	0.001 [-0.002 ; 0.005]	0.49
	FHS EUR	0.001 [-0.005 ; 0.006]	0.9	-0.001 [-0.005 ; 0.003]	0.61
	CARDIA EUR	-0.048 [-0.026 ; 0.0314]	0.87	0.011 [-0.013 ; 0.035]	0.37
	CHS EUR	0.002 [-0.006 ; 0.011]	0.55	0.001 [-0.006 ; 0.008]	0.82
MC3R: rs3827103	EPN	-0.014 [-0.029 ; 0.001]	0.06	-0.001 [-0.025 ; 0.002]	0.1
	EPZ	0.021 [-0.002 ; 0.045]	0.06	0.004 [-0.014 ; 0.022]	0.66
	WP	-0.001 [-0.020 ; 0.019]	0.95	0.0001 [-0.017 ; 0.018]	0.98
	EPWP	0.012 [-0.032 ; 0.057]	0.57	-0.001 [-0.042 ; 0.041]	0.95
	NPH	0.028 [0.006 ; 0.049]	0.01	0.016 [-0.004 ; 0.036]	0.11
	NZ EUR	0.012 [-0.009 ; 0.034]	0.27	0.009 [-0.0001 ; 0.027]	0.29
	ARIC EUR	-0.00004 [-0.005 ; 0.005]	0.98	-0.001 [-0.005 ; 0.003]	0.6
	FHS EUR	-0.003 [-0.009 ; 0.002]	0.28	-0.001 [-0.004 ; 0.003]	0.86
	CARDIA EUR	-0.006 [-0.037 ; 0.024]	0.67	-0.004 [-0.030 ; 0.021]	0.72
	CHS EUR	0.006 [-0.002 ; 0.015]	0.16	0.004 [-0.003 ; 0.011]	0.29
MC4R: rs17700633	EPN	0.008 [-0.006 ; 0.023]	0.26	0.0063 [-0.006 ; 0.019]	0.34
	EPZ	-0.017 [-0.041 ; 0.006]	0.14	0.000[-0.018 ; 0.019]	0.93
	WP	0.008 [-0.010 ; 0.027]	0.36	0.004 [-0.012 ; 0.021]	0.58
	EPWP	0.008 [-0.030 ; 0.048]	0.65	0.011 [-0.023 ; 0.046]	0.5
	NPH	0.006 [-0.016 ; 0.029]	0.58	0.008 [-0.013 ; 0.029]	0.46
	NZ EUR	0.001 [-0.012 ; 0.015]	0.84	-0.002 [0.001 ; 0.009]	0.67
	ARIC EUR	0.0002 [-0.002 ; 0.003]	0.86	0.001 [-0.002 ; 0.002]	0.77
	FHS EUR	0.001 [-0.002 ; 0.005]	0.41	0.002 [-0.001 ; 0.004]	0.09
	CARDIA EUR	0.006 [-0.011 ; 0.023]	0.47	0.004 [-0.009 ; 0.018]	0.52
	CHS EUR	0.003 [-0.001 ; 0.008]	0.15	0.002 [-0.002 ; 0.006]	0.41
MC4R: rs17782313	EPN	0.001 [-0.028 ; 0.031]	0.92	0.006 [-0.021 ; 0.033]	0.65
	EPZ	-0.01 [-0.042 ; 0.022]	0.53	0.001 [-0.024 ; 0.026]	0.93
	WP	0.002 [-0.018 ; 0.022]	0.84	0.008 [-0.009 ; 0.026]	0.35
	EPWP	0.02 [-0.022 ; 0.064]	0.34	0.013 [-0.024 ; 0.052]	0.47

	NPH	-0.00005 [-0.044 ; 0.043]	0.99	-0.002 [-0.06 ; 0.017]	0.18
	NZ EUR	-0.0001 [-0.016 ; 0.016]	0.99	0.001 [-0.002 ; 0.014]	0.87
	ARIC EUR	0.001 [-0.001 ; 0.005]	0.3	0.001 [-0.001 ; 0.003]	0.47
	FHS EUR	0.005 [0.001 ; 0.009]	0.01	0.0037 [0.0008 ; 0.006]	0.01
	CARDIA EUR	-0.006 [-0.024 ; 0.012]	0.52	-0.006 [-0.021 ; 0.009]	0.42
	CHS EUR	0.006 [0.002 ; 0.012]	0.008	0.004 [-0.001 ; 0.008]	0.07
ADTRP: rs6903956	EPN	0.007 [-0.014 ; 0.029]	0.48	0.0085 [-0.010 ; 0.028]	0.38
	EPZ	-0.017 [-0.042 ; 0.006]	0.15	-0.015 [-0.034 ; 0.002]	0.09
	WP	-0.011 [-0.035 ; 0.012]	0.36	-0.012 [-0.034 ; 0.009]	0.27
	EPWP	-0.035 [-0.086 ; 0.015]	0.16	-0.005 [-0.055 ; 0.045]	0.83
	NPH	-0.036 [-0.062 ; -0.011]	0.004	-0.035 [-0.059 ; -0.011]	0.004
	NZ EUR	0.006 [-0.007 ; 0.020]	0.35	0.006 [-0.004 ; 0.018]	0.24
	ARIC EUR	0.002 [-0.001 ; 0.005]	0.15	0.002 [-0.00003; 0.005]	0.05
	FHS EUR	0.001 [-0.002 ; 0.004]	0.45	0.0004 [-0.002 ; 0.002]	0.75
	CARDIA EUR	0.0003 [-0.014 ; 0.015]	0.96	-0.00005 [-0.0127 ; 0.011]	0.92
	CHS EUR	0.0006 [-0.004 ; 0.005]	0.78	0.0001 [-0.004 ; 0.004]	0.94

*All values are adjusted for age, sex and BMI for NZ European plus with the estimates of grand-parental ancestry for NZ Polynesian datasets. EP; East Polynesian, WP; West Polynesian, EPWP; Mixture of East and West Polynesian, NPH; Ngati Porou Hauora, NZ EUR; NZ Europeans, ARIC EUR; Europeans from Atherosclerosis risk in community study cohort, FHS; Europeans from Framingham Heart study Cohort (Offspring and Generation 3), CARDIA EUR; Europeans from Coronary Artery Risk Development in (Young) Adults Study Cohort, CHS EUR; Europeans from Cardiovascular Health Study Cohort, rs ID; Reference SNP cluster ID, ß; Beta/effect estimates for the alternate allele, 95% CI; 95% confidence interval, *P*; *p*-values for effect estimates.

Table 4.9: Meta-analysis results of MetS-related variants for association with gout and serum urate

rs ID	Population	Gout (case/control)			Serum urate (mmol L^{-1})		
		OR [95% CI]	P_{OR}	P_{Het}	ß [95% CI]	$P_{ß}$	P_{Het}
rs4994	POLY	0.98 [0.74 ; 1.28]	0.88	0.1	0.008 [-0.003 ; 0.019]	0.16	0.08
	EUR	0.89 [0.80 ; 0.99]	0.03	0.22	0.0003 [-0.0002 ; 0.003]	0.84	0.62
	POLY & EUR	0.91 [0.82 ; 0.99]	0.04	0.14	0.0007 [-0.002 ; 0.003]	0.61	0.19
rs3827103	POLY	0.85 [0.72 ; 1.02]	0.07	0.22	-0.001 [-0.009 ; 0.007]	0.88	0.25
	EUR	0.92 [0.84 ; 1.02]	0.14	0.27	0.0001 [-0.003 ; 0.002]	0.95	0.62
	POLY & EUR	0.92 [0.83 ; 0.99]	0.03	0.27	0.0001 [-0.003 ; 0.002]	0.99	0.53
rs17700633	POLY	0.97 [0.82 ; 1.15]	0.76	0.29	0.005 [-0.003 ; 0.013]	0.18	0.97
	EUR	1.03 [0.97 ; 1.09]	0.24	0.76	0.001 [-0.0003 ; 0.003]	0.12	0.79
	POLY & EUR	1.03 [0.97 ; 1.09]	0.31	0.48	0.001 [-0.0001 ; 0.003]	0.07	0.96
rs17782313	POLY	0.96 [0.76 ; 1.21]	0.77	0.23	0.004 [-0.008 ; 0.015]	0.52	0.57
	EUR	1.07 [1.08 ; 1.14]	0.02	0.11	0.002 [0.001 ; 0.004]	0.007	0.47
	POLY & EUR	1.06 [1.01 ; 1.13]	0.03	0.23	0.002 [0.001 ; 0.003]	0.006	0.69
rs6903956	POLY	1.02 [0.83 ; 1.26]	0.83	0.49	-0.011 [-0.021 ; -0.002]	0.02	0.08
	EUR	0.98 [0.93 ; 1.04]	0.66	0.95	0.001 [-0.0003 ; 0.002]	0.11	0.61
	POLY & EUR	0.99 [0.94 ; 1.04]	0.71	0.74	0.001 [-0.001 ; 0.002]	0.22	0.05

POLY; NZ Polynesian, EUR; NZ and UK Biobank Cohort Europeans for gout analysis and NZ, ARIC, FHS, CHS and CARDIA Cohort Europeans for serum urate analysis, rs ID; Reference SNP cluster ID, OR/ß; Odds ratio/beta or effect estimates for the alternate allele, 95% CI; 95% confidence interval, *P*; *p*-values, $P_{OR/ß}$; *p*-value for odds ratio/beta estimates, Het; Heterogeneity.

SECTION 4.3: DISCUSSION

4.3.1 *ADRB3*

Note: The below findings have recently been published in the *Rheumatology International Journal* (Fatima *et al.*, 2016) (Appendix C).

This study reports association of the Arg64 (G) allele with increased serum urate in WP individuals (β (mmol L^{-1}) = 0.036, P = 0.004). This finding is consistent with the Morcillo *et al.* (2010) study in Spanish individuals that associated Arg64 allele with development of hyperuricaemia. The finding is also in line with the studies that demonstrated an association of the Arg64 allele with increased serum urate and risk of gout in Asian subjects (Huang *et al.*, 2013; Rho *et al.*, 2007; Wang *et al.*, 2011a). The consistent positive association observed in these findings collectively increases the support for a causal role of *ADRB3* and adrenergic system in urate control. However, it was not possible to meta-analyse serum urate findings in the NZ data set with previously published studies as they were either described as secondary findings in conjunction with other metabolic conditions, from a population subgroup or as a binary outcome (Rho *et al.*, 2007; Strazzullo *et al.*, 2001; Wang *et al.*, 2002). My study also reports a protective association of the Arg64 (G) allele of *rs4994* polymorphism with gout in European individuals (OR = 0.89, P = 0.03). The G allele was also found to show a protective association in the WP sample set (OR = 0.62, $P_{\text{Unadjusted}}$ = 0.04) (Table 4.6), which did not change when adjusted for comorbidities, T2DM, hypertension and renal dysfunction (Table 4.7). The direction of association with gout in European and WP individuals conflicted with Wang *et al.* (2011a), who reported the G allele to be associated with increased susceptibility to gout in Han Chinese individuals. The same allele was found to increase the risk of gout by 86% in the EPZ sample set when adjusted for renal dysfunction (OR = 1.86, P = 0.05) (Table 4.6), which is consistent with the Wang *et al.* (2011a) findings in Han Chinese. However, for the WP sample set, the direction of association of the Arg64 allele with gout was opposite to what was observed earlier for serum urate. A possible explanation for the opposite effect of Agr64 allele with gout and hyperuricaemia could be its pleiotropic effect in the WP population, perhaps having a role both in determining hyperuricaemia and in the inflammatory processes leading to gout. The opposing effect of the Arg64 allele in the EPZ population may reflect different

ancestral haplotypes, which is consistent with *ABCG2* having differential effects between EP and WP sample sets (Phipps-Green *et al.*, 2010).

The *ADRB3* gene encodes for beta-3 adrenergic receptor, primarily via its expression in adipose tissue (human visceral fat). The receptor activation, in turn, induces lipolysis in adipose tissue and thermogenesis in skeletal muscles. The receptor is also responsible for delivery of free fatty acids into the portal vein (Emorine *et al.*, 1994), which can disrupt the insulin receptor signaling pathway, thereby leading to increased insulin resistance (Savage *et al.*, 2005). The Arg64 allele has been reported to be associated with increased BMI, obesity and higher IR (Kurokawa *et al.*, 2008; Park *et al.*, 2008; Zhan and Ho, 2005). A decrease in the activity of the receptor due to the Arg64 polymorphism (Ahles and Engelhardt, 2014) could lead to a decrease in lipolysis as well as an increase in fat deposition in adipose tissue. The beta-3 adrenergic receptor is proposed to be a part of 'leptin-sympathetic-leptin-feedback loop'; decrease in the activity of this receptor causes an increase in leptin secretion from the adipose tissue (Mark *et al.*, 2003). Leptin is secreted in the body as an appetite suppressor, thereby opposing the effect of the hormone ghrelin that increases hunger (Chen *et al.*, 2015c). The increase in leptin level has been positively associated with obesity (Pan *et al.*, 2014), while a decrease in leptin has been associated with an improved insulin sensitivity (Wang *et al.*, 2013). An increase in the levels of leptin has also been observed in hyperuricaemic patients (Bedir *et al.*, 2003), which is consistent with association of Arg64 with hyperuricaemia. Obesity and hyperuricaemia can be possibly linked via IR (Li *et al.*, 2013; Modan *et al.*, 1987). My findings for *ADRB3* in relation to the Arg64 allele, supports the study hypothesis that the IR and BMI associated variant (*rs4994*) is involved in determining hyperuricaemia and gout.

4.3.2 *MC3R*

This study also reports a protective association of the Ile81 (A) allele of the *rs3827103* polymorphism of the *MC3R* gene with gout in combined Māori, Pacific Island and European data set (OR = 0.92, P = 0.03) (Table 4.9). Melanocortin 3 receptor (MC3R) is a major receptor for ACTH and α and ß-MSH and is thereby involved in anti-inflammatory response (Daoussis *et al.*, 2014; Getting *et al.*, 2006), weight regulation, energy metabolism and regulation of the cardiovascular system (Feng *et al.*, 2005;

Getting *et al.*, 2006; Tao and Segaloff, 2004). This receptor is located on chromosome 20q13 and Val81Ile or *rs3827103* is a missense variant in the first transmembrane helix of the MC3R protein, involved in the binding of melanocortin peptides. Variations in this domain are predicted to affect melanocortin receptor functions (Schiöth *et al.*, 1998). This variant has been related with *in vitro* diminished functionality and expression of the receptor, showing a significant association with childhood obesity in a case-control study (Feng *et al.*, 2005). However, more recent genetic epidemiological studies have reported that the predisposing effect of the Ile81 allele to common forms of obesity in adults is unlikely (Cieslak *et al.*, 2013; Wannaiampikul *et al.*, 2015). This variant has been reported to be positively associated with developing insulin and leptin resistance in obese Greek Caucasians (Yiannakouris *et al.*, 2004) and increased BMI in Belgian individuals (Zegers *et al.*, 2010) while a number of studies including relatively larger data sets from different populations were unable to find any association between the Ile81 allele (or 'AA' homozygosity) and childhood/adolescent obesity (Cieslak *et al.*, 2013; Obregón *et al.*, 2010) or increased leptin levels (Alsmadi *et al.*, 2014). One possible explanation for these inconsistencies could be the more pronounced damaging effects of the *MC3R* polymorphisms during childhood. The variant, Ile81Val, is a conservative substitution and has been found in homologous regions of MC3/4/5 receptors of normally functioning mammalian models (mice, rat). No association of this variant has been reported for the obesity phenotype in extremely obese African American and Caucasian females (Li *et al.*, 2000), which suggests that Ile81Val polymorphism may not be involved in altering the functionality of human *MC3R*. In addition to these, a study in Māori kindred failed to report a role of the *MC3R* coding region variations in the development of T2D but suggested that other variations in regulatory regions of this gene may possibly be involved in insulin secretion and T2D development (Wong *et al.*, 2002). These inconsistencies are also supportive for non-significant findings of the Ile81 allele with serum urate in my study, especially in context of the hypothesis that obesity-related variants should also be associated with increased serum urate (Table 4.8 and Table 4.9). My findings are also in line with the largest urate GWA study to now, which failed to report an association between the Ile81 allele with serum urate (P = 0.61) in data from ~110,000 European individuals (Köttgen *et al.*, 2013). Differential epistatic interactions between the *MC3R* alleles and other loci between the populations also could be a possible cause for the discrepancy between previous studies and my findings.

While admitting the fact that the role of this variant in disrupting receptor functionality is controversial, it is suggested that the presence of this variant may play a role in increasing the binding between MC3R and ACTH in gout patients through modifying the overall receptor structure or it's binding site(s). This may in turn be involved in enhanced anti-inflammatory effects of ACTH through modulating proinflammatory cytokines, attenuation of the effect(s) of inflammatory response genes and subsequent leukocyte extravasation in gouty arthritis.

4.3.3 *MC4R*

Another important finding of this study is the association of the C allele of the variant *rs17782313* with increased risk of gout in European participants (OR = 1.03, *P* = 0.02) (Table 4.9). A similar association of the C allele of *rs17782313* was also found in EPZ (OR = 1.51, $P_{\text{Unadjusted}}$ = 0.03), which was not retained when age, sex, BMI and Polynesian grandparental ancestry were added as potential confounders (OR = 1.56, *P* = 0.11) (Table 4.9). A consistent direction of association of the C allele with increased risk of gout was also observed in most of the Polynesian data sets (EPN, EPZ and NPH) (Table 4.6), which became significant when all Polynesian and European data sets were analysed together (OR = 1.06, *P* = 0.03) (Table 6.9). However, my study found no association of the A allele of *rs17700633* with gout or serum urate (Table 4.6, 4.8 and 4.9).

These findings point towards an unknown underpinning pathway that depicts obesity as a possible risk factor for gout, especially in Europeans. Melanocortin 4 receptor (MC4R) is the fourth member of the MCR family. It is located on chromosome 18q21 and essentially acts as a receptor for α- and ß-MSH. The expression of *MC4R* is restricted to food regulating centres in brain (specifically within arcuate nuclei of hypothalamus) where it controls food intake and energy expenditure by integrating an agonist (satiety) signal from α-MSH and antagonist (orexigenic) signal from AgRP (Huszar *et al.*, 1997; Lu *et al.*, 1994). A large GWAS study including 16,876 individuals of European descent has reported two intergenic variants near *MC4R*, *rs17782313* and *rs17700633*, to be associated with the onset of obesity (Loos *et al.*, 2008). Other GWAS and replication studies also reported *MC4R* variants to be significantly associated with early onset of obesity and increased BMI (Beckers *et al.*, 2011; Chambers *et al.*, 2008;

Zobel *et al.*, 2009). Moreover, using *MC4R* and *TMEM18* variants (*rs17782313, rs6548238*) as instruments for BMI in an MR study, Palmer *et al.* (2013) reported increased BMI to be causal for hyperuricaemia. Similarly, another MR study reported fat mass explained by genetic variants of the *FTO, MC4R* and *TMEM18* genes to be positively associated with hyperuricaemia (Lyngdoh *et al.*, 2012). The findings from previous studies being in line with my study is supportive of the idea that *rs17782313* (C allele) is a genuine signal for obesity and, could in turn contribute to the aetiology of gout. A positive association of the C allele of *rs17782313* with serum urate in the combined European (β (mmol L^{-1}) = 0.002, *P* = 0.007) data set is also consistent with the Köttgen *et al.* (2013) urate GWAS findings that found a significant association (*P* = 2.6E-05) of the same allele with urate in a large population including ~110,000 European adults. A positive association of the C allele with serum urate in combined European plus Polynesian meta-analysis (β (mmol L^{-1}) = 0.002, *P* = 0.006) (Table 4.9) was also consistent with these findings and suggested the contribution of this variant (*rs17782313*: C allele) in gout pathophysiology through hyperuricaemia.

As two selected *MC4R* variants, *rs17782313* and *rs17700633* are intergenic, it is not possible to describe any functionality for these two. However, their association with obesity-related phenotypes in large GWAS datasets from several populations may indicate their involvement in regulating/manipulating the expression and/or response of other obesity-associated genes. The minor alleles of these two variants have been shown to possess differential, and sometimes opposite, correlation with the same obesity-related phenotype(s) (explained with references hereafter) in one population. The contradictory non-significant association results for the A allele of *rs17700633* in both NZ Polynesian and European populations can be explained in light of these findings. In a population of 2,265 from the Nurse's Health Study (NHS) cohort, Qi *et al.* (2008) found the C allele of *rs17782313* to be associated with a higher intake of total dietary fat and other adiposity measures but was unable to find any association for the A allele of *rs17700633* with the same phenotypic traits. Another more recent analysis, including data from men and women, identified a relationship between *rs17782313* (C allele) and morbid obesity in North Indian individuals, but could not find similar associations for *rs17700633* (A allele) (Srivastava *et al.*, 2014). These findings are supportive of the differential association patterns observed for the two *MC4R* variants (*rs17782313* and *rs17700633*) among different populations in my study.

It is possible that the presence of the obesity-associated C allele of *rs17782313* in gout patients leads to an impaired signal transduction in food regulating nuclei in central nervous system (CNS). Peripheral enhanced food intake and reduced energy expenditure appear as effective responses to these abnormal signals. A recent study indicated that increased food craving and emotional over-eating in a set of adult European individuals may account for an association between *rs17782313* (C allele) and increased BMI (Yilmaz *et al.*, 2015). Evidences of abnormal regulation of the satiety responsiveness and hyperphagia (excessive eating) add to this explanation (Valette *et al.*, 2013). Tschritter *et al.* (2011) provided evidence of leptin and insulin resistance in the presence of the minor C allele of *rs17782313*. Leptin and insulin resistance are strong risk factors for T2DM, renal dysfunction and hyperuricaemia (Vuorinen-Markkola and Yki-Järvinen, 1994). Given that hyperuricaemia, gout, obesity (McAdams DeMarco *et al.*, 2011) and the above mentioned comorbidities often coexist, a consistent trend of positive association after including these comorbidities in the regression model (Table 4.7) strongly suggests a role of the C allele (*rs17782313*) as a possible etiologic link between obesity, elevated serum urate and gout.

4.3.4 *ADTRP*

The study reports no association of the A allele of *rs6903956* with gout in either the European (OR = 0.98, *P* = 0.66) and NZ Polynesian populations (OR = 1.02, *P* = 0.83) (Table 4.9). However, the A allele was found to be negatively associated with serum urate concentrations in the combined Polynesian population (β (mmol L^{-1}) = -0.011, *P* = 0.02), especially in Māori individuals in the NPH dataset (β (mmol L^{-1}) = -0.035, *P* = 0.004) (Table 6.8 and 6.9). In addition, the A allele of *rs6903956* indicated a nominally positive association with serum urate in the ARIC European data set (β (mmol L^{-1}) = 0.002, *P* = 0.05) (Table 4.8).

Rs6903956 is located in intron 1 of the *ADTRP* gene (classically known as the *C6orf105* gene). The variant was first identified as a CHD-associated signal in a GWAS in the Han Chinese population (Wang *et al.*, 2011c). Following this, a number of studies confirmed this association with CHD in other population data sets (Nikpay *et al.*, 2015; Tayebi *et al.*, 2013). Although, a large scale study by Schunkert *et al.* (2011) was not able to find an association between the A allele and the risk of CHD in Europeans. The A

allele of *rs6903956* was recently found to be associated with asymptomatic hyperuricaemia in Han Chinese individuals (Meng *et al.*, 2015). Although, functionally uncharacterised, the expression of the *C6orf105* (*ADTRP*) gene has largely been reported in heart, skin, stomach and kidney, and to a lesser extent in leukocytes (Wang *et al.*, 2011c). The A allele of *rs6903956* has been suggested to decrease the expression of TFPI (Tissue factor pathway inhibitor: a natural inhibitor of coagulation), which in turn causes increased coagulation and thrombosis leading to atherosclerosis (Lupu *et al.*, 2011; Wang *et al.*, 2011c), indicating the predictable role of *rs6903956* in CAD. However, a recent study by Chang *et al.* (2017) reported the A allele not to be associated with plasma coagulation factors in Singaporean Chinese adults and neonates, which contradicts to the idea of the A allele being involved in thrombosis. Additionally, the mechanism(s) involved in the negative association of *rs6903956* (A allele) and serum urate in my study remains unclear.

A possible explanation for association of the A allele with decreased urate concentration could be given by considering a study by Chen *et al.* (2016), who found better prognosis of acute coronary heart stroke in men with higher urate levels. There are a number of recent clinical studies indicating a beneficial role of using urate as a potential therapy agent for patients with acute ischemic heart stroke (Amaro *et al.*, 2016; Llull *et al.*, 2016). It is suggested that the presence of the A allele may be involved in decreasing serum urate concentration in NZ Polynesian individuals and, thereby, increasing the risk of developing CHD via increased TFPI production. The contradictory findings in my study to Meng *et al.* (2015) can be explained by a couple of reasons. First, the associations described by Meng *et al.* (2015) could be false positive due to a small sample size being used. Second, the negative association may be present exclusively in the Polynesian population due to different haplotypic background and may not be comparable with the positive association findings of Meng *et al.* (2015) in the Han Chinese population.

4.3.5 Strengths and limitations of the study

This is the first study to provide evidence for a direct genetic association of variants in the melanocortin system involved in obesity and anti-inflammatory response with serum urate and gout. The study also represents the first example describing the

involvement of adrenergic system TFP inhibitor (ADTRP) in reducing the serum urate concentration in NZ Polynesian individuals. It is also important to note that two MetS-related variants within *ADRB3* and *ADTRP* analysed in my study were not present in the largest serum urate GWAS (Köttgen *et al.*, 2013). The reason could be the absence of any data for these two variants in the HapMap2 platform (www.hapmap.ncbi.nlm.nih.gov) that was used for imputing genotypes in Köttgen *et al.* (2013) GWAS. The findings in my study, therefore, would be a valuable addition to the knowledge through associations of these loci with serum urate and gout in European as well as Polynesian populations. Moreover, this study is one of the few studies reporting associations of more than one MetS-related variant in the European population as well as variants exclusive for NZ-based Polynesian subgroups. Although all data sets used in my study had adequate power to determine an intermediate effect at a low allele frequency (> 0.2), still the small size of individual Polynesian datasets to detect a lower effect represents a study limitation. In addition, it is hard to comment on functional roles of the variant *rs17782313* (and *rs17700633*) in context of *MC4R* due to the absence of any data reporting its association with expression of *MC4R* gene in currently available gene-tissue expression platforms (GTEx portal: www.gtexportal.org and Haploreg4: www.archive.broadinstitute.org).

4.3.6 Conclusion and future directions

The study reports effects at four different MetS-related loci, *ADRB3, MC3R, MC4R* and *ADTRP*, predictable for increase/decrease in serum urate in European and NZ Polynesian individuals. It is not abnormal to find differences in direction of associations for the same genetic variant while dealing with two or more populations with an ancestrally different background. Even within the same population set, an allele could be expected to be positively associated with increased risk of gout while showing a negative association with urate levels. In addition to the difference between the genetic history of two populations, the involvement of different pathways controlling increased urate and establishment of gout (via inflammation) can help explain intra-group association inconsistencies. In my study, two of the loci, *ADRB3* and *MC4R*, also indicated population-specific associations with gout, while negative effects of the minor allele of *ADTRP* with serum urate were exclusive to Polynesian individuals only. Results that were contradictory to the previous literature still warrant testing in larger data sets from diverse populations to eliminate the chances of any false positive/negative findings this study

may have found. Overall, my study reports a positive association of components of MetS in determining the risk of hyperuricaemia and gout in European and NZ Polynesian individuals. If replicated successfully, the proposed mechanism(s) may also have a significant impact on increasing our understanding of hyperuricaemia and gout pathophysiology that coexist with metabolic syndrome or some of its components. Use of *in vitro* techniques to study gene expression and function in early childhood, adulthood and elderly may also provide a better idea of the roles these MetS-related variants play in these age groups. The *in vitro* gene functional studies exploring the role of these loci in up- and down-regulating the expression of other genes would also help to better understand the involvement of these genes in gout pathophysiology via urate production, its transport and renal impairment or other conditions associated with renal impairment e.g., insulin resistance. Availability of customary human/animal cell lines, tissue cultures and mammalian model animals (for gene expression studies in intact animals) have made these functional studies relatively easy to follow. Studying individual sequence variant's functionality and expression may require a panel of genetically matched 'normal cells' alongside the 'mutated cells'. This isogenic system can be obtained via engineered nucleases (e.g., zinc finger or transcription-activator-like effector nucleases) to generate pluripotent stem cells, introducing specific mutations and to study specific disease mechanisms.

Chapter 5

Rare Variants Contributing to Gout Susceptibility

SECTION 5.1 RARE VARIANTS IN COMPLEX DISEASE: AN INTRODUCTION

5.1.1 Background

Understanding the genetic basis of complex diseases has been the subject of interest for many decades. A number of genetic factors have been recognised that contribute to the risk of common diseases, with an open question of 'how many more genetic factors to be found?' Genome-wide association studies (GWAS) in humans have successfully identified thousands of common genetic variants contributing to quite a large number of more common phenotypes/traits. Despite these successes, the genetic makeup of complex traits has not been fully elucidated. As most of the genome-wide significant variants in GWAS exist in intronic or non-coding regions of the genome with weaker effect, it is hard to predict any functionality or to explain a link between disease and the variant through a biological mechanism. The robust associations explained by common genetic variants in GWAS usually explain a relatively small proportion of the disease heritability, with a large proportion of the heritability still 'missing' or unexplained (Goldstein, 2009; Pritchard, 2001). It is possible that some of this so-called 'missing heritability' is accounted by the genetic variants that are either rare or occur in low-frequency (Manolio *et al.*, 2009; McClellan and King, 2010). A second possibility is from studies that have suggested replacing the term 'missing' with 'hidden' by taking into account all SNPs being genotyped in a GWAS to explain variance in disease heritability (Morris *et al.*, 2012; Yang *et al.*, 2010a). For example, Yang *et al.* (2010a) reported that a large proportion (45%) of the heritability for human height can be explained by considering all common SNPs in the genome-wide data simultaneously rather than considering significant signals only. Consistently, the most recent urate GWAS provided evidence to support the possibility of ~27-41% of this 'hidden' heritability to be explainable when all common variants in the GWAS are considered together (Köttgen *et al.*, 2013). Still, that a substantial proportion of variance remains unexplained by common variants requires looking into other possible culprits i.e., rare variants. The idea that these less common variants could possess considerably larger effect on disease phenotype(s) led researchers to discover high-throughput methodologies that precisely capture these variants in the genome to analyse their association with phenotypes, although the key

challenge in the analyses of such variants is strictly their 'rarity'. High-tech advances in molecular biology using high-throughput sequencing (HTS) and exome arrays have provided detailed insights into the entire genome of individuals in contrast to focusing on some predefined loci. This not only opens up new opportunities for the genetic research on common diseases but also raises some new challenges e.g., how to exploit these data in the most efficient way. These human genome resequencing methodologies have explored many different types of study designs and analytical techniques (as explained in section 5.1.5) that can be applied for rare variant association studies with greater accuracy (Auer and Lettre, 2015).

5.1.2 Common and rare variants and common disease

Simple (or Mendelian) diseases are the results of mutation(s) in single genes, which are causal and may have a variable degree of penetrance and expressivity (Auer and Lettre, 2015). In contrast, complex (or common) human diseases result from a combined effect of multiple genetic modifications and environmental factors (Craig, 2008). These genetic variants may or may not be causal, which means that they might be involved in increasing the risk of developing the disease rather than individually causing the disease itself. Explaining the genetic architecture of common disease susceptibility relies on two model patterns i.e., common disease-common variant/CDCV and common disease-rare variant (Schork *et al.*, 2009). The CDCV explains that more frequent, low penetrance variants are the major contributors in increasing the risk to common disease in a general population. This model is the underlying rationale to most of the GWA studies. The common disease-rare variant model argues for the presence of multiple, rather rare, high penetrance variants that increase the genetic susceptibility for a complex trait in specific individuals (Reich and Lander, 2001).

5.1.3 Defining a rare variant

Given that the idea of studying rare variants is nascent, there is little consensus to separate rare from common variants on the basis of their allele frequency in any population. Assigning the rarity is rather reliant on their insufficient frequency and failure of GWAS genotyping assays to capture these variants (Pritchard, 2001). It is also not possible to detect such variants by classical linkage analysis in family studies, as they usually don't carry sufficiently large effect sizes (McCarthy and Hirschhorn, 2008).

Though the meaning differs from one study to the other, 'rarity' is usually differentiated from 'commonality' based on the frequency of a minor allele in a said population. For example, a MAF of < 1% has been described as rare by Frazer *et al.* (2009), whereas Gorlov *et al.* (2011) set a threshold of 5% for rare variants. Bodmer and Bonilla (2008) described an upper limit of 1% and a lower limit of ~0.1% to differentiate the rare variants from their third category of 'clearly deleterious' mutations. The authors also described the possibility of an overlap between 'low-frequency' and 'rare' variants. A classical categorisation assigned the terms very common, less common and rare for any variant with MAF of 5-50%, 1-5% and < 1%, respectively (Cirulli and Goldstein, 2010). McCarthy *et al.* (2008) and Manolio *et al.* (2009), however, provided a more precise classification for assigning the rarity to any variant on the basis of allele frequency and disease susceptibility or penetrance (Figure 5.1). Common variants, on the contrary, are those that are employed in most GWAS. Zuk *et al.* (2014) have recently defined common variants as those occurring with a count that is frequent enough to allow their testing in affected and healthy subjects.

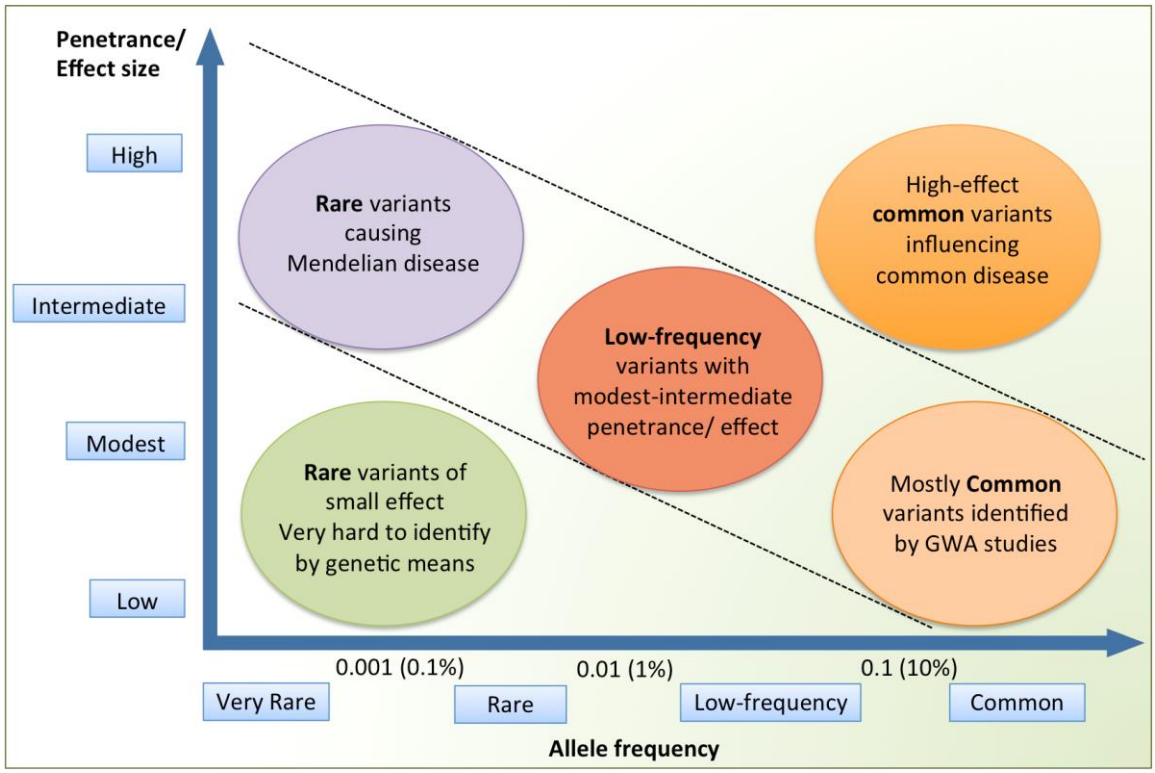


Figure 5.1: Feasibility of identifying genetic variants by risk allele frequency and effect size. Modified and redrawn from McCarthy *et al.* (2008) and Manolio *et al.* (2009).

5.1.4 Types of rare genetic variants

Genetic variants are broadly categorised as single nucleotide polymorphisms and structural variants (Frazer *et al.*, 2009). A change at a single nucleotide position on the DNA sequence is referred to as single nucleotide variant. Although most SNPs are biallelic, the presence of more than one allele or multiallelic SNPs have also been observed (Hüebner *et al.*, 2007). Structural variants primarily include indel (insertion-deletion) polymorphisms, where an addition or removal of one to a few nucleotides in the DNA sequence occurs. Other variants are block substitutions and inversions, which lead to a change in numerous adjacent nucleotides or their order. Copy number variants (CNVs) are DNA sequences of ≥ 1kb found in variable number in contrast to the reference genome (Feuk *et al.*, 2006). Feuk *et al.* (2006) differentiated CNV from copy-number polymorphism (CNP) by defining CNP as "CNV that ensues in more than 1% of the population".

5.1.5 Evaluation of rare variants for a complex trait

As rare variants are hard to identify because of their extremely low frequency, assessing these variants is a major challenge. One of the current approaches used to investigate the role of a rare variant in a complex disease is direct genotyping. The Exome chip (www.genome.sph.umich.edu/wiki/Exome_Chip_Design), metabochip (Voight *et al.*, 2012) and the Immunochip (Cortes and Brown, 2011) are a few examples of targeted and custom designed arrays used for direct genotyping of such variants. Genotyping on such arrays usually requires a GWAS as a supportive platform for imputation of low-frequency variants and a sequenced reference panel (e.g., 1000 Genomes Project) to provide a base to imputing genotypes (Abecasis *et al.*, 2012). Another method efficiently used to assess rare variants is sequencing – either sequencing of whole genome (or exome) or targeted sequencing of specific regions or candidate genes within the genome, using again a reference panel genome to align and arrange the small reads obtained in re-sequencing. Figure 5.2 illustrates a schematic overview of the steps involved in rare-variants' association studies for a complex trait.

5.1.5.1 Association analysis approach

Development and implementation of an efficient method to analyse rare-variant associations is vital. An underpowered single-point based association analysis is becoming obsolete. It is suggested that combining information across multiple variants (or burden) within a gene should be used as a viable alternative to isolated rare-variant association analyses. A number of statistical models are now available to carry out such locus-based analyses (explained hereafter). An example of widely used method is the Kernel Based Associations Test and Sequence Kernel Association Test/SKAT that analyse data for the contribution of both rare and common genetic variants to risk of complex diseases (Mukhopadhyay *et al.*, 2010; Wu *et al.*, 2011). Other methods broadly involve allele-collapsing approaches based on summary statistics [such as Cohort Allelic Sum Test (Morgenthaler and Thilly, 2007), Weighted Sum Test (Madsen and Browning, 2009) and Combined Multivariate and Collapsing Test (Li and Leal, 2008)] and regression models (Han and Pan, 2010; Morris and Zeggini, 2010; Zhou *et al.*, 2010). The associations obtained from the above statistical approaches are known as first-step or discovery-phase associations.

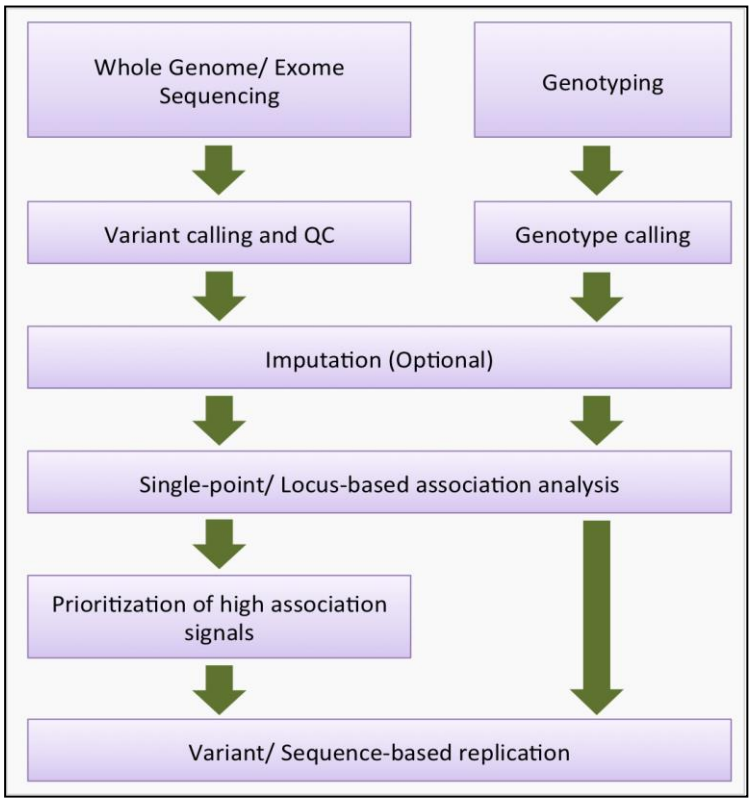


Figure 5.2: An overview of the steps involved in low-frequency and rare-variant association studies. Modified and redrawn from Panoutsopoulou *et al.* (2013) and Lee *et al.* (2014).

5.1.5.2 Functional annotation

In addition to choosing a suitable statistical approach to test for rare variant association, it is also crucial to decide if all of the discovered rare variants are to be included for a combined analysis. For instance, the sequencing of a particular gene might generate a range of variants in both coding (exonic) as well as non-coding (intronic, regulatory and un-translated) regions. It is ideal to pick those variants in coding regions that more likely result in structural and functional change of the protein (Auer and Lettre, 2015). Table 5.1 presents a list of some of the many resources available to interpret functional consequences of genetic variants. These annotations are exceedingly helpful in prioritising the variants to be included and/or excluded from group-based analysis (e.g., burden analysis) or replication analysis. Usually, high-signal, non-synonymous variants with large effect sizes are chosen or prioritised from the discovery-phase to follow-up in replication (Lee *et al.*, 2014).

Table 5.1: List of tools and resources for functional annotation

Tool/Resource	Brief description	URL (Reference)
CADD	CADD (Combined Annotation Dependent Depletion) is a framework that integrates multiple annotations into one metric by contrasting variants that survived natural selection with simulated mutations.	http://cadd.gs.washington.edu/ (Kircher *et al.*, 2014)
ENCODE	ENCODE (The Encyclopedia of DNA Elements) is a platform to build a comprehensive parts list of functional elements in the human genome using multiple cell lines.	https://www.encodeproject.org/ (Mitchell, 2012)
PolyPhen-2	PolyPhen-2 **(**Polymorphism Phenotyping version 2) is a tool that predicts the possible impact of an amino acid substitution on the structure and function of a human protein using straightforward physical and comparative considerations.	http://genetics.bwh.harvard.edu/pph2/ (Adzhubei *et al.*, 2010)
SIFT	SIFT (Scale-Invariant Feature Transform) predicts whether an amino acid substitution affects protein function. SIFT prediction is based on the degree of conservation of amino acid residues in sequence alignments derived from closely related sequences.	http://sift.jcvi.org/ (Kumar *et al.*, 2009)

HaploReg v1-4	HaploReg is a tool for exploring annotations of the noncoding genome at variants on haplotype blocks, such as candidate regulatory SNPs at disease-associated loci in mammals.	http://www.broadinstitute.org/mammals/haploreg/haploreg.php (Ward and Kellis, 2012)
VEP	VEP (Variant Effect Predictor) determines the effect of given variants (SNPs, insertions, deletions, CNVs or structural variants) on genes, transcripts and protein sequence, as well as regulatory regions.	http://useast.ensembl.org/info/docs/tools/vep/index.html?redirect=no (McLaren *et al.*, 2010)
PROVEAN	PROVEAN (Protein Variation Effect Analyser) is a software tool, which predicts whether an amino acid substitution or indel has an impact on the biological function of a protein.	http://provean.jcvi.org/index.php (Choi, 2015)

5.1.5.3 Replication of genetic signals

5.1.5.3.1 Foreword

It is essential to prioritise and replicate any signals of interest identified in the discovery-phase of a rare variant analysis/project in another large and independent cohort of individuals (replication cohort). The criteria of prioritisation may vary according to the study, population genetics, functional consequences, direction of association, rarity and novelty of the variants. The replication, however, totally relies on the allelic architecture of the associated locus and may require either genotyping or resequencing in the replication cohort (Lee *et al.*, 2014). For example, for a single rare-variant signal, genotyping of the said variant in an independent sample set would suffice (Beaudoin *et al.*, 2012). For multiple variants with high signals, a variant-based (genotyping of discovery-phase variants) and locus-based (resequencing of whole region) follow-up replication would be ideal (Figure 5.2).

5.1.5.3.2 Why we need replication?

As described briefly above, it is important to validate any association found in sequencing analysis by replicating it in an independent set of individuals called a replication cohort. Replication of the genome-wide level ($P < 5 \times 10^{-8}$) signals from GWAS is now considered mandatory to accept, standardise and publish a variant as a novel phenotype/disease associated variant (Barsh *et al.*, 2012). For sequencing studies, however, replication of the rare or low-frequency variants is more challenging than for

the variants identified by GWAS. The main factor that makes replication difficult for rare variants is indeed their extremely low frequency and population-specificity. A rare variant in the *LDLR* (low-density lipoprotein receptor; MAF = 0.5%) gene, associated with LDL-cholesterol, is an example of one such variant found to be polymorphic only in the Sardinian population (Sanna *et al.*, 2011). Discovering other variants in the same gene associated with the same trait and evaluating their functionality to describe possible metabolic pathways could be the only approach to replicate such association in other populations.

In addition to the rare and low-frequency signals, it is also crucial to confirm any common and population-specific associations in the discovery-phase of the sequencing study by replicating them in an independent cohort. The first step in the replication-phase is to set a criterion (which may vary from study to study) and prioritise a few variants accordingly. While there are no definite criteria described in the literature, the prioritisation could be based on the direction and significance of association, frequency or population statistics of the alternate allele, estimated effect size, apparent biological relevance and functionality and novelty of the variant (Lee *et al.*, 2014).

Once prioritised, the replication of the associated locus may require either genotyping or resequencing in the replication cohort (Lee et al., 2014). For a few or single variant signals, genotyping of the prioritised variants in an independent sample set would suffice (Beaudoin et al., 2012). This could be performed in a multitude of ways using a different population or an independent set of individuals within the same population, particularly for the population-specific variants. The strategies could be genotyping of the selected (prioritised) variants, selected genotyping of a perfect tag SNP or *in silico* genotyping (imputation; considered as a suboptimal approach for replication) of the variants of interest (Auer and Lettre, 2015). For multiple variants with high signals, a locus-based follow-up replication or resequencing is recommended as ideal. Adapting both, genotyping and resequencing, has been successful to replicate a discovery association in practice. Replication of a *DCTN4* (Dynactin subunit 4) gene rare variant by Sanger sequencing (Emond et al., 2012) and of four loss-of-function variants in *APOC3* (Apolipoprotein C-III) by Illumi-Exome Chip genotyping (Crosby et al., 2014) are two such successful examples. A statistically significant ($P \leq 0.05$) association signal in the replication cohort indicates that the discovery has been replicated and validated.

Sample size in the replication-phase is the most important determinant of power. While designing a replication study, it is important to select a large group of individuals to gain adequate power, especially if it involves a rare or low-frequency variant. Liu and Leal (2010) indicated that in comparison to single variant genotyping, sequencing of a region could be a better and more powerful approach to validate the discovery for rare variants. They described that sequencing could also increase the power of the study by recognising additional variants that might have been missed in discovery-phase. Whichever strategy (sequencing or genotyping) results in the largest sample size and the higher rare variant frequency should be the preferred approach (Auer and Lettre, 2015).

5.1.6 Rare variants contributing to complex traits

The number of rare variant association studies is growing continuously with advancement in the technology and analytical tools. Success stories like *PCSK9* (proprotein convertase subtilisin kexin type 9) in hypercholesterolemia are further motivating researchers to characterise the role of rare variants in other complex phenotypes. Rare variants have been associated with increased risk of major diseases in large GWAS and resequencing data sets. Targeted resequencing studies found multiple rare variants independently associated with type 1 (T1D) (Nejentsev *et al.*, 2009) and type 2 diabetes (T2D) (Bonnefond *et al.*, 2012; Lohmueller *et al.*, 2013). A large resequencing study of 25 GWAS including genes for autoimmune diseases identified a total of 2,990 protein coding variants with 97.1% having frequency of < 0.5%, along with a large proportion of variants (68.9%) only reported in one or two individuals (Hunt *et al.*, 2013). The study, however, did not formally test variants for association with autoimmune diseases (Hunt *et al.*, 2013). Five discrete rare variants in the *NOD2* gene have been reported to be associated with Crohn's disease (Hugot *et al.*, 2001; Ogura *et al.*, 2001; Rivas *et al.*, 2011). There are also examples of several rare variants associated with complex diseases with large effect sizes. A rare missense variant in the *MYH6* gene was found to increase the risk of sick sinus syndrome by 12-fold (Holm *et al.*, 2011). Huyghe *et al.* (2013) identified five independent low-frequency variants associated with fasting insulin levels in a study based on exome array genotyping.

5.1.7 Rare genetics of serum urate and gout

Genome-wide association studies have identified a number of loci explaining the genetic architecture of urate and gout (reviewed in Dalbeth (2016)). Still, it is challenging to describe the functional consequences of a number of high-signal variants in non-coding regions. Like many other complex phenotypes, the rare genetics' studies are new for serum urate and gout and not much has been described in the literature. Li *et al.* (2007) used the term 'rare allele' to describe the association of a variant (*rs6855911*) in the *GLUT9* gene for an allele frequency as high as 0.26 (Allele G). However, a whole-genome sequencing study of 16 million SNPs from 457 Icelanders identified a low-frequency missense variant (risk-allele (G) frequency = 0.019) in the *ALDH16A1* (aldehyde dehydrogenase 16 family member A1) gene to be associated with gout (OR = 3.12, $P = 1.5 \times 10^{-16}$) and serum urate (effect = 0.36, $P = 4.5 \times 10^{-21}$) at genome-wide threshold levels (Sulem *et al.*, 2011). They further confirmed the association through Sanger sequencing on 6,017 Icelanders and reported it to be consistent with a stronger effect in males than females. A recent report in 622 Han Chinese males described two intronic variants as 'rare' for their population. One of the two variants in *SLC2A9* (*rs734553*) was associated with gout (case-control allele (C) frequency = 0.008, 0.018, *P* = 0.028, OR = 0.45), while the association for other rare variant in *SLC16A9* (*rs12356193*) was neutral (case-control allele (G) frequency = 0.002, *P* = 0.659, OR = 1.43) (Zhou *et al.*, 2015). Given that no data are available to explain the possible 'missing heritability' of gout until now, the next section of this chapter was designed to carry out a rare-variant association analysis of exon sequencing data of two lipid-related genes, *LRP2* and *A1CF*, for serum urate and gout. The specific aims of this study were;

1. To characterise the variants obtained from exon sequencing of *LRP2* and *A1CF* genes in NZ Polynesian and European individuals.
2. To identify common, low-frequency and rare variants and test for an association with hyperuricaemia and gout.
3. To identify all non-synonymous and rare variants and test for their burden for hyperuricaemia in Europeans NZ Polynesians.
4. To replicate any interesting findings in larger NZ ancestry based population cohorts (Polynesian and European).

SECTION 5.2 EXON SEQUENCING OF *LRP2* AND *A1CF* GENES: ASSOCIATION ANALYSES

5.2.1 Background

Quantitative traits underlying common diseases including urate control have a significant degree of heritability. The most recent genome-wide association study from Köttgen *et al.* (2013), that combined data from > 110,000 European-ancestry individuals, provided evidence for association of 28 loci that collectively explain 7.0% of the variance in urate. Besides the loci involved in urate transport, an understanding of non-urate transporting loci (those controlling formation of MSU crystals and the immune reaction) is important to characterise the shared genetic basis of gout or urate with other metabolic conditions.

As described in detail in Chapter 1 (Section 1.1.7.3), lipoprotein receptor-related protein 2 (*LRP2*/megalin) is a non-urate transport locus that has been identified for its major role in lipid metabolism (Cabezas *et al.*, 2011; Christensen and Birn, 2002). The megalin (LRP2) protein is principally located on the apical surface of the proximal renal tubule (Cue et al, 1996, Nielsen 1998) and known for its role in reabsorption and metabolism of glomerular-filtered substances (Hosaka *et al.*, 2009). Clusters of rare variants have been identified within *LRP2* to be associated with pathologic conditions like ASD (Autism spectrum disorders) (Ionita-Laza *et al.*, 2012) and renal injury (McMahon *et al.*, 2014). Apolipoprotein B mRNA-editing enzyme 1 (APOBEC1) complementation factor (*A1CF*) is another emerging candidate in the list of non-urate transport genes. This gene is known for its role in the production of two different protein isomers, apo B-48 and apo B-100 from one nuclear gene (Chen *et al.*, 1987; Powell *et al.*, 1987). In the recent GWAS, the *A1CF* variant (*rs10821905*) was associated with serum urate levels in Europeans (Köttgen *et al.*, 2013), along with significant association with gout risk in New Zealand Europeans although not in Māori and Pacific Islanders (Phipps-Green *et al.*, 2014). In terms of data from fine mapping and whole genome or exonic sequencing, no study has yet reported any disease-based association for other variants within *A1CF*. Both *LRP2* and *A1CF* have been reported to be associated with increased eGFR (a marker for kidney function) in a recent GWAS based on European population (Appendix B Figure 5.2) (Pattaro *et al.*, 2016).

Despite evidence for involvement of *LRP2* and *A1CF* in a number of physiological processes, little is known about the effect of their possible causal variants on biological pathways, particularly those leading to gout. It was, therefore, hypothesised that other population-specific variants within *LRP2* and *A1CF* coding regions could contribute to gout risk in New Zealand (NZ) Māori and Pacific Islanders and Europeans, which may prove helpful in explaining some part of the 'unexplained heritability' for gout in these populations. This particular section of my thesis was aimed towards carrying out association analyses to determine if other common and rare variants that reside in exonic regions of *LRP2* and *A1CF* are associated with gout and hyperuricaemia (HU).

5.2.2 Methods

A replication based rare-variant analysis approach was adopted to analyse the association of variants within exonic regions of *LRP2* and *A1CF* genes with hyperuricaemia and gout. Genotype information for these two genes was extracted from the ReSequencing Cohort followed by discovery and replication analyses. Below are the major steps that were sequentially followed for the completion of rare-variant-analysis in this chapter:

1. Sequencing protocol: This section describes the overall selection of gene regions and the methods that were used to carry out sequencing in European and NZ Polynesian subjects (The ReSequencing Cohort).
2. Association analyses – Discovery phase: This section details the extraction of the *LRP2* and *A1CF* exonic regions from the ReSequencing Cohort followed by;
    i. Single variant analysis (individual association analysis of each variant within two genes for hyperuricaemia (European and NZ Polynesian) and gout (NZ Polynesian only).
    ii. Burden analysis (combined analysis for rare and non-synonymous variants to test for a burden of risk/protection for hyperuricaemia and gout (European and NZ Polynesian).
    iii. Functional annotation of variants (to check the functional status of each exonic variant within *LRP2* and *A1CF* genes).

3. Association analyses – Replication phase: This section provides details for a list of variants that were prioritised from the Discovery-phase analysis to replicate (via TaqMan genotyping) in an independent cohort (NZ Polynesian and European subjects) and their subsequent association analyses.

5.2.2.1 Sequencing protocol

5.2.2.1.1 Study participants

The ReSequencing Cohort consisted of 819 individuals from three different data sets selected according to urate level. The first set included 526 European and Polynesian individuals from New Zealand. This data set was a sub-set of New Zealand Gout Cohort (Section 1.2). Two additional sets of European individuals were selected from Health Professional Follow-up Study (HPFS: n = 169) and Nurses' Health Study (NHS: n = 125)[4]. The ReSequencing Cohort was then stratified on the basis of serum urate levels with a cut-off value of $\geq$ 0.41mmol L^{-1} between hyperuricaemic cases (n = 427: serum urate $\geq$ 0.41 mmol L^{-1}) and normouricaemic controls (n = 392: serum urate $\leq$ 0.37mmol L^{-1}). Hyperuricaemic cases were selected irrespective of their gout status and urate-lowering therapy while controls self-reported no history of gout. Self-reported ancestry was used to further split this cohort into European (n = 376) and Polynesian (n = 443) sub-sets. As it is not ideal to perform rare-variant analyses in smaller groups due to a drastic reduction in the power of analysis, the Polynesian sub-set was not further split into Eastern and Western Polynesian subgroups. Table 5.2 represents a summary of demographic details for the ReSequencing Cohort.

[4] Samples from the HPFS and NHS were kindly provided by Professor Hyon K Choi (Gout and Crystal Arthropathy Center, Massachusetts General Hospital, Boston, MA).

Table 5.2: Demographic details of the individuals in the ReSequencing Cohort

Population	**NZ Polynesian**		**European***	
Group	**Hyperuricaemic**	**Normouricaemic**	**Hyperuricaemic**	**Normouricaemic**
		Baseline Information		
Total number (n)	229	214	198	178
Males, n (%)	192 (83.84)	92 (42.99)	165 (83.33)	73 (41.01)
Age (years)^	45 ± 14	45 ± 14	60 ± 15	-
BMI (kg/m^2)^	36.97 ± 7.80	31.56 ± 6.69	30.16 ± 3.83	-
Serum Urate ($mmol\ L^{-1}$)^	0.53 ± 0.64	0.30 ± 0.40	0.52 ± 0.55	0.22 ± 0.33
		Co-morbidities		
Type II Diabetes, n (%)	33 (14.40)	42 (19.63)	15 (7.57)	-
Dyslipidaemia, n (%)	26 (11.35)	75 (35.04)	31 (15.65)	-
Heart problems, n (%)	20 (8.73)	18 (8.41)	19 (23.17)	-
Hypertension, n (%)	83 (36.24)	39 (18.22)	44 (53.66)	-
Kidney problems, n (%)	22 (9.60)	5 (2.34)	13 (15.85)	-
		Gout Characteristics		
Gout cases, n (%)	170 (74.24)	-	76 (92.68)	-
Age at onset gout (years)^	34 ± 12	-	47 ± 18	-
On Diuretics/ULT, n (%)	120 (52.41)	-	38 (46.34)	-
Gout attacks (per year)^	11.56 ± 37.05	-	4.33 ± 7.56	-
Gout tophus, n (%)	57 (24.89)	-	20 (24.39)	-

NZ: New Zealand, BMI: Body mass index, n (%): total number (percentage), ULT: Urate lowering therapy. ^Data are shown as mean ± standard deviation. Data for the co-morbidities are self-reported. *Demographic data for the European cohort is based only on the NZ samples (excluding sex and serum urate levels).

5.2.2.1.2 Selection of the genome regions for sequencing

The project was based on 'targeted resequencing' of candidate genes and the genomic regions of interest. The idea behind this specific project was to sequence the genetic loci that have been associated with hyperuricaemia and/or gout in the earlier literature, and to further analyse these associations through fine mapping of these loci in European and Polynesian populations. For this purpose, several criteria as described below were followed while selecting the genomic regions with prior evidence for an association with gout or serum urate concentrations.[5]

1. Gene/gene regions identified in Köttgen *et al.* (2013) in the largest serum urate GWAS including more than 110,000 Europeans.
2. Sixty-seven biallelic genomic control markers genotyped in Hollis-Moffatt *et al.* (2012b) to calculate estimates of Eastern Polynesian ancestry (Appendix A Table 5.1).
3. Gene/gene regions coding for those proteins that contribute to allopurinol metabolism.
4. Urate-transport genes that had not previously been associated with urate or gout.
5. Gene/gene regions influencing renal function, body mass index or weight.
6. Candidate gout risk genes known to determine the risk of gout through other metabolic pathways e.g., inflammation, lipid metabolism, or some unknown mechanisms.
7. Deoxyribonuclease (DNAse) hypersensitivity regions were idenetified from their expression data in 19 renal cell samples (Appendix A Table 5.2). The expression data was accessed through Gene Expression Omnibus website (www.ncbi.nlm.nih.gov/geo).
8. The genes/regions within or around the above chosen genic regions were also selected for sequencing.

[5] Professor Tony R Merriman (PhD supervisor), Assistant Professor Eli Stahl (Icahn School of Medicine, Mount Sinai, New York), Professor Hyon K Choi (Gout and Crystal Arthropathy Center, Massachusetts General Hospital, Boston, MA) and Assisstant Professor David B Mount (Brigham and Women's Hospital, Boston) selected the regions for targeted sequencing. Two staff members of the Merriman Laboratory identified the specific genome positions to be sequenced: Murray Cadzow (Assistant Research Fellow) and Ruth Topless (Assistant Research Fellow).

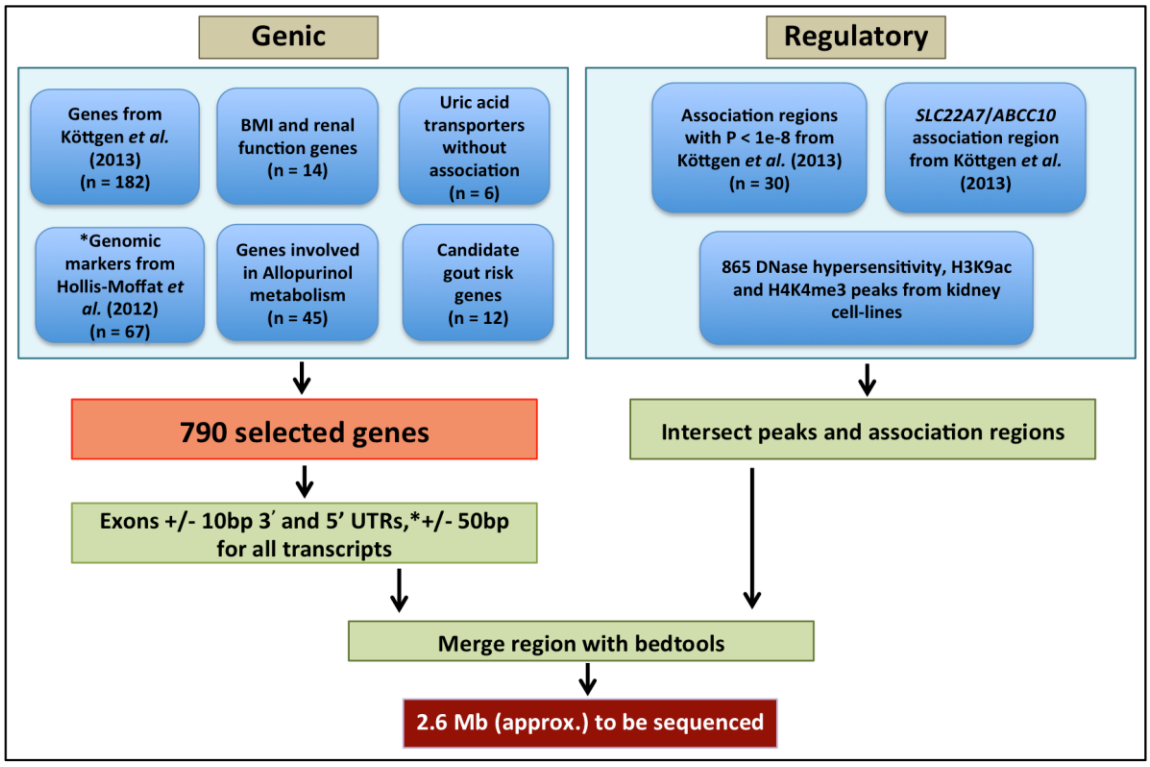


Figure 5.3: An overview of the detail of region selection for sequencing

Using the above selection criteria ended up in a list of 790 different genes. This selection essentially included both, genic and regulatory genes/regions. In detail, this list comprised 182 gout and urate associated genes encompassing 28 loci from the Köttgen *et al.* (2013) GWAS study, all Polynesian ancestry informative genes from Hollis-Moffatt *et al.* (2012b), 14 BMI and renal function genes, 45 genes involved in allopurinol metabolism, 12 candidate gout risk genes and 6 urate transporters not previously known for an association with gout. Selected regulatory regions included *SLC22A7/ABCC10* and 30 other regions that indicated association in Köttgen *et al.* (2013) at GWAS level (with $P < 5x10^{-8}$). The regulatory regions also included 865 DNAse hypersensitivity peaks (H3K9ac and H4K4me3) from kidney cell lines. Overlapping data from these peaks were combined with association regions to reduce the sequencing primer design. For each gene on this list, exon sequence was extracted for +/- 10 bp (base pairs) and 3' and 5' un-translated regions (UTRs). Moreover, regions of 800 bp upstream of the transcription start site(s) were also sequenced for each of these genes. Markers of Polynesian ancestry information were sequenced +/- 50 bp upstream/downstream and entire peaks were sequenced for DNAse hypersensitivity regions. Merging these regions ended up in sequencing of nearly 2.6 Mb of the whole genome (Figure 5.3).

5.2.2.1.3 Library preparation and sequencing[6]

For library preparation and sequencing, 250 ng of the genomic DNA was sent to McDonnell Genome Institute (Washington University, St. Louis) for each participant included in the Sequencing Cohort. Construction of Ilmumina indexed libraries was done following a modified version of the manufacturer's protocol and Nextera DNA Sample Prep kit (Illumina Inc., San Diego). The protocol includes fragmentation of DNA in 100 to 400 bp strands using a Covaris E220 DNA Sonicator (Covaris Inc., Woburn) followed by an addition of a sequence adapter to either end of the fragmented DNA. The amplification was done in eight cycles adding sequencing primers and indices to the adapter ligated DNA fragments. Library purification and size selection was carried out via the Solid Phase Reversible Immobilisation (SPRI) technique that uses AMPure XP beads specified for targeting 300 to 500 bp DNA fragments. Next, fragment hybridisation and amplification was completed using custom Roche NimbleGen SeqCap kit (Roche NimbleGen Inc., Madison). All instructions were followed as provided in the manufacturer's protocol. Sequencing was done on Illumina HiSeq 2500 (Illumina, San Diego) with 90 dual-indexed samples combined and captured as a pool to run in a single sequence lane. The sequencing was completed with an average of 51.2x coverage across 2.6 Mb of the targeted sequenced genome.

5.2.2.1.4 Variant calling[7]

The Genome Analysis Tool Kit (GATK) best practice pipeline (Auwera *et al.*, 2013) was followed for alignment of the raw sequence data and variant calling. All procedures for sequence alignment and variant calling were carried out in the Merriman Laboratory. As a first step of variant calling, Sequence Alignment/Map (SAM) tools 1.1.2 was used to extract FastQ files from the raw sequence data files. These files were then aligned to human reference genome (build GRCh37) using the 'mem' (memory) command of the Burrows-Wheeler Aligner (BWA 0.7.12) to create binary alignment/map (BAM) files (Li and Durbin, 2009; Li *et al.*, 2009). To mark the reads from PCR duplication, the above BAM files were processed using Picard 1.114

[6] A staff member of the Merriman Laboratory, Amanda Phipps Green (Assistant Research Fellow), managed the sample shipping.

[7] All sequence alignment, variant calling and library preparation was performed by James Boocock (Research Assistant) at the Merriman Laboratory.

(www.github.com/broadinstitute/picard), followed by insertion/deletion (indel) alignment and base recalibration in GATK 3.3.0 (McKenna *et al.*, 2010). Genomic variant call format (gVCF) files were created using GNU Parallel command line and the GATK HaplotypeCaller (DePristo *et al.*, 2011; Tange, 2011). Prior to variant calling, these gVCF files were merged in batches (100 files per batch) using GATK. The SNPs in the resulting variant call format (VCF) file were processed using the GATK variant quality score recalibration utility. This utility calculated the quality of the SNPs by depth, Fisher score, mapping quality rank sum and read position rank sum. Next, any SNP with a truth sensitivity of less than 99.00 was removed during recalibration of the VCF file. Ensembl variant effect predictor or V*e*!P (www.grch37.ensembl.org/Tools/VEP) and GEMINI 0.17.0 (www.gemini.readthedocs.org) platforms were used to annotate the resultant SNPs. The annotations mainly included assigning the reference SNP cluster ID (rs number), type of the variant (exonic, intronic or intergenic), genes and transcripts (with length of the transcript) possibly affected by each variant and location of each variant.

5.2.2.2 Association analyses – Discovery-phase

For the purpose of analysis in this study, the detailed exon (coding) region data of *LRP2* and *A1CF* genes were extracted from the targeted sequencing of 2.6 Mb of the genome (Figure 5.3). A systematic flow was followed to analyse for an association of variants within the *LRP2* and *A1CF* genes with hyperuricaemia (NZ Polynesian and European) and gout (NZ Polynesian). It was not possible to assess the European data set for an association with gout due to the unavailability of gout diagnosis status in HPFS and NHS datasets. After extracting out all coding (exon) region variants (SNPs) for both genes, each of these variants was analysed separately to assess an association with hyperuricaemia and gout. Non-synonymous (missense), low-frequency and rare variants were then selected and analysed for a burden for hyperuricaemia and gout in respective populations.

5.2.2.2.1 Single variant association analysis

The study focused only on variants that fell within the *LRP2* (2:169,983,619 to 170,219,195) and *A1CF* (10:52,559,169 to 52,645,435) transcribed gene regions (ENST00000263816 and ENST00000374001 coordinates). Those variants that passed the quality control criteria were selected within these regions. Any variant for which an rs ID

was not available was annotated by its location on the chromosome with a prefix 'var' (var = variant). A logistic regression based loop model was created in R version 3.3.2 (RCore, 2016) to calculate the effect estimates and allelic odds ratio (OR) for each variant (explanatory variable) in the NZ Polynesian and European data sets. The regression was applied to test for association with hyperuricaemia and gout (binary response variables) and estimated glomerular filtration rate (eGFR: continuous response variable). The regression model was adjusted for sex, age and BMI (wherever possible) to obtain the odds ratios for both populations. Estimates for grand-parental ancestry and collection study were added as additional adjustors for NZ Polynesian and European data sets, respectively. In order to calculate the combined effect of variants, meta-analysis was done using the 'meta package' (www.CRAN.R-project.org/package=meta) within R (version 3.3.2) using a fixed-effect model. For analyses showing heterogeneity ($P_{Het} < 0.05$), the fixed-effect model was replaced with a random-effect model. All allelic odds ratio calculated for the ReSequencing Cohort were relevant to the minor allele. Any individual with missing or incomplete data was excluded from the various analyses. Power to detect a $P_{OR} < 0.05$ was calculated in NZ Polynesian and European data sets following the Johnson *et al.* (2001) methodology.

5.2.2.2.2 Burden analysis

Burden analysis is designed to calculate the effect of all variants in a gene together in contrast to analysing each variant separately. This type of analysis considers the gene or region of interest a single unit rather than taking the effect of each variant alone. The analysis could, therefore, mitigate the inherent lack of power in single variant association analyses (Lee *et al.*, 2014).

Allele frequencies obtained from single variant analyses were combined separately for each variant, in each sample set, for both *LRP2* and *A1CF* genes. A combined allele frequency of < 0.05 (5%) was set as cut-off to determine the rare (< 0.05 - 0.01) and very rare (< 0.01) variants in the genes regardless of their functional cosequences. The allele collapsing methodology of Li and Leal (2008) was used to classify the individuals as carriers or non-carriers of the rare allele. Following this methodology, the individuals were classified as having or not having a rare allele at any site, i.e.,

$$X_n = \left\{ \begin{matrix} 1 \text{ rare allele present} \\ 0 \text{ rare allele absent} \end{matrix} \right\}$$

The number of individuals carrying a rare allele, it's frequency and the difference between allele-frequency was then calculated for hyperuricaemic (cases) and normouricaemic (controls) via Pearson Chi-square test (Li and Leal, 2008).

A weighted sum statistic was used to calculate the burden of rare variants for hyperuricaemia in the NZ Polynesian and European sample sets. The specific methodology followed to calculate the weighted sum (adaptive burden test) of variants was first proposed by Ionita-Laza *et al.* (2011). This test is less sensitive to the presence of both risk and protective variants in a genetic region of interest and calculates the burden according to the frequency of variants in cases and controls. The test was performed by following these basic steps;

1. Variants in question were classified into those that have a higher observed frequency in cases (k') and those that have a higher observed frequency in controls (k).
2. A two-sided weighted (SNP effect-weighted) summary statistic (S^+ and S^-) was calculated that represents the possible burden of risk and protective variants in the gene or region.
3. A one-sided *p*-value was calculated for risk and protective variants through standard permutation. The permutation involves random permuting (shuffling) of case/control label followed by recalculating *S* statistic.
4. A two-sided *p*-value was calculated to assess combined burden of risk and protective variants ($S^c = S^+ + S^-$) via the same permutation.

The source code/C++ for the analysis was downloaded from www.columbia.edu. The code was separately run for both study populations, each tested for 1,000; 10,000; 50,000 and 100,000 turns of permutations. The above strategy was applied to calculate the burden of low-frequency and rare (< 5% MAF), only rare (≤ 1% MAF) and non-synonymous variants in both genes.

5.2.2.2.3 Functional annotation of variants

Functional attributes for each variant were identified and added to the VCF during the 'variant calling' phase via GEMINI and V*e*!P. This phase added synonymous and non-synonymous annotations to all variants in the VCF along with the indication of the codon and the amino acid these variants could change. The functional annotations were assessed using four of the several tools specified in the Table 5.1. These tools are designed to provide a prediction score with a cut-off value that indicates if a variant is deleterious or neutral for a change in protein. For this study, the impact of each exonic variant within *LRP2* and *A1CF* was evaluated online by applying corresponding binary classification criteria (deleterious or damaging and neutral). The online platforms used to assess the scores were; CADD (www.cadd.gs.washington.edu: score ≥ 10 indicates deleteriousness), SIFT (www.sift.jcvi.org: score ≤ 0.05 indicates damaging affect), PolyPhen-2 (www.genetics.bwh.harvard.edu: score ≥ 0.43 indicates damaging affect) and PROVEAN (www.provean.jcvi.org: score ≤ -2.5 is deleterious).

5.2.2.3 Association analyses – Replication-phase

The distinct association signals in the discovery-phase analysis were replicated in NZ European and Polynesian datasets. The replication was carried out to 1) confirm and validate the putative associations found in the 'discovery-phase', and 2) to describe the possible biological significance of the validated associations for hyperuricaemia and gout.

5.2.2.3.1 Prioritisation of variants

Prioritisation of the discovery-phase variants within *LRP2* and *A1CF* was done based on several criteria. The same selection criteria were applied for NZ Polynesian and European populations. These criteria were:

1. Level of significance – All variants having a nominally significant ($P < 0.05$) association for hyperuricaemia for at least one population were included.
2. Direction of association – The prioritisation was done independent of direction of association i.e., rare variants showing either risk or protective association were selected from both populations.

3. Population-specificity – Any variant found in only one of the two populations (NZ Polynesian and European) was included.
4. Novelty – Any significant variant with no earlier record in 1000 Genome database was also included.

In order to rule out the possibility of complete linkage disequilibrium, an LD (R-squared limit = 0.60) was calculated between all prioritised variants using information available in the 1000 Genome database (www.browser.1000genomes.org).

5.2.2.3.2 Study participants

A sub-set of individuals was included from New Zealand Gout Cohort (Section 1.2.1) to create the Replication Cohort. All subjects who were part of the ReSequencing Cohort (Discovery-phase) were excluded from the Replication Cohort. The Replication Cohort was categorised into two major ancestral groups, NZ Polynesian and European (728 cases and 446 controls). The New Zealand Polynesian group (1001 cases and 1134 controls) was further divided into three sub-groups, East Polynesian (EP: 396 cases and 613 controls), West Polynesian (WP: 310 cases and 270 controls) and a mixture of East and West Polynesian (EPWP: 26 cases and 59 controls). The categorisation was done as described by Hollis-Moffatt *et al.* (2012b). Data from the Ngati Porou Hauora (NPH) individuals were also included as a separate Māori sample set (NPH: 269 cases and 192 controls) (See details of each Polynesian sub-group in Chapter 1, Section 1.2.1). Table 5.3 provides demographic details of the Replication Cohort.

Of the above, control (non-gout) subjects were included to test for an association with serum urate in NZ Polynesian and NZ European individuals in the Replication Cohort. All subjects who self-reported as taking diuretic medication, or had renal failure or gout, or who were first-degree relatives with gout patients, were removed from serum urate association analysis.

5.2.2.3.3 Genotyping

Two Polynesian specific variants, *rs111360923* and var170115626, were genotyped over the NZ Polynesian samples only. Genotyping was performed by an allelic discrimination assay i.e., TaqMan® SNP Genotyping assay (C_2215549_20; Applied

Biosystems, Foster City, USA) following the instruction provided in their protocol. The TaqMan SNP genotyping were auto-called via reporter dye signal plots on Lightcycler® 480 Real-Time Polymerase Chain Reaction (RT-PCR) System (Roche Applied Science, Indianapolis, IN, USA). These plots were then used to analyse and export correct genotype clustering. Genotyping data for *rs2302694, rs4667596, rs2075252, rs4667591* and *rs41268685* were sourced from the Illumina Chip CoreExome dataset. The genotyping details for this dataset are already described in Chapter 2 (Section 2.2.2.2).

5.2.2.3.4 Statistical analysis

A multiply adjusted regression analysis approach was used to measure an association of the seven prioritised variants (explanatory variables) with hyperuricaemia, gout (binary response variables) and serum urate (continuous response variable). All logistic and linear regression analyses were done using statistical software R version 3.3.2 (RCore, 2016). Any participant with missing data for any variable was excluded from the various analyses. Allelic ORs and β-estimates were calculated for each variant including age, sex and BMI as primary adjustors in the regression model. For Polynesian individuals, self-reported grandparental ancestry was included as an additional adjustor in the various analyses. To increase the power of analysis, all NZ Polynesian subgroups were combined separately and with the European data set in meta-analysis using the Meta package within R (http://CRAN.R-project.org/package=meta, 2014) using a fixed-effect model. For a meta analysis showing heterogeneity ($P_{\text{Het}} < 0.05$), the fixed-effect model was replaced with a random-effect model. A $P \leq 0.05$ was used to indicate the threshold for nominal statistical significance between response and explanatory variables in regression and meta models. Power to detect a $P_{\text{OR}} < 0.05$ was calculated in NZ Polynesian and European data sets following the Johnson *et al.* (2001) methodology. For any variants with Polynesian-specific outcomes, an intra-variant LD (R-squared limit = 0.60) was calculated in order to rule out the possibility of complete linkage disequilibrium, using genotype information in NZ Polynesian population.

Table 5.3: Demographic details of the Replication Cohort

Population	NZ Polynesian								NZ European	
Sub-population	East Polynesian/EP		West Polynesian/WP		EPWP		Ngati Porou Hauora/NPH			
Group	Gout	Non-gout	Gout	Non-gout	Gout	Non-gout	Gout	Non-gout	Gout	Non-gout
Baseline Information										
Total number (n)	396	613	310	270	26	59	269	192	728	446
Males, n (%)	334 (84.34)	250 (40.78)	308 (99.35)	154 (57.03)	25 (96.15)	24 (40.67)	224 (83.27)	99 (51.56)	654 (89.83)	308 (69.05)
Age (years)^	57.02 ± 12.36	43.19 ± 15.43	49.92 ± 11.86	37.37 ± 14.98	45.25 ± 11.78	35.29 ± 17.11	58.41 ± 12.66	42.15 ± 14.64	63.88 ± 12.87	54.40 ± 16.68
BMI (kg/m2)^	35.06 ± 7.68	32.11 ± 8.20	36.80 ± 8.33	33.90 ± 6.56	36.76 ± 8.14	33.52 ± 7.57	36.49 ± 8.01	30.71 ± 5.89	30.29 ± 5.30	27.65 ± 5.63
Serum Urate (mmol L^{-1})^	0.38 ± 0.10	0.36 ± 0.08	0.41 ± 0.11	0.40 ± 0.08	0.43 ± 0.07	0.40 ± 0.10	0.43 ± 0.11	0.36 ± 0.08	0.39 ± 0.11	0.34 ± 0.11
Co-morbidities										
Type II Diabetes, n (%)	122 (28.57)	68 (11.48)	83 (24.85)	35 (13.31)	9 (32.14)	4 (6.77)	82 (31.41)	12 (6.41)	118 (15.14)	30 (6.83)
Dyslipidaemia, n (%)	202 (50.88)	77 (13.65)	166 (51.55)	32 (12.40)	14 (51.85)	5 (8.92)	137 (52.69)	37 (20.11)	370 (48.23)	86 (20.57)
Heart problems, n (%)	192 (44.44)	71 (11.79)	80 (23.52)	18 (6.69)	9 (33.33)	4 (6.66)	92 (35.38)	19 (10.05)	288 (37.02)	65 (14.61)
Hypertension, n (%)	277 (64.58)	152 (25.81)	173 (52.26)	40 (14.92)	15 (53.57)	8 (13.79)	177 (67.30)	35 (18.61)	424 (54.85)	104 (23.63)
Kidney problems, n (%)	114 (26.76)	19 (3.59)	72 (21.95)	8 (3.18)	8 (29.62)	2 (4.00)	32 (12.40)	5 (3.04)	176 (22.76)	22 (5.21)
Gout Characteristics										
Age at onset gout (years)^	41.57 ± 15.35	-	37.11 ± 13.05	-	29.39 ± 10.84	-	40.49 ± 15.24	-	48.82 ± 16.93	-
On Diuretics/ULT, n (%)	143 (45.11)	-	80 (27.02)	-	12 (48.00)	-	15 (19.23)	-	199 (71.84)	-
Gout attacks (per year)^	9.53 ± 34.66	-	11.30 ± 28.35	-	7.53 ± 10.34	-	2.89 ± 4.85	-	7.97 ± 35.36	-
Gout tophus, n (%)	164 (38.95)	-	155 (46.68)	-	12 (46.15)	-	25 (9.54)	-	271 (65.38)	-

NZ: New Zealand, BMI: Body mass index, n (%): total number (percentage: calculated for the number of individuals who reported for any category), ULT: Urate lowering therapy. ^Data are shown as mean ± standard deviation (for the number of individuals who reported for any category). Data for the co-morbidities are self-reported. Data for NZ Polynesian are represented as an average for individuals genotyped through TaqMan PCR and Illumina Core Exome Chip genotyping (EP: 370 cases and 463 controls, WP: 282 cases and 192 controls, EPWP: 24 cases and 45 controls, NPH: 144 cases and 65 controls), while demographic data for NZ European represents demographic of indiviuals genotyped through Illumina Core Exome Chip genotyping only.

5.2.3 Results – Discovery-phase

Power calculations inidcated European and NZ Polynesian data sets to be only adequately powered (> 80%) to detect a significant association (OR = 2.0) in the variants with a minor allele frequency > 0.2 (Figure 5.4). A formal multiple-testing correction was avoided due to the low power of the study and due to replication being included in the study design.

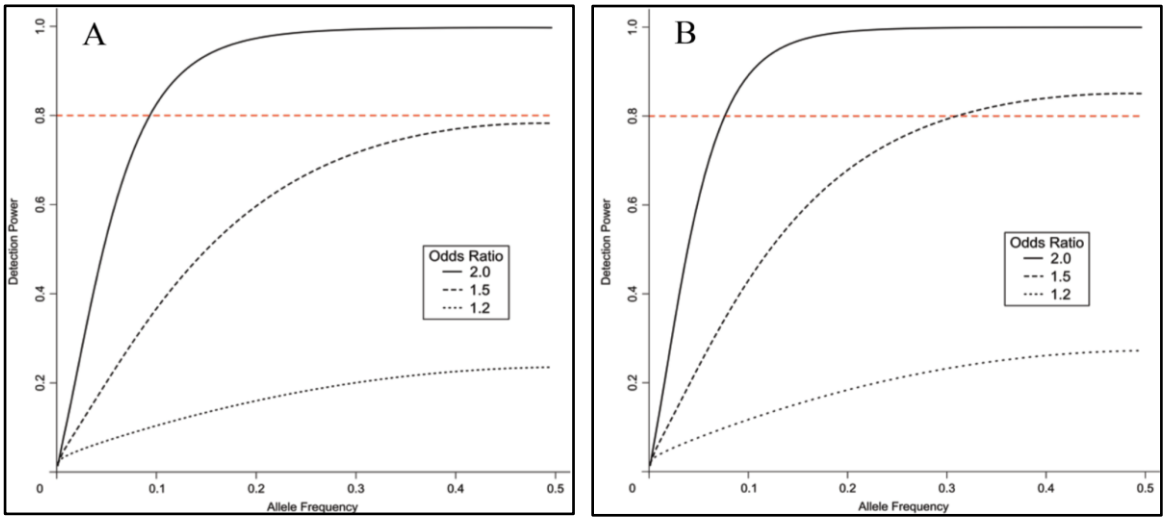


Figure 5.4: Association detection power in the (A) European and (B) Polynesian ReSequencing Cohorts across a range of odds ratio effect sizes and minor allele frequencies. The broken red line indicates an adequate detection power ≥ 80%.

Extracting the data of coding regions ended up in a list of 134 variants for *LRP2* and 69 for *A1CF* (Appendix A Table 5.3 and Appendix B Figure 5.1). Within *LRP2*, 81 variants were identified to be non-synonymous (missense), while only 4 non-synonymous (missense) variants were found in *A1CF*. Table 5.4 provides a summary of variable nucleotide sites for both genes.

Table 5.4: Summary of coding region variants of *LRP2* and *A1CF*

Ex	Len	Ts	Tv	Ts/Tv	Sn	NS	Ex	Len	Ts	Tv	Ts/Tv	Sn	NS
A1CF													
5' UTR	-	3	0	-	-	-	8	274	1	0	-	1	0
1	95	0	0	-	0	0	9	182	1	0	-	0	1
2	144	1	0	-	1	0	10	137	0	0	-	0	0
3	135	2	0	-	2	0	11	149	0	0	-	0	0
4	131	1	0	-	1	0	12	7,472	1	1	1	2	0
5	239	3	0	-	1	2	3' UTR	-	31	20	1.55	-	-
6	165	2	0	-	1	1	Total	9,221	48	21	2.28	11	4
7	98	2	0	-	2	0							
LRP2													
1	365	0	0	-	0	0	41	159	0	1	-	1	0
2	108	1	1	1.00	0	2	42	290	5	0	-	4	1
3	123	1	0	-	0	1	43	204	0	0	-	0	0
4	117	1	2	0.50	2	1	44	243	0	1	-	0	1
5	111	0	0	-	0	0	45	126	1	0	-	0	1
6	114	1	0	-	1	0	46	120	1	0	-	0	1
7	117	1	0	-	0	1	47	132	1	0	-	0	1
8	153	2	1	2.00	2	1	48	202	1	0	-	1	0
9	120	0	0	-	0	0	49	296	1	0	-	0	1
10	129	0	0	-	0	0	50	517	4	0	-	2	2
11	170	2	0	-	2	0	51	153	2	1	2.00	1	2
12	224	2	0	-	0	2	52	171	1	0	-	1	0
13	207	4	0	-	1	3	53	224	1	0	-	0	1
14	203	4	0	-	2	2	54	178	1	0	-	1	0
15	141	5	0	-	2	3	55	197	3	0	-	1	2
16	204	2	3	0.67	1	4	56	246	2	1	2.00	1	2
17	193	0	0	-	0	0	57	129	2	0	-	1	1
18	126	0	1	-	0	1	58	120	1	0	-	1	0
19	131	5	2	2.50	5	2	59	117	1	1	1.00	1	1
20	138	2	0	-	1	1	60	117	1	0	-	0	1
21	282	1	0	-	0	1	61	138	1	0	-	1	0
22	240	2	0	-	1	1	62	123	1	0	-	0	1
23	120	2	1	2.00	2	1	63	129	0	0	-	0	0
24	117	0	0	-	0	0	64	132	1	1	1.00	0	2
25	378	0	0	-	0	0	65	132	1	0	-	0	1
26	249	1	1	1.00	1	1	66	144	1	0	-	1	0
27	212	0	1	-	0	1	67	166	1	1	1.00	0	2
28	185	0	1	-	1	0	68	129	0	0	-	0	0
29	229	4	0	-	2	2	69	221	1	1	1.00	0	2
30	178	2	0	-	1	1	70	177	1	1	1.00	1	1
31	129	1	3	0.33	1	3	71	120	0	0	-	0	0
32	167	1	0	-	0	1	72	91	1	0	-	0	1
33	144	0	0	-	0	0	73	68	2	0	-	0	2
34	110	0	0	-	0	0	74	121	1	1	1.00	1	1
35	178	0	0	-	0	0	75	130	0	1	-	1	0
36	214	2	0	-	1	1	76	102	0	0	-	0	0
37	240	2	1	2.00	0	3	77	108	1	0	-	0	1
38	189	1	0	-	0	1	78	72	0	0	-	0	0
39	921	7	3	2.33	2	8	79	1,722	1	0	-	1	0
40	166	0	0	-	0	0	Total	15,808	102	32	3.19	53	81

Ex: Exon number, Len: Length of the exon in base pairs (bps) according to the longest transcript reported in Genome Reference consortium human genome build 37 (GRCh37) on Ensembl (http://grch37.ensembl.org/), Ts: Transitions, Tv: Transversions, Ts/Tv: Ratio of transitions to transversions, Sn: Total synonymous variants, NS: Total non-synonymous (missense) variants, UTR: Untranslated region.

5.2.3.1 Single variant association analyses

Single variant association analysis for each variant within *LRP2* and *A1CF* exonic regions indicated population-specific variants for both genes. Only 31.86% of the variants were present in one out of two populations with 18 European-specific (12 for *LRP2*; 6 for *A1CF*) and 11 Polynesian-specific variants (8 for *LRP2*; 3 for *A1CF*). A total of 25 (19 for *LRP2*; 6 for *A1CF*) variants were found to be common (combined altered allele frequency > 5%) for both genes.

For *LRP2*, overall, 10 variants (i.e., the altered allele of each variant) were significantly associated with hyperuricaemia in European (*rs34693334, rs41268685, rs2075252, rs4667591*) and NZ Polynesian (*rs111360923, rs13397109, rs2302694, rs4667596, rs830994,* var170115626) (Table 5.5). Out of these, two non-synonymous (missense) variants were found only in the Polynesian population with almost equal allele frequencies (*rs111360923*: C-allele frequency = 0.071; var170115626: C-allele frequency = 0.074) (Table 5.5). The C allele of *rs111360923* and the C allele of var170115626 were protectively associated with hyperuricaemia (*rs111360923*: OR = 0.55, P_{OR} = 0.02; var170115626: OR = 0.57, P_{OR} = 0.03) (Table 5.5). It was also not possible to calculate linkage disequilibrium (LD) between these two variants in any population due to unavailability of the information in 1000 Genome database (www.browser.1000genomes.org). The altered alleles of all other variants indicated a protective association with hyperuricaemia, except two missense variants, *rs34693334* and *rs34693334.* The G allele of *rs34693334* and T allele of *rs41268685* were found to be associated with an increased risk of hyperuricaemia in Europeans only (*rs34693334*: OR = 2.33, P_{OR} = 0.02; *rs41268685*: OR = 3.85, P_{OR} = 0.04) (Table 5.5). However, similar effects were not observed for NZ Polynesians for these two variants (Table 5.5). Another variant of interest was *rs4667591* (missense), where the G allele was found to be protectively associated with hyperuricaemia in Europeans (OR = 0.61, P_{OR} = 0.05) and indicated a similar trend of association in NZ Polynesians (OR = 0.75, P_{OR} = 0.08) (Table 5.5). Of the above 10 significant variants (except 2 Polynesian-specific SNPs), none was in LD with any other in the Caucasian or Han Chinese populations (Figure 5.5). Two exceptions were *rs2075252* and *rs4667591* having 82% LD and *rs4667596* and *rs2302694* having 100% LD only in Han Chinese population (Figure 5.5). No association

was found for any of the *A1CF* variants with hyperuricaemia in either population (Table 5.5).

When the European and NZ Polynesian data sets were combined together in meta-analysis, only few variants within *LRP2* indicated some interesting results (Appendix A Table 5.3). The A allele of *rs2302694* (OR = 0.54, P_{OR} = 0.001, P_{Het} = 0.29), the T allele of *rs4667596* (OR = 0.36, P_{OR} = 0.03, P_{Het} = 0.31), the C allele of *rs2075252* (OR = 0.75, P_{OR} = 0.03, P_{Het} =0.41) and the G allele of *rs4667591* (OR = 0.72, P_{OR} = 0.01, P_{Het} = 0.43) all indicated a protective association with hyperuricaemia (Appendix A Table 5.3). The C allele of *rs2229265* was associated with 32% increase in risk of developing hyperuricaemia (OR = 1.32, P_{OR} = 0.03, P_{Het} = 0.81) (Appendix A Table 5.3).

Single variant regression analysis was also carried out with gout and eGFR (estimated glomerular filtration rate) to assess an association with gout and kidney function. However, these two analyses were limited to NZ Polynesians only due to unavailability of information for both traits, gout status and eGFR, in HPFS and NHS data sets. Owing to the significance of outcomes, the results for association analysis for gout are provided below, while those for eGFR are provided in Appendix A Table 5.4.

The C alleles of two Polynesian-specific non-synonymous variants within *LRP2* were also protectively associated with gout (*rs111360923*: OR = 0.44, P_{OR} = 0.004; var170115626: OR = 0.46, P_{OR} = 0.006) (Table 5.6). The consistent pattern of protective association with gout was observed for the G allele of *rs4667591* (OR = 0.72, P_{OR} = 0.04), the G allele of *rs830994* (OR = 0.55, P_{OR} = 0.003), the A allele of *rs2302694* (OR = 0.44, P_{OR} = 0.001) and the C allele of *rs13397109* (OR = 0.45, P_{OR} = 0.0009) (Table 5.6). In addition, the T allele of *rs2075249* (OR = 1.42, P_{OR} = 0.05) and the C allele of *rs2229265* (OR = 1.55, P_{OR} = 0.019) indicated a susceptible association with gout in NZ Polynesians (Table 5.6). Neither of these two variants were in LD with each other in European or Hans Chinese populations on 1000 Genomes database (Figure 5.5). In addition, only the A allele of *rs149367019* (within *LRP2*) indicated a negative association with eGFR (ß (mL/min/1.73m^2) = -44.15, $P_{ß}$ = 0.044) (Appendix A Table 5.4).

Two variants within *A1CF* indicated interesting results (Table 5.6). The G allele of *rs16751* was found to be associated with an increased risk of gout (OR = 1.38, P_{OR} = 0.01), while the A allele of *rs184644838* showed an association with decreased risk of

gout (OR = 0.59, P_{OR} = 0.006) in the NZ Polynesian population (Table 5.6). It was not possible to calculate the LD for these variants in European and Han Chinese populations as no information was available for *rs16751* on any selected array on 1000 Genome database.

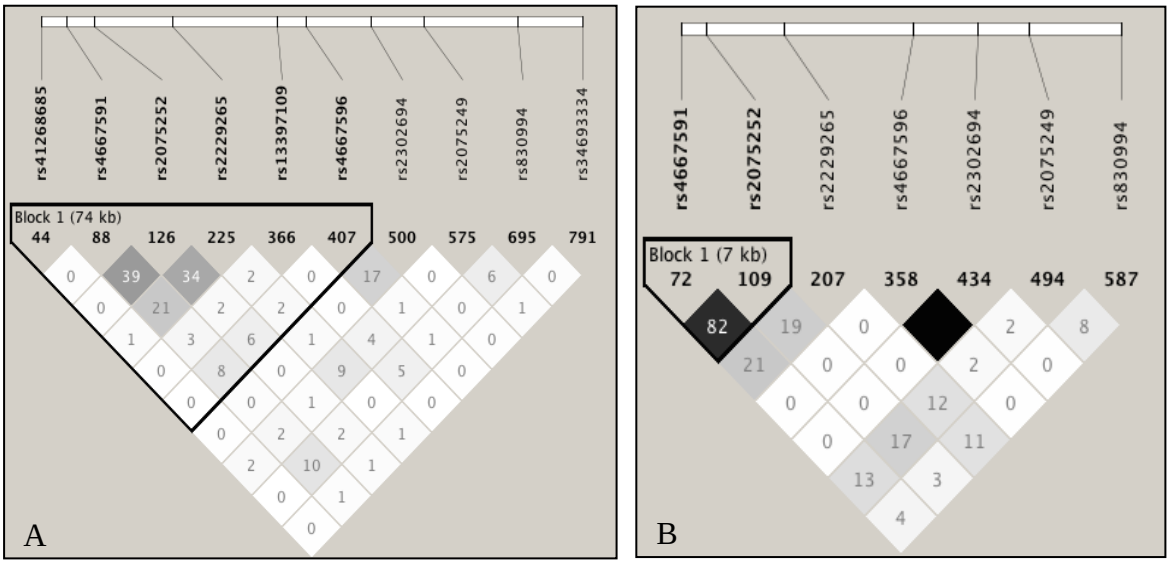


Figure 5.5: Linkage disequilibrium (LD) plot indicating 'R-squared/r^2' values between the 12 *LRP2* variants in (A) Europeans and (B) Chinese populations. An $r^2 \geq 60$ in the above plot indicates LD between the particular variants. Information for variant location, rs ID and LD values are from 1000 Genome database (http://browser.1000genomes.org/). The plots were generated using Haploview v4.2.

Table 5.5: Summary of single variant association analysis within *LRP2* and *A1CF* for hyperuricaemia

Exon	Position	rs ID/var position	Allele		Codon	European				NZ Polynesian			
						Allele Frequency		^OR [95% CI]	^P_{OR}	Allele Frequency		*OR [95% CI]	*P_{OR}
			Ref	Alt	Ref|Alt	Normo	Hyper			Normo	Hyper		
						LRP2							
2	2:170177382	*rs144829356*	G	A	gGg|gAg	-	-	-	-	70 (0.164)	96 (0.211)	1.30 [0.90; 1.87]	0.20
3	2:170175334	*rs2229263*	T	C	aTc|aCc	118 (0.341)	140 (0.357)	1.05 [0.71; 1.56]	0.81	66 (0.155)	65 (0.143)	0.87 [0.58; 1.31]	0.59
4	2:170163816	*rs34104660*	G	T	ccG|ccT	35 (0.101)	32 (0.082)	0.62 [0.30; 1.27]	0.19	3 (0.007)	0 (0.000)	-	-
6	2:170150671	*rs2229266*	G	A	gaG|gaA	102 (0.295)	102 (0.260)	0.79 [0.52; 1.21]	0.29	154 (0.363)	165 (0.361)	0.96 [0.70; 1.32]	0.80
8	2:170147502	*rs34693334*	C	G	Cgt|Ggt	23 (0.066)	24 (0.061)	2.33 [1.11; 4.87]	0.02	4 (0.009)	4 (0.009)	0.58 [0.13; 2.50]	0.47
14	2:170131548	*rs111360923*	T	C	tTt|tCt	-	-	-	-	47 (0.110)	32 (0.070)	0.55 [0.33; 0.92]	0.02
14	2:170129474	*rs145709922*	G	A	ttG|ttA	1 (0.003)	2 (0.005)	0.76 [0.05; 12.42]	0.85	-	-	-	-
14	2:170129547	*rs34291900*	C	T	gCc|gTc	11 (0.032)	15 (0.038)	1.58 [0.57; 4.38]	0.38	2 (0.005)	1 (0.002)	0.40 [0.03; 5.05]	0.47
14	2:170129528	*rs830994*	A	G	gtA|gtG	123 (0.355)	135 (0.344)	1.11 [0.74; 1.68]	0.62	95 (0.223)	76 (0.167)	0.63 [0.44; 0.92]	0.02
15	2:170115588	*rs2241190*	T	C	acT|acC	155 (0.448)	183 (0.467)	1.17 [0.79; 1.72]	0.44	169 (0.397)	176 (0.388)	0.97 [0.72; 1.32]	0.87
15	2:170115672	*rs33954745*	A	G	gaA|gaG	26 (0.075)	28 (0.071)	0.92 [0.44; 1.91]	0.82	7 (0.016)	5 (0.011)	0.62 [0.17; 2.29]	0.57
16	2:170127559	*rs141180155*	G	A	acG|acA	3 (0.009)	6 (0.015)	2.45 [0.45; 13.23]	0.30	0 (0.000)	1 (0.002)	-	-
17	2:170115626	var170115626	T	C	Ttc|Ctc	-	-	-	-	46 (0.108)	34 (0.075)	0.57 [0.34; 0.96]	0.03

19	2:170103471	*rs17848143*	C	T	acC|acT	-	-	-	-	5 (0.012)	4 (0.009)	0.73 [0.17; 3.10]	0.58
19	2:170103351	*rs2075249*	G	T	acG|acT	154 (0.445)	196 (0.500)	1.06 [0.72; 1.55]	0.77	101 (0.237)	120 (0.264)	1.20 [0.85; 1.69]	0.28
19	2:170103336	*rs831043*	T	C	acT|acC	148 (0.428)	181 (0.462)	1.22 [0.82; 1.82]	0.32	169 (0.397)	175 (0.385)	0.97 [0.71; 1.31]	0.83
21	2:170100011	*rs150552608*	G	A	cGt|cAt	2 (0.006)	1 (0.003)	0.38 [0.03; 4.23]	0.43	-	-	-	-
22	2:170099473	*rs831042*	T	C	gcT|gcC	146 (0.422)	180 (0.459)	1.26 [0.85; 1.87]	0.26	169 (0.397)	175 (0.385)	0.97 [0.71; 1.31]	0.83
23	2:170097707	*rs17848149*	T	G	gTc|gGc	7 (0.020)	12 (0.031)	0.94 [0.29; 3.10]	0.92	26 (0.061)	29 (0.064)	1.19 [0.65; 2.18]	0.62
26	2:170096095	*rs34915742*	C	G	cgC|cgG	2 (0.006)	1 (0.003)	0.38 [0.03; 4.23]	0.43	0 (0.000)	1 (0.002)	-	-
27	2:170094756	*rs146289506*	C	A	Ctc|Atc	1 (0.003)	1 (0.003)	0.76 [0.05; 12.42]	0.85	-	-	-	-
28	2:170093726	var170093726	T	G	acT|acG	-	-	-	-	5 (0.076)	2 (0.004)	0.32 [0.05; 1.93]	0.25
29	2:170092439	*rs151079411*	G	A	Gtg|Atg	1 (0.003)	1 (0.003)	1.97 [0.09; 43.53]	0.67	-	-	-	-
29	2:170092395	*rs2229267*	A	G	tgA|tgG	77 (0.223)	94 (0.240)	1.07 [0.71; 1.63]	0.74	128 (0.300)	129 (0.284)	0.94 [0.68; 1.31]	0.69
29	2:170092504	var170092504	C	T	cCc|cTc	-	-	-	-	5 (0.012)	10 (0.022)	1.85 [0.56; 6.10]	0.35
30	2:170089934	*rs145384264*	C	T	tcC|tcT	4 (0.012)	7 (0.018)	1.62 [0.30; 8.80]	0.57	2 (0.005)	1 (0.002)	0.41 [0.03; 5.28]	0.47
31	2:170088351	*rs2302694*	G	A	tcG|tcA	37 (0.107)	32 (0.082)	0.71 [0.37; 1.37]	0.31	64 (0.150)	43 (0.095)	0.46 [0.29; 0.74]	0.01
32	2:170082936	*rs138070797*	T	C	aTt|aCt	1 (0.003)	3 (0.008)	1.54 [0.14; 17.36]	0.73	-	-	-	-
36	2:170070348	*rs11886219*	T	C	cgT|cgC	15 (0.043)	18 (0.046)	0.78 [0.30; 2.06]	0.62	7 (0.106)	5 (0.011)	0.76 [0.20; 2.85]	0.77
36	2:170070172	*rs4667596*	C	T	aCa|aTa	8 (0.023)	5 (0.013)	0.65 [0.15; 2.78]	0.56	10 (0.023)	4 (0.009)	0.22 [0.07; 0.77]	0.02
39	2:170063263	*rs149367019*	G	A	Gta|Ata	1 (0.003)	4 (0.010)	0.76 [0.05; 12.42]	0.85	1 (0.002)	0 (0.000)	-	-

39	2:170063471	*rs35114151*	A	G	gaA|gaG	8 (0.023)	14 (0.036)	1.18 [0.40; 3.49]	0.77	26 (0.061)	28 (0.062)	1.15 [0.63; 2.12]	0.69
39	2:170062977	*rs61995915*	T	C	gTa|gCa	4 (0.012)	7 (0.018)	2.00 [0.42; 9.62]	0.39	0 (0.000)	1 (0.002)	-	-
41	2:170062078	*rs13397109*	G	C	cgG|cgC	17 (0.049)	23 (0.059)	1.26 [0.53; 2.98]	0.60	72 (0.169)	55 (0.121)	0.63 [0.41; 0.96]	0.03
42	2:170060603	*rs17848169*	T	C	Tat|Cat	13 (0.038)	17 (0.043)	1.27 [0.48; 3.38]	0.64	2 (0.005)	1 (0.002)	0.40 [0.03; 5.05]	0.47
44	2:170058345	var170058345	C	A	Ccc|Acc	-	-	-	-	5 (0.076)	3 (0.007)	0.39 [0.09; 1.81]	0.25
46	2:170053505	*rs2228171*	T	C	Tca|Cca	251 (0.725)	266 (0.679)	0.74 [0.48; 1.12]	0.16	182 (0.427)	187 (0.412)	0.92 [0.68; 1.23]	0.64
48	2:170048482	*rs149148763*	C	T	agC|agT	4 (0.012)	1 (0.003)	0.45 [0.04; 4.79]	0.51	-	-	-	-
50	2:170042245	*rs35734447*	T	C	Tac|Cac	1 (0.003)	1 (0.003)	1.97 [0.09; 43.53]	0.67	-	-	-	-
54	2:170032989	*rs2229265*	T	C	caT|caC	157 (0.454)	207 (0.528)	1.28 [0.89; 1.83]	0.18	320 (0.751)	362 (0.797)	1.40 [0.99; 1.98]	0.08
55	2:170031824	*rs17848184*	C	T	ccC|ccT	-	-	-	-	85 (0.200)	99 (0.218)	1.19 [0.82; 1.72]	0.40
57	2:170029657	*rs34355135*	C	T	Ctg|Ttg	2 (0.006)	4 (0.010)	1.98 [0.22; 17.85]	0.54	-	-	-	-
58	2:170028529	*rs199528723*	G	A	aaG|aaA	-	-	-	-	5 (0.076)	5 (0.011)	0.57 [0.16; 2.10]	0.38
60	2:170026248	var170026248	C	T	Cct|Tct	1 (0.003)	1 (0.003)	1.97 [0.09; 43.53]	0.67	-	-	-	-
61	2:170025083	*rs2229268*	A	G	tgA|tgG	71 (0.205)	85 (0.217)	1.06 [0.69; 1.63]	0.79	28 (0.066)	39 (0.086)	1.37 [0.77; 2.45]	0.21
64	2:170013904	*rs79723119*	A	C	gAt|gCt	6 (0.017)	8 (0.020)	1.71 [0.38; 7.60]	0.48	1 (0.002)	0 (0.000)	-	-
66	2:170010985	*rs2075252*	T	C	Tag|Cag	272 (0.786)	280 (0.714)	0.65 [0.42; 1.01]	0.05	139 (0.326)	131 (0.289)	0.78 [0.56; 1.07]	0.19
67	2:170009390	*rs142934522*	C	T	cCc|cTc	2 (0.006)	1 (0.003)	0.38 [0.03; 4.23]	0.43	-	-	-	-
69	2:170003432	*rs4667591*	T	G	Ttc|Gtc	284 (0.821)	301 (0.768)	0.61 [0.38; 0.99]	0.05	170 (0.399)	155 (0.341)	0.75 [0.55; 1.02]	0.08

70	2:170002412	var170002412	T	C	gTc|gCc	-	-	-	-	1 (0.002)	1 (0.002)	0.49 [0.03; 7.92]	0.61
72	2:169997051	*rs990626*	G	A	atG|atA	272 (0.786)	281 (0.717)	0.75 [0.48; 1.16]	0.19	180 (0.423)	179 (0.394)	0.89 [0.66; 1.19]	0.51
73	2:169996070	*rs41268685*	C	T	gCc|gTc	5 (0.014)	11 (0.028)	3.85 [1.03; 14.38]	0.04	1 (0.002)	0 (0.000)	-	-
77	2:169989127	*rs142245618*	A	G	gAt|gGt	3 (0.009)	2 (0.005)	0.67 [0.10; 4.45]	0.68	-	-	-	-
79	2:169985338	*rs34564141*	C	T	atC|atT	3 (0.009)	4 (0.010)	1.38 [0.24; 8.11]	0.72	-	-	-	-
							A1CF						
5' UTR	10:52645424	*rs10994860*	C	T	-	58 (0.168)	65 (0.166)	0.95 [0.55; 1.63]	0.85	49 (0.115)	56 (0.123)	1.36 [0.85; 2.16]	0.20
2	10:52623804	var52623804	A	G	-	-		-		9 (0.021)	8 (0.018)	0.62 [0.22; 1.74]	0.36
3	10:52603754	*rs35967725*	A	G	tgA|tgG	1 (0.003)	6 (0.015)	6.49 [0.49; 86.26]	0.16	3 (0.007)	2 (0.004)	0.61 [0.09; 4.39]	0.63
3	10:52603874	*rs142969066*	T	C	ggT|ggC	5 (0.014)	4 (0.010)	0.65 [0.14; 2.99]	0.58	1 (0.002)	0 (0.000)	-	-
6	10:52587964	*rs143315865*	G	A	atG|atA	1 (0.003)	1 (0.003)	0.76 [0.05; 12.42]	0.85	-	-	-	
7	10:52576068	*rs4245008*	A	G	-	49 (0.142)	56 (0.143)	0.95 [0.55; 1.65]	0.86	1 (0.002)	0 (0.000)	-	-
7	10:52580318	*rs372408821*	A	G	aaA|aaG	0 (0.000)	1 (0.003)	-	1.00	1 (0.002)	1 (0.002)	1.26 [0.06; 28.03]	0.88
8	10:52576025	*rs142026324*	G	A	tcG|tcA	3 (0.009)	2 (0.005)	0.90 [0.08; 10.72]	0.93	-	-	-	-
9	10:52573772	*rs41274050*	C	T	Cgc|Tgc	2 (0.006)	4 (0.010)	3.40 [0.40; 29.06]	0.26	71 (0.167)	87 (0.192)	1.31 [0.89; 1.93]	0.18
12	10:52566594	*rs34190540*	G	C	ctG|ctC	1 (0.003)	1 (0.003)	0.51 [0.02; 11.25]	0.67	-	-	-	0.20
3' UTR	10:52559596	var52559596	C	T	-	1 (0.003)	1 (0.003)	0.76 [0.05; 12.42]	0.85	-	-	-	-
3' UTR	10:52559634	*rs10994507*	A	G	-	31 (0.090)	32 (0.082)	0.99 [0.51; 1.91]	0.98	22 (0.052)	18 (0.040)	0.83 [0.41; 1.67]	0.60

3' UTR	10:52559843	*rs7084132*	G	T	-	50 (0.145)	56 (0.143)	0.91 [0.53; 1.57]	0.74	71 (0.167)	87 (0.192)	1.31 [0.89; 1.93]	0.18
3' UTR	10:52560557	*rs6479731*	A	G	-	50 (0.145)	56 (0.143)	0.91 [0.53; 1.57]	0.74	1 (0.002)	3 (0.007)	1.46 [0.15; 14.23]	0.74
3' UTR	10:52560658	*rs10994521*	A	T	-	1 (0.003)	1 (0.003)	0.76 [0.05; 12.42]	0.85	-	-	-	-
3' UTR	10:52561803	*rs74352101*	A	G	-	30 (0.087)	31 (0.079)	1.01 [0.52; 1.96]	0.99	71 (0.167)	87 (0.192)	1.31 [0.89; 1.93]	0.18
3' UTR	10:52561829	*rs7072584*	A	C	-	50 (0.145)	56 (0.143)	0.91 [0.53; 1.57]	0.74	71 (0.167)	87 (0.192)	1.31 [0.89; 1.93]	0.18
3' UTR	10:52561919	*rs12571156*	C	T	-	2 (0.006)	1 (0.003)	-	0.99	99 (0.232)	110 (0.242)	1.11 [0.79; 1.57]	0.54
3' UTR	10:52561920	var52561920	G	A	-	0 (0.000)	1 (0.003)	-	1.00	3 (0.007)	6 (0.013)	2.07 [0.44; 9.70]	0.36
3' UTR	10:52562434	var52562434	A	C	-	1 (0.003)	1 (0.003)	0.51 [0.02; 11.25]	0.67	1 (0.002)	0 (0.000)	-	-
3' UTR	10:52563051	*rs4282939*	C	T	-	86 (0.254)	86 (0.223)	0.91 [0.60; 1.37]	0.64	71 (0.167)	87 (0.192)	1.31 [0.89; 1.93]	0.18
3' UTR	10:52563196	*rs16751*	A	G	-	177 (0.512)	201 (0.513)	1.05 [0.74; 1.50]	0.79	111 (0.261)	141 (0.311)	1.10 [0.80; 1.52]	0.54
3' UTR	10:52563845	var52563862	C	T	-	1 (0.003)	1 (0.003)	0.76 [0.05; 12.42]	0.85	0 (0.000)	1 (0.002)	-	-
3' UTR	10:52563898	*rs112824128*	C	A	-	2 (0.006)	3 (0.008)	1.17 [0.14; 9.71]	0.88	-	-	-	-
3' UTR	10:52563981	*rs150545950*	C	T	-	2 (0.006)	2 (0.005)	0.61 [0.05; 7.93]	0.71	0 (0.000)	1 (0.002)	-	-
3' UTR	10:52564065	*rs4619096*	A	G	-	50 (0.145)	56 (0.143)	0.91 [0.53; 1.57]	0.74	200 (0.481)	175 (0.403)	0.78 [0.58; 1.04]	0.08
3' UTR	10:52564421	*rs80080606*	G	T	-	1 (0.003)	0 (0.000)	-	1.00	99 (0.232)	110 (0.242)	1.11 [0.79; 1.57]	0.54
3' UTR	10:52564524	var52564700	A	T	-	-	-	-	-	4 (0.009)	1 (0.002)	0.47 [0.04; 5.00]	0.53
3' UTR	10:52564700	*rs74874346*	G	A	-	29 (0.084)	31 (0.079)	0.93 [0.48; 1.82]	0.83	22 (0.052)	18 (0.040)	0.83 [0.41; 1.67]	0.60
3' UTR	10:52565100	var52565132	A	C	-	1 (0.003)	2 (0.005)	1.31 [0.08; 21.30]	0.85	0 (0.000)	1 (0.002)	-	-

3' UTR	10:52565132	*rs185182715*	C	T	-	1 (0.003)	0 (0.000)	-	1.00	103 (0.242)	85 (0.187)	0.72 [0.50; 1.03]	0.07
3' UTR	10:52565377	*rs75583477*	C	T	-	1 (0.003)	0 (0.000)	-	1.00	22 (0.052)	18 (0.040)	0.83 [0.41; 1.67]	0.60
3' UTR	10:52565903	*rs4078160*	C	A	-	49 (0.142)	56 (0.143)	1.05 [0.61; 1.81]	0.86	-	-	-	-
3' UTR	10:52565951	*rs4619097*	T	C	-	2 (0.006)	1 (0.003)	0.38 [0.03; 4.23]	0.43	71 (0.167)	87 (0.192)	1.31 [0.89; 1.93]	0.18
3' UTR	10:52566049	*rs10821846*	A	G	-	177 (0.512)	201 (0.513)	1.05 [0.74; 1.50]	0.79	111 (0.261)	141 (0.311)	1.10 [0.80; 1.52]	0.54
3' UTR	10:52566333	var52566417	G	A	-	-	-	-		25 (0.059)	19 (0.042)	0.64 [0.32; 1.28]	0.21

UTR: Un-translated region, rs ID: Reference SNP cluster ID, Chr: Chromosome, var: Variant, Ref: Reference, Alt: Alternate. Normo: Normouricaemic, Hyper: Hyperuricaemic, OR: Odds ratio for the alternate allele, 95% CI: 95% confidence interval, P_{OR}: *p*-value for odds ratio. *Odds ratios are adjusted for sex, age, BMI and estimates of grand-parental ancestry. ^Odds ratios are adjusted for sex and collection study. Exon numbers and variant positions are sourced from the Genome Reference Consortium human genome build 37 (GRCh37) on Ensembl (http://grch37.ensembl.org/).

Table 5.6: Summary of single variant association analysis within *LRP2* and *A1CF* for gout in NZ Polynesians

Exon	Chr:Position	rs ID/var position	Allele		Codon	Allele Frequency		*OR [95% CI]	*P_{OR}
			Ref	Alt	Ref|Alt	Control	Case		
					LRP2				
2	2:170177382	*rs144829356*	G	A	gGg|gAg	93 (0.171)	73 (0.221)	1.27 [0.88 ; 1.82]	0.20
3	2:170175334	*rs2229263*	T	C	aTc|aCc	86 (0.158)	44 (0.133)	0.78 [0.51 ; 1.20]	0.26
4	2:170163816	*rs34104660*	G	T	ccG|ccT	3 (0.006)	0 (0.000)	-	0.98
4	2:170163808	var170163808	G	T	gGt|gTt	0 (0.000)	1 (0.003)	-	0.98
6	2:170150671	*rs2229266*	G	A	gaG|gaA	190 (0.349)	125 (0.379)	1.19 [0.87 ; 1.65]	0.28
7	2:170148871	var170148871	T	C	Tcc|Ccc	0 (0.000)	1 (0.003)	-	0.98
8	2:170147502	*rs34693334*	C	G	Cgt|Ggt	5 (0.009)	3 (0.009)	0.62 [0.14 ; 2.75]	0.53
8	2:170147408	var170147408	A	G	tAa|tGa	0 (0.000)	1 (0.003)	-	0.98
14	2:170131548	*rs111360923*	T	C	tTt|tCt	60 (0.110)	20 (0.061)	0.44 [0.25 ; 0.78]	0.004
14	2:170129547	*rs34291900*	C	T	gCc|gTc	2 (0.004)	1 (0.003)	0.77 [0.06 ; 9.86]	0.84
14	2:170129528	*rs830994*	A	G	gtA|gtG	121 (0.222)	49 (0.148)	0.55 [0.37 ; 0.82]	0.003
15	2:170115588	*rs2241190*	T	C	acT|acC	219 (0.403)	123 (0.373)	0.91 [0.67 ; 1.24]	0.56
15	2:170115672	*rs33954745*	A	G	gaA|gaG	8 (0.015)	4 (0.012)	0.76 [0.19 ; 2.96]	0.69
16	2:170127559	*rs141180155*	G	A	acG|acA	0 (0.000)	1 (0.003)	-	0.98
16	2:170113671	*rs147621120*	T	A	Tct|Act	1 (0.002)	0 (0.000)	-	0.98
16	2:170127497	var170127497	A	G	aAt|aGt	0 (0.000)	1 (0.003)	-	0.98
17	2:170115626	var170115626	T	C	Ttc|Ctc	59 (0.108)	20 (0.061)	0.46 [0.26 ; 0.81]	0.006
18	2:170104017	var170104017	A	C	Att|Ctt	2 (0.004)	0 (0.000)	-	0.98
19	2:170103471	*rs17848143*	C	T	acC|acT	6 (0.011)	3 (0.009)	0.81 [0.18 ; 3.67]	0.79
19	2:170103351	*rs2075249*	G	T	acG|acT	126 (0.232)	93 (0.282)	1.42 [1.00 ; 2.01]	0.05
19	2:170103336	*rs831043*	T	C	acT|acC	218 (0.401)	123 (0.373)	0.92 [0.68 ; 1.26]	0.61
19	2:170103219	var170103219	T	A	acT|acA	1 (0.002)	0 (0.000)	-	0.98
22	2:170099474	*rs144723964*	G	A	gGa|gAa	1 (0.002)	0 (0.000)	-	0.98

22	2:170099473	*rs831042*	T	C	gcT|gcC	218 (0.401)	123 (0.373)	0.92 [0.68 ; 1.26]	0.61
23	2:170097655	*rs138030034*	G	A	tgG|tgA	1 (0.002)	0 (0.000)	-	0.98
23	2:170097707	*rs17848149*	T	G	gTc|gGc	35 (0.064)	183 (0.555)	1.12 [0.61 ; 2.06]	0.72
26	2:170096095	*rs34915742*	C	G	cgC|cgG	1 (0.002)	0 (0.000)	-	0.98
28	2:170093726	var170093726	T	G	acT|acG	5 (0.009)	2 (0.006)	0.60 [0.10 ; 3.74]	0.59
29	2:170092395	*rs2229267*	A	G	tgA|tgG	164 (0.301)	90 (0.273)	0.90 [0.64 ; 1.27]	0.56
29	2:170092504	var170092504	C	T	cCc|cTc	6 (0.011)	9 (0.027)	2.59 [0.81 ; 8.33]	0.11
30	2:170089934	*rs145384264*	C	T	tcC|tcT	2 (0.004)	1 (0.003)	0.79 [0.06 ; 10.35]	0.86
31	2:170088296	*rs144054579*	C	G	Cgg|Ggg	1 (0.002)	0 (0.000)	-	0.98
31	2:170088242	*rs149469954*	G	A	Gtc|Atc	0 (0.000)	1 (0.003)	-	0.98
31	2:170088351	*rs2302694*	G	A	tcG|tcA	78 (0.143)	28 (0.085)	0.44 [0.27 ; 0.72]	0.001
36	2:170070348	*rs11886219*	T	C	cgT|cgC	8 (0.015)	4 (0.012)	0.93 [0.23 ; 3.74]	0.92
36	2:170070172	*rs4667596*	C	T	aCa|aTa	10 (0.018)	4 (0.012)	0.43 [0.12 ; 1.46]	0.17
39	2:170063263	*rs149367019*	G	A	Gta|Ata	1 (0.002)	0 (0.000)	-	0.98
39	2:170063471	*rs35114151*	A	G	gaA|gaG	35 (0.064)	17 (0.052)	1.07 [0.57 ; 1.98]	0.84
39	2:170063380	*rs35413340*	T	C	Tct|Cct	1 (0.002)	0 (0.000)	-	0.98
39	2:170062977	*rs61995915*	T	C	gTa|gCa	0 (0.000)	1 (0.003)	-	0.98
41	2:170062078	*rs13397109*	G	C	cgG|cgC	95 (0.175)	31 (0.094)	0.45 [0.28 ; 0.72]	0.0009
42	2:170060603	*rs17848169*	T	C	Tat|Cat	2 (0.004)	1 (0.003)	0.77 [0.06 ; 9.86]	0.84
42	2:170060619	*rs199593393*	C	T	caC|caT	0 (0.000)	1 (0.003)	-	0.98
44	2:170058345	var170058345	C	A	Ccc|Acc	5 (0.009)	3 (0.009)	0.75 [0.16 ; 3.46]	0.71
45	2:170055385	var170055385	T	C	cTt|cCt	1 (0.002)	0 (0.000)	-	0.98
46	2:170053505	*rs2228171*	T	C	Tca|Cca	234 (0.430)	132 (0.400)	0.88 [0.65 ; 1.19]	0.40
48	2:170048482	var170038806	C	T	cCc|cTc	0 (0.000)	1 (0.003)	-	0.98
54	2:170032989	*rs2229265*	T	C	caT|caC	409 (0.752)	268 (0.812)	1.55 [1.07 ; 2.23]	0.019
55	2:170031824	*rs17848184*	C	T	ccC|ccT	107 (0.197)	76 (0.230)	1.31 [0.90 ; 1.92]	0.16
56	2:170030506	*rs142549310*	C	T	cCc|cTc	1 (0.002)	0 (0.000)	-	0.98
56	2:170030556	var170030556	G	A	caG|caA	0 (0.000)	3 (0.009)	-	0.98

58	2:170028529	*rs199528723*	G	A	aaG|aaA	5 (0.009)	5 (0.015)	1.08 [0.29 ; 3.99]	0.91
61	2:170025083	*rs2229268*	A	G	tgA|tgG	34 (0.063)	33 (0.100)	1.72 [0.96 ; 3.09]	0.06
62	2:170022511	var170022511	T	C	Ttt|Ctt	1 (0.002)	0 (0.000)	-	0.98
64	2:170013904	*rs79723119*	A	C	gAt|gCt	1 (0.002)	0 (0.000)	-	0.98
64	2:170013979	var170013979	A	G	aAa|aGa	1 (0.002)	0 (0.000)	-	0.98
66	2:170010985	*rs2075252*	T	C	Tag|Cag	173 (0.318)	96 (0.291)	0.81 [0.58 ; 1.13]	0.21
69	2:170003432	*rs4667591*	T	G	Ttc|Gtc	214 (0.393)	110 (0.333)	0.72 [0.53 ; 0.99]	0.04
70	2:170002412	var170002412	T	C	gTc|gCc	1 (0.002)	1 (0.003)	0.90 [0.06 ; 14.56]	0.94
72	2:169997051	*rs990626*	G	A	atG|atA	231 (0.425)	127 (0.385)	0.81 [0.60 ; 1.11]	0.19
73	2:169996070	*rs41268685*	C	T	gCc|gTc	1 (0.002)	0 (0.000)	-	0.98
74	2:169995880	*rs370978040*	G	A	acG|acA	0 (0.000)	1 (0.003)	-	0.98
					A1CF				
5' UTR	10:52645424	*rs10994860*	C	T	-	67 (0.124)	38 (0.113)	1.03 [0.63 ; 1.67]	0.91
5' UTR	10:52619722	var52619722	A	G	-	1 (0.002)	0 (0.000)	-	0.98
2	10:52623804	var52623804	A	G	-	11 (0.020)	5 (0.015)	0.64 [0.22 ; 1.88]	0.42
3	10:52603874	*rs142969066*	T	C	ggT|ggC	1 (0.002)	0 (0.000)	-	0.98
3	10:52603754	*rs35967725*	A	G	tgA|tgG	1 (0.002)	1 (0.003)	2.40 [0.11 ; 51.57]	0.58
5	10:52595864	var52595864	C	T	Cct|Tct	1 (0.002)	0 (0.000)	-	0.98
6	10:52588045	*rs146662131*	C	T	caC|caT	1 (0.002)	0 (0.000)	-	0.98
7	10:52576068	*rs4245008*	A	G	-	100 (0.185)	57 (0.170)	0.94 [0.63 ; 1.39]	0.74
9	10:52573772	*rs41274050*	C	T	Cgc|Tgc	1 (0.002)	0 (0.000)	-	0.98
12	10:52566594	*rs34190540*	G	C	ctG|ctC	3 (0.006)	2 (0.006)	1.32 [0.18 ; 9.81]	0.78
3' UTR	10:52559291	*rs61856570*	A	G	-	1 (0.002)	3 (0.009)	1.84 [0.18 ; 18.82]	0.61
3' UTR	10:52559634	*rs10994507*	A	G	-	27 (0.050)	13 (0.039)	0.85 [0.41 ; 1.76]	0.20
3' UTR	10:52559843	*rs7084132*	G	T	-	100 (0.185)	57 (0.170)	0.94 [0.63 ; 1.39]	0.74
3' UTR	10:52559853	var52559853	T	G	-	1 (0.002)	0 (0.000)	-	0.98
3' UTR	10:52560181	var52560181	T	C	-	1 (0.002)	0 (0.000)	-	0.98
3' UTR	10:52560476	*rs75907017*	G	A	-	124 (0.229)	85 (0.253)	1.26 [0.89 ; 1.81]	0.20

3' UTR	10:52560557	*rs6479731*	A	G	-	100 (0.185)	57 (0.170)	0.94 [0.63 ; 1.39]	0.74
3' UTR	10:52561178	var52561178	G	C	-	0 (0.000)	1 (0.003)	-	0.98
3' UTR	10:52561680	var52561680	T	G	-	4 (0.007)	5 (0.015)	2.30 [0.50 ; 10.55]	0.28
3' UTR	10:52561803	*rs74352101*	A	G	-	27 (0.050)	13 (0.039)	0.85 [0.41 ; 1.76]	0.65
3' UTR	10:52561829	*rs7072584*	A	C	-	100 (0.185)	57 (0.170)	0.94 [0.63 ; 1.39]	0.74
3' UTR	10:52561919	*rs12571156*	C	T	-	124 (0.229)	85 (0.253)	1.26 [0.89 ; 1.81]	0.98
3' UTR	10:52562339	var52562339	C	T	-	1 (0.002)	0 (0.000)	-	0.98
3' UTR	10:52563051	*rs4282939*	C	T	-	250 (0.473)	125 (0.391)	0.76 [0.57 ; 1.02]	0.07
3' UTR	10:52563196	*rs16751*	A	G	-	136 (0.251)	115 (0.342)	1.38 [1.00 ; 1.90]	0.01
3' UTR	10:52563707	*rs184644838*	G	A	-	133 (0.245)	55 (0.164)	0.59 [0.40 ; 0.86]	0.006
3' UTR	10:52563725	var52563845	A	C	-	0 (0.000)	1 (0.003)	-	0.98
3' UTR	10:52563904	var52563981	G	C	-	5 (0.009)	0 (0.000)	-	0.98
3' UTR	10:52563981	*rs150545950*	C	T	-	1 (0.002)	0 (0.000)	-	0.05
3' UTR	10:52564065	*rs4619096*	A	G	-	100 (0.185)	57 (0.170)	0.94 [0.63 ; 1.39]	0.74
3' UTR	10:52564524	var52564700	A	T	-	0 (0.000)	1 (0.003)	-	0.98
3' UTR	10:52564700	*rs74874346*	G	A	-	27 (0.050)	13 (0.039)	0.85 [0.41 ; 1.76]	0.65
3' UTR	10:52564768	var52565100	G	A	-	0 (0.000)	1 (0.003)	-	0.98
3' UTR	10:52565903	*rs4078160*	C	A	-	100 (0.185)	57 (0.170)	0.94 [0.63 ; 1.39]	0.74
3' UTR	10:52565940	var52565951	T	C	-	30 (0.055)	14 (0.042)	0.72 [0.35 ; 1.47]	0.36
3' UTR	10:52566049	*rs10821846*	A	G	-	136 (0.251)	115 (0.342)	1.38 [1.00 ; 1.90]	0.65
3' UTR	10:52566333	var52566417	G	A	-	0 (0.000)	1 (0.003)	-	0.98

UTR: Un-translated region, rs ID: Reference SNP cluster ID, Chr: Chromosome, var: Variant, Ref: Reference, Alt: Alternate. Normo: Normouricaemic, Hyper: Hyperuricaemic, OR: Odds ratio for the alternate allele, 95% CI: 95% confidence interval, P_{OR}: *p*-value for odds ratio. *Odds ratios are adjusted for sex, age, BMI and estimates of grand-parental ancestry. ^Odds ratios are adjusted for sex and collection study. Exon numbers and variant positions are sourced from the Genome Reference Consortium human genome build 37 (GRCh37) on Ensembl (http://grch37.ensembl.org/).

5.2.3.2 Burden analysis

Combining the allele frequencies obtained from single variant analyses with a cut-off combined allele frequency of < 0.05 (5%) to determine the rarity resulted in a list of 111 variants (11 rare and 100 very rare variants) within *LRP2* and 53 variants (1 rare 52 very rare variants) within *A1CF* (Appendix A Table 5.4). Non-synonymous variants were also identified for both genes regardless of their allele frequency. This ended up in a list of 81 missense variants in *LRP2* and 4 missense variants in *A1CF*. Out of 81 missense variants in *LRP2*, 74 were still very rare (< 0.01), while only 7 variants had a combined allele frequency < 0.05 to 0.01. One out of 4 missense variants in *A1CF* was rare and 3 were very rare (Appendix A Table 5.4). Non-synonymous deleterious variants were identified on the basis of information gained via functional annotations for the variants in all four annotation platforms being used (see details in Sections 5.2.2.2.3 and 5.2.3.3).

Analysing the burden of rare exonic variants of *LRP2* and *A1CF* did not signify an effect associated with hyperuricaemia in both NZ Polynesian and European individuals (Table 5.7). The rare variants within *A1CF* indicated a slight trend for protection against hyperuricaemia only in European individuals ($P_{\text{Burden (Protection)}} = 0.06$) but not Polynesians ($P_{\text{Burden (Protection)}} = 0.51$) (Table 5.7). The frequency of individuals carrying a rare allele was almost equal when compared between hyperuricaemic cases and normouricaemic controls for both genes. All results from allele-collapsing and adaptive weighted sum tests were insignificant for a rare variant burden in *LRP2* and *A1CF* (Table 5.7). Changing the number of permutations also did not indicate a considerable change for burden associations at any gene (Table 5.7).

Non-synonymous variants within *LRP2* and *A1CF* with combined allele-frequency of < 0.05 were also tested for having a burden for hyperuricaemia in NZ Polynesian and Europeans. The analysis indicated a significant burden of risk variants within *LRP2* only in the European population ($P_{\text{Burden (Risk)}} = 0.009$) (Table 5.8), which retained when burden for risk and protective variants was collectively calculated ($P_{\text{Burden (Combined)}} = 0.023$) (Table 5.8). The number of rare allele carriers for non-synonymous variants was significantly higher in hyperuricaemic than normouricaemic European individuals ($P = 0.04$) (Table 5.8). The rare non-synonymous variants did not indicate a burden for hyperuricaemia in NZ Polynesian ($P_{\text{Burden (Risk)}} = 0.14$ and $P_{\text{Burden (Protection)}} =$

0.37) with almost equal frequency of rare non-synonymous variants between the hyperuricaemic cases and nonrmouricaemic controls (P = 0.98) (Table 5.8). No notable burden effects or difference in frequency were observed in *A1CF* for both Europeans and Polynesians (Table 5.8).

Table 5.7: Rare variant burden analysis of *LRP2* and *A1CF* for hyperuricaemia

Permut	NZ Polynesian					European				
	Frequency		P_{Burden}			Frequency		P_{Burden}		
	Normo	Hyper	Comb	Risk	Protect	Normo	Hyper	Comb	Risk	Protect
					LRP2					
Allele-collapsing	43 (0.09)	46 (0.11)	0.998	-	-	115 (0.31)	131 (0.35)	0.997	-	-
1,000	-	-	0.543	0.649	0.375	-	-	0.565	0.183	0.843
10,000	-	-	0.519	0.666	0.349	-	-	0.555	0.171	0.859
50,000	-	-	0.521	0.664	0.349	-	-	0.565	0.168	0.861
100,000	-	-	0.525	0.669	0.346	-	-	0.564	0.166	0.861
					A1CF					
Allele-collapsing	28 (0.06)	26 (0.05)	0.692	-	-	36 (0.09)	41 (0.11)	0.999	-	-
1,000	-	-	0.811	0.778	0.528	-	-	0.077	0.407	0.077
10,000	-	-	0.828	0.796	0.503	-	-	0.066	0.396	0.066
50,000	-	-	0.819	0.789	0.508	-	-	0.068	0.401	0.068
100,000	-	-	0.817	0.787	0.510	-	-	0.066	0.401	0.066

Frequency: number and frequency of individulas carrying a rare allele. Permut: Permutations, Normo: Normouricaemic controls, Hyper: Hyperuricaemic cases, Comb: Combined, Protect: Protective.

Table 5.8: Rare and non-synonymous variant burden analysis of *LRP2* and *A1CF* for hyperuricaemia

Permut	NZ Polynesian					European				
	Frequency		P_{Burden}			Frequency		P_{Burden}		
	Normo	Hyper	Comb	Risk	Protect	Normo	Hyper	Comb	Risk	Protect
					LRP2					
Allele-collapsing	49 (0.11)	52 (0.12)	0.985	-	-	68 (0.18)	107 (0.28)	0.045	-	-
1,000	-	-	0.255	0.141	0.372	-	-	0.024	0.008	0.150
10,000	-	-	0.270	0.151	0.365	-	-	0.023	0.009	0.151
50,000	-	-	0.259	0.144	0.373	-	-	0.023	0.009	0.151
100,000	-	-	0.256	0.144	0.370	-	-	0.024	0.009	0.150
					A1CF					
Allele-collapsing	0 (0.00)	1 (0.002)	1	-	-	3 (0.008)	1 (0.003)	0.529	-	-
1,000	-	-	1	0.546	1	-	-	0.432	0.432	1
10,000	-	-	1	0.550	1	-	-	0.458	0.458	1
50,000	-	-	1	0.537	1	-	-	0.444	0.444	1
100,000	-	-	1	0.543	1	-	-	0.444	0.444	1

All values are calculated at combined allel-frequency threshold of < 0.05. Frequency; number and frequency of individulas carrying a rare allele. Permut: Permutations, Normo: Normouricaemic controls, Hyper: Hyperuricaemic cases, Com: Combined, Protect: Protective.

5.2.3.3 Functional annotation of variants

All non-synonymous (missense) variants within *LRP2* and *A1CF* were classified on the basis of having deleterious (damaging) and neutral effects at the protein level. The binary prediction scores indicated some consistency in all four 'functional annotation platforms' namely, CADD, SIFT, PolyPhen-2 and PROVEAN (Table 5.1). The four scores were consistent for a damaging effect on protein function in 52.63% of the cases (Figure 5.6).

Overall, the prediction score was calculated for 72 (84.70%) missense variants for either a deleterious or neutral effect on protein function. Ten variants (*LRP2*: 9; *A1CF*: 1) were identified to have damaging effect in all prediction programs, while 21 variants were identified to be deleterious in at least two out of the four prediction programs. There were 27 (37.5%) missense variants that were found to have a neutral effect on protein function when all four scores were obtained. Both genes had 64 (*LRP2*: 53; *A1CF*: 11) synonymous variants, with 54 untranslated region/UTR variants (5'UTR: 3; 3'UTR: 51) in *A1CF*. A detail of significant non-synonymous annotations is provided in Table 5.9 below (full details provided in Appendix A, Table 5.6). A CADD score of ≥ 10 indicates deleteriousness of the particular variant at protein level. Likewise, a score of ≤ 0.05, ≥ 0.43 and ≤ -2.5 indicate a damaging effect as per SIFT, PolyPhen-2 and PROVEAN, respectively.

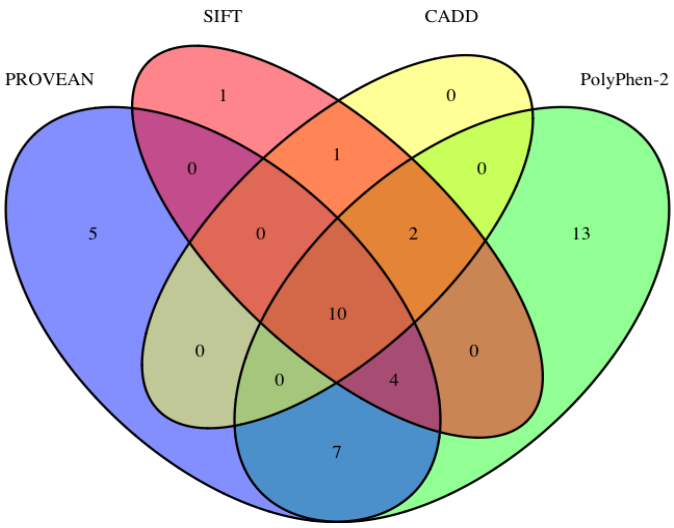


Figure 5.6: Consistency between four methods predicting the deleterious functional consequences of non-synonymous variants in *LRP2* and *A1CF*.

Table 5.9: Predicted consequences of ten *LRP2* and *A1CF* non-synonymous variants on protein function

Exon	Chr: location	rs ID/var position	CADD	PolyPhen-2		SIFT		PROVEAN	
			Score	Score	Prediction	Score	Prediction	Score	Prediction
73	2:169996058	var169996058	29.7	2.1	Deleterious	0.03	Deleterious	-3.15	Deleterious
59	2:170027153	*rs200475391*	32	2.19	Deleterious	0	Deleterious	-6.94	Deleterious
56	2:170030458	var170030458	28.1	2.02	Deleterious	0	Deleterious	-7.67	Deleterious
51	2:170038795	var170038795	25	2.19	Deleterious	0.01	Deleterious	-3.22	Deleterious
50	2:170042245	*rs35734447*	27.8	1.63	Deleterious	0.01	Deleterious	-4.96	Deleterious
37	2:170068592	var170068592	34	3.62	Deleterious	0.02	Deleterious	-3.34	Deleterious
15	2:170115652	var170115652	28.7	3.46	Deleterious	0	Deleterious	-8.28	Deleterious
14	2:170129529	*rs116332504*	24	1.42	Deleterious	0	Deleterious	-3.8	Deleterious
13	2:170134385	*rs201490492*	31	3.58	Deleterious	0.02	Deleterious	-4.29	Deleterious
5	10:52595978	*rs143123872*	10.7	1.54	Deleterious	0.03	Deleterious	-2.58	Deleterious

Chr: Chromosome, rs ID: Reference SNP cluster ID, var: variant, CADD: Combined annotation dependent depletion, PolyPhen-2: Polymorphism phenotyping verision 2, SIFT: Scale-invariant feature transform, PROVEAN: Protein variantion effect analyser.

5.2.4 Results – Replication-phase

Power calculations inidcated the NZ Polynesian dataset to be highly powered (> 90%) to detect a moderate effect (OR = 1.5) at an altered allele frequency > 0.1 (Figure 5.7). The European dataset was only adequately powered (> 80%) to detect a moderate effect (OR = 1.5) at an altered allele frequency > 0.1 (Figure 5.7).

Applying the criteria for prioritisation described in Section 5.2.2.3.1 ended up in a list of seven common variants within exonic regions of *LRP2* only. Table 5.10 provides a detailed summary of these variants while Figure 5.8 illustrates an overview of exonic regions (n (exons) = 79) of the gene *LRP2* with positions of all 7 variants indicated within their respective exons. Calculation of LD showed no in between variant LD in Europeans, while there was a complete LD between *rs2302694* and *rs4667596* and between *rs4667591* and *rs2075252* only in the Han Chinese population. However, all four variants were selected for replication in NZ datasets. (Table 5.11). Only one variant, *rs41268685*, was found to be monomorphic for European and South Asian populations when population-genetic data were used from the Genome Reference consortium human genome build 37 (GRCh37) on Ensembl (www.grch37.ensembl.org), while all other six variants were polymorphic.

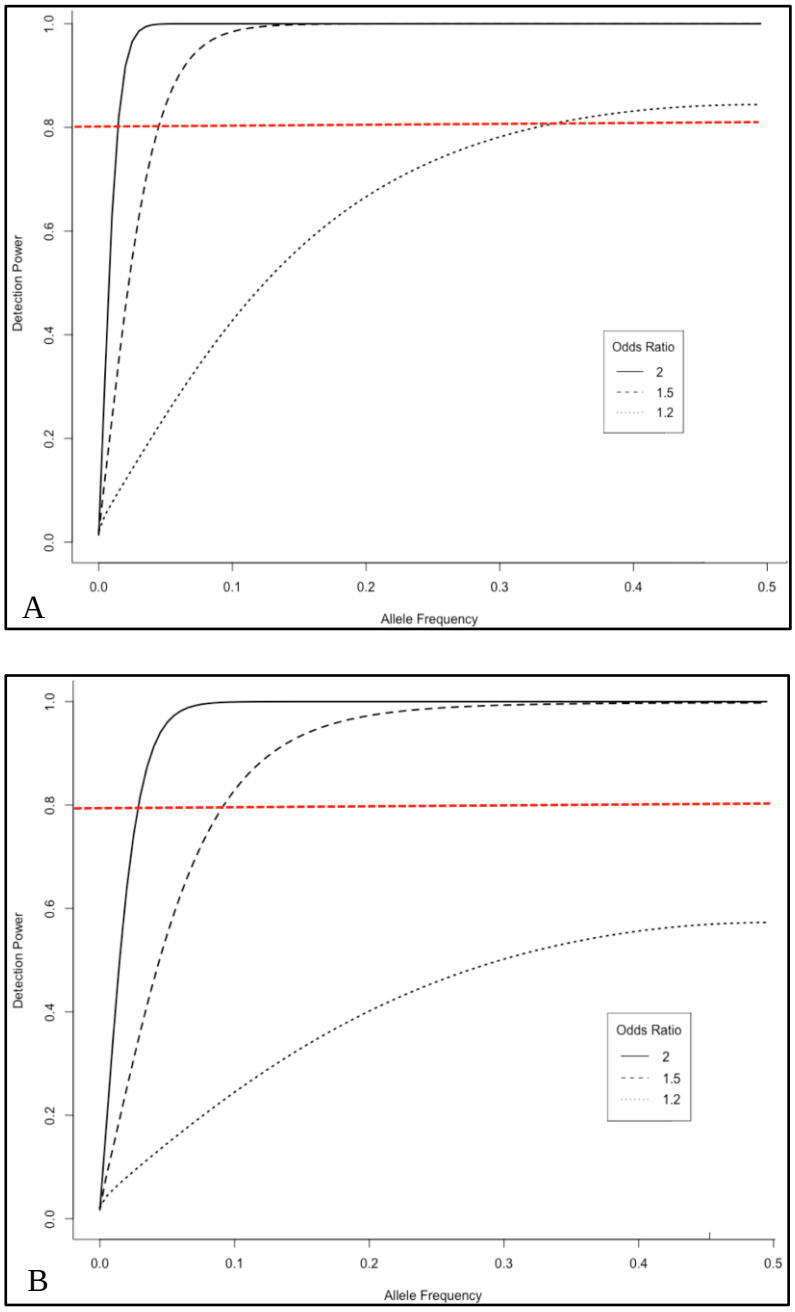


Figure 5.7: Association detection power in the (A) NZ Polynesian and (B) NZ European subsets in the Replication Cohorts across a range of odds ratio effect sizes and minor allele frequencies. The broken red line indicates an adequate detection power ≥ 80%.

Table 5.10: Summary detail of variants prioritised for replication

Exon	Location	rs ID/var position	Functional consequence	OR (P_{OR})			Direction of association	Population specificity	Alternative-allele frequency		Novelty
				Hyperuricaemia		Gout			NZ Polynesian	European	
				Poly	Eur	Poly					
14	2:170131548	*rs111360923*	Missense	0.55 (0.02)	-	0.44 (0.004)	Protective	Polynesian	0.070	-	Not in 1000 Genome database
17	2:170115626	var170115626	Missense	0.57 (0.03)	-	0.46 (0.006)	Protective	Polynesian	0.075	-	Not in 1000 Genome database
31	2:170088351	*rs2302694*	Synonymous	0.46 (0.01)	0.71 (0.31)	0.44 (0.001)	Protective	Both	0.095	0.082	Not novel
36	2:170070172	*rs4667596*	Missense	0.22 (0.02)	0.65 (0.56)	0.43 (0.17)	Protective	Both	0.009	0.013	Not novel
66	2:170010985	*rs2075252*	Stop-lost (nonsense)	0.78 (0.19)	0.65 (0.05)	0.18 (0.21)	Protective	Both	0.289	0.714	Not novel
69	2:170003432	*rs4667591*	Missense	0.75 (0.08)	0.61 (0.05)	0.72 (0.05)	Protective	Both	0.341	0.768	Not novel
73	2:169996070	*rs41268685*	Missense	-	3.85 (0.04)	-	Risk	Both	0.002	0.028	Not novel

rs ID: Reference SNP cluster ID, var: Variant, Poly: NZ Polynesian, Eur: European, OR: Odds ratio, P_{OR}: p-value of odds ratio.

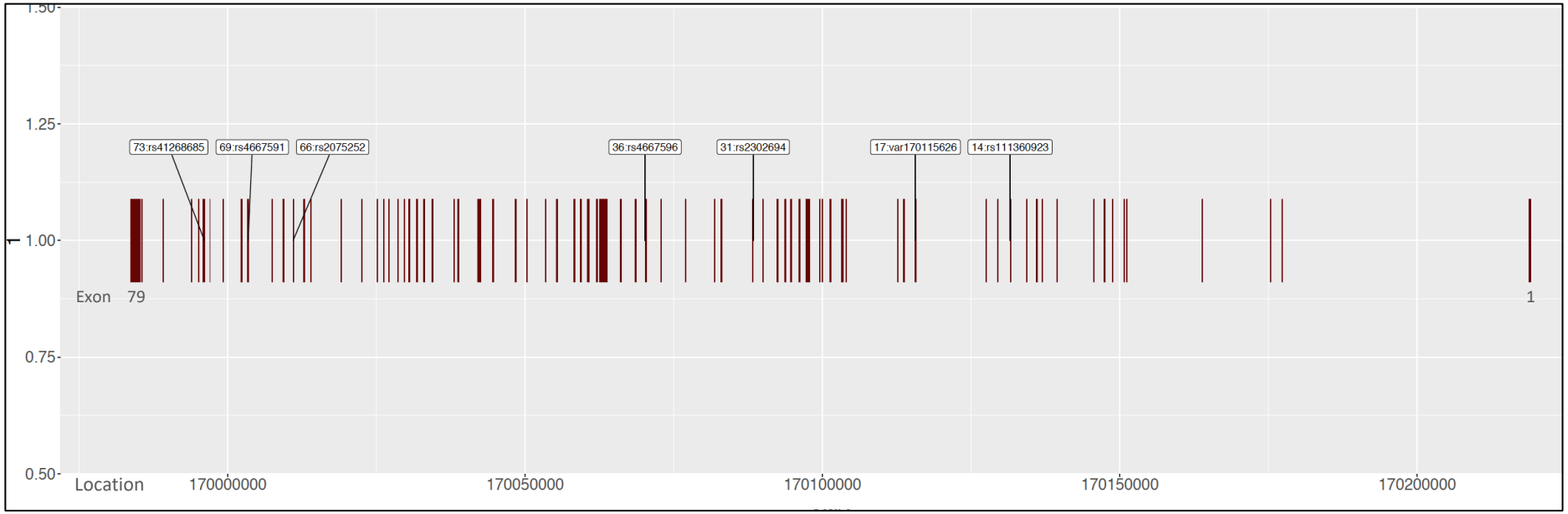


Figure 5.8: An overview of *LRP2* showing 79 exons (dark red blocks) and the position of each prioritised variant on the respective exons. Labels on the x-axis indicates location of variants in each exon, while values on y-axis are only presenting graphics in the plot. Plot was created using exon sequencing data in European and NZ Polynesian individuals in R (v3.3.2). Exons numbers and variant positions are sourced from the Genome Reference Consortium human genome build 37 (GRCh37) on Ensembl (http://grch37.ensembl.org/).

Table 5.11: R-squared (r^2) LD values between prioritised variants in Caucasian and Chinese/Japanese populations

Variant/SNP		Distance	R-squared (r^2)	
Query	**Proxy**		**Caucasian**	**Han Chinese**
rs111360923	No record on any array	-	-	-
var170115626	No record on any array	-	-	-
rs4667591	*rs2075252*	7553	0.39	0.829
	rs4667596	66740	0.087	0.001
	rs2302694	84919	0.009	0.001
rs2302694	*rs4667596*	18179	0.174	1
	rs2075252	77366	0.007	0.003
	rs4667591	84919	0.009	0.001
rs2075252	*rs4667591*	7553	0.39	0.829
	rs4667596	59187	0.064	0.003
	rs2302694	77366	0.007	0.003
rs4667596	*rs2302694*	18179	0.174	1
	rs4667591	66740	0.087	0.001
	rs2075252	59187	0.39	0.003
rs41268685	*rs4667591*	7362	0.004	-
	rs2302694	92281	0.156	-
	rs2075252	14915	0.004	-
	rs4667596	74102	0	-

SNP: Single nucleotide polymorphism, Query: SNP for which the LD was calculated, Proxy: SNP with which the LD was calculated. R-squared ≥ 0.60 indicates an LD. All values are as provided in 1000 Genome database (http://browser.1000genomes.org/).

The association results in resequencing for *rs111360923* were successfully replicated for gout in a larger NZ Polynesian dataset (Table 5.12). The altered allele (C) of the Polynesian-specific missense variant *rs111360923* was significantly associated with gout in both EP (OR = 0.59, P = 0.01) and WP (OR = 0.57, P = 0.02) individuals (Table 5.12). The association was not significant for both EPWP (OR = 1.15, P = 0.80) and NPH (OR = 0.89, P = 0.76) individuals (Table 5.12). When all Polynesian subgroups were combined together in a meta-analysis, the C allele of *rs111360923* indicated a strong protective association for gout with no heterogeneity between the subgroups (OR = 0.64, P_{OR} = 0.002, P_{Het} = 0.52) (Table 5.14). The Replication-phase failed to follow similar trend of association between *rs111360923* and hyperuricaemia (Table 5.12 and Table 5.14). Also, no association was observed between the C allele of *rs111360923* and serum urate concentrations in NZ Polynesian controls (ß (mmol L^{-1}) = -0.01, $P_{ß}$ = 0.12, P_{Het} = 0.61) (Table 5.14).

The association for the C allele of another Polynesian-specific variant, var170115626, was also successfully replicated and was found to be nominally associated with a decreased risk of gout in WP individuals (OR = 0.73, P = 0.05) (Table 5.12). No

significant association was found for any other population subgroup, although the direction of effect was consistent in EP, WP and NH subgoups (Table 5.12). Combining the Polynesian subgroups in meta-analysis indicated a significant protective association between the C allele of var170115625 and gout with no heterogeneity intra populations (OR = 0.73, P_{OR} = 0.02, P_{Het} = 0.65) (Table 5.14). The C allele also indicated a trend of negative association with serum urate concentrations (ß (mmol L^{-1}) = -0.012, $P_{ß}$ = 0.06, P_{Het} = 0.59) (Table 5.14). However, the association results from the Discovery-phase were not replicated for hyperuricaemia in any Polynesian subgroups (Table 5.12 and Table 5.14).

Given the data for above two variants indicated Polynesian-specific effects and the evidence of previously reported variant within *LRP2* (*rs2544390*) that also indicated to have Polynesian-specific effects (Rasheed *et al.*, 2013a), an LD between all three variants was calculated using the genotype data in NZ Polynesian population using Haploview (v4.2) platform. While the results from the LD calculation showed a strong (83%) LD between *rs111360923* and var170115626, both of the variants were found to be independent of the signals at *rs2544390* with only 5% in between LD (Figure 5.9).

The A allele of the synonymous variant *rs2302694* was not associated with either hyperuricaemia or gout when tested in the NZ Polynesian subgroups and NZ Europeans separately (Table 5.12), although meta-analysis successfully replicated the protective association against hyperuricaemia (OR = 0.79, P_{OR} = 0.04, P_{Het} = 0.35) (Table 5.14). The A allele also indicated a trend of protective association with gout (OR = 0.80, P_{OR} = 0.06, P_{Het} = 0.89) in NZ Polynesians only (Table 5.14). The allele A was also found to be associated with decreasing the serum urate concentrations in EP (ß (mmol L^{-1}) = -0.014, P = 0.04) (Table 5.13) and combined Polynesian healthy controls (ß (mmol L^{-1}) = -0.015, $P_{ß}$ = 0.01, P_{Het} = 0.63) (Table 5.14).

The association results for another synonymous variant, *rs2075252*, were successfully replicated for hyperuricaemia. The C allele of *rs2075252* was associated with hyperuricaemia with a protective effect in EP (OR = 0.81, P = 0.05), while it showed an opposite susceptible association in EPWP (OR = 2.54, P = 0.05) individuals (Table 5.12). Meta-analysis of all Polynesian group together indicated a protective association of the C allele against hyperuricaemia (OR = 0.83, P_{OR} = 0.01, P_{Het} = 0.11) (Table 5.12), which followed the similar trend of association when NZ Polynesian and European

populations were combined together (OR = 0.87, P_{OR} = 0.03, P_{Het} = 0.12) (Table 5.14). However, the C allele was not associated with serum urate in either NZ Polynesian (ß (mmol L^{-1}) = -0.005, $P_{ß}$ = 0.23, P_{Het} = 0.24) or NZ European (ß (mmol L^{-1}) = -0.007, P = 0.30) individuals (Table 5.13 and Table 5.14).

The missense variant *rs4667591* represents another example of successful replication of Discovery-phase associations in this study. The G allele of *rs4667591* showed a significant protective association with hyperuricaemia in both EP (OR = 0.74, P = 0.006) and NPH (OR = 0.67, P = 0.04) sample sets (Table 5.12). The same G allele was also found to be associated with increased risk of developing gout in NZ Europeans (OR = 1.26, P = 0.02) (Table 5.12). Combining the NZ Polynesian sample sets together indicated association of the G allele with hyperuricaemia (OR = 0.79, P_{OR} = 0.002, P_{Het} = 0.38). Meta-analysis also showed association with hyperuricaemia when NZ Polynesian and European populations were combined together (OR = 0.82, P_{OR} = 0.002, P_{Het} = 0.39) (Table 5.14). The G allele was not associated with serum urate concentrations in any population group (Table 5.13 and Table 5.14).

The Replication-phase failed to replicate the association for two missense variants, *rs4667596* and *rs41268685*, in NZ Polynesian and European populations (Table 5.12). However, the T allele of *rs4667596* (ß (mmol L^{-1}) = 0.136, $P_{ß}$ = 0.03) and the G allele of *rs41268685* (ß (mmol L^{-1}) = 0.192, $P_{ß}$ = 0.01) were found to be associated with increased serum urate concentrations in NPH individuals only (Table 5.13).

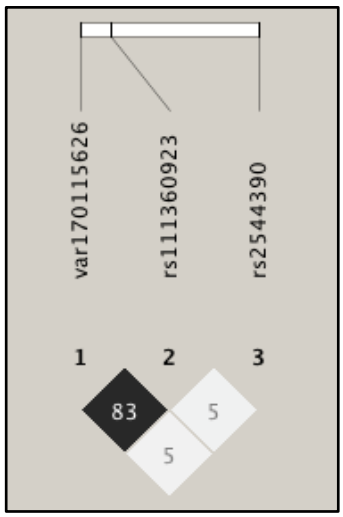


Figure 5.8: Linkage disequilibrium (LD) plot indicating 'R-squared/r^2' values between the 3 Polynesian-specific *LRP2* variants. An $r^2 \geq 60$ in the above plot indicates LD between the particular variants. Information for variant location, rs ID and LD values are from NZ Polynesian data. The plot was generated using Haploview v4.2.

Table 5.12: Genotype/allele frequencies and association analysis results for the *LRP2* variants in the Replication Cohort for gout and hyperuricaemia

Population/rs ID/var position	Group	Genotype/ Altered Allele Frequency				Gout (case/control)				Hyperuricaemia (NU/HU)			
						Unadjusted		Adjusted*		Unadjusted		Adjusted*	
						OR [95% CI]	*P*	OR [95% CI]	*P*	OR [95% CI]	*P*	OR [95% CI]	*P*
rs111360923		**TT**	**CT**	**CC**	**C**								
EP	Case	339 (0.856)	55 (0.138)	2 (0.005)	59 (0.074)	0.84 [0.62 ; 1.14]	0.27	0.59 [0.40 ; 0.88]	0.01	0.96 [0.69 ; 1.35]	0.83	0.84 [0.58 ; 1.21]	0.35
	Control	522 (0.851)	85 (0.138)	6 (0.009)	97 (0.079)								
WP	Case	271 (0.874)	39 (0.125)	0 (0.000)	39 (0.062)	0.70 [0.47 ; 1.05]	0.08	0.57 [0.35 ; 0.93]	0.02	0.84 [0.56 ; 1.28]	0.42	0.79 [0.51 ; 1.22]	0.29
	Control	234 (0.866)	32 (0.118)	4 (0.015)	40 (0.074)								
EPWP	Case	21 (0.807)	5 (0.192)	0 (0.000)	5 (0.096)	1.39 [0.53 ; 3.45]	0.49	1.15 [0.37 ; 3.49]	0.80	1.55 [0.60 ; 4.36]	0.38	1.32 [0.45 ; 4.24]	0.62
	Control	49 (0.830)	10 (0.169)	0 (0.000)	10 (0.085)								
NPH	Case	230 (0.855)	36 (0.134)	3 (0.011)	42 (0.078)	1.19 [0.72 ; 2.01]	0.50	0.89 [4.34 ; 1.92]	0.76	0.98 [0.61 ; 1.55]	0.92	0.77 [0.45 ; 1.29]	0.33
	Control	167 (0.869)	25 (0.130)	0 (0.000)	25 (0.065)								
var170115626		**TT**	**CT**	**CC**	**C**								
EP	Case	337 (0.851)	56 (0.141)	3 (0.007)	62 (0.078)	0.92 [0.68 ; 1.24]	0.60	0.72 [0.49 ; 1.07]	0.11	0.87 [0.61 ; 1.21]	0.41	0.79 [0.55 ; 1.14]	0.22
	Control	529 (0.863)	77 (0.125)	7 (0.014)	91 (0.074)								
WP	Case	268 (0.864)	42 (0.135)	0 (0.000)	42 (0.067)	0.78 [0.53 ; 1.16]	0.22	0.73 [0.39 ; 1.02]	0.05	0.79 [0.52 ; 1.19]	0.27	0.71 [0.46 ; 1.10]	0.13
	Control	235 (0.870)	31 (0.115)	4 (0.015)	39 (0.072)								
EPWP	Case	22 (0.846)	4 (0.154)	0 (0.000)	4 (0.077)	1.40 [0.50 ; 3.64]	0.50	0.63 [0.41 ; 4.28]	0.63	1.47 [0.54 ; 4.44]	0.46	1.40 [0.45 ; 4.91]	0.57
	Control	49 (0.830)	10 (0.169)	0 (0.000)	10 (0.084)								

NPH	Case	232 (0.862)	36 (0.134)	1 (0.004)	38 (0.071)	1.13 [0.66 ; 1.95]	0.66	1.33 [4.15 ; 1.97]	0.76	1.18 [0.72 ; 1.91]	0.51	0.95 [0.54 ; 1.64]	0.85
	Control	168 (0.875)	24 (0.125)	0 (0.000)	24 (0.062)								
rs2302694		**GG**	**AG**	**AA**	**A**								
EP	Case	278 (0.751)	85 (0.229)	7 (0.019)	99 (0.134)	0.90 [0.69 ; 1.15]	0.40	0.85 [0.62 ; 1.16]	0.30	0.83 [0.62 ; 1.09]	0.19	0.79 [0.58 ; 1.06]	0.12
	Control	351 (0.758)	102 (0.220)	10 (0.022)	122 (0.132)								
WP	Case	224 (0.794)	56 (0.198)	2 (0.007)	60 (0.106)	0.85 [0.59 ; 1.23]	0.39	0.71 [0.47 ; 1.10]	0.12	0.80 [0.55 ; 1.16]	0.24	0.71 [0.48 ; 1.05]	0.08
	Control	156 (0.812)	33 (0.172)	3 (0.015)	39 (0.101)								
EPWP	Case	31 (0.688)	14 (0.311)	0 (0.000)	14 (0.155)	1.11 [0.44 ; 2.69]	0.82	0.99 [0.33 ; 2.82]	0.98	1.61 [0.67 ; 4.12]	0.30	2.13 [0.73 ; 7.03]	0.19
	Control	18 (0.750)	6 (0.250)	0 (0.000)	6 (0.125)								
NPH	Case	105 (0.729)	38 (0.264)	1 (0.007)	40 (0.138)	1.00 [0.54 ; 1.94]	0.99	0.70 [2.76 ; 1.83]	0.46	0.93 [0.54 ; 1.56]	0.77	0.84 [0.46 ; 1.49]	0.55
	Control	47 (0.723)	18 (0.276)	0 (0.000)	18 (0.138)								
EUR	Case	586 (0.804)	131 (0.179)	11 (0.015)	153 (0.105)	0.98 [0.75 ; 1.29]	0.89	0.96 [0.71 ; 1.31]	0.81	1.03 [0.77 ; 1.38]	0.82	0.98 [0.73 ; 1.33]	0.92
	Control	362 (0.812)	80 (0.179)	4 (0.009)	88 (0.098)								
rs4667596		**CC**	**CT**	**TT**	**T**								
EP	Case	451 (0.974)	12 (0.026)	0 (0.000)	12 (0.129)	1.73 [0.93 ; 3.25]	0.08	1.76 [0.82 ; 3.92]	0.15	0.85 [0.41 ; 1.68]	0.65	0.79 [0.37 ; 1.63]	0.53
	Control	348 (0.941)	22 (0.059)	0 (0.000)	22 (0.029)								
WP	Case	268 (0.950)	13 (0.046)	1 (0.003)	15 (0.026)	1.46 [0.68 ; 3.41]	0.35	0.98 [0.42 ; 2.47]	0.97	0.86 [0.36 ; 2.04]	0.72	0.69 [0.29 ; 1.66]	0.40
	Control	186 (0.968)	6 (0.031)	0 (0.000)	6 (0.015)								
EPWP	Case	22 (0.916)	2 (0.083)	0 (0.000)	2 (0.042)	4.32 [0.40 ; 95.09]	0.24	3.77 [0.28 ; 94.81]	0.33	0.38 [0.02 ; 4.07]	0.43	0.51 [0.02 ; 6.69]	0.61
	Control	44 (0.978)	1 (0.022)	0 (0.000)	1 (0.011)								
NPH	Case	139	5	0	5	0.55	0.38	0.68	0.69	1.27	0.70	1.51	0.58

		(0.965)	(0.034)	(0.000)	(0.017)	[0.14 ; 2.28]		[1.10 ; 5.75]		[0.36 ; 4.31]		[0.35 ; 6.85]	
	Control	61 (0.938)	4 (0.061)	0 (0.000)	4 (0.031)								
	Case	690 (0.947)	38 (0.052)	0 (0.000)	38 (0.026)	0.94 [0.56 ; 1.60]	0.81	0.82 [0.46 ; 1.50]	0.51	1.25 [0.71 ; 2.17]	0.43	1.14 [0.63 ; 2.02]	0.66
EUR	Control	422 (0.946)	24 (0.054)	0 (0.000)	24 (0.027)								
rs2075252		**TT**	**CT**	**CC**	**C**								
	Case	132 (0.356)	172 (0.464)	66 (0.178)	304 (0.411)	0.90 [0.75 ; 1.07]	0.24	0.86 [0.69 ; 1.07]	0.17	0.79 [0.66 ; 0.97]	0.02	0.81 [0.65 ; 1.00]	0.05
EP	Control	145 (0.313)	232 (0.501)	86 (0.185)	404 (0.436)								
	Case	132 (0.468)	120 (0.425)	30 (0.106)	180 (0.319)	1.09 [0.86 ; 1.38]	0.46	1.27 [0.95 ; 1.70]	0.11	0.85 [0.66 ; 1.09]	0.19	0.85 [0.65 ; 1.10]	0.22
WP	Control	102 (0.531)	69 (0.359)	21 (0.109)	111 (0.289)								
	Case	13 (0.542)	11 (0.458)	0 (0.000)	11 (0.229)	0.69 [0.33 ; 1.38]	0.30	0.57 [0.21 ; 1.45]	0.25	1.56 [0.79 ; 3.22]	0.21	2.54 [1.03 ; 6.78]	0.05
EPWP	Control	13 (0.288)	27 (0.600)	5 (0.111)	37 (0.411)								
	Case	44 (0.305)	71 (0.493)	29 (0.201)	129 (0.447)	1.13 [0.76 ; 1.71]	0.55	1.07 [5.76 ; 1.99]	0.84	0.76 [0.54 ; 1.06]	0.11	0.75 [0.51 ; 1.08]	0.12
NPH	Control	25 (0.384)	26 (0.400)	14 (0.215)	54 (0.415)								
	Case	48 (0.065)	276 (0.379)	404 (0.555)	1048 (0.744)	1.13 [0.94 ; 1.36]	0.20	1.12 [0.92 ; 1.38]	0.26	0.99 [0.81 ; 1.20]	0.89	0.97 [0.79 ; 1.19]	0.78
EUR	Control	32 (0.072)	191 (0.428)	223 (0.500)	637 (0.714)								
rs4667591		**TT**	**GT**	**GG**	**G**								
	Case	112 (0.303)	171 (0.462)	87 (0.235)	345 (0.466)	0.93 [0.79 ; 1.10]	0.41	0.87 [0.70 ; 1.07]	0.19	0.77 [0.64 ; 0.93]	0.006	0.74 [0.60 ; 0.92]	0.006
EP	Control	135 (0.292)	216 (0.466)	112 (0.242)	440 (0.475)								
	Case	110 (0.390)	137 (0.485)	35 (0.124)	252 (0.446)	0.78 [0.62 ; 0.98]	0.03	0.94 [0.71 ; 1.24]	0.66	0.83 [0.65 ; 1.05]	0.13	0.89 [0.69 ; 1.15]	0.39
WP	Control	62 (0.324)	91 (0.476)	38 (0.199)	167 (0.437)								
EPWP	Case	11 (0.458)	13 (0.5412)	0 (0.000)	13 (0.271)	0.57 [0.28 ; 1.09]	0.10	0.76 [0.34 ; 1.66]	0.49	1.04 [0.56 ; 1.94]	0.91	1.21 [0.56 ; 2.70]	0.62

NPH	Control	8 (0.178)	29 (0.644)	8 (0.178)	45 (0.50)								
	Case	36 (0.250)	76 (0.527)	32 (0.222)	140 (0.486)	1.04 [0.69 ; 1.57]	0.86	0.79 [4.16 ; 1.48]	0.46	0.73 [0.52 ; 1.02]	0.07	0.67 [0.45 ; 0.99]	0.04
	Control	20 (0.307)	28 (0.431)	17 (0.261)	62 (0.476)								
EUR	Case	25 (0.034)	248 (0.341)	455 (0.625)	1158 (0.795)	1.29 [1.06 ; 1.57]	0.01	1.26 [1.02 ; 1.56]	0.02	0.94 [0.76 ; 1.16]	0.57	0.91 [0.73 ; 1.14]	0.43
	Control	26 (0.058)	174 (0.396)	246 (0.552)	666 (0.746)								
rs41268685		**AA**	**AG**	**GG**	**G**								
EP	Case	367 (0.992)	3 (0.008)	0 (0.000)	3 (0.004)	1.05 [0.21 ; 4.78]	0.95	0.50 [0.09 ; 2.84]	0.42	1.52 [0.28 ; 8.25]	0.61	1.78 [0.31 ; 10.61]	0.51
	Control	460 (0.993)	3 (0.006)	0 (0.000)	3 (0.003)								
WP	Case	281 (0.996)	1 (0.003)	0 (0.000)	1 (0.002)	0.78 [0.03 ; 19.8]	0.86	0.50 [0.02 ; 16.01]	0.66	-	-	-	-
	Control	191 (0.994)	1 (0.005)	0 (0.000)	1 (0.002)								
EPWP	Case	24 (100)	0 (0.000)	0 (0.000)	0 (0.000)	-	-	-	-	-	-	-	-
	Control	44 (0.978)	1 (0.022)	0 (0.000)	1 (0.011)								
NPH	Case	144 (100)	0 (0.000)	0 (0.000)	0 (0.000)	-	-	-	-	-	-	-	-
	Control	64 (0.984)	1 (0.015)	0 (0.000)	1 (0.007)								
EUR	Case	687 (0.943)	41 (0.056)	0 (0.000)	41 (0.028)	1.49 [0.86 ; 2.71]	0.17	1.64 [0.88 ; 3.19]	0.13	0.88 [0.48 ; 1.56]	0.67	0.90 [0.48 ; 1.67]	0.75
	Control	430 (0.964)	16 (0.036)	0 (0.000)	16 (0.018)								

*All values are adjusted for age, sex and BMI for NZ European plus with the estimates of grand-parental ancestry for NZ Polynesian. EP: East Polynesian, WP: West Polynesian, EPWP: Mixture of East and West Polynesian, NPH: Ngati Porou Hauora, EUR: NZ Europeans, NU: Normouricaemic, HU: Hyperuricaemic, rs ID: Reference SNP cluster ID, var: Variant, OR: Odds ratio for the alternate allele, 95% CI: 95% confidence interval, *P*: *p*-value for odds ratio.

Table 5.13 Association analysis results for the variants in the Replication Cohort for serum urate

rs ID/var position	Population	Unadjusted		Adjusted*	
		ß [95% CI]	*P*	ß [95% CI]	*P*
rs111360923	EP	0.003 [-0.017 ; 0.022]	0.77	-0.011 [-0.028 ; 0.005]	0.17
	WP	-0.001 [-0.027 ; 0.025]	0.94	-0.008 [-0.031 ; -0.015]	0.49
	EPWP	0.046 [-0.024 ; 0.117]	0.19	0.028 [-0.034 ; 0.090]	0.37
	NPH	-0.020 [-0.061 ; 0.022]	0.35	-0.020 [-0.059 ; -0.019]	0.32
var170115626	EP	-0.001 [-0.020 ; -0.018]	0.93	-0.014 [-0.031 ; 0.002]	0.08
	WP	-0.002 [-0.029 ; 0.024]	0.86	-0.01 [-0.033 ; -0.013]	0.39
	EPWP	0.049 [-0.027 ; 0.125]	0.21	0.033 [-0.034 ; 0.100]	0.33
	NPH	-0.014 [-0.054 ; 0.026]	0.48	-0.015 [-0.053 ; 0.024]	0.45
rs2302694	EP	-0.012 [-0.028 ; 0.005]	0.17	-0.014 [-0.015 ; 0.005]	0.04
	WP	-0.009 [-0.036 ; -0.018]	0.52	-0.021 [-0.045 ; 0.004]	0.10
	EPWP	0.027 [-0.036 ; 0.090]	0.40	0.022 [-0.037 ; 0.081]	0.46
	NPH	-0.014 [-0.069 ; 0.041]	0.62	-0.015 [-0.076 ; 0.046]	0.62
	EUR	0.008 [-0.015 ; 0.031]	0.51	0.007 [-0.012 ; 0.025]	0.48
rs4667596	EP	-0.022 [-0.067 ; 0.022]	0.32	-0.029 [-0.066 ; 0.008]	0.12
	WP	0.004 [-0.077 ; 0.085]	0.92	-0.032 [-0.104 ; 0.040]	0.38
	EPWP	0.003 [-0.214 ; 0.219]	0.98	-0.011 [-0.207 ; 0.185]	0.91
	NPH	0.104 [-0.010 ; 0.198]	0.03	0.136 [-0.013 ; 0.258]	0.03
	EUR	-0.013 [-0.033 ; 0.059]	0.58	0.006 [-0.031 ; 0.043]	0.74
rs2075252	EP	-0.005 [-0.017 ; 0.006]	0.36	-0.005 [-0.015 ; 0.005]	0.33
	WP	-0.015 [-0.033 ; 0.003]	0.11	-0.014 [-0.030 ; 0.003]	0.11
	EPWP	0.008 [-0.039 ; 0.054]	0.76	0.029 [-0.020 ; 0.078]	0.26
	NPH	-0.014 [-0.045 ; -0.017]	0.37	-0.017 [-0.022 ; 0.056]	0.37
	EUR	-0.008 [-0.024 ; 0.007]	0.28	-0.007 [-0.019 ; 0.006]	0.30
rs4667591	EP	-0.002 [-0.013 ; 0.009]	0.75	-0.002 [-0.012 ; 0.007]	0.62
	WP	0.001 [-0.016 ; -0.018]	0.93	0.008 [-0.008 ; 0.023]	0.34
	EPWP	-0.017 [-0.062 ; 0.028]	0.46	-0.005 [-0.047 ; 0.036]	0.80
	NPH	-0.006 [-0.039 ; 0.026]	0.69	-0.012 [-0.024 ; 0.048]	0.50
	EUR	-0.016 [-0.032 ; -0.001]	0.04	-0.008 [-0.021 ; 0.004]	0.20
rs41268685	EP	-0.018 [-0.098 ; 0.135]	0.76	-0.013 [-0.110 ; 0.084]	0.79
	WP	-	-	-	-
	EPWP	-0.120 [-0.334 ; 0.095]	0.28	-	-
	NPH	0.171 [-0.012 ; 0.331]	0.04	0.192 [0.047 ; 0.337]	0.01
	EUR	-0.021 [-0.072 ; 0.029]	0.41	-0.019 [-0.061 ; 0.024]	0.39

EP: East Polynesian, WP: West Polynesian, EPWP: Mixture of East and West Polynesian, NPH: Ngati Porou Hauora, EUR: NZ Europeans, rs ID: Reference SNP cluster ID, var: Variant, ß: Beta/effect estimates for the alternate allele, 95% CI: 95% confidence interval, *P*: *p*-values for effect estimates. *All values are adjusted for age, sex and BMI for NZ European plus with the estimates of grand-parental ancestry for NZ Polynesian.

Table 5.14 Meta-analysis results for the variants in the Replication Cohort for gout, hyperuricaemia and serum urate

rs ID/var position	Population	Gout (case/control)			Hyperuricaemia (NU/HU)			Serum urate (mmol L^{-1})		
		OR [95% CI]	P_{OR}	P_{Het}	OR [95% CI]	P_{OR}	P_{Het}	ß [95% CI]	$P_{ß}$	P_{Het}
rs111360923	POLY	0.646 [0.491; 0.851]	0.002	0.52	0.828 [0.650; 1.054]	0.13	0.85	-0.010 [-0.022; 0.003]	0.12	0.61
var170115626	POLY	0.736 [0.561; 0.966]	0.03	0.65	0.814 [0.638; 1.039]	0.10	0.69	-0.012 [-0.023; 0.001]	0.06	0.59
rs2302694	POLY & EUR	0.859 [0.712; 1.035]	0.11	0.82	0.857 [0.718; 1.022]	0.08	0.34	-0.015 [-0.025; -0.003]	0.01	0.78
	POLY	0.801 [0.632; 1.015]	0.06	0.89	0.797 [0.641; 0.991]	0.04	0.35	-0.015 [-0.026; -0.003]	0.01	0.63
rs4667596	POLY & EUR	1.069 [0.718; 1.591]	0.74	0.48	0.936 [0.637; 1.373]	0.74	0.78	-0.008 [-0.031; 0.016]	0.53	0.09
	POLY	1.339 [0.778; 2.301]	0.29	0.56	0.803 [0.480; 1.341]	0.40	0.81	-0.018 [-0.049; 0.013]	0.27	0.07
rs2075252	POLY & EUR	1.035 [0.909; 1.176]	0.61	0.15	0.879 [0.777; 0.993]	0.03	0.12	-0.006 [-0.012; 0.001]	0.12	0.37
	POLY	0.979 [0.828; 1.156]	0.79	0.12	0.833 [0.715; 0.969]	0.01	0.11	-0.005 [-0.013; 0.003]	0.23	0.24
rs4667591	POLY & EUR	1.003 [0.883; 1.139]	0.95	0.10	0.828 [0.732; 0.936]	0.002	0.39	-0.002 [-0.008; 0.004]	0.61	0.54
	POLY	0.880 [0.751; 1.032]	0.11	0.92	0.794 [0.685; 0.919]	0.002	0.38	0.001 [-0.007; 0.008]	0.86	0.65
rs41268685	POLY & EUR	-	-	-	-	-	-	0.036 [-0.067; 0.139]	0.50	0.02
	POLY	-	-	-	-	-	-	0.083 [-0.117; 0.283]	0.42	0.02

POLY: NZ Polynesian, EUR: NZ European, NU: Normouricaemic, HU: Hyperuricaemic, rs ID: Reference SNP cluster ID, var: Variant, OR/ß: Odds ratio/beta or effect estimates for the alternate allele, 95% CI: 95% confidence interval, *P*; *p*-values, $P_{OR/ß}$: p-value for odds ratio/beta estimates, Het: Heterogeneity. All values are adjusted for age, sex and BMI for NZ European plus with the estimates of grand-parental ancestry for NZ Polynesian.

SECTION 5.3 DISCUSSION

The findings of this study are four fold. First, the characterisation of exonic variants within *LRP2* and *A1CF* identified 11 Polynesian-specific variants in total, with 7 variants having an allele frequency > 0.05 (common) for both genes (Table 5.5). The Discovery-phase association analyses revealed the presence of common, low-frequency and rare variants within both genes, carrying both risk and protection for hyperuricaemia and gout (Table 5.5 and Table 5.6). Two Polynesian-specific variants showing protection against hyperuricaemia and gout were also identified in this analysis (*rs111360923* and var170115626) (Table 5.5 and Table 5.6). Second, calculating the burden for rare and non-synonymous variants within *LRP2* and *A1CF* indicated a burden of hyperuricaemia risk variants in European individuals within *LRP2* (Table 5.8). Third, assigning functional annotations to exonic variants via four different annotation programs showed that at least 53% of the variants within coding regions of *LRP2* could have damaging (functional) effects on the proteins (Table 5.8). Overall, the Discovery-phase results from *LRP2* came out to be more interesting than *A1CF*, with a number of variants showing an association with hyperuricaemia and gout in Europeans and Polynesians (Table 5.5 and Table 5.6). Several variants were, therefore, selected from *LRP2* for validation of their association with hyperuricaemia and/or gout in a larger independent Replication Cohort, which represents the fourth finding of this study (detailed below).

Multiply-adjusted regression analysis of the seven prioritised variants from the Discovery-phase successfully replicated five variants for an association with either hyperuricaemia or gout (Table 5.12 to Table 5.14). The genotype and allele frequencies in the Replication-phase were similar to that observed in the Discovery-phase for all replicated variants (Table 5.12). Two Polynesian-specific variants, *rs111360923* and var170115626 were replicated for an association with gout (*rs111360923*; OR = 0.64, P_{OR} = 0.002, var170115626 = OR = 0.73, P_{OR} = 0.03) but not hyperuricaemia (Table 5.14), while var170115626 indicated a trend of negative association with urate in otherwise healthy (control) Polynesian individuals (ß (mmol L^{-1}) = -0.012, P = 0.06) (Table 5.14). The similar results and population specificity for these two variants were validated via LD that indicated both of these variants to be in strong LD (r^2 = 83%) (Figure 5.8). Similar but very low LD (5%) with the previously reported variant at *LRP2: rs2544390* that showed Polynesian-specific association with gout (Rasheed *et al.*, 2013a)

further indicated that association signals from these two variants in my study are independent of *rs2544390*. Both of these variants have been reported as having extremely low frequencies in the ExAC (The Exome Aggregation Consortium) database (http://exac.broadinstitute.org/) for other populations. However, no homozygotes have been reported for either variant. The C allele frequency reported for *rs111360923* was 8.9E-04 (108 out of 121,016 alleles), predominantly reported in Africans (C allele frequency = 0.008) and South Asians (C allele frequency = 0.001). The C allele frequency for the var170115626 was 6.05E-05 (1 out of 165,08 alleles), reported only in South Asian population in ExAC. However, this variant was not reported for any other population in ExAC (total allele record: 121,232). Moreover, there is no previous report available for both of these variants in the context of a pathologic condition (to my knowledge). The absence of these variants in other populations, their rarity in South Asians and frequent occurence in NZ Polynesian (*rs111360923* = 0.07, var170115626 = 0.07) is likely linked to the settlement processes in the Polynesian region. While people moving out of East Asia and speaking Austronesian languages (reaching Polynesia with the Lapita culture around 3000 years ago) seem to have contributed to a great extent to the genetic variation present in Polynesians, there was also admixture with the earliest inhabitants of the New Guinea and Solomon Islands en route. Most of the Y chromosome present in Polynesian populations can be traced back to original inhabitants of the Sahul continent (which comprised what is now Australia and New Guinea) rather than the East Asian ancestors, which indicates some significant admixture (Kayser, 2010). The first people to settle the Sahul continent probably arrived via South Asia as part of one of the earliest movements of people out of Africa (Hudjashov *et al.*, 2007; Kayser, 2010; O'Connell and Allen, 2004; Summerhayes *et al.*, 2010). This is likely where these variants originated and became higher in frequency in Polynesians as a result of founder effect, genetic drift and population expansion. Due to the scarcity of information available in the literature and no functional annotations assigned on any platform for *rs111360923* and due to the var170115626 being annotated as having a neutral effect on protein function in three (PolyPhen-2, SIFT and PROVEAN) out of four annotation programs, it is hard to elucidate an effect on protein function. However, the protective effect (for hyperuricaemia and gout) of *rs111360923* and var170115626 variants in the resequencing data and its successful replication through TaqMan genotyping (for gout only) represents a novel Polynesian-specific variant known to influence gout risk.

Another common missense variant that showed significant associations with hyperuricaemia and gout is *rs4667591*. The G allele was only nominally associated with hyperuricaemia in Europeans (OR = 0.61, *P* = 0.05) and gout in Polynesians (OR = 0.72, *P* = 0.04) with a protective effect in Discovery-phase analysis (Table 5.5 and Table 5.6). Following replication through TaqMan PCR genotyping, the G allele successfully replicated the strong protective association with hyperuricaemia in the combined NZ Polynesian group (OR = 0.79, *P* = 0.002) but not NZ Europeans (OR = 0.91, *P* = 0.43) (Table 5.12 and Table 5.14). However, the Replication-phase revealed that addition of each G allele increased the risk of gout in NZ Europeans by 26% (OR = 1.26, *P* = 0.02) (Table 5.12). The variant has been reported to be common in other populations in Genome Reference Consortium human genome build 37 (GRCh37) on Ensembl (http://grch37.ensembl.org/) with the G allele frequency of 0.49 for East Asian and 0.79 for European (consistent with this study) populations. The frequency of the G allele was little different between the East and West Polynesians and Māori individuals from Ngati Porou Hauora (0.46, 0.44 and 0.48, respectively). The association of this variant with increased risk of gout in NZ European data set is consistent with the recent GWAS data from Pattaro *et al.* (2016) that reported the other (T) allele of *rs4667591* to be associated with increased eGFR (ß (mL/min/1.73m^2) = 0.003, *P* = 0.006) in the European population (Appendix B Figure 5.3). The fact that a decrease in eGFR has been associated with an increase in urate concentration (Johnson *et al.*, 2013a; Mohandas and Johnson, 2008; Suliman *et al.*, 2006) and higher prevalence of gout (Krishnan, 2012), the allele T should be ideally associated with decreasing a risk of gout by decreasing urate concentrations. So, the G allele can be expected to have opposite effects on urate concentration and, thus, the risk of developing gout in Europeans.

Another common (nonsense/stop) variant within *LRP2* that indicated significant associations in this study is *rs2075252*. The analysis of sequencing data in the Discovery-phase showed a protective association of the C allele of *rs2075252* with hyperuricaemia in Europeans only (OR = 0.65, *P* = 0.05) (Table 5.5). However, genotyping in the larger Replication Cohort did not follow the same trend in NZ Europeans (OR = 0.97, *P* = 0.78) (Table 5.12) but the C allele indicated a protection against hyperuricaemia in the combined NZ Polynesian dataset (OR = 0.83, *P* = 0.01) (Table 5.14). Similar to the results for *rs4667591*, the data presented in large European GWAS from Pattaro *et al.* (2016) indicated a positive association of the opposite (T) allele of the *rs2075252* with

increased eGFR (ß (mL/min/1.73m^2) = 0.004, *P* = 4.5E-06) (Appendix B Figure 5.3). The protective effect of C alleles of both of these variants (*rs4667591* and *rs2075252*) for hyperuricaemia in the NZ Polynesian dataset may be supported by ancestral differences between these populations. However, in the absence of literature reporting these two variants in the context of Polynesian population, it would be naive to comment on protective findings in my study.

Both *rs4667591* and *rs2075252* are in the same haplotype block with LD (r^2) values of 0.39 in European and 0.82 in Han Chinese populations (Table 5.11). In a recent GWAS including 2,640 European individuals from the FHS Offspring Cohort, the A alleles of *rs2075252* and *rs4667591* were found to be positively associated with urinary levels of Trefoil Factor-3 (*rs2075252*: ß = 0.22, *P* = 1.62E-16; *rs4667591*: ß = 0.19, *P* = 0.01) (McMahon *et al.*, 2014). Higher urinary levels of TFF3 have been recognised as indicators of renal tubular injury (Yu *et al.*, 2010), acute and/or chronic kidney disease (Lebherz-Eichinger *et al.*, 2015) and/or process of kidney damage repair (Astor *et al.*, 2011; Taupin and Podolsky, 2003). Given that *rs2075252* and *rs4667591* (the A alleles) have been positively associated with higher urinary TFF3 levels, these variants might be playing a role in determining the change in urate concentrations. In addition, *rs2075252* and the haplotype GA of *rs4667591* were found to be associated with variation in bone mineral density in Chinese females (Wang *et al.*, 2011b). Higher bone mineral density has been associated with higher urate levels and has been attributed to protect bone health in primary osteoporotic patients (Chen *et al.*, 2015b). Although this may present a connecting role of these two variants in determining bone health via a change in serum urate concentrations, an opposite relationship has also been reported between the two metabolic pathways (Zhang *et al.*, 2015). Although none of these two variants have been functionally characterised so far, *rs4667591* was annotated as having a damaging effect on protein function in PolyPhen-2 and SIFT with a high CADD score (31) (Appendix A Table 5.6), which may possibly be one of the explanations describing the role of this variant in influencing hyperuricaemia and gout in NZ Polynesian and European population.

The Replication-phase analysis successfully replicated *rs2302694* for an association of the A allele with hyperuricaemia in NZ Polynesian (OR = 0.79, *P* = 0.04), showing a trend of protective association with gout (OR = 0.80, *P* = 0.06) and negative

association with serum urate concentrations (ß (mmol L^{-1}) = -0.015, P = 0.01) (Table 5.14). The same allele A was also reported to be negatively associated with increased eGFR in European individuals (ß (mL/min/1.73m^2) = -0.004, P = 0.005) in a recent GWAS (Pattaro *et al.*, 2016) (Appendix B Figure 5.4). However, the A allele did not show any association with either hyperuricaemia (OR = 1.14, P = 0.66) or serum urate concentrations (ß (mmol L^{-1}) = -0.007, P = 0.30) in the NZ European population (Table 5.12 and Table 5.13). One reason that can expalin this difference could be the small sample size in my study compared to the Pattaro *et al.* (2016) GWAS. In addition, the individuals included in the Replication Cohort were not from the extreme spectrum of serum urate (Table 5.3), which may have caused the insignificant outcome in NZ European dataset.

The study has also found a significant burden of rare non-synonymous variants within *LRP2* that increases the risk of developing hyperuricaemia in Europeans ($P_{Burden\ (Risk)}$ = 0.009) (Table 5.8). The allele collapsing method revealed a significantly higher (P = 0.04) frequency of the altered allele in hyperuricaemic cases (allele frequency = 0.28) compared to normouricaemic controls (allele frequency = 0.18) in the European population (Table 5.8). Previous research indicated that more common variants within *LRP2* is inconsistently associated with serum urate and gout in other populations (Kamatani *et al.*, 2010; Nakayama *et al.*, 2014; Rasheed *et al.*, 2013a) but no such common variant risk associations have been replicated in Europeans. Only a GWAS including data from European individuals (FHS Offspring Cohort: n = 2,640) found common variants within *LRP2* to be associated with urinary biomarker TFF3. An exon sequencing follow-up burden analysis in their study also found that rare SNPs within *LRP2* together explain 3.1% of the variance ($P = 2.9\times10^{-4}$) in TFF3 levels (McMahon *et al.*, 2014). The findings of the present study are consistent with the fact that rare (and non-synonymous) variant clusters within *LRP2* are involved in determining the risk of various kidney-related conditions in European individuals. This study presents the first evidence of rare variants within *LRP2* conferring a risk of hyperuricaemia in the NZ European population.

5.3.1 Strengths and Limitations of the Study

The exon sequencing of *LRP2* and *A1CF* and their follow-up replication indicated some interesting results for two populations under study. However, the study also has strengths and limitations presented below as a summary.

The ReSequencing Cohort was only adequately powered to detect a moderate or weaker effect for both European and Polynesian populations (Figure 5.2). This means that results from the Discovery-phase analysis could have small study biases in terms of having false positive associations and elevated effect sizes. Such underpowered studies have low reproducibility and replication success rate (Button *et al.*, 2013). The underlying genetic architecture of any trait determines the best-powered design to detect an association (Laird and Lange, 2006). For complex traits like gout and hyperuricameia, where the genetic architecture is partially known, it is not possible to predict *a priori* the most powerful study design (Auer and Lettre, 2015). However, the study had several other complementary factors which could overcome the low detection power to some extent. First, the phenotypes assessed were extreme or binary i.e., hyperuricaemic vs normouricaemic and gout case vs control. The binary characterisation is described to be better than having a continuous phenotype (e.g., serum urate) while carrying out an analysis of sequencing data, especially rare variant analysis (Auer and Lettre, 2015; Button *et al.*, 2013; Do *et al.*, 2012). Second, a burden analysis approach was used to combine the effect of all rare and non-synonymous variants within *LRP2* and *A1CF*. This analysis is specialised to calculate the effect of all variants together considering them as one unit and is more powerful than testing the effect of each variant separately (Lee *et al.*, 2014).

Follow-up genotyping of the interesting results from the Discovery-phase represents the major strength of this study. It is essential to replicate the outcomes of a whole genome/exome sequencing analysis via one of several follow-up replication strategies (Auer and Lettre, 2015) to confirm the true and false associations. The study Replication Cohort was designed by choosing larger and well-characterised data sets from two populations. Also, the Replication-phase analysis was able to reproduce similar associations for at least one of the two phenotypes (hyperuricaemia and gout) with a total success rate of 75%. However, there were some findings from Discovery-analysis that

were not replicable and/or were not significant in the Replication-phase. One potential limitation could be the average serum urate levels in hyperuricaemic group in the Replication Cohort were only marginally hyperuricaemic and were not from extreme ends of the urate spectrum as for the Discovery Cohort. Use of individuals with average urate levels matching the Discovery Cohort may provide better replication results, but such individuals were not available.

Another strength of the study is the selection of two different approaches to calculate a burden of rare and non-synonymous variants. The selection, however, could also be a likely limitation of the study. Falling within five major categories, many different burden testing designs are now available (Lee *et al.*, 2014). The burden testing approach used in this study was based on the simplicity of computing and the ability of the weighted sum statistic test to calculate the burden in two (risk and protective) directions. Other approaches (e.g., SKAT) may or may not be more efficient in picking up rare variants than the present approach. Additionally, no other methodolodgy is yet available to test the burden for risk and protective variants separately.

5.3.2 Conclusion and future directions

The work described in this section provides a detailed characterisation of exonic regions of two genes, *LRP2* and *A1CF*. Overall in both genes, 12 common, low frequency and rare variants were detected to be associated with hyperuricaemia in Europeans and hyperuricaemia or gout or both in NZ Polynesians. Two non-synonymous Polynesian-specific variants (i.e., alleles) have been identified within the coding region of the gene *LRP2* showing protection for gout. The confirmation of this association through follow-up genotyping in a larger cohort presents a novel finding of this study exploring population-specific variants in NZ Polynesian individuals indicating a real effect associated with gout. The study for the first-time reports rare non-synonymous variants within *LRP2* to carry a significant risk burden for hyperuricaemia in European individuals. The study, however, was unable to find any potentially interesting associations for the gene *A1CF* in the context of hyperuricaemia and gout for European and Polynesian populations. A few inconsistent results in the discovery-phase analysis for an association with hyperuricaemia and gout in NZ Polynesian population may indicate an independent involvement of these variants in either urate transport or the inflammatory pathway. The

failure of replication of sequencing associations could be a result of the small sample size (low power) of the study (Button *et al.*, 2013). Increasing the sample size of Discovery and Replication Cohorts and broadening by including hyperuricaemic individuals with urate levels at extreme ends of the hyperuricaemia spectrum would be helpful in explaining the associations that were not replicable in my study. The replicated non-synonymous variants can be further tested for their functional consequences using gene cloning and expression approaches in human/animal cell lines. Using such in vitro techniques, the effect of variants on the expression of the respective gene, mRNA expression and protein activity can be tested. Following the expression assays, the proteins encoded by these genes (especially *LRP2*) can be purified via chromatography to further test for their binding to appropriate ligands (e.g., LDL for LRP2), which would be beneficial in explaining the role of these genes in the underpinning pathway between lipid metabolism and gout aetiology.

Chapter 6

Conclusions

SECTION 6.1: CONCLUSIONS

6.1.1 Summary

The need to disentangle the intricate pathophysiology of gout in a country like New Zealand with an isolated geographical location and markedly high prevalence of this complex phenotype is discernible. My study was an attempt to shed light on the potential relationship of gout with other coexisting metabolic conditions with a focus of "which is driving which". Through the implementation of cross-disciplinary research methods, I have been able to explore and report as novel findings the genetic/causal associations of urate, hyperuricaemia and gout with a set of several metabolic conditions that has more often than not been ignored in previous studies of the kind.

The first two research chapters (Chapter 2 and Chapter 3) of this thesis aimed out to explore a possible causal relationship between gout and imbalanced iron metabolism as its co-existing complication. Chapter 2 reports a successful replication of a previously reported (Ghio *et al.*, 2005) positive association of urate with serum levels of ferritin (iron storage protein that directly reflects total body iron) and transferrin saturation in Europeans and/or African Americans using the data in the NHANES III and the JHS cohorts (Table 2.5). The ferritin-urate associations were extended from Europeans and African Americans to New Zealand Polynesian males (P = 2.53E-04) (Table 2.5). To ensure that the observed positive associations were not aggravated by an inflammatory condition e.g., hyperuricaemia and/or gout, the robustness of association was validated via adjusting for serum levels of CRP as an indicatory marker for acute inflammation. This study for the first time reported an association of increased serum ferritin with the risk for gout (OR = 1.03, P = 1.76E-03) and frequency of gout flares (ß (flares/year) = 0.09, P = 0.04) in the New Zealand Polynesian and European male populations, respectively (Table 2.7 and 2.8). Results of this study also showed that the levels of ferritin were significantly higher in New Zealand Polynesian gout cases (P = 2.29E-04) compared to non-gout controls (Appendix A Table 2.4).

The observational study in Chapter 2 was primarily conducted to elaborate a biochemical association and was not aimed to elucidate the exact mechanism of iron-urate relationship. However, the results of this study along with the data reported in the previous literature suggested a possible role of change in blood levels of ferritin in

increasing the risk of gout via hyperuricaemia, especially in New Zealand Māori and Pacific Island individuals. Evidence reporting a possible relationship of iron and urate are not recent (Green and Mazur, 1957; Mazur *et al.*, 1958; Muirden and Senator, 1968). Urate being an antioxidant in humans has been reported as reducing the oxidative stress of free floating iron by acting as a metal chelator (Aust *et al.*, 1985; Davies *et al.*, 1986). Increased exposure to iron has been shown to upregulate the activity of xanthine oxidase resulting in the increased production of urate (Ghio *et al.*, 2002; Martelin *et al.*, 2002; Stonehuerner *et al.*, 1998). Additionally, the presence of iron deposits in the synovial membrane in people with rheumatoid arthritis (Muirden and Senator, 1968), and the reduction in the severity of gout flares following the depletion of the metal via phlebotomy are supportive of my observational findings in Chapter 2 (Facchini, 2003). My study also reports a preliminary finding of increased serum ferritin in New Zealand Polynesians ($P = 0.04$) who self-reported the consumption of iron-rich food (seafood, fish and red meat) as a trigger for gout (Table 2.10). This finding, along with the results from previous research conducted in the Merriman Laboratory (Flynn *et al.*, 2015), further provoked the idea for a possible role of iron-rich food in increasing, both the frequency and severity of gout flares.

Albeit the outcomes of Chapter 2 presented an interesting theme of the involvement of iron metabolism in gout aetiology via hyperuricaemia, although the results were still prone to biases due to the presence of possible confounders (Mann and Wood, 2012). To confirm the observational data an initial genetic association analysis was done using two iron-related variants (*rs1799945* and *rs1800562*) within the *HFE* (haemochromatosis) gene. In the adjusted analyses (adjusted for age, body mass index/BMI and estimates of grand-parental ancestry in Polynesians), the G allele of *rs1799945* was found to be positively associated with serum urate in the New Zealand population sub-groups and with a 28% increase in the risk of developing gout in New Zealand European females (OR = 2.28, $P = 0.03$) (Table 2.12). These results not only provided support to my observational data but also a logical base to design the research in Chapter 3.

The research in Chapter 3 was designed to validate the observational findings in Chapter 2 using the robust ‘Two-sample Mendelian randomisation’ approach to infer the causal association between iron and urate metabolism/metabolic pathways in the larger

population datasets of European descent. My study, for the first time, exploited summary statistics data from two large GWAS (Benyamin *et al.*, 2014; Köttgen *et al.*, 2013) to investigate the iron-urate causality in a population. Mendelian randomisation is a statistical approach to infer a cause-effect relationship via using the appropriate genetic instrument(s) (Smith and Ebrahim, 2003). The analysis provided an evidence for a causal role for serum iron (ß (SD) = 0.11, *P* = 1.96E-04) and ferritin (ß (SD) = 0.14, *P* = 0.03) to raise serum urate levels in European individuals (Table 3.4). Given that the study included effect estimate data from > 240,000 Europeans, it was presumed to have high power to detect a weaker effect. Since the outcomes of Mendelian randomisation studies require selection of instruments with 'no pleiotropy', a thorough search via literature review and the use of bioinformatics tools/platforms was carried out to rule out this possibility of any such pleiotropic effects. The results were reported only after removing all possible ambiguities (to my knowledge). However, it is impossible to remove the effect from any unknown confounders and such studies thus merit further exploration. Overall, result from my study in Chapter 3 provides evidence of a causal role of exposure iron and ferritin in increasing urate concentration as an outcome. However, my study does not report a causal effect of exposure of urate to influence iron homeostasis, which in turn, suggested that an increase in blood levels iron (or ferritin) is causative for an increase in urate concentration as an outcome.

The prevalence of gout is high in the New Zealand population (~7% in Polynesians and ~3% in Europeans) (Winnard *et al.*, 2012) and so is the prevalence of its comorbidities (detailed in Chapter 1). Most of these comorbidities (IR, T2D, hypertension and CVDs) are collectively referred to as components of or 'metabolic syndrome (MetS)' per se. Seldom has MetS been studied for its genetic relationship with gout and hyperuricaemia with most of the studies reporting only its prevalence in the New Zealand population. The research in Chapter 4 was, therefore, designed to explore the shared genetic basis between gout and the components of MetS in the New Zealand based ancestral groups. TaqMan genotyping was used to collect genotype information for > 4,600 New Zealanders. The data from large publicly-available European populations were also sourced from the UK Biobank, ARIC, FHS, CHS and CARDIA cohorts. The variants that have primarily been reported for their association with the components of MetS and/or serum urate in other populations were selected including *ADRB3: rs4994, MC3R: rs3827103, MC4R: rs17700633* and *rs17782313* and *ADTRP: rs6903956*. The

genotype association analyses were adjusted for baseline confounders (age, sex, BMI and estimates of grand-parental ancestry for New Zealand Polynesian datasets) plus with other comorbid conditions (hypertension, renal dysfunction, T2D and dyslipidaemia) of hyperuricaemia and gout. The analysis successfully replicated the previously reported positive associations (Huang *et al.*, 2013; Morcillo *et al.*, 2010; Rho *et al.*, 2007; Wang *et al.*, 2011a) of the insulin resistance-related G allele of *rs4994* with increased serum urate in the New Zealand Polynesians (WP: β (mmol L^{-1}) = 0.036, P = 0.004) (Table 4.8). The analysis further revealed a possible pleiotropic effect of the G allele in Western Polynesian sub-group, showing a protective effect against gout (OR = 0.62, $P_{Unadjusted}$ = 0.04) (Table 4.6). The protective association of the G allele with gout was also observed in the much larger European dataset from the UK Biobank Cohort (OR = 0.88, P = 0.04) (Table 4.6) and was consistent with my previously published data for the G allele of *ADRB3: rs4994* in Western Polynesian (described above) (Fatima *et al.*, 2016) and suggests a possible positive relationship of insulin-resistance and hyperuricaemia.

The research in Chapter 4 also reports interesting results for variants within/near the obesity-related melanocortin receptor genes (*MC3R* and *MC4R*) with gout in the New Zealand Polynesian and European populations. While the A allele of *MC3R: rs3827103* indicated a protective association with gout in the overall New Zealand population (OR = 0.92, P = 0.03) (Table 4.9), the C allele of *MC4R: rs17782313* was found to be associated with an increased risk for gout in New Zealand Europeans (OR = 1.03, P = 0.02) and the combined NZ European and Polynesian datasets (OR = 1.06, P = 0.03) (Table 6.9). The C allele of *rs17782313* was also associated with increased serum urate in the non-gout European plus Polynesian individuals (β (mmol L^{-1}) = 0.002, P = 0.006) (Table 4.9). Collectively these findings are consistent with the previously reported genetic/causal association of *MC3/4R* (for the same alleles of variants under study) with serum urate in Europeans (Köttgen *et al.*, 2013; Lyngdoh *et al.*, 2012; Palmer *et al.*, 2013). The C allele of *rs17782313* has been abundantly reported as a genuine signal for obesity and increased BMI in several populations (Beckers *et al.*, 2011; Chambers *et al.*, 2008; Loos *et al.*, 2008; Zobel *et al.*, 2009). Given that the melanocortin system is involved in the anti-inflammatory response (via *MC3R*) and in food uptake regulation (via *MC4R*) (Getting *et al.*, 2002; Huszar *et al.*, 1997; Lu *et al.*, 1994), my findings in Chapter 4 not only depicts obesity and insulin resistance to have causal relationship with gout but also are suggestive of an unknown shared metabolic pathway between gout and MetS. Put together, these

findings hence constructed a bridge between gout and MetS via the genetic association analyses in New Zealand-based populations.

The research in Chapter 5 dealt with 'dyslipidaemia' as the third and last comorbid condition for gout among the comorbidities included in my PhD project via replication-based rare variant association analysis approach. Two lipid-related genes, *LRP2* and *A1CF*, were assessed for their association with hyperuricaemia and/or gout in the European and NZ Polynesian individuals using exon sequencing data for these particular genes from the information available in the ReSequencing Cohort. Exploring the genetic association of non-urate transporter genes with hyperuricaemia and gout was the major purpose of this research. The results of the Discovery-phase showed several novel population-specific association signals for gout and hyperuricaemia from common (via single variant analyses) and rare (via non-synonymous burden) variants within the coding regions of the *LRP2* gene (Table 5.5 and Table 5.6). The C alleles of two Polynesian-specific variants, *rs111360923* and var170115626, were successfully replicated for an association with gout (*rs111360923*; OR = 0.64, *P* = 0.002, var170115626 = OR = 0.73, *P* = 0.03) when validated in a larger Replication Cohort that included the data from > 3,300 New Zealand-based Polynesian and European individuals (Table 5.12 to Table 5.14). Owing to the presence of similar allele frequencies of *rs111360923* (MAF = 0.07) and var170115626 (MAF = 0.07) LD was calculated, which indicated a strong LD (r^2 = 83%) between *rs111360923* and var170115626 in the New Zealand Polynesian population (Figure 5.8). The association signals of these two variants (the C alleles) were, however, found to be independent of the T allele of *LRP2: rs2544390* (r^2 = 5%) that was previously reported for its Polynesian-specific association with gout (Rasheed *et al.*, 2013a). Both of these variants have been reported as having extremely low frequencies in other populations (*rs111360923* MAF = for Africans (0.008) and South Asians (0.001); var170115626 MAF = 6.05E-05 only in South Asians). A thorough literature search also revealed that these two variants have never been described in relation to any pathophysiologic condition, which further ascribed the novelty of my findings in Chapter 5. The allelic associations of three more common variants, *rs4667591*, *rs2075252* and *rs2302694*, were also replicated for gout and/or hyperuricaemia in the New Zealand European and/or Polynesian ancestral groups. The G allele of *rs4667591* was significantly associated with hyperuricaemia, albeit in opposite direction, in the New Zealand Polynesian (OR = 0.79, *P* = 0.002) and European (OR =

1.26, P = 0.02) groups (Table 5.14). The C allele of *rs2075252* and the A allele of *rs2302694* were found to have a protective effect for hyperuricaemia in the New Zealand Polynesians (*rs2075252*: OR = 0.83, P = 0.01; *rs2302694:* OR = 0.79, P = 0.04) (Table 5.14). Further to these, the rare non-synonymous variants within the *LRP2* gene showed a significant burden risk of developing hyperuricaemia in European individuals ($P_{\text{Burden (Risk)}}$ = 0.009) (Table 5.8). The combined rare variant allele frequency was significantly higher (P = 0.04) in hyperuricaemic cases compared to controls (Table 5.8). While the previous literature provided evidence for the involvement of more common variants within the *LRP2* gene in increasing the risk of hyperuricaemia and/or gout in other populations (Kamatani *et al.*, 2010; Nakayama *et al.*, 2014; Rasheed *et al.*, 2013a), my findings in Chapter 5 are suggestive for a protective role of common variants within the *LRP2* gene for hyperuricaemia. The replicated protective findings of *LRP2* in my study are of interest and are not supportive for a role of *LRP2* in increasing the risk of hyperuricaemia and/or gout in Polynesian individuals. It is also notable that the sample size for the Discovery-phase indicated a lower power and the individuals in the Replication Cohort did not have their urate levels from extreme spectrum of hyperuricaemia. Admitting that these limitations could have caused a difference in the direction of association for *LRP2* variants, the protective findings in my study merits further exploration e.g., use of larger data set with hyperuricaemic individuals having higher average urate levels than the Replication Cohort.

6.1.2 Study limitations

The limitations of the work presented in this thesis are detailed below;

1. The observational data provided in Chapter 2 suggested a positive association of imbalanced iron homeostasis with urate and gout with an initial suggestion of an involvement of iron-rich food in gout pathophysiology. However, it was not possible to further investigate the relationship between iron-rich diet and gout and frequency of gout flares due to the small sample size used.
2. Mendelian randomisation analyses in Chapter 3 represent three major limitations of my study. First, the causal effect estimates described by the single variant as well as the sensitivity analyses were of limited magnitude and therefore may not retain causality if summary data encompassing a wider range of iron biomarkers

were used. Second, the causal role of iron and ferritin in increasing serum urate was confirmed in Europeans only due to the unavailability of GWAS data for Polynesian individuals. These findings, therefore, cannot guarantee the existence of a similar causal relationship between iron and urate in the New Zealand Polynesian individuals. And finally, the causality was described for ferritin/iron-urate only, again, due to the absence of GWAS data for gout-related loci, which means the causality may or may not retain significance if measured for an iron-gout relationship.

3. The associations presented in Chapter 4 were collectively assessed in large New Zealand Polynesian and European datasets. However, the data presented for individual Polynesian sub-groups could possibly have small-study biases.
4. The rare variant association analyses using sequencing data resulted in several positive significant associations that were not replicable in the larger cohort. The failure of replication presents a limitation of these analyses i.e., the Discovery-Cohort was only adequately powered to detect an intermediate effect (Figure 5.4) due to its small size and may have overrepresented some false-positive associations.

6.1.3 Future directions

The limitations detailed above need to be addressed appropriately to further confirm the false positive association(s) in my study, if any. A few suggestions to extend this research are detailed below;

1. The observational associations between ferritin and gout should be replicated in a larger data set from New Zealand individuals. In addition, detailed information about the involvement of iron-rich diet in triggering gout flares and the frequency of flares should be gained to elaborate the aspects of and further extend the research presented in Chapter 2.
2. A GWAS study detecting gout-related loci in the New Zealand European and Polynesian populations will facilitate analysing and validating of the iron/ferritin-urate causal associations via Mendelian randomisation. Also, the data covering a wider range of serum iron biomarkers and urate would be beneficial to assess if the causality still exists.

3. The recruitment of a larger New Zealand cohort to replicate the association in Chapter 4 and the Discovery-phase findings in resequencing analyses (Chapter 5) will be beneficial. In addition to increasing the sample size, choosing individuals from extreme ranges of hyperuricaemia would be desirable to investigate as to why replication of some variants failed.

6.1.4 Concluding remarks

The initially stated overarching aim of this research was to identify the potential causal relationship between gout and its comorbidities in the New Zealand population. While recognising the limitations, I believe that my study represents an original contribution to fill in the substantial gap in literature explaining the cause-effect relationship between gout, iron homeostasis, metabolic syndrome and dyslipidaemia. My study for the first time reports the association of ferritin with gout and the frequency of gout flares in the New Zealand individuals. It is also the first to describe a causal contribution of iron homeostasis in hyperuricaemia and gout pathophysiology in Europeans. My study is one of a few that addresses several components of the metabolic syndrome simultaneously while explaining their genetic relationship in gout. Its is also so far the only study reporting two Polynesian-specific common variants within the *LRP2* gene that influence any pathophysiological condition per se, hyperuricaemia and gout in this case. Finally, the study is the only one of its kind to report rare variants within the coding regions of *LRP2* to carry a significant burden of risk of hyperuricaemia in the European population.

REFERENCES

Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., and McVean, G.A. (2012). 1000 Genomes Project Consortium: An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56-65.

Acheson, R., and Florey, C. (1969). Body-weight, ABO blood-groups, and altitude of domicile as determinants of serum-uric-acid in military recruits in four countries. The Lancet *294*, 391-395.

Adams, P.C., Reboussin, D.M., Barton, J.C., McLaren, C.E., Eckfeldt, J.H., McLaren, G.D., Dawkins, F.W., Acton, R.T., Harris, E.L., and Gordeuk, V.R. (2005). Hemochromatosis and iron-overload screening in a racially diverse population. New England Journal of Medicine *352*, 1769-1778.

Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. Nature Methods *7*, 248-249.

AHA (2005). Metabolic syndrome: New guidance for prevention and treatment. Retrieved March 5, 2009.

Ahles, A., and Engelhardt, S. (2014). Polymorphic variants of adrenoceptors: pharmacology, physiology, and role in disease. Pharmacological reviews *66*, 598-637.

Akira, S., and Takeda, K. (2004). Toll-like receptor signalling. Nature Reviews Immunology *4*, 499-511.

Alberti, K.G.M., Zimmet, P., and Shaw, J. (2005). The metabolic syndrome-a new worldwide definition. The Lancet *366*, 1059.

Allison, D., Heo, M., Faith, M., and Pietrobelli, A. (1998). Meta-analysis of the association of the Trp64Arg polymorphism in the β 3 adrenergic receptor with body mass index. International Journal of Obesity & Related Metabolic Disorders *22*.

Alsmadi, O., Melhem, M., Hebbar, P., Thareja, G., John, S.E., Alkayal, F., Behbehani, K., and Thanaraj, T.A. (2014). Leptin in association with common variants of MC3R mediates hypertension. American journal of hypertension, hpt285.

Altemtam, N., Russell, J., and El Nahas, M. (2011). A study of the natural history of diabetic kidney disease (DKD). Nephrology Dialysis Transplantation, gfr561.

Altès, A., Bach, V., Ruiz, A., Esteve, A., Remacha, A.F., Sardà, M.P., Felez, J., and Baiget, M. (2009). Does the SLC40A1 gene modify HFE-related haemochromatosis phenotypes? Annals of hematology *88*, 341.

Amaro, S., Laredo, C., Renú, A., Llull, L., Rudilosso, S., Obach, V., Urra, X., Planas, A.M., and Chamorro, Á. (2016). Uric Acid Therapy Prevents Early Ischemic Stroke Progression. Stroke *47*, 2874-2876.

Ames, B.N., Cathcart, R., Schwiers, E., and Hochstein, P. (1981). Uric acid provides an antioxidant defense in humans against oxidant-and radical-caused aging and cancer: a hypothesis. Proceedings of the National Academy of Sciences *78*, 6858-6862.

Anagnostopoulos, I., Zinzaras, E., Alexiou, I., Papathanasiou, A.A., Davas, E., Koutroumpas, A., Barouta, G., and Sakkas, L.I. (2010). The prevalence of rheumatic diseases in central Greece: a population survey. BMC musculoskeletal disorders *11*, 98.

Anderson, E.R., and Shah, Y.M. (2013). Iron homeostasis in the liver. Comprehensive Physiology.

Anderson, G.J., and Vulpe, C.D. (2009). Mammalian iron transport. Cellular and Molecular Life Sciences *66*, 3241-3261.

Andrews, F., Morris, C., Kondratowicz, G., and Blake, D. (1987). Effect of iron chelation on inflammatory joint disease. Annals of the rheumatic diseases *46*, 327-333.

Anzai, N., and Endou, H. (2011). Urate transporters: an evolving field. Paper presented at: Seminars in Nephrology (Elsevier).

Arosio, P., and Levi, S. (2010). Cytosolic and mitochondrial ferritins in the regulation of cellular iron homeostasis and oxidative damage. Biochimica et Biophysica Acta (BBA)-General Subjects *1800*, 783-792.

Aryal, N., and Wasti, S.P. (2016). The prevalence of metabolic syndrome in South Asia: a systematic review. International Journal of Diabetes in Developing Countries *36*, 255-262.

Astor, B.C., Köttgen, A., Hwang, S.-J., Bhavsar, N., Fox, C.S., and Coresh, J. (2011). Trefoil factor 3 predicts incident chronic kidney disease: a case-control study nested within the Atherosclerosis Risk in Communities (ARIC) study. American journal of nephrology *34*, 291-297.

Auer, P.L., and Lettre, G. (2015). Rare variant association studies: considerations, challenges and opportunities. Genome medicine *7*, 1.

Aust, S.D., Morehouse, L.A., and Thomas, C.E. (1985). Role of metals in oxygen radical reactions. Journal of free radicals in biology & medicine *1*, 3-25.

Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., del Angel, G., Levy- Moonshine, A., Jordan, T., Shakir, K., Roazen, D., and Thibault, J. (2013). From FastQ data to high- confidence variant calls: the genome analysis toolkit best practices pipeline. Current protocols in bioinformatics, 11.10. 11-11.10. 33.

Badley, E., and DesMeules, M. (2003). Arthritis in Canada: an ongoing challenge. Ottawa (Canada): Health Canada.

Bădulescu, M., Macovei, L., and Rezuş, E. (2013). Acute gout attack with normal serum uric acid levels. Revista medico-chirurgicala a Societatii de Medici si Naturalisti din Iasi *118*, 942-945.

Baker, J.F., and Ghio, A.J. (2009). Iron homoeostasis in rheumatic disease. Rheumatology *48*, 1339-1344.

Baker, J.F., Schumacher, H.R., and Krishnan, E. (2007). Serum uric acid level and risk for peripheral arterial disease: analysis of data from the multiple risk factor intervention trial. Angiology *58*, 450-457.

Baldwin, W., McRae, S., Marek, G., Wymer, D., Pannu, V., Baylis, C., Johnson, R.J., and Sautin, Y.Y. (2011). Hyperuricemia as a mediator of the proinflammatory endocrine imbalance in the adipose tissue in a murine model of the metabolic syndrome. Diabetes *60*, 1258-1269.

Barsh, G.S., Copenhaver, G.P., Gibson, G., and Williams, S.M. (2012). Guidelines for genome-wide association studies. PLoS Genetics *8*, e1002812.

Barton, J.C., Lee, P.L., West, C., and Bottomley, S.S. (2006). Iron overload and prolonged ingestion of iron supplements: Clinical features and mutation analysis of hemochromatosis- associated genes in four cases. American journal of hematology *81*, 760-767.

Basu, S., Yoffe, P., Hills, N., and Lustig, R.H. (2013). The relationship of sugar to population-level diabetes prevalence: an econometric analysis of repeated cross-sectional data. PloS One *8*, e57873.

Beaudoin, M., Lo, K.S., N'Diaye, A., Rivas, M.A., Dubé, M.-P., Laplante, N., Phillips, M.S., Rioux, J.D., Tardif, J.-C., and Lettre, G. (2012). Pooled DNA Re-Sequencing of 68 Myocardial Infaction Candidate Genes in French Canadians. Circulation: Cardiovascular Genetics, CIRCGENETICS. 112.963165.

Beckers, S., Zegers, D., de Freitas, F., Mertens, I.L., Van Gaal, L.F., and Van Hul, W. (2011). Association study of MC4R with complex obesity and replication of the rs17782313 association signal. Molecular genetics and metabolism *103*, 71-75.

Bedir, A., Topbas, M., Tanyeri, F., Alvur, M., and Arik, N. (2003). Leptin might be a regulator of serum uric acid concentrations in humans. Japanese heart journal *44*, 527-536.

Bellamy, N., Downie, W., and Buchanan, W. (1987). Observations on spontaneous improvement in patients with podagra: implications for therapeutic trials of non- steroidal anti- inflammatory drugs. British journal of clinical pharmacology *24*, 33-36.

Bellia, A., Giardina, E., Lauro, D., Tesauro, M., Di Fede, G., Cusumano, G., Federici, M., Rini, G., Novelli, G., and Lauro, R. (2009). "The Linosa Study": Epidemiological and heritability data of the metabolic syndrome in a Caucasian genetic isolate. Nutrition, Metabolism and Cardiovascular Diseases *19*, 455-461.

Beltrán-Sánchez, H., Harhay, M.O., Harhay, M.M., and McElligott, S. (2013). Prevalence and trends of metabolic syndrome in the adult US population, 1999–2010. Journal of the American College of Cardiology *62*, 697-703.

Benyamin, B., Tonu, E., Ried, J.S., Radhakrishnan, A., Vermeulen, S.H., Traglia, M., Gögele, M., Anderson, D., Broer, L., Podmore, C*., et al.* (2014). Novel loci affecting iron homeostasis and their effects in individuals at risk for hemochromatosis. Nature communications *5*, 4926.

Bezwoda, W., Bothwell, T., Charlton, R., Torrance, J., MacPhail, A., Derman, D., and Mayet, F. (1983). The relative dietary importance of haem and non-haem iron. South African Medical Journal *64*, 552-556.

Billiet, L., Doaty, S., Katz, J.D., and Velasquez, M.T. (2014). Review of hyperuricemia as new marker for metabolic syndrome. ISRN rheumatology *2014*.

Björn-Rasmussen, E., Hallberg, L., Isaksson, B., and Arvidsson, B. (1974). Food Iron Absorption in Man Applications of the Two-Pool Extrinsic Tag Method to Measure Heme and Nonheme Iron Absorption from the Whole Diet. Journal of Clinical Investigation *53*, 247.

Blake, D.R., Hall, N.D., Bacon, P., Dieppe, P., Halliwell, B., and Gutteridge, J. (1983). Effect of a specific iron chelating agent on animal models of inflammation. Annals of the rheumatic diseases *42*, 89-93.

Blocker, B. (2013). Beta-3 adrenergic receptor. Sciences *77*, 2051-2058.

Bobulescu, I.A., and Moe, O.W. (2012). Renal transport of uric acid: evolving concepts and uncertainties. Advances in chronic kidney disease *19*, 358-371.

Bodmer, W., and Bonilla, C. (2008). Common and rare variants in multifactorial susceptibility to common diseases. Nature Genetics *40*, 695-701.

Boffetta, P., Nordenvall, C., Nyrén, O., and Ye, W. (2009). A prospective study of gout and cancer. European Journal of Cancer Prevention *18*, 127-132.

Bolann, B.J., and Ulvik, R. (1987). Release of iron from ferritin by xanthine oxidase. Role of the superoxide radical. Biochemical Journal *243*, 55-59.

Bonnefond, A., Clément, N., Fawcett, K., Yengo, L., Vaillant, E., Guillaume, J.-L., Dechaume, A., Payne, F., Roussel, R., and Czernichow, S. (2012). Rare MTNR1B variants impairing melatonin receptor 1B function contribute to type 2 diabetes. Nature Genetics *44*, 297-301.

Borghi, C., Verardi, F.M., Pareo, I., Bentivenga, C., and Cicero, A.F. (2014). Hyperuricemia and cardiovascular disease risk. Expert review of cardiovascular therapy *12*, 1219-1225.

Bos, M.J., Koudstaal, P.J., Hofman, A., Witteman, J.C., and Breteler, M.M. (2006). Uric acid is a risk factor for myocardial infarction and stroke. Stroke *37*, 1503-1507.

Bowden, J., Davey Smith, G., Haycock, P.C., and Burgess, S. (2016). Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. Genetic epidemiology *40*, 304-314.

Bowden, J., Del Greco, M., Minelli, C., Davey Smith, G., Sheehan, N., and Thompson, J. (2017). A framework for the investigation of pleiotropy in two- sample summary data Mendelian randomization. Statistics in Medicine.

Bowden, J., Smith, G.D., and Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. International journal of epidemiology *44*, 512-525.

Brauer, G., and Prior, I. (1978). A prospective study of gout in New Zealand Maoris. Annals of the rheumatic diseases *37*, 466-472.

Bresgen, N., and Eckl, P.M. (2015). Oxidative stress and the homeodynamics of iron metabolism. Biomolecules *5*, 808-847.

Briggs, J., Levitt, M., and Abramson, R. (1977). Renal excretion of allantoin in rats: a micropuncture and clearance study. American Journal of Physiology-Renal Physiology *233*, F373-F381.

Brondino, C.D., Romão, M.J., Moura, I., and Moura, J.J. (2006). Molybdenum and tungsten enzymes: the xanthine oxidase family. Current opinion in chemical biology *10*, 109-114.

Bullen, C., and Beaglehole, R. (1997). Ethnic differences in coronary heart disease case fatality rates in Auckland. Australian and New Zealand journal of public health *21*, 688-693.

Burgess, S., Bowden, J., Fall, T., Ingelsson, E., and Thompson, S.G. (2017). Sensitivity analyses for robust causal inference from mendelian randomization analyses with multiple genetic variants. Epidemiology (Cambridge, Mass) *28*, 30-42.

Burgess, S., Butterworth, A., and Thompson, S.G. (2013). Mendelian randomization analysis with multiple genetic variants using summarized data. Genetic epidemiology *37*, 658-665.

Burgess, S., Davies, N.M., and Thompson, S.G. (2016). Bias due to participant overlap in two- sample Mendelian randomization. Genetic epidemiology *40*, 597-608.

Burgess, S., Dudbridge, F., and Thompson, S.G. (2015a). Combining information on multiple instrumental variables in Mendelian randomization: comparison of allele score and summarized data methods. Statistics in Medicine.

Burgess, S., Small, D.S., and Thompson, S.G. (2015b). A review of instrumental variable estimators for Mendelian randomization. Statistical Methods in Medical Research, 0962280215597579.

Burgess, S., and Thompson, S.G. (2013). Use of allele scores as instrumental variables for Mendelian randomization. International journal of epidemiology *42*, 1134-1144.

Burgess, S., and Thompson, S.G. (2015). Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. American journal of epidemiology *181*, 251-260.

Button, K.S., Ioannidis, J.P., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S., and Munafò, M.R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. Nature Reviews Neuroscience *14*, 365-376.

Cabezas, F., Lagos, J., Céspedes, C., Vio, C.P., Bronfman, M., and Marzolo, M.a.-P. (2011). Megalin/LRP2 expression is induced by peroxisome proliferator-activated receptor-alpha and-gamma: implications for PPARs' roles in renal function. PloS one *6*, e16794.

Cadzow, M., Merriman, T., and Dalbeth, N. (2016). Precision of Gout Definitions for Population-Based Genetic Studies: Analysis of the UK Biobank [abstract]. Arthritis Rheumatol *68*.

Camaschella, C. (2006). Diagnosis and treatment of non-HFE-haemochromatosis. Disorders of iron homeostasis, erythrocytes, erythropoesis Paris: ESH and Club du Globule Rouge et du Fer, 467-473.

Camaschella, C. (2015). Iron-deficiency anemia. N Engl J Med *2015*, 1832-1843.

Cameron, A.J., Magliano, D.J., Zimmet, P.Z., Welborn, T., and Shaw, J.E. (2007). The metabolic syndrome in Australia: prevalence using four definitions. Diabetes research and clinical practice *77*, 471-478.

Carmelli, D., Cardon, L.R., and Fabsitz, R. (1994). Clustering of hypertension, diabetes, and obesity in adult male twins: same genes or same environments? American journal of human genetics *55*, 566.

Carocho, M., and Ferreira, I.C. (2013). A review on antioxidants, prooxidants and related controversy: natural and synthetic compounds, screening and analysis methodologies and future perspectives. Food and Chemical Toxicology *51*, 15-25.

Caron, L., Rousseau, F., Gagnon, É., and Isenring, P. (2000). Cloning and functional characterization of a cation-Cl− cotransporter-interacting protein. Journal of Biological Chemistry *275*, 32027-32036.

Carpenter, C.E., and Mahoney, A.W. (1992). Contributions of heme and nonheme iron to human nutrition. Critical Reviews in Food Science & Nutrition *31*, 333-367.

Carpenter, M.A., Crow, R., Steffes, M., Rock, W., Heilbraun, J., Evans, G., Skelton, T., Jensen, R., and Sarpong, D. (2004). Laboratory, reading center, and coordinating center data management methods in the Jackson Heart Study. The American journal of the medical sciences *328*, 131-144.

CDC (2006). Recommendations to prevent and control iron deficiency in the United States. MMWR. 1998; 47 (RR-3): 1-36.

Chambers, J.C., Elliott, P., Zabaneh, D., Zhang, W., Li, Y., Froguel, P., Balding, D., Scott, J., and Kooner, J.S. (2008). Common genetic variation near MC4R is associated with waist circumference and insulin resistance. Nature Genetics *40*, 716-718.

Chan, J.C., Cho, N.H., Tajima, N., and Shaw, J. (2014). Diabetes in the western pacific region—past, present and future. Diabetes research and clinical practice *103*, 244-255.

Chang, H.-Y., Pan, W.-H., Yeh, W.-T., and Tsai, K.-S. (2001). Hyperuricemia and gout in Taiwan: results from the Nutritional and Health Survey in Taiwan (1993-96). The Journal of rheumatology *28*, 1640-1646.

Chang, X., Chin, H.-L., Quek, S.-C., Goh, D.Y., Dorajoo, R., Friedlander, Y., and Heng, C.-K. (2017). The genetic variation rs6903956 in the novel androgen-dependent tissue factor pathway inhibitor regulating protein (ADTRP) gene is not associated with levels of plasma coagulation factors in the Singaporean Chinese. Thrombosis Journal *15*, 1.

Chen, J.-H., Lan, J.-L., Cheng, C.-F., Liang, W.-M., Lin, H.-Y., Tsay, G.J., Yeh, W.-T., and Pan, W.-H. (2015a). Effect of urate-lowering therapy on all-cause and cardiovascular mortality in hyperuricemic patients without Gout: A case-matched cohort study. PloS One *10*, e0145193.

Chen, J.Ä., Chuang, S.Ä., Chen, H.Ä., Yeh, W.Ä., and Pan, W.Ä. (2009a). Serum uric acid level as an independent risk factor for all,Äêcause, cardiovascular, and ischemic stroke mortality: A chinese cohort study. Arthritis Care & Research *61*, 225-232.

Chen, J.H., Chuang, S.Y., Chen, H.J., Yeh, W.T., and Pan, W.H. (2009b). Serum uric acid level as an independent risk factor for all- cause, cardiovascular, and ischemic stroke mortality: A chinese cohort study. Arthritis Care & Research *61*, 225-232.

Chen, L., Peng, Y., Fang, F., Chen, J., Pan, L., and You, L. (2015b). Correlation of serum uric acid with bone mineral density and fragility fracture in patients with primary osteoporosis: a single-center retrospective study of 253 cases. International journal of clinical and experimental medicine *8*, 6291.

Chen, L.-H., Zhong, C., Xu, T., Xu, T., Peng, Y., Wang, A., Wang, J., Peng, H., Li, Q., and Ju, Z. (2016). Sex-specific Association Between Uric Acid and Outcomes After Acute Ischemic Stroke: A Prospective Study from CATIS Trial. Scientific Reports *6*.

Chen, L.-y., Zhu, W.-h., Chen, Z.-w., Dai, H.-l., Ren, J.-j., Chen, J.-h., Chen, L.-q., and Fang, L.-z. (2007). Relationship between hyperuricemia and metabolic syndrome. Journal of Zhejiang University Science B *8*, 593.

Chen, M.-Y., Zhao, C.-C., Li, T.-T., Zhu, Y., Yu, T.-P., Bao, Y.-Q., Li, L.-X., and Jia, W.-P. (2017). Serum uric acid levels are associated with obesity but not cardio-cerebrovascular events in Chinese inpatients with type 2 diabetes. Scientific Reports *7*.

Chen, S.-H., Habib, G., Yang, C.-Y., Gu, Z.-W., Lee, B.R., Weng, S.-A., Cai, S., Deslypere, J., and Rosseneu, M. (1987). Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in-frame stop codon. Science *238*, 363-366.

Chen, S.-J., Yen, C.-H., Huang, Y.-C., Lee, B.-J., Hsia, S., and Lin, P.-T. (2012). Relationships between inflammation, adiponectin, and oxidative stress in metabolic syndrome. PloS One *7*, e45693.

Chen, Y., Lin, Y.-C., Kuo, T.-W., and Knight, Z.A. (2015c). Sensory detection of food rapidly modulates arcuate feeding circuits. Cell *160*, 829-841.

Cherfane, C.E., Hollenbeck, R.D., Go, J., and Brown, K.E. (2013). Hereditary hemochromatosis: missed diagnosis or misdiagnosis? The American journal of medicine *126*, 1010-1015.

Chhana, A., Lee, G., and Dalbeth, N. (2015). Factors influencing the crystallization of monosodium urate: a systematic literature review. BMC musculoskeletal disorders *16*, 1.

Choi, H., De Vera, M., and Krishnan, E. (2008). Gout and the risk of type 2 diabetes among men with a high cardiovascular risk profile. Rheumatology *47*, 1567-1570.

Choi, H., and Ford, E. (2008). Haemoglobin A1c, fasting glucose, serum C-peptide and insulin resistance in relation to serum uric acid levels—the Third National Health and Nutrition Examination Survey. Rheumatology *47*, 713-717.

Choi, H.K., Atkinson, K., Karlson, E.W., Willett, W., and Curhan, G. (2004a). Purine-rich foods, dairy and protein intake, and the risk of gout in men. N Engl J Med *2004*, 1093-1103.

Choi, H.K., Atkinson, K., Karlson, E.W., Willett, W., and Curhan, G. (2004b). Purine-rich foods, dairy and protein intake, and the risk of gout in men. New England Journal of Medicine *350*, 1093-1103.

Choi, H.K., and Curhan, G. (2007a). Coffee, tea, and caffeine consumption and serum uric acid level: the third national health and nutrition examination survey. Arthritis Care & Research *57*, 816-821.

Choi, H.K., and Curhan, G. (2007b). Independent impact of gout on mortality and risk for coronary heart disease. Circulation *116*, 894-900.

Choi, H.K., and Ford, E.S. (2007). Prevalence of the metabolic syndrome in individuals with hyperuricemia. The American journal of medicine *120*, 442-447.

Choi, H.K., Ford, E.S., Li, C., and Curhan, G. (2007). Prevalence of the metabolic syndrome in patients with gout: the Third National Health and Nutrition Examination Survey. Arthritis Care & Research *57*, 109-115.

Choi, H.K., Liu, S., and Curhan, G. (2005a). Intake of purine- rich foods, protein, and dairy products and relationship to serum levels of uric acid: the Third National Health and Nutrition Examination Survey. Arthritis & Rheumatism *52*, 283-289.

Choi, H.K., Mount, D.B., and Reginato, A.M. (2005b). Pathogenesis of gout. Annals of internal medicine *143*, 499-516.

Choi, Y., Chan (2015). PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics *31*, 2745-2747.

Choi, Y.-J., Shin, H.-S., Choi, H.S., Park, J.-W., Jo, I., Oh, E.-S., Lee, K.-Y., Lee, B.-H., Johnson, R.J., and Kang, D.-H. (2014). Uric acid induces fat accumulation via generation of endoplasmic reticulum stress and SREBP-1c activation in hepatocytes. Laboratory investigation *94*, 1114-1125.

Chou, C., and Lai, J. (1998). The epidemiology of hyperuricaemia and gout in Taiwan aborigines. Rheumatology *37*, 258-262.

Christensen, E.I., and Birn, H. (2002). Megalin and cubilin: multifunctional endocytic receptors. Nature Reviews Molecular Cell Biology *3*, 258-268.

Cicerchi, C., Li, N., Kratzer, J., Garcia, G., Roncal-Jimenez, C.A., Tanabe, K., Hunter, B., Rivard, C.J., Sautin, Y.Y., and Gaucher, E.A. (2014). Uric acid-dependent inhibition of AMP kinase induces hepatic glucose production in diabetes and starvation: evolutionary implications of the uricase loss in hominids. The FASEB Journal *28*, 3339-3350.

Cieslak, J., Majewska, K., Tomaszewska, A., Skowronska, B., Fichna, P., and Switonski, M. (2013). Common polymorphism (81Val> Ile) and rare mutations (257Arg> Ser and 335Ile> Ser) of the MC3R gene in obese Polish children and adolescents. Molecular biology reports *40*, 6893-6898.

Cirulli, E.T., and Goldstein, D.B. (2010). Uncovering the roles of rare variants in common disease through whole-genome sequencing. Nature Reviews Genetics *11*, 415-425.

Clarson, L., Chandratre, P., Hider, S., Belcher, J., Heneghan, C., Roddy, E., and Mallen, C. (2013). Increased cardiovascular mortality associated with gout: a systematic review and meta-analysis. European journal of preventive cardiology, 2047487313514895.

Coelho, A., Antunes- Lopes, T., Gillespie, J., and Cruz, F. (2017). Beta- 3 adrenergic receptor is expressed in acetylcholine- containing nerve fibers of the human urinary bladder: An immunohistochemical study. Neurourology and Urodynamics.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences Lawrence Earlbaum Associates. Hillsdale, NJ, 20-26.

Colhoun, H.M., McKeigue, P.M., and Smith, G.D. (2003). Problems of reporting genetic associations with complex outcomes. The Lancet *361*, 865-872.

Collins, J. (2010). Kidney disease in Maori and Pacific people in New Zealand. Clinical nephrology *74*.

Collins, J., and Anderson, G. (2012). Intestinal iron absorption, 5th edn (New York: Elsevier).

Coppell, K.J., Mann, J.I., Williams, S.M., Jo, E., Drury, P.L., Miller, J.C., and Parnell, W.R. (2013). Prevalence of diagnosed and undiagnosed diabetes and prediabetes in New Zealand: findings from the 2008/09 Adult Nutrition Survey. The New Zealand Medical Journal (Online) *126*.

Corbin, L.J., Richmond, R.C., Wade, K.H., Burgess, S., Bowden, J., Smith, G.D., and Timpson, N.J. (2016). Body mass index as a modifiable risk factor for type 2 diabetes: Refining and understanding causal estimates using Mendelian randomisation. Diabetes, db160418.

Cortes, A., and Brown, M.A. (2011). Promise and pitfalls of the Immunochip. Arthritis research & therapy *13*, 1.

Courville, P., Chaloupka, R., and Cellier, M. (2006). Recent progress in structure-function analyses of Nramp proton-dependent metal-ion transporters This paper is one of a selection of papers published in this Special Issue, entitled CSBMCB-Membrane Proteins in Health and Disease. Biochemistry and Cell Biology *84*, 960-978.

Craig, J. (2008). Complex diseases: Research and applications. Nature Education *1*, 184.

Crosby, Peloso GM, Auer PL, Crosslin DR, and NO, S. (2014). TG and HDL Working Group of the Exome Sequencing Project, National Heart, Lung, and Blood Institute: Loss-of-function mutations in APOC3, triglycerides, and coronary disease. The New England journal of medicine *371*, 22.

Crownover, B.K., and Covey, C.J. (2013). Hereditary hemochromatosis. American family physician *87*.

Dabbagh, A., Blake, D., and Morris, C. (1992). Effect of iron complexes on adjuvant arthritis in rats. Annals of the rheumatic diseases *51*, 516-521.

Dalbeth, N., Chen, P., White, M., Gamble, G.D., Barratt-Boyes, C., Gow, P.J., and Orr-Walker, B. (2014). Impact of bariatric surgery on serum urate targets in people with morbid obesity and diabetes: a prospective longitudinal study. Annals of the rheumatic diseases *73*, 797-802.

Dalbeth, N., Merriman, T.R., and Stamp, L.K. (2016). Gout. The Lancet: Seminar, 1-14.

Dalbeth, N., Merriman, TR, Stamp, LK (2016). Gout. The Lancet: Seminar, 1-14.

Daoussis, D., Antonopoulos, I., and Andonopoulos, A.P. (2014). ACTH as a treatment for acute crystal-induced arthritis: Update on clinical evidence and mechanisms of action. Paper presented at: Seminars in Arthritis and Rheumatism (Elsevier).

Davies, K., Sevanian, A., Muakkassah-Kelly, S., and Hochstein, P. (1986). Uric acid-iron ion complexes. A new aspect of the antioxidant functions of uric acid. Biochemical Journal *235*, 747-754.

De Domenico, I., Ward, D.M., di Patti, M.C.B., Jeong, S.Y., David, S., Musci, G., and Kaplan, J. (2007). Ferroxidase activity is required for the stability of cell surface ferroportin in cells expressing GPI-ceruloplasmin. The EMBO journal *26*, 2823-2831.

De Luis, D., Gonzalez Sagrado, M., Aller, R., Izaola, O., and Conde, R. (2009). Influence of Trp64Arg polymorphism of beta 3-adrenoreceptor gene on insulin resistance, adipocytokines and weight loss secondary to two hypocaloric diets. Annals of Nutrition and Metabolism *54*, 104-110.

Dehghan, A., Köttgen, A., Yang, Q., Hwang, S.-J., Kao, W.L., Rivadeneira, F., Boerwinkle, E., Levy, D., Hofman, A., and Astor, B.C. (2008a). Association of three genetic loci with uric acid concentration and risk of gout: a genome-wide association study. The Lancet *372*, 1953-1961.

Dehghan, A., Van Hoek, M., Sijbrands, E.J., Hofman, A., and Witteman, J.C. (2008b). High serum uric acid as a novel risk factor for type 2 diabetes. Diabetes care *31*, 361-362.

del Greco, F., Foco, L., Pichler, I., Eller, P., Eller, K., Benyamin, B., Whitfield, J.B., Pramstaller, P.P., Thompson, J.R., and Pattaro, C. (2016). Serum iron level and kidney function: a Mendelian randomization study. Nephrology Dialysis Transplantation, gfw215.

DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., Del Angel, G., Rivas, M.A., and Hanna, M. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature Genetics *43*, 491-498.

Desai, R.V., Ahmed, M.I., Fonarow, G.C., Filippatos, G.S., White, M., Aban, I.B., Aronow, W.S., and Ahmed, A. (2010). Effect of serum insulin on the association between hyperuricemia and incident heart failure. The American journal of cardiology *106*, 1134-1138.

Després, J.-P., and Lemieux, I. (2006). Abdominal obesity and metabolic syndrome. Nature *444*, 881-887.

Didelez, V., and Sheehan, N. (2007). Mendelian randomization as an instrumental variable approach to causal inference. Statistical Methods in Medical Research *16*, 309-330.

Dinour, D., Gray, N.K., Campbell, S., Shu, X., Sawyer, L., Richardson, W., Rechavi, G., Amariglio, N., Ganon, L., and Sela, B.-A. (2010). Homozygous SLC2A9 mutations cause severe renal hypouricemia. Journal of the American Society of Nephrology *21*, 64-72.

Dixon, S.J., and Stockwell, B.R. (2014). The role of iron and reactive oxygen species in cell death. Nature chemical biology *10*, 9-17.

Dlouhy, A., and Outten, C. (2013). Iron Uptake, Trafficking and Storage. In Metallomics and the Cell Metal Ions in Life Sciences, L. In Banci, ed. (Springer), pp. 241–278.

Do, R., Kathiresan, S., and Abecasis, G.R. (2012). Exome sequencing and complex disease: practical aspects of rare variant association studies. Human molecular genetics *21*, R1-R9.

Dong, Z., Zhao, D., Yang, C., Zhou, J., Qian, Q., Ma, Y., He, H., Ji, H., Yang, Y., and Wang, X. (2015). Common Variants in LRP2 and COMT Genes Affect the Susceptibility of Gout in a Chinese Population. PloS One *10*, e0131302.

Dores, R.M. (2009). Adrenocorticotropic Hormone, Melanocyte- stimulating Hormone, and the Melanocortin Receptors: Revisiting the Work of Robert Schwyzer. Annals of the New York Academy of Sciences *1163*, 93-100.

Doria, A., and Krolewski, A.S. (2011). Diabetes: Lowering serum uric acid levels to prevent kidney failure. Nature Reviews Nephrology *7*, 495-496.

Döring, A., Gieger, C., Mehta, D., Gohlke, H., Prokisch, H., Coassin, S., Fischer, G., Henke, K., Klopp, N., and Kronenberg, F. (2008). SLC2A9 influences uric acid concentrations with pronounced sex-specific effects. Nature Genetics *40*, 430-436.

Du, F., Qian, C., Ming Qian, Z., Wu, X.M., Xie, H., Yung, W.H., and Ke, Y. (2011). Hepcidin directly inhibits transferrin receptor 1 expression in astrocytes via a cyclic AMP- protein kinase a pathway. Glia *59*, 936-945.

Du, X., She, E., Gelbart, T., Truksa, J., Lee, P., Xia, Y., Khovananth, K., Mudd, S., Mann, N., and Moresco, E.M.Y. (2008). The serine protease TMPRSS6 is required to sense iron deficiency. Science *320*, 1088-1092.

Duckworth, D. (1889). A treatise on gout (Griffin).

Durward, W. (1976). Gout and hypothyroidism in males. Arthritis & Rheumatology *19*, 123-123.

Emmerson, B. (1966). Effects of allopurinol on iron metabolism in man. Annals of the rheumatic diseases *25*, 700-703.

Emond, M.J., Louie, T., Emerson, J., Zhao, W., Mathias, R.A., Knowles, M.R., Wright, F.A., Rieder, M.J., Tabor, H.K., and Nickerson, D.A. (2012). Exome sequencing of extreme phenotypes identifies DCTN4 as a modifier of chronic Pseudomonas aeruginosa infection in cystic fibrosis. Nature Genetics *44*, 886-889.

Emorine, L., Blin, N., and Strosberg, A. (1994). The human β3-adrenoceptor: the search for a physiological function. Trends in Pharmacological Sciences *15*, 3-7.

Ervin, R.B. (2009). Prevalence of metabolic syndrome among adults 20 years of age and over, by sex, age, race and ethnicity, and body mass index: United States. National health statistics reports *13*, 1-8.

Evans, J.G., Prior, I., and Harvey, H. (1968). Relation of serum uric acid to body bulk, haemoglobin, and alcohol intake in two South Pacific Polynesian populations. Annals of the rheumatic diseases *27*, 319.

Fabbrini, E., Sullivan, S., and Klein, S. (2010). Obesity and nonalcoholic fatty liver disease: biochemical, metabolic, and clinical implications. Hepatology *51*, 679-689.

Facchini, C., Malfatto, G., Giglio, A., Facchini, M., Parati, G., and Branzi, G. (2016). Lung ultrasound and transthoracic impedance for noninvasive evaluation of pulmonary congestion in heart failure. Journal of cardiovascular medicine *17*, 510-517.

Facchini, F., Chen, Y.-D.I., Hollenbeck, C.B., and Reaven, G.M. (1991). Relationship between resistance to insulin-mediated glucose uptake, urinary uric acid clearance, and plasma uric acid concentration. Jama *266*, 3008-3011.

Facchini, F.S. (2003). Near-iron deficiency-induced remission of gouty arthritis. Rheumatology *42*, 1550-1555.

Fam, A.G. (2002). Gout, diet, and the insulin resistance syndrome. The Journal of Rheumatology *29*, 1350-1355.

Fatima, T., Altaf, S., Phipps-Green, A., Topless, R., Flynn, T.J., Stamp, L.K., Dalbeth, N., and Merriman, T.R. (2016). Association analysis of the beta-3 adrenergic receptor Trp64Arg (rs4994) polymorphism with urate and gout. Rheumatology International *36*, 255-261.

Fauci, A.S. (2008). Harrison's principles of internal medicine, Vol 2, 17th edn (McGraw-Hill, Medical Publishing Division).

Feder, J., Gnirke, A., Thomas, W., Tsuchihashi, Z., Ruddy, D., Basava, A., Dormishian, F., Domingo Jr, R., Ellis, M., and Fullan, A. (1996). A novel MHC class I-like gene is mutated in patients with hereditary. Nat genet *13*, 399-408.

Feigin, V., Carter, K., Hackett, M., Barber, P.A., McNaughton, H., Dyall, L., Chen, M.-h., Anderson, C., and Group, A.R.C.S.S. (2006). Ethnic disparities in incidence of stroke subtypes: Auckland Regional Community Stroke Study, 2002–2003. The Lancet Neurology *5*, 130-139.

Feinstein, A.R. (1970). The pre-therapeutic classification of co-morbidity in chronic disease. Journal of chronic diseases *23*, 455-468.

Feng, N., Young, S.F., Aguilera, G., Puricelli, E., Adler-Wailes, D.C., Sebring, N.G., and Yanovski, J.A. (2005). Co-occurrence of two partially inactivating polymorphisms of MC3R is associated with pediatric-onset obesity. Diabetes *54*, 2663-2667.

Ference, B.A., Yoo, W., Alesh, I., Mahajan, N., Mirowska, K.K., Mewada, A., Kahn, J., Afonso, L., Williams, K.A., and Flack, J.M. (2012). Effect of long-term exposure to lower low-density lipoprotein cholesterol beginning early in life on the risk of coronary heart disease. Journal of the American College of Cardiology *60*, 2631-2639.

Feuk, L., Carson, A.R., and Scherer, S.W. (2006). Structural variation in the human genome. Nature Reviews Genetics *7*, 85-97.

Fewell, Z., Smith, G.D., and Sterne, J.A. (2007). The impact of residual and unmeasured confounding in epidemiologic studies: a simulation study. American Journal of Epidemiology *166*, 646-655.

Fischbach, F., Gregory, D., Harrison, P.M., Hoy, T., and Williams, J. (1971). On the structure of hemosiderin and its relationship to ferritin. Journal of ultrastructure research *37*, 495-503.

Fisher, C.E., and Howie, S.E. (2006). The role of megalin (LRP-2/Gp330) during development. Developmental biology *296*, 279-297.

Flynn, T.J., Cadzow, M., Dalbeth, N., Jones, P.B., Stamp, L.K., Hindmarsh, J.H., Todd, A.S., Walker, R.J., Topless, R., and Merriman, T.R. (2015). Positive association of tomato consumption with serum urate: support for tomato consumption as an anecdotal trigger of gout flares. BMC musculoskeletal disorders *16*, 1.

Fourcroy, A. (1804). Système des Connaissances Chimiques: , Vol 11 (Paris, 1800, X: Translated into English in London under the title 'General system of chemical knowledge').

Fox, E.R., Benjamin, E.J., Sarpong, D.F., Rotimi, C.N., Wilson, J.G., Steffes, M.W., Chen, G., Adeyemo, A., Taylor, J.K., and Samdarshi, T.E. (2008). Epidemiology, heritability, and genetic linkage of C-reactive protein in African Americans (from the Jackson Heart Study). The American journal of cardiology *102*, 835-841.

Frayling, T.M., Timpson, N.J., Weedon, M.N., Zeggini, E., Freathy, R.M., Lindgren, C.M., Perry, J.R., Elliott, K.S., Lango, H., and Rayner, N.W. (2007). A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science *316*, 889-894.

Frazer, K.A., Murray, S.S., Schork, N.J., and Topol, E.J. (2009). Human genetic variation and its contribution to complex traits. Nature Reviews Genetics *10*, 241-251.

Fu, C.-C., Wu, D.-A., Wang, J.-h., Yang, W.-C., and Tseng, C.-H. (2009). Association of C-reactive protein and hyperuricemia with diabetic nephropathy in Chinese type 2 diabetic patients. Acta diabetologica *46*, 127.

Fujihara, M., Muroi, M., Tanamoto, K.-i., Suzuki, T., Azuma, H., and Ikeda, H. (2003). Molecular mechanisms of macrophage activation and deactivation by lipopolysaccharide: roles of the receptor complex. Pharmacology & therapeutics *100*, 171-194.

Fuqua, B.K., Vulpe, C.D., and Anderson, G.J. (2012). Intestinal iron absorption. Journal of Trace Elements in Medicine and Biology *26*, 115-119.

Füredi-Milhofer, H., Babić-Ivaniĉić, V., Milat, O., Brown, W.E., and Gregory, T.M. (1987). Precipitation of sodium acid urate from electrolyte solutions. Journal of crystal growth *83*, 572-580.

Fushimi, K., Shichiri, M., and Marumo, F. (1990). Decreased fractional excretion of urate as an indicator of prerenal azotemia. American journal of nephrology *10*, 489-494.

Gage, S.H., Jones, H.J., Burgess, S., Bowden, J., Smith, G.D., Zammit, S., and Munafò, M.R. (2016). Assessing causality in associations between cannabis use and schizophrenia risk: a two-sample Mendelian randomization study. Psychological Medicine, 1-10.

Ganz, T. (2005). Cellular iron: ferroportin is the only way out. Cell Metabolism *1*, 155-157.

Ganz, T. (2013). Systemic iron homeostasis. Physiological reviews *93*, 1721-1741.

Gao, C., Patel, C.J., Michailidou, K., Peters, U., Gong, J., Schildkraut, J., Schumacher, F.R., Zheng, W., Boffetta, P., and Stucker, I. (2016). Mendelian randomization study of adiposity-related traits and risk of breast, ovarian, prostate, lung and colorectal cancer. International Journal of Epidemiology, dyw129.

Gentles, D., Metcalf, P., Dyall, L., Sundborn, G., Schaaf, D., Black, P., Scragg, R., and Jackson, R. (2007). Metabolic syndrome prevalence in a multicultural population in Auckland, New Zealand. The New Zealand Medical Journal (Online) *120*.

Gertler, M.M., Garn, S.M., and Levine, S.A. (1951). Serum uric acid in relation to age and physique in health and in coronary heart disease. Annals of internal medicine *34*, 1421-1431.

Getting, S.J., Christian, H.C., Flower, R.J., and Perretti, M. (2002). Activation of melanocortin type 3 receptor as a molecular mechanism for adrenocorticotropic hormone efficacy in gouty arthritis. Arthritis Rheum *46*, 2765-2775.

Getting, S.J., Lam, C.W., Chen, A.S., Grieco, P., and Perretti, M. (2006). Melanocortin 3 receptors control crystal-induced inflammation. FASEB J *20*, 2234-2241.

Ghio, A.J., Ford, E.S., Kennedy, T.P., and Hoidal, J.R. (2005). The association between serum ferritin and uric acid in humans. Free radical research *39*, 337-342.

Ghio, A.J., Kennedy, T.P., Rao, G., Cooke, C.L., Miller, M.J., and Hoidal, J.R. (1994). Complexation of iron cation by sodium urate crystals and gouty inflammation. Archives of biochemistry and biophysics *313*, 215-221.

Ghio, A.J., Kennedy, T.P., Stonehuerner, J., Carter, J.D., Skinner, K.A., Parks, D.A., and Hoidal, J.R. (2002). Iron regulates xanthine oxidase activity in the lung. American Journal of Physiology-Lung Cellular and Molecular Physiology *283*, L563-L572.

Gibson, T., Waterworth, R., Hatfield, P., Robinson, G., and Bremner, K. (1984). Hyperuricaemia, gout and kidney function in New Zealand Maori men. Rheumatology *23*, 276-282.

Gjesing, A., Andersen, G., Borch-Johnsen, K., Jørgensen, T., Hansen, T., and Pedersen, O. (2008). Association of the β 3-adrenergic receptor Trp64Arg polymorphism with common metabolic traits: Studies of 7605 middle-aged white people. Molecular genetics and metabolism *94*, 90-97.

Go, A.S., Chertow, G.M., Fan, D., McCulloch, C.E., and Hsu, C.-y. (2004). Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. New England Journal of Medicine *351*, 1296-1305.

Goldstein, D.B. (2009). Common genetic variation and human traits. New England Journal of Medicine *360*, 1696.

Goldstein, D.B., Allen, A., Keebler, J., Margulies, E.H., Petrou, S., Petrovski, S., and Sunyaev, S. (2013). Sequencing studies in human genetics: design and interpretation. Nature Reviews Genetics *14*, 460-470.

Gorlov, I.P., Gorlova, O.Y., Frazier, M.L., Spitz, M.R., and Amos, C.I. (2011). Evolutionary evidence of the effect of rare variants on disease etiology. Clinical genetics *79*, 199-206.

Gosling, A.L., Matisoo-Smith, E., and Merriman, T.R. (2014). Hyperuricaemia in the Pacific: why the elevated serum urate levels? Rheumatology International *34*, 743-757.

Grayson, P.C., Kim, S.Y., LaValley, M., and Choi, H.K. (2011). Hyperuricemia and incident hypertension: a systematic review and meta- analysis. Arthritis care & research *63*, 102-110.

Green, S., and Mazur, A. (1957). Relation of uric acid metabolism to release of iron from hepatic ferritin. Journal of Biological Chemistry *227*, 653-668.

Grimaldi-Bensouda, L., Alpérovitch, A., Aubrun, E., Danchin, N., Rossignol, M., Abenhaim, L., Richette, P., Fromage, P., Furber, A., and Mery, D. (2014). Impact of allopurinol on risk of myocardial infarction. Annals of the rheumatic diseases, annrheumdis-2012-202972.

Gropper, S., and Smith, J. (2013). Advanced Nutrition and Human Metabolism, 6th edn (Belmont, CA: Wadsworth).

Grundy, S.M., Brewer, H.B., Cleeman, J.I., Smith, S.C., and Lenfant, C. (2004). Definition of metabolic syndrome. Circulation *109*, 433-438.

Grundy, S.M., Cleeman, J.I., Daniels, S.R., Donato, K.A., Eckel, R.H., Franklin, B.A., Gordon, D.J., Krauss, R.M., Savage, P.J., and Smith, S.C. (2005). Diagnosis and management of the metabolic syndrome. Circulation *112*, 2735-2752.

Gunter, E.L., BL, Koncikowski, SM (1996). Laboratory procedures used for the Third National Health and Nutrition Examination Survey (NHANES III), 1988–1994. In NHANES III Reference Manuals and Reports, U.S.D.O.H.A.H.S.C.f.D.C.a. Prevention, ed. (Hyattsville, MD), pp. VII-R-1.

Guo, L., Cheng, Y., Wang, X., Pan, Q., Li, H., Zhang, L., and Wang, Y. (2012). Association between microalbuminuria and cardiovascular disease in type 2 diabetes mellitus of the Beijing Han nationality. Acta diabetologica *49*, 65-71.

Guo, Y., He, J., Zhao, S., Wu, H., Zhong, X., Sheng, Q., Samuels, D.C., Shyr, Y., and Long, J. (2014). Illumina human exome genotyping array clustering and quality control. nature protocols *9*, 2643-2662.

Hachem, C.Y., and El-Zimaity, H. (2007). A Man With Rheumatoid Arthritis and Iron-Deficiency Anemia. Medscape General Medicine *9*, 64.

Hadley, M.E., and Dorr, R.T. (2006). Melanocortin peptide therapeutics: historical milestones, clinical studies and commercialization. Peptides *27*, 921-930.

Hadley, M.E., and Haskell, L., CARRIE (1999). The proopiomelanocortin system. Annals of the New York Academy of Sciences *885*, 1-21.

Haig, A. (1889). On uric acid and arterial tension. British medical journal *1*, 288.

Hamajima, N., Naito, M., Okada, R., Kawai, S., Yin, G., Morita, E., Higashibata, T., Tamura, T., Nakagawa, H., and Matsuo, H. (2012). Significant interaction between LRP2 rs2544390 in intron 1 and alcohol drinking for serum uric acid levels among a Japanese population. Gene *503*, 131-136.

Han, F., and Pan, W. (2010). A data-adaptive sum test for disease association with multiple common or rare variants. Human heredity *70*, 42-54.

Han, G., Gonzalez, S., and DeVries, D. (2014). Combined effect of hyperuricemia and overweight/obesity on the prevalence of hypertension among US adults: result from the National Health and Nutrition Examination Survey. Journal of human hypertension *28*, 579-586.

Hani, E.H., Dupont, S., Durand, E., Dina, C., Gallina, S., Gantz, I., and Froguel, P. (2001). Naturally Occurring Mutations in the Melanocortin Receptor 3 Gene Are Not Associated with Type 2 Diabetes Mellitus in French Caucasians 1. The Journal of Clinical Endocrinology & Metabolism *86*, 2895-2898.

Hanova, P., Pavelka, K., Dostal, C., Holcatova, I., and Pikhart, H. (2006). Epidemiology of rheumatoid arthritis, juvenile idiopathic arthritis and gout in two regions of the Czech Republic in a descriptive population-based survey in 2002–2003. Clin Exp Rheumatol *24*, 499-507.

Hanson, E.H., Imperatore, G., and Burke, W. (2001). HFE gene and hereditary hemochromatosis: a HuGE review. American journal of epidemiology *154*, 193-206.

Harris, M.F. (2013). The metabolic syndrome. Australian Family Physician *42*, 524-527.

Harrison, P.M., Treffry, A., and Lilley, T.H. (1986). Ferritin as an iron-storage protein: mechanisms of iron uptake. Journal of inorganic biochemistry *27*, 287-293.

Haslam, D.B., and Baenziger, J.U. (1996). Expression cloning of Forssman glycolipid synthetase: a novel member of the histo-blood group ABO gene family. Proceedings of the National Academy of Sciences *93*, 10697-10702.

Haycock, P.C., Burgess, S., Wade, K.H., Bowden, J., Relton, C., and Smith, G.D. (2016). Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. The American journal of clinical nutrition *103*, 965-978.

Hayden, J.A., van der Windt, D.A., Cartwright, J.L., Côté, P., and Bombardier, C. (2013). Assessing bias in studies of prognostic factors. Annals of internal medicine *158*, 280-286.

Hemani, G., Zheng, J., Wade, K.H., Laurin, C., Elsworth, B., Burgess, S., Bowden, J., Langdon, R., Tan, V., and Yarmolinsky, J. (2016). MR-Base: a platform for systematic causal inference across the phenome using billions of genetic associations. bioRxiv, 078972.

Hentze, M.W., Muckenthaler, M.U., and Andrews, N.C. (2004). Balancing acts: molecular control of mammalian iron metabolism. cell *117*, 285-297.

Hentze, M.W., Muckenthaler, M.U., Galy, B., and Camaschella, C. (2010). Two to tango: regulation of Mammalian iron metabolism. Cell *142*, 24-38.

Hider, R.C., and Kong, X. (2013). Iron: effect of overload and deficiency. In Interrelations between Essential Metal Ions and Human Diseases (Springer), pp. 229-294.

Higgins, P., Walters, M., Murray, H., McArthur, K., McConnachie, A., Lees, K., and Dawson, J. (2014). Allopurinol reduces brachial and central blood pressure, and carotid intima-media thickness progression after ischaemic stroke and transient ischaemic attack: a randomised controlled trial. Heart, heartjnl-2014-305683.

Hingorani, A., and Humphries, S. (2005). Nature's randomised trials. The Lancet *366*, 1906-1908.

Hollis-Moffatt, J.E., Phipps-Green, A.J., Chapman, B., Jones, G.T., van Rij, A., Gow, P.J., Harrison, A.A., Highton, J., Jones, P.B., and Montgomery, G.W. (2012a). The renal urate transporter SLC17A1 locus: confirmation of association with gout. Arthritis research & therapy *14*, R92.

Hollis-Moffatt, J.E., Phipps-Green, A.J., Chapman, B., Jones, G.T., van Rij, A., Gow, P.J., Harrison, A.A., Highton, J., Jones, P.B., and Montgomery, G.W. (2012b). The renal urate transporter SLC17A1 locus: confirmation of association with gout. Arthritis research & therapy *14*, 1.

Hollis- Moffatt, J.E., Xu, X., Dalbeth, N., Merriman, M.E., Topless, R., Waddell, C., Gow, P.J., Harrison, A.A., Highton, J., and Jones, P.B. (2009). Role of the urate transporter SLC2A9 gene in susceptibility to gout in New Zealand Māori, Pacific Island, and Caucasian case–control sample sets. Arthritis & Rheumatology *60*, 3485-3492.

Holm, H., Gudbjartsson, D.F., Sulem, P., Masson, G., Helgadottir, H.T., Zanon, C., Magnusson, O.T., Helgason, A., Saemundsdottir, J., and Gylfason, A. (2011). A rare variant in MYH6 is associated with high risk of sick sinus syndrome. Nature Genetics *43*, 316-320.

Hosaka, K., Takeda, T., Iino, N., Hosojima, M., Sato, H., Kaseda, R., Yamamoto, K., Kobayashi, A., Gejyo, F., and Saito, A. (2009). Megalin and nonmuscle myosin heavy chain IIA interact with the adaptor protein Disabled-2 in proximal tubule cells. Kidney international *75*, 1308-1315.

Howell, R.R., and Wyngaarden, J.B. (1960). On the mechanism of peroxidation of uric acids by hemoproteins. J Biol Chem *235*, 3544-3550.

Howie, B., and Marchini, J., M (2011). Genotype imputation with thousands of genomes G3: genes, genomes. Genetics *1*, 457-470.

Huang, H., Huang, B., Li, Y., Huang, Y., Li, J., Yao, H., Jing, X., Chen, J., and Wang, J. (2014). Uric acid and risk of heart failure: a systematic review and meta- analysis. European journal of heart failure *16*, 15-24.

Huang, Q., Zhang, L.-F., Cheng, Y., Zhao, Y.-C., Si, L., Gao, Y., and Wei, W. (2013). Trp64Arg (rs4994) polymorphism of β3-adrenergic receptor gene is associated with hyperuricemia in a Chinese male population. Clinical chemistry and laboratory medicine *51*, 1755-1760.

Hudjashov, G., Kivisild, T., Underhill, P.A., Endicott, P., Sanchez, J.J., Lin, A.A., Shen, P., Oefner, P., Renfrew, C., and Villems, R. (2007). Revealing the prehistoric settlement of Australia by Y chromosome and mtDNA analysis. Proceedings of the National Academy of Sciences *104*, 8726-8730.

Hüebner, C., Petermann, I., Browning, B.L., Shelling, A.N., and Ferguson, L.R. (2007). Triallelic single nucleotide polymorphisms and genotyping error in genetic epidemiology studies: MDR1 (ABCB1) G2677/T/A as an example. Cancer Epidemiology Biomarkers & Prevention *16*, 1185-1192.

Hughes, K., Flynn, T., De Zoysa, J., Dalbeth, N., and Merriman, T.R. (2014). Mendelian randomization analysis associates increased serum urate, due to genetic variation in uric acid transporters, with improved renal function. Kidney international *85*, 344-351.

Hugot, J.-P., Chamaillard, M., Zouali, H., Lesage, S., Cézard, J.-P., Belaiche, J., Almer, S., Tysk, C., O'Morain, C.A., and Gassull, M. (2001). Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature *411*, 599-603.

Hunt, J.R. (2002). Moving toward a plant-based diet: are iron and zinc at risk? Nutrition reviews *60*, 127-134.

Hunt, J.R., Zito, C.A., and Johnson, L.K. (2009). Body iron excretion by healthy men and women. The American journal of clinical nutrition *89*, 1792-1798.

Hunt, K.A., Mistry, V., Bockett, N.A., Ahmad, T., Ban, M., Barker, J.N., Barrett, J.C., Blackburn, H., Brand, O., and Burren, O. (2013). Negligible impact of rare autoimmune-locus coding-region variants on missing heritability. Nature *498*, 232-235.

Hurrell, R., and Egli, I. (2010). Iron bioavailability and dietary reference values. The American journal of clinical nutrition *91*, 1461S-1467S.

Huszar, D., Lynch, C.A., Fairchild-Huntress, V., Dunmore, J.H., Fang, Q., Berkemeier, L.R., Gu, W., Kesterson, R.A., Boston, B.A., and Cone, R.D. (1997). Targeted disruption of the melanocortin-4 receptor results in obesity in mice. Cell *88*, 131-141.

Huyghe, J.R., Jackson, A.U., Fogarty, M.P., Buchkovich, M.L., Stančáková, A., Stringham, H.M., Sim, X., Yang, L., Fuchsberger, C., and Cederberg, H. (2013). Exome array analysis identifies new loci and low-frequency variants influencing insulin processing and secretion. Nature Genetics *45*, 197-201.

Illumina (2014). Infinium Genotyping Data Analysis: A Guide for Analyzing Infinium Genotyping Data Using the GenomeStudio Genotyping Module. 1-10.

Inoue, A., and Solon, G. (2010). Two-sample instrumental variables estimators. The Review of Economics and Statistics *92*, 557-561.

Ionita-Laza, I., Buxbaum, J.D., Laird, N.M., and Lange, C. (2011). A new testing strategy to identify rare variants with either risk or protective effect on disease. PLoS Genetics *7*, e1001289.

Ionita-Laza, I., Makarov, V., Buxbaum, J.D., and Consortium, A.A.S. (2012). Scan-statistic approach identifies clusters of rare disease variants in LRP2, a gene linked and associated with autism spectrum disorders, in three datasets. The American Journal of Human Genetics *90*, 1002-1013.

Ishizaka, N., Ishizaka, Y., Toda, E.-I., Hashimoto, H., Nagai, R., and Yamakado, M. (2007). Higher serum uric acid is associated with increased arterial stiffness in Japanese individuals. Atherosclerosis *192*, 131-137.

Ito, H., Abe, M., Mifune, M., Oshikiri, K., Antoku, S., Takeuchi, Y., and Togane, M. (2011). Hyperuricemia is independently associated with coronary heart disease and renal dysfunction in patients with type 2 diabetes mellitus. PLoS One *6*, e27817.

Iturrieta-Zuazo, I., and Walter, S. (2015). Mendelian randomization: present and future of epidemiological studies in cardiology. Revista Española de Cardiología *68*, 87-91.

Jacobs, A., Miller, F., Worwood, M., Beamish, M., and Wardrop, C. (1972). Ferritin in the serum of normal subjects and patients with iron deficiency and iron overload. Br med J *4*, 206-208.

Janssens, H.J., Fransen, J., Van de Lisdonk, E.H., van Riel, P.L., van Weel, C., and Janssen, M. (2010). A diagnostic rule for acute gouty arthritis in primary care without joint fluid analysis. Archives of internal medicine *170*, 1120-1126.

Jeong, J., Rao, A.U., Xu, J., Ogg, S.L., Hathout, Y., Fenselau, C., and Mather, I.H. (2009). The PRY/SPRY/B30. 2 Domain of Butyrophilin 1A1 (BTN1A1) Binds to Xanthine Oxidoreductase implications for the function of BTN1A1 in the mammary gland and other tissues. Journal of Biological Chemistry *284*, 22444-22456.

Jeppesen, J., Hansen, T.W., Rasmussen, S., Ibsen, H., Torp-Pedersen, C., and Madsbad, S. (2007). Insulin resistance, the metabolic syndrome, and risk of incident cardiovascular disease: a population-based study. Journal of the American College of Cardiology *49*, 2112-2119.

Jing, J., Kielstein, J.T., Schultheiss, U.T., Sitter, T., Titze, S.I., Schaeffner, E.S., McAdams-DeMarco, M., Kronenberg, F., Eckardt, K.-U., and Köttgen, A. (2014). Prevalence and correlates of gout in a large cohort of patients with chronic kidney disease: the German Chronic Kidney Disease (GCKD) study. Nephrology Dialysis Transplantation, gfu352.

Johnson, G.C., Esposito, L., Barratt, B.J., Smith, A.N., Heward, J., Di Genova, G., Ueda, H., Cordell, H.J., Eaves, I.A., and Dudbridge, F. (2001). Haplotype tagging for the identification of common disease genes. Nature Genetics *29*, 233-237.

Johnson, R.J., Gaucher, E.A., Sautin, Y.Y., Henderson, G.N., Angerhofer, A.J., and Benner, S.A. (2008). The planetary biology of ascorbate and uric acid and their relationship with the epidemic of obesity and cardiovascular disease. Medical hypotheses *71*, 22-31.

Johnson, R.J., Kivlighn, S.D., Kim, Y.-G., Suga, S., and Fogo, A.B. (1999). Reappraisal of the pathogenesis and consequences of hyperuricemia in hypertension, cardiovascular disease, and renal disease. American journal of kidney diseases *33*, 225-234.

Johnson, R.J., Merriman, T., and Lanaspa, M.A. (2015). Causal or noncausal relationship of uric acid with diabetes. Diabetes *64*, 2720-2722.

Johnson, R.J., Nakagawa, T., Jalal, D., Sánchez-Lozada, L.G., Kang, D.-H., and Ritz, E. (2013a). Uric acid and chronic kidney disease: which is chasing which? Nephrology Dialysis Transplantation, gft029.

Johnson, R.J., Nakagawa, T., Sanchez-Lozada, L.G., Shafiu, M., Sundaram, S., Le, M., Ishimoto, T., Sautin, Y.Y., and Lanaspa, M.A. (2013b). Sugar, uric acid, and the etiology of diabetes and obesity. Diabetes *62*, 3307-3315.

Johnson, R.J., Perez-Pozo, S.E., Sautin, Y.Y., Manitius, J., Sanchez-Lozada, L.G., Feig, D.I., Shafiu, M., Segal, M., Glassock, R.J., and Shimada, M. (2009a). Hypothesis: could excessive fructose intake and uric acid cause type 2 diabetes? Endocrine reviews *30*, 96-116.

Johnson, R.J., Sautin, Y.Y., Oliver, W.J., Roncal, C., Mu, W., Sanchez-Lozada, L.G., Rodriguez-Iturbe, B., Nakagawa, T., and Benner, S.A. (2009b). Lessons from comparative physiology: could uric acid represent a physiologic alarm signal gone awry in western society? Journal of comparative physiology B *179*, 67-76.

Johnson, R.J., Segal, M.S., Sautin, Y., Nakagawa, T., Feig, D.I., Kang, D.-H., Gersch, M.S., Benner, S., and Sánchez-Lozada, L.G. (2007). Potential role of sugar (fructose) in the epidemic of hypertension, obesity and the metabolic syndrome, diabetes, kidney disease, and cardiovascular disease. The American journal of clinical nutrition *86*, 899-906.

Johnson, R.J., Stenvinkel, P., Martin, S.L., Jani, A., Sánchez- Lozada, L.G., Hill, J.O., and Lanaspa, M.A. (2013c). Redefining metabolic syndrome as a fat storage condition based on studies of comparative physiology. Obesity *21*, 659-664.

Johnson, R.J., and Tuttle, K.R. (2000). Much ado about nothing, or much to do about something? The continuing controversy over the role of uric acid in cardiovascular disease. Hypertension *35*, e10-e10.

Jung, U.J., and Choi, M.-S. (2014). Obesity and its metabolic complications: the role of adipokines and the relationship between obesity, inflammation, insulin resistance, dyslipidemia and nonalcoholic fatty liver disease. International journal of molecular sciences *15*, 6184-6223.

Kamatani, Y., Matsuda, K., Okada, Y., Kubo, M., Hosono, N., Daigo, Y., Nakamura, Y., and Kamatani, N. (2010). Genome-wide association study of hematological and biochemical traits in a Japanese population. Nature genetics *42*, 210-215.

Kanbay, M., Jensen, T., Solak, Y., Le, M., Roncal-Jimenez, C., Rivard, C., Lanaspa, M.A., Nakagawa, T., and Johnson, R.J. (2016). Uric acid in metabolic syndrome: from an innocent bystander to a central player. European journal of internal medicine *29*, 3-8.

Kanbay, M., Yilmaz, M.I., Sonmez, A., Solak, Y., Saglam, M., Cakir, E., Unal, H.U., Arslan, E., Verim, S., and Madero, M. (2012). Serum uric acid independently predicts cardiovascular events in advanced nephropathy. American journal of nephrology *36*, 324-331.

Kang, D.-H., and Ha, S.-K. (2014). Uric acid puzzle: dual role as anti-oxidantand pro-oxidant. Electrolytes & Blood Pressure *12*, 1-6.

Kao, M.P., Ang, D.S., Gandy, S.J., Nadir, M.A., Houston, J.G., Lang, C.C., and Struthers, A.D. (2011). Allopurinol benefits left ventricular mass and endothelial dysfunction in chronic kidney disease. Journal of the American Society of Nephrology *22*, 1382-1389.

Kaur, J. (2014). A comprehensive review on metabolic syndrome. Cardiology research and practice *2014*.

Kawabata, H., Germain, R.S., Vuong, P.T., Nakamaki, T., Said, J.W., and Koeffler, H.P. (2000). Transferrin receptor 2-α supports cell growth both in iron-chelated cultured cells and in vivo. Journal of Biological Chemistry *275*, 16618-16625.

Kayser, M. (2010). The human genetic history of Oceania: near and remote views of dispersal. Current Biology *20*, R194-R201.

Keenan, T., Zhao, W., Rasheed, A., Ho, W.K., Malik, R., Felix, J.F., Young, R., Shah, N., Samuel, M., and Sheikh, N. (2016). Causal assessment of serum urate levels in cardiometabolic diseases through a Mendelian randomization study. Journal of the American College of Cardiology *67*, 407-416.

Kell, D.B., and Pretorius, E. (2014). Serum ferritin is an important inflammatory disease marker, as it is mainly a leakage product from damaged cells. Metallomics *6*, 748-773.

Kellogg, E.W., and Fridovich, I. (1977). Liposome oxidation and erythrocyte lysis by enzymically generated superoxide and hydrogen peroxide. Journal of Biological Chemistry *252*, 6721-6728.

Kim, S.Y., Guevara, J.P., Kim, K.M., Choi, H.K., Heitjan, D.F., and Albert, D.A. (2009). Hyperuricemia and risk of stroke: a systematic review and meta- analysis. Arthritis Care & Research *61*, 885-892.

Kim, S.Y., Guevara, J.P., Kim, K.M., Choi, H.K., Heitjan, D.F., and Albert, D.A. (2010). Hyperuricemia and coronary heart disease: a systematic review and meta- analysis. Arthritis care & research *62*, 170-180.

Kippen, I., Klinenberg, J.R., Weinberger, A., and Wilcox, W.R. (1974). Factors affecting urate solubility in vitro. Annals of the rheumatic diseases *33*, 313.

Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. Nature Genetics *46*, 310.

Kleber, M.E., Delgado, G., Grammer, T.B., Silbernagel, G., Huang, J., Krämer, B.K., Ritz, E., and März, W. (2015). Uric acid and cardiovascular events: a Mendelian randomization study. Journal of the American Society of Nephrology *26*, 2831-2838.

Klemp, P., Stansfield, S.A., Castle, B., and Robertson, M.C. (1997). Gout is on the increase in New Zealand. Annals of the rheumatic diseases *56*, 22-26.

Knutson, M., and Wessling-Resnick, M. (2003). Iron metabolism in the reticuloendothelial system. Critical reviews in biochemistry and molecular biology *38*, 61-88.

Knutson, M.D. (2010). Iron-sensing proteins that regulate hepcidin and enteric iron absorption. Annual review of nutrition *30*, 149-171.

Kohgo, Y., Ikuta, K., Ohtake, T., Torimoto, Y., and Kato, J. (2008). Body iron metabolism and pathophysiology of iron overload. International journal of hematology *88*, 7-15.

Kolz, M., Johnson, T., Sanna, S., Teumer, A., Vitart, V., Perola, M., Mangino, M., Albrecht, E., Wallace, C., and Farrall, M. (2009). Meta-analysis of 28,141 individuals identifies common variants within five new loci that influence uric acid concentrations. PLoS Genetics *5*, e1000504.

Komoda, F., Sekine, T., Inatomi, J., Enomoto, A., Endou, H., Ota, T., Matsuyama, T., Ogata, T., Ikeda, M., and Awazu, M. (2004). The W258X mutation in SLC22A12 is the predominant cause of Japanese renal hypouricemia. Pediatric Nephrology *19*, 728-733.

Kono, H., Rusyn, I., Bradford, B.U., Connor, H.D., Mason, R.P., and Thurman, R.G. (2000). Allopurinol prevents early alcohol-induced liver injury in rats. Journal of Pharmacology and Experimental Therapeutics *293*, 296-303.

Köttgen, A., Albrecht, E., Teumer, A., Vitart, V., Krumsiek, J., Hundertmark, C., Pistis, G., Ruggiero, D., O'Seaghdha, C.M., and Haller, T. (2013). Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. Nature Genetics *45*, 145-154.

Kovesdy, C.P., and Kalantar-Zadeh, K. (2012). Observational studies versus randomized controlled trials: avenues to causal inference in nephrology. Advances in chronic kidney disease *19*, 11-18.

Kraja, A.T., Vaidya, D., Pankow, J.S., Goodarzi, M.O., Assimes, T.L., Kullo, I.J., Sovio, U., Mathias, R.A., Sun, Y.V., and Franceschini, N. (2011). A bivariate genome-wide approach to metabolic syndrome. Diabetes *60*, 1329-1339.

Kramer, H.J., Choi, H.K., Atkinson, K., Stampfer, M., and Curhan, G.C. (2003). The association between gout and nephrolithiasis in men: The Health Professionals' Follow-Up Study. Kidney international *64*, 1022-1026.

Kramer, H.M., and Curhan, G. (2002). The association between gout and nephrolithiasis: the National Health and Nutrition Examination Survey III, 1988-1994. American Journal of Kidney Diseases *40*, 37-42.

Kratzer, J.T., Lanaspa, M.A., Murphy, M.N., Cicerchi, C., Graves, C.L., Tipton, P.A., Ortlund, E.A., Johnson, R.J., and Gaucher, E.A. (2014). Evolutionary history and metabolic insights of ancient mammalian uricases. Proceedings of the National Academy of Sciences *111*, 3763-3768.

Krief, S., Lönnqvist, F., Raimbault, S., Baude, B., Van Spronsen, A., Arner, P., Strosberg, A.D., Ricquier, D., and Emorine, L.J. (1993). Tissue distribution of beta 3-adrenergic receptor mRNA in man. Journal of Clinical Investigation *91*, 344.

Kring, S.I.I., Holst, C., Toubro, S., Astrup, A., Hansen, T., Pedersen, O., and Sørensen, T.I. (2010). Common variants near MC4R in relation to body fat, body fat distribution, metabolic traits and energy expenditure. International Journal of Obesity *34*, 182-189.

Krishnan, E. (2012). Reduced glomerular function and prevalence of gout: NHANES 2009–10. PLoS One *7*, e50046.

Krishnan, E., Baker, J.F., Furst, D.E., and Schumacher, H.R. (2006). Gout and the risk of acute myocardial infarction. Arthritis & Rheumatism *54*, 2688-2696.

Krishnan, E., Pandya, B.J., Chung, L., and Dabbous, O. (2011). Hyperuricemia and the risk for subclinical coronary atherosclerosis-data from a prospective observational cohort study. Arthritis research & therapy *13*, R66.

Krishnan, E., Pandya, B.J., Chung, L., Hariri, A., and Dabbous, O. (2012). Hyperuricemia in young adults and risk of insulin resistance, prediabetes, and diabetes: a 15-year follow-up study. American journal of epidemiology *176*, 108-116.

Kristiansson, K., Perola, M., Tikkanen, E., Kettunen, J., Surakka, I., Havulinna, A.S., Stančáková, A., Barnes, C., Widen, E., and Kajantie, E. (2012). Genome-Wide Screen for Metabolic Syndrome Susceptibility Loci Reveals Strong Lipid Gene Contribution But No Evidence for Common Genetic Basis for Clustering of Metabolic Syndrome TraitsCLINICAL PERSPECTIVE. Circulation: Cardiovascular Genetics *5*, 242-249.

Kumar, P., Henikoff, S., and Ng, P.C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nature protocols *4*, 1073-1081.

Kuo, C.-F., Grainge, M.J., Mallen, C., Zhang, W., and Doherty, M. (2014a). Comorbidities in patients with gout prior to and following diagnosis: case-control study. Annals of the rheumatic diseases, annrheumdis-2014-206410.

Kuo, C.-F., Grainge, M.J., Mallen, C., Zhang, W., and Doherty, M. (2014b). Rising burden of gout in the UK but continuing suboptimal management: a nationwide population study. Annals of the rheumatic diseases, annrheumdis-2013-204463.

Kuo, C.-F., Grainge, M.J., Zhang, W., and Doherty, M. (2015). Global epidemiology of gout: prevalence, incidence and risk factors. Nature Reviews Rheumatology *11*, 649-662.

Kuo, C.-F., Luo, S.-F., See, L.-C., Chou, I.-J., Fang, Y.-F., and Yu, K.-H. (2012). Increased risk of cancer among gout patients: a nationwide population study. Joint Bone Spine *79*, 375-378.

Kuo, C.-F., See, L.-C., Luo, S.-F., Ko, Y.-S., Lin, Y.-S., Hwang, J.-S., Lin, C.-M., Chen, H.-W., and Yu, K.-H. (2009). Gout: an independent risk factor for all-cause and cardiovascular mortality. Rheumatology, kep364.

Kuo, C.-F., Yu, K.-H., See, L.-C., Chou, I.-J., Ko, Y.-S., Chang, H.-C., Chiou, M.-J., and Luo, S.-F. (2013). Risk of myocardial infarction among patients with gout: a nationwide population-based study. Rheumatology *52*, 111-117.

Kurokawa, N., Young, E.H., Oka, Y., Satoh, H., Wareham, N.J., Sandhu, M., and Loos, R.J. (2008). The ADRB3 Trp64Arg variant and BMI: a meta-analysis of 44 833 individuals. International journal of obesity *32*, 1240-1249.

Kuzell, W.C., Schaffarzick, R.W., Naugler, W.E., Koets, P., Mankle, E.A., Brown, B., and Champlin, B. (1955). Some observations on 520 gouty patients. Journal of chronic diseases *2*, 645-669.

Laird, N.M., and Lange, C. (2006). Family-based designs in the age of large-scale gene-association studies. Nature Reviews Genetics *7*, 385-394.

Lanaspa, M.A., Cicerchi, C., Garcia, G., Li, N., Roncal-Jimenez, C.A., Rivard, C.J., Hunter, B., Andrés-Hernando, A., Ishimoto, T., and Sánchez-Lozada, L.G. (2012a). Counteracting roles of AMP deaminase and AMP kinase in the development of fatty liver. PLoS One *7*, e48801.

Lanaspa, M.A., Epperson, L.E., Li, N., Cicerchi, C., Garcia, G.E., Roncal-Jimenez, C.A., Trostel, J., Jain, S., Mant, C.T., and Rivard, C.J. (2015). Opposing activity changes in AMP deaminase and AMP-activated protein kinase in the hibernating ground squirrel. PloS One *10*, e0123509.

Lanaspa, M.A., Sanchez-Lozada, L.G., Choi, Y.-J., Cicerchi, C., Kanbay, M., Roncal-Jimenez, C.A., Ishimoto, T., Li, N., Marek, G., and Duranay, M. (2012b). Uric acid induces hepatic steatosis by generation of mitochondrial oxidative stress potential role in fructose-dependent and-independent fatty liver. Journal of Biological Chemistry *287*, 40732-40744.

Lanaspa, M.A., Sanchez-Lozada, L.G., Cicerchi, C., Li, N., Roncal-Jimenez, C.A., Ishimoto, T., Le, M., Garcia, G.E., Thomas, J.B., and Rivard, C.J. (2012c). Uric acid stimulates fructokinase and accelerates fructose metabolism in the development of fatty liver. PloS one *7*, e47948.

Lane, D., Merlot, A., Huang, M.-H., Bae, D.-H., Jansson, P., Sahni, S., Kalinowski, D., and Richardson, D. (2015). Cellular iron uptake, trafficking and metabolism: key molecules and mechanisms and their roles in disease. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research *1853*, 1130-1144.

Lawlor, D.A., Harbord, R.M., Sterne, J.A., Timpson, N., and Davey Smith, G. (2008). Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Statistics in Medicine *27*, 1133-1163.

Lebherz-Eichinger, D., Tudor, B., Ankersmit, H.J., Reiter, T., Haas, M., Roth-Walter, F., Krenn, C.G., and Roth, G.A. (2015). Trefoil Factor 1 Excretion Is Increased in Early Stages of Chronic Kidney Disease. PloS One *10*, e0138312.

Lee, I.R., Yang, L., Sebetso, G., Allen, R., Doan, T.H., Blundell, R., Lui, E.Y., Morrow, C.A., and Fraser, J.A. (2013). Characterization of the complete uric acid degradation pathway in the fungal pathogen Cryptococcus neoformans. PloS One *8*, e64292.

Lee, S., Abecasis, G.R., Boehnke, M., and Lin, X. (2014). Rare-variant association analysis: study designs and statistical tests. The American Journal of Human Genetics *95*, 5-23.

Lee, Y.S., Poh, L.K.S., Kek, B.L.K., and Loke, K.Y. (2007). The role of melanocortin 3 receptor gene in childhood obesity. Diabetes *56*, 2622-2630.

Lennane, G., Rose, B., and Isdale, I. (1960). Gout in the Maori. Annals of the rheumatic diseases *19*, 120.

Leong, W., and Lonnerdal, B. (2012). Iron nutrition (New York: Springer).

Li, B., and Leal, S.M. (2008). Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. The American Journal of Human Genetics *83*, 311-321.

Li, C., Hsieh, M.-C., and Chang, S.-J. (2013). Metabolic syndrome, diabetes, and hyperuricemia. Current opinion in rheumatology *25*, 210-216.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics *25*, 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078-2079.

Li, J., Lange, L.A., Duan, Q., Lu, Y., Singleton, A.B., Zonderman, A.B., Evans, M.K., Li, Y., Taylor, H.A., and Willis, M.S. (2015). Genome-wide admixture and association study of serum iron, ferritin, transferrin saturation and total iron binding capacity in African Americans. Human molecular genetics *24*, 572-581.

Li, M., Hou, W., Zhang, X., Hu, L., and Tang, Z. (2014). Hyperuricemia and risk of stroke: a systematic review and meta-analysis of prospective studies. Atherosclerosis *232*, 265-270.

Li, M., Hu, X., Fan, Y., Li, K., Zhang, X., Hou, W., and Tang, Z. (2016). Hyperuricemia and the risk for coronary heart disease morbidity and mortality a systematic review and dose-response meta-analysis. Scientific Reports *6*.

Li, S., Sanna, S., Maschio, A., Busonero, F., Usala, G., Mulas, A., Lai, S., Dei, M., Orrù, M., and Albai, G. (2007). The GLUT9 gene is associated with serum uric acid levels in Sardinia and Chianti cohorts. PLoS Genetics *3*, e194.

Li, W., Joo, E., Furlong, E., Galvin, M., Abel, K., Bell, C., and Price, R.A. (2000). Melanocortin 3 receptor (MC3R) gene variants in extremely obese women. International journal of obesity *24*, 206.

Lim, K.-H., and Staudt, L.M. (2013). Toll-like receptor signaling. Cold Spring Harbor perspectives in biology *5*, a011247.

Lin, H.-F., Boden-Albala, B., Juo, S., Park, N., Rundek, T., and Sacco, R. (2005). Heritabilities of the metabolic syndrome and its components in the Northern Manhattan Family Study. Diabetologia *48*, 2006-2012.

Lindsay, K., Gow, P., Vanderpyl, J., Logo, P., and Dalbeth, N. (2011). The experience and impact of living with gout: a study of men with chronic gout using a qualitative grounded theory approach. JCR: Journal of Clinical Rheumatology *17*, 1-6.

Linhart, C., Tivollier, J.-M., Taylor, R., Barguil, Y., Magliano, D.J., Bourguignon, C., and Zimmet, P. (2016). Changes in cardiovascular disease risk factors over 30 years in Polynesians in the French Pacific Territory of Wallis Island. European journal of preventive cardiology *23*, 856-864.

Lioté, F., Merriman, T., Nasi, S., and So, A. (2013). Workshop report: 4th European crystal network meeting. Arthritis research & therapy *15*, 304.

Lippi, G., Targher, G., Montagnana, M., Salvagno, G.L., and Guidi, G.C. (2008). High serum uric acid as a novel risk factor for type 2 diabetes. Diabetes care *31*, e68-e68.

Lipschitz, D.A., Cook, J.D., and Finch, C.A. (1974). A clinical evaluation of serum ferritin as an index of iron stores. New England Journal of Medicine *290*, 1213-1216.

Lipsitch, M., Jha, A., and Simonsen, L. (2016). Observational studies and the difficult quest for causality: lessons from vaccine effectiveness and impact studies. International journal of epidemiology, dyw124.

Lipsitch, M., Tchetgen, E.T., and Cohen, T. (2012). Negative control exposures in epidemiologic studies. Epidemiology *23*, 351-352.

Liu, D.J., and Leal, S.M. (2010). Replication strategies for rare variant complex trait association studies via next-generation sequencing. The American Journal of Human Genetics *87*, 790-801.

Liu, Z., Que, S., Zhou, L., and Zheng, S. (2015). Dose-response relationship of serum uric acid with metabolic syndrome and non-alcoholic fatty liver disease incidence: a meta-analysis of prospective studies. Scientific Reports *5*, 14325.

Livrea, M., Tesoriere, L., Pintaudi, A., Calabrese, A., Maggio, A., Freisleben, H., D'arpa, D., D'anna, R., and Bongiorno, A. (1996). Oxidative stress and antioxidant status in beta-thalassemia major: iron overload and depletion of lipid-soluble antioxidants. Blood *88*, 3608-3614.

Llull, L., Amaro, S., and Chamorro, Á. (2016). Administration of uric acid in the emergency treatment of acute ischemic stroke. Current neurology and neuroscience reports *16*, 4.

Lohmueller, K.E., Sparsø, T., Li, Q., Andersson, E., Korneliussen, T., Albrechtsen, A., Banasik, K., Grarup, N., Hallgrimsdottir, I., and Kiil, K. (2013). Whole-exome sequencing of 2,000 Danish individuals and the role of rare coding variants in type 2 diabetes. The American Journal of Human Genetics *93*, 1072-1086.

Loos, R.J., Lindgren, C.M., Li, S., Wheeler, E., Zhao, J.H., Prokopenko, I., Inouye, M., Freathy, R.M., Attwood, A.P., and Beckmann, J.S. (2008). Common variants near MC4R are associated with fat mass, weight and risk of obesity. Nature genetics *40*, 768-775.

Lopez, A., Cacoub, P., Macdougall, I.C., and Peyrin-Biroulet, L. (2016). Iron deficiency anaemia. The Lancet *387*, 907-916.

Lu, D., Willard, D., Patel, I.R., Kadwell, S., Overton, L., Kost, T., Luther, M., Chen, W., Woychik, R.P., and Wilkison, W.O. (1994). Agouti protein is an antagonist of the melanocyte-stimulating-hormone receptor. Nature *371*, 799.

Luo, C., Wang, F., Qin, S., Chen, Q., and Wang, Q.K. (2016). Coronary artery disease susceptibility gene ADTRP regulates cell cycle progression, proliferation, and apoptosis by global gene expression regulation. Physiological genomics *48*, 554-564.

Lupu, C., Zhu, H., Popescu, N.I., Wren, J.D., and Lupu, F. (2011). Novel protein ADTRP regulates TFPI expression and function in human endothelial cells in normal conditions and in response to androgen. Blood *118*, 4463-4471.

Lyngdoh, T., Vuistiner, P., Marques-Vidal, P., Rousson, V., Waeber, G., Vollenweider, P., and Bochud, M. (2012). Serum uric acid and adiposity: deciphering causality using a bidirectional Mendelian randomization approach. PloS one *7*, e39321.

Macdougall, I.C., Bircher, A.J., Eckardt, K.-U., Obrador, G.T., Pollock, C.A., Stenvinkel, P., Swinkels, D.W., Wanner, C., Weiss, G., and Chertow, G.M. (2016). Iron management in chronic kidney disease: conclusions from a "Kidney Disease: Improving Global Outcomes"(KDIGO) Controversies Conference. Kidney international *89*, 28-39.

Madero, M., Sarnak, M.J., Wang, X., Greene, T., Beck, G.J., Kusek, J.W., Collins, A.J., Levey, A.S., and Menon, V. (2009). Uric acid and long-term outcomes in CKD. American Journal of Kidney Diseases *53*, 796-803.

Madsen, B.E., and Browning, S.R. (2009). A groupwise association test for rare mutations using a weighted sum statistic. PLoS Genetics *5*, e1000384.

Magoulas, C., and Fried, M. (2000). Isolation and genomic analysis of the human surf-6 gene: a member of the Surfeit locus. Gene *243*, 115-123.

Mainous, A.G., Knoll, M.E., Everett, C.J., Matheson, E.M., Hulihan, M.M., and Grant, A.M. (2011). Uric acid as a potential cue to screen for iron overload. The Journal of the American Board of Family Medicine *24*, 415-421.

Malik, V.S., Pan, A., Willett, W.C., and Hu, F.B. (2013). Sugar-sweetened beverages and weight gain in children and adults: a systematic review and meta-analysis. The American journal of clinical nutrition *98*, 1084-1102.

Malik, V.S., Popkin, B.M., Bray, G.A., Després, J.-P., Willett, W.C., and Hu, F.B. (2010). Sugar-sweetened beverages and risk of metabolic syndrome and type 2 diabetes. Diabetes care *33*, 2477-2483.

Mandal, A.K., and Mount, D.B. (2015). The molecular physiology of uric acid homeostasis. Annual review of physiology *77*, 323-345.

Mann, B., and Wood, E. (2012). Confounding in observational studies explained. Open Epidemiol J *5*, 18-20.

Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., and Chakravarti, A. (2009). Finding the missing heritability of complex diseases. Nature *461*, 747-753.

Marcadenti, A., Fuchs, F.D., Matte, U., Sperb, F., Moreira, L.B., and Fuchs, S.C. (2013). Effects of FTO RS9939906 and MC4R RS17782313 on obesity, type 2 diabetes mellitus and blood pressure in patients with hypertension. Cardiovascular diabetology *12*, 103.

Mark, A., Rahmouni, K., Correia, M., and Haynes, W. (2003). A leptin- sympathetic- leptin feedback loop: potential implications for regulation of arterial pressure and body fat. Acta physiologica scandinavica *177*, 345-349.

Markadieu, N., and Delpire, E. (2014). Physiology and pathophysiology of SLC12A1/2 transporters. Pflügers Archiv-European Journal of Physiology *466*, 91-105.

Márquez-Sandoval, F., Macedo-Ojeda, G., Viramontes-Hörner, D., Ballart, J.F., Salvadó, J.S., and Vizmanos, B. (2011). The prevalence of metabolic syndrome in Latin America: a systematic review. Public Health Nutrition *14*, 1702-1713.

Martelin, E., Lapatto, R., and Raivio, K.O. (2002). Regulation of xanthine oxidoreductase by intracellular iron. American Journal of Physiology-Cell Physiology *283*, C1722-C1728.

Martín, N.E., and Nieto, V.G. (2011). Hypouricemia and tubular transport of uric acid. Nefrologia *31*, 44-50.

Masako, O., Yoko, S., Osamu, T., and Naoyuki, T. (2013). Loss of urate oxidase activity in hominoids and its evolutionary implications. Molecular Biology and Evolution.

Masuo, K., Kawaguchi, H., Mikami, H., Ogihara, T., and Tuck, M.L. (2003). Serum uric acid and plasma norepinephrine concentrations predict subsequent weight gain and blood pressure elevation. Hypertension *42*, 474-480.

Mazur, A., Green, S., Saha, A., and Carleton, A. (1958). Mechanism of release of ferritin iron in vivo by xanthine oxidase. Journal of Clinical Investigation *37*, 1809.

McAdams DeMarco, M.A., Maynard, J.W., Huizinga, M.M., Baer, A.N., Köttgen, A., Gelber, A.C., and Coresh, J. (2011). Obesity and younger age at gout onset in a community- based cohort. Arthritis care & research *63*, 1108-1114.

McAdams-DeMarco, M.A., Maynard, J.W., Coresh, J., and Baer, A.N. (2012). Anemia and the onset of gout in a population-based cohort of adults: Atherosclerosis Risk in Communities study. Arthritis research & therapy *14*, R193.

McAdams-DeMarco, M.A., Maynard, J.W., Huizinga, M.M., Baer, A.N., Köttgen, A., Gelber, A.C., and Coresh, J. (2011). Obesity and younger age at gout onset in a community- based cohort. Arthritis care & research *63*, 1108-1114.

McCarthy, M.I., Abecasis, G.R., Cardon, L.R., Goldstein, D.B., Little, J., Ioannidis, J.P., and Hirschhorn, J.N. (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nature reviews genetics *9*, 356-369.

McCarthy, M.I., and Hirschhorn, J.N. (2008). Genome-wide association studies: potential next steps on a genetic journey. Human molecular genetics *17*, R156-R165.

McClellan, J., and King, M.-C. (2010). Genetic heterogeneity in human disease. Cell *141*, 210-217.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., and Daly, M. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research *20*, 1297-1303.

McKie, A.T. (2008). The role of Dcytb in iron metabolism: an update. Biochemical Society Transactions *36*, 1239-1241.

McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P., and Cunningham, F. (2010). Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics *26*, 2069-2070.

McMahon, G.M., Olden, M., Garnaas, M., Yang, Q., Liu, X., Hwang, S.-J., Larson, M.G., Goessling, W., Fox, C.S., and Consortium, C. (2014). Sequencing of LRP2 reveals multiple rare variants associated with urinary trefoil factor-3. Journal of the American Society of Nephrology *25*, 2896-2905.

McNabb, R.A., and McNabb, F.A. (1980). Physiological chemistry of uric acid: solubility, colloid and ion-binding properties. Comparative Biochemistry and Physiology Part A: Physiology *67*, 27-34.

Meigs, J.B., Mittleman, M.A., Nathan, D.M., Tofler, G.H., Singer, D.E., Murphy-Sheehy, P.M., Lipinska, I., D'agostino, R.B., and Wilson, P.W. (2000). Hyperinsulinemia, hyperglycemia, and impaired hemostasis: the Framingham Offspring Study. Jama *283*, 221-228.

Meng, J., Tan, W., Zhu, Y., Wang, F., Li, X., and Zhang, M. (2015). A coronary artery disease-associated SNP rs6903956 contributed to asymptomatic hyperuricemia susceptibility in Han Chinese. Lipids in health and disease *14*, 33.

Merriman, T.R., Cadzow, M., Tanner, C., Brown, M.A., Cremin, K., Janssen, M., Jansen, T., Joosten, L.A., Radstake, T., and Riches, P.L. (2015). A Genome-Wide Association Study Reveals Association of the Transferrin Receptor Locus with Gout. Arthritis & Rheumatology *67*

Merryweather-Clarke, A.T., Pointon, J.J., Shearman, J.D., and Robson, K. (1997). Global prevalence of putative haemochromatosis mutations. Journal of medical genetics *34*, 275-278.

Meshkani, R., Zargari, M., and Larijani, B. (2011). The relationship between uric acid and metabolic syndrome in normal glucose tolerance and normal fasting glucose subjects. Acta diabetologica *48*, 79-88.

Minaur, N., Sawyers, S., Parker, J., and Darmawan, J. (2004). Rheumatic disease in an Australian Aboriginal community in North Queensland, Australia. A WHO-ILAR COPCORD survey. The Journal of rheumatology *31*, 965-972.

Mitchell, B.D. (2012). An integrated encyclopedia of DNA elements in the human genome. Diabetes *61*, 3335-3336.

Modan, M., Halkin, H., Karasik, A., and Lusky, A. (1987). Elevated serum uric acid—a facet of hyperinsulinaemia. Diabetologia *30*, 713-718.

Mohandas, R., and Johnson, R.J. (2008). Uric acid levels increase risk for new-onset kidney disease. Journal of the American Society of Nephrology *19*, 2251-2253.

Mokdad, A.H., Ford, E.S., Bowman, B.A., Dietz, W.H., Vinicor, F., Bales, V.S., and Marks, J.S. (2003). Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. Jama *289*, 76-79.

Moos, T. (2002). Brain iron homeostasis. Danish medical bulletin *49*, 279-301.

Morcillo, S., Rojo-Martínez, G., Martín-Núñez, G.M., Gómez-Zumaquero, J.M., García-Fuentes, E., de ADANA, M.R., de la CRUZ ALMARAZ, M., and Soriguer, F. (2010). Trp64Arg polymorphism of the ADRB3 gene predicts hyperuricemia risk in a population from southern Spain. The Journal of rheumatology *37*, 417-421.

Moreno- Carralero, M.I., Muñoz- Muñoz, J.A., Cuadrado- Grande, N., López- Rodríguez, R., José Hernández- Alfaro, M., del- Castillo- Rueda, A., Enríquez- de- Salamanca, R., Méndez, M., and Morán-

Jiménez, M.J. (2014). A novel mutation in the SLC40A1 gene associated with reduced iron export in vitro. American journal of hematology *89*, 689-694.

Morgenthaler, S., and Thilly, W.G. (2007). A strategy to discover genes that carry multi-allelic or mono-allelic risk for common diseases: a cohort allelic sums test (CAST). Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis *615*, 28-56.

Moriarity, J.T., Folsom, A.R., Iribarren, C., Nieto, F.J., and Rosamond, W.D. (2000). Serum uric acid and risk of coronary heart disease: Atherosclerosis Risk in Communities (ARIC) Study. Annals of epidemiology *10*, 136-143.

Morris, A.P., Voight, B.F., Teslovich, T.M., Ferreira, T., Segre, A.V., Steinthorsdottir, V., Strawbridge, R.J., Khan, H., Grallert, H., and Mahajan, A. (2012). Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nature Genetics *44*, 981.

Morris, A.P., and Zeggini, E. (2010). An evaluation of statistical approaches to rare variant analysis in genetic association studies. Genetic epidemiology *34*, 188-193.

Moulin, S.R., Baldo, M.P., Souza, J.B., Luchi, W.M., Capingana, D.P., Magalhães, P., and Mill, J.G. (2017). Distribution of serum uric acid in black africans and its association with cardiovascular risk factors. The Journal of Clinical Hypertension *19*, 45-50.

Mozumdar, A., and Liguori, G. (2011). Persistent increase of prevalence of metabolic syndrome among US adults: NHANES III to NHANES 1999–2006. Diabetes care *34*, 216-219.

Muirden, K., and Senator, G. (1968). Iron in the synovial membrane in rheumatoid arthritis and other joint diseases. Annals of the rheumatic diseases *27*, 38.

Mukhopadhyay, I., Feingold, E., Weeks, D.E., and Thalamuthu, A. (2010). Association tests using kernel-based measures of multi- locus genotype similarity between individuals. Genetic epidemiology *34*, 213-221.

Muntane, J., Puig-Parellada, P., Fernandez, Y., Mitjavila, S., and Mitjavila, M. (1993). Antioxidant defenses and its modulation by iron in carrageenan-induced inflammation in rats. Clinica chimica acta *214*, 185-193.

Muscelli, E., Natali, A., Bianchi, S., Bigazzi, R., Galvan, A.Q., Sironi, A.M., Frascerra, S., Ciociaro, D., and Ferrannini, E. (1996). Effect of insulin on renal sodium and uric acid handling in essential hypertension. American journal of hypertension *9*, 746-752.

Mutombo, P.B.w.B., Yamasaki, M., Hamano, T., Isomura, M., Nabika, T., and Shiwaku, K. (2014). MC4R rs17782313 gene polymorphism was associated with glycated hemoglobin independently of its effect on BMI in Japanese: the Shimane COHRE study. Endocrine research *39*, 115-119.

Nagase, T., Aoki, A., Yamamoto, M., Yasuda, H., Kado, S., Nishikawa, M., Kugai, N., Akatsu, T., and Nagata, N. (1997). Lack of association between the Trp64Arg mutation in the β3-adrenergic receptor gene and obesity in Japanese men: a longitudinal analysis. The Journal of Clinical Endocrinology & Metabolism *82*, 1284-1287.

Nai, A., Pagani, A., Silvestri, L., Campostrini, N., Corbella, M., Girelli, D., Traglia, M., Toniolo, D., and Camaschella, C. (2011). TMPRSS6 rs855791 modulates hepcidin transcription in vitro and serum hepcidin levels in normal individuals. Blood *118*, 4459-4462.

Nakagawa, T., Hu, H., Zharikov, S., Tuttle, K.R., Short, R.A., Glushakova, O., Ouyang, X., Feig, D.I., Block, E.R., and Herrera-Acosta, J. (2006). A causal role for uric acid in fructose-induced metabolic syndrome. American Journal of Physiology-Renal Physiology *290*, F625-F631.

Nakanishi, N., Okamoto, M., Yoshida, H., Matsuo, Y., Suzuki, K., and Tatara, K. (2003). Serum uric acid and risk for development of hypertension and impaired fasting glucose or Type II diabetes in Japanese male office workers. European journal of epidemiology *18*, 523-530.

Nakayama, A., Matsuo, H., Shimizu, T., Takada, Y., Nakamura, T., Shimizu, S., Chiba, T., Sakiyama, M., Naito, M., and Morita, E. (2014). Common variants of a urate-associated gene LRP2 are not associated with gout susceptibility. Rheumatology International *34*, 473-476.

Ndrepepa, G., Braun, S., Haase, H.-U., Schulz, S., Ranftl, S., Hadamitzky, M., Mehilli, J., Schömig, A., and Kastrati, A. (2012). Prognostic value of uric acid in patients with acute coronary syndromes. The American journal of cardiology *109*, 1260-1265.

Nejatinamini, S., Ataie-Jafari, A., Qorbani, M., Nikoohemat, S., Kelishadi, R., Asayesh, H., and Hosseini, S. (2015). Association between serum uric acid level and metabolic syndrome components. Journal of Diabetes & Metabolic Disorders *14*, 70.

Nejentsev, S., Walker, N., Riches, D., Egholm, M., and Todd, J.A. (2009). Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. Science *324*, 387-389.

Nemeth, E., and Ganz, T. (2009). The role of hepcidin in iron metabolism. Acta haematologica *122*, 78-86.

Nemeth, E., Tuttle, M.S., Powelson, J., Vaughn, M.B., Donovan, A., Ward, D.M., Ganz, T., and Kaplan, J. (2004). Hepcidin regulates cellular iron efflux by binding to ferroportin and inducing its internalization. Science *306*, 2090-2093.

Ng, M., Fleming, T., Robinson, M., Thomson, B., Graetz, N., Margono, C., Mullany, E.C., Biryukov, S., Abbafati, C., and Abera, S.F. (2014). Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013. The Lancet *384*, 766-781.

Nieto, F.J., Iribarren, C., Gross, M.D., Comstock, G.W., and Cutler, R.G. (2000). Uric acid and serum antioxidant capacity: a reaction to atherosclerosis? Atherosclerosis *148*, 131-139.

Nikpay, M., Anuj, G., Hong-Hee, W., Hall, L.M., Willenborg, C., Kanoni, S., Saleheen, D., Kyriakou, T., Nelson, C.P., Hopewell, J.C*., et al.* (2015). A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nature Genetics.

Nishida, Y. (1992). Relation between creatinine and uric acid excretion. Annals of the rheumatic diseases *51*, 101-102.

Njajou, O., Alizadeh, B., Aulchenko, Y., Zillikens, M., Pols, H., Oostra, B., Swinkels, D., and Van Duijn, C. (2006). Heritability of serum iron, ferritin and transferrin saturation in a genetically isolated population, the Erasmus Rucphen Family (ERF) Study. Human heredity *61*, 222-228.

Noman, A., Ang, D.S., Ogston, S., Lang, C.C., and Struthers, A.D. (2010). Effect of high-dose allopurinol on exercise in patients with chronic stable angina: a randomised, placebo controlled crossover trial. The Lancet *375*, 2161-2167.

Nuki, G. (2014). An appraisal of the 2012 American College of Rheumatology Guidelines for the Management of Gout. Current opinion in rheumatology *26*, 152-161.

NZMH (2014). Managing Chronic Kidney Disease in Primary Care: A National Consensus Statement. New Zealand Ministry of Health.

O'Connell, J.F., and Allen, J. (2004). Dating the colonization of Sahul (Pleistocene Australia–New Guinea): a review of recent research. Journal of Archaeological Science *31*, 835-853.

O'Donnell, S., Rusu, C., Hawker, G.A., Bernatsky, S., McRae, L., Canizares, M., MacKay, C., and Badley, E.M. (2015). Arthritis has an impact on the daily lives of Canadians young and old: results from a population-based survey. BMC musculoskeletal disorders *16*, 230.

O'Neill, S., and O'Driscoll, L. (2015). Metabolic syndrome: a closer look at the growing epidemic and its associated pathologies. obesity reviews *16*, 1-12.

Obregón, A., Amador, P., Valladares, M., Weisstaub, G., Burrows, R., and Santos, J. (2010). Melanocortin-3 receptor gene variants: association with childhood obesity and eating behavior in Chilean families. Nutrition *26*, 760-765.

Oda, M., Satta, Y., Takenaka, O., and Takahata, N. (2002). Loss of urate oxidase activity in hominoids and its evolutionary implications. Molecular biology and evolution *19*, 640-653.

Ogino, K., Kato, M., Furuse, Y., Kinugasa, Y., Ishida, K., Osaki, S., Kinugawa, T., Igawa, O., Hisatome, I., and Shigemasa, C. (2009). Uric acid lowering treatment with benzbromarone in patients with heart failure: a double-blind placebo-controlled cross-over preliminary study. Circulation: Heart Failure, CIRCHEARTFAILURE. 109.868604.

Ogryzlo, M.A. (1965). Hyperuricemia induced by high fat diets and starvation. Arthritis & Rheumatism *8*, 799-822.

Ogura, Y., Bonen, D.K., Inohara, N., Nicolae, D.L., Chen, F.F., Ramos, R., Britton, H., Moran, T., Karaliuskas, R., and Duerr, R.H. (2001). A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature *411*, 603-606.

Ohno, I., Hosoya, T., Gomi, H., Ichida, K., Okabe, H., and Hikita, M. (2001). Serum uric acid and renal prognosis in patients with IgA nephropathy. Nephron *87*, 333-339.

Okafor, C.I. (2012). The metabolic syndrome in Africa: Current trends. Indian journal of endocrinology and metabolism *16*, 56.

Okura, T., Higaki, J., Kurata, M., Irita, J., Miyoshi, K.-i., Yamazaki, T., Hayashi, D., Kohro, T., Nagai, R., and Investigators, J.S. (2009). Elevated serum uric acid is an independent predictor for cardiovascular events in patients with severe coronary artery stenosis. Circulation Journal *73*, 885-891.

Ollier, W., Sprosen, T., and Peakman, T. (2005). UK Biobank: from concept to reality.

Orlov, S.N., Koltsova, S.V., Kapilevich, L.V., Gusakova, S.V., and Dulin, N.O. (2015). NKCC1 and NKCC2: The pathogenetic role of cation-chloride cotransporters in hypertension. Genes & Diseases *2*, 186-196.

Öztürk, M.A., Kaya, A., Şenel, S., Dönmez, S., Balkarlı, A., Çobankara, V., Erhan, Ç., Sayarlıoğlu, M., Ugan, Y., and Tunç, Ş.E. (2013). Demographic and clinical features of gout patients in Turkey: a multicenter study. Rheumatology International *33*, 847-852.

Pal, S., and Ellis, V. (2010). The chronic effects of whey proteins on blood pressure, vascular function, and inflammatory markers in overweight individuals. Obesity *18*, 1354-1359.

Palmer, T.M., Nordestgaard, B.G., Benn, M., Tybjærg-Hansen, A., Smith, G.D., Lawlor, D.A., and Timpson, N.J. (2013). Association of plasma uric acid with ischaemic heart disease and blood pressure: mendelian randomisation analysis of two large cohorts. Bmj *347*, f4262.

Pan, H., Guo, J., and Su, Z. (2014). Advances in understanding the interrelations between leptin resistance and obesity. Physiology & behavior *130*, 157-169.

Panoutsopoulou, K., Tachmazidou, I., and Zeggini, E. (2013). In search of low-frequency and rare variants affecting complex traits. Human molecular genetics *22*, R16-R21.

Park, H.S., Shin, E.S., and Lee, J.E. (2008). Genotypes and haplotypes of β 2-adrenergic receptor and parameters of the metabolic syndrome in Korean adolescents. Metabolism *57*, 1064-1070.

Pattaro, C., Teumer, A., Gorski, M., Chu, A.Y., Li, M., Mijatovic, V., Garnaas, M., Tin, A., Sorice, R., and Li, Y. (2016). Genetic associations at 53 loci highlight cell types and biological pathways relevant for kidney function. Nature communications *7*.

Peden, D.B., Hohman, R., Brown, M.E., Mason, R.T., Berkebile, C., Fales, H.M., and Kaliner, M.A. (1990). Uric acid is a major antioxidant in human nasal airway secretions. Proceedings of the National Academy of Sciences *87*, 7638-7642.

Pei, S.-N., Ma, M.-C., You, H.-L., Fu, H.-C., Kuo, C.-Y., Rau, K.-M., Wang, M.-C., and Lee, C.-T. (2014). TMPRSS6 rs855791 polymorphism influences the susceptibility to iron deficiency anemia in women at reproductive age. International journal of medical sciences *11*, 614.

Peláez-Ballestas, I., Cuevas, C.H., Burgos-Vargas, R., Roque, L.H., Terán, L., Espinoza, J., Esquivel-Valerio, J.A., Goycochea-Robles, M.V., Aceves, F.J., and Bernard, A.G. (2010). Diagnosis of chronic gout: evaluating the american college of rheumatology proposal, European league against rheumatism recommendations, and clinical judgment. The Journal of rheumatology *37*, 1743-1748.

Pérez-Magán, E., de Lope, Á.R., Ribalta, T., Ruano, Y., Campos-Martín, Y., Pérez-Bautista, G., García, J.F., García-Claver, A., Fiaño, C., and Hernández-Moneo, J.-L. (2010). Differential expression profiling analyses identifies downregulation of 1p, 6q, and 14q genes and overexpression of 6p histone cluster 1 genes as markers of recurrence in meningiomas. Neuro-oncology, noq081.

Perlstein, T.S., Gumieniak, O., Williams, G.H., Sparrow, D., Vokonas, P.S., Gaziano, M., Weiss, S.T., and Litonjua, A.A. (2006). Uric acid and the development of hypertension. Hypertension *48*, 1031-1036.

Pfister, R., Barnes, D., Luben, R., Forouhi, N., Bochud, M., Khaw, K.-T., Wareham, N., and Langenberg, C. (2011). No evidence for a causal link between uric acid and type 2 diabetes: a Mendelian randomisation approach. Diabetologia *54*, 2561-2569.

Phipps-Green, A., Merriman, M., Topless, R., Altaf, S., Montgomery, G., Franklin, C., Jones, G., van Rij, A., White, D., and Stamp, L. (2014). Twenty-eight loci that influence serum urate levels: analysis of association with gout. Annals of the rheumatic diseases, annrheumdis-2014-205877.

Phipps-Green, A., Merriman, M., Topless, R., Altaf, S., Montgomery, G., Franklin, C., Jones, G., Van Rij, A., White, D., and Stamp, L. (2016). Twenty-eight loci that influence serum urate levels: analysis of association with gout. Annals of the rheumatic diseases *75*, 124-130.

Phipps-Green, A.J., Hollis-Moffatt, J.E., Dalbeth, N., Merriman, M.E., Topless, R., Gow, P.J., Harrison, A.A., Highton, J., Jones, P.B., and Stamp, L.K. (2010). A strong role for the ABCG2 gene in susceptibility to gout in New Zealand Pacific Island and Caucasian, but not Māori, case and control sample sets. Human molecular genetics, ddq412.

Picavet, H., and Hazes, J. (2003). Prevalence of self reported musculoskeletal diseases is high. Annals of the rheumatic diseases *62*, 644-650.

Pichler, I., Fabiola Del Greco, M., Gögele, M., Lill, C.M., Bertram, L., Do, C.B., Eriksson, N., Foroud, T., Myers, R.H., and Nalls, M. (2013). Serum iron levels and the risk of Parkinson disease: a mendelian randomization study. PLoS Medicine *10*, e1001462.

Pierce, B.L., and Burgess, S. (2013). Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. American journal of epidemiology, kwt084.

Polašek, O., Marušić, A., Rotim, K., Hayward, C., Vitart, V., Huffman, J., Campbell, S., Janković, S., Boban, M., and Biloglav, Z. (2009). Genome-wide association study of anthropometric traits in Korčula Island, Croatia. Croatian medical journal *50*, 7-16.

Powell, L., and Emmerson, B. (1966). Haemosiderosis associated with xanthine oxidase inhibition. The Lancet *287*, 239-240.

Powell, L.M., Wallis, S.C., Pease, R.J., Edwards, Y.H., Knott, T.J., and Scott, J. (1987). A novel form of tissue-specific RNA processing produces apolipoprotein-B48 in intestine. Cell *50*, 831-840.

Powell, L.W. (1970). Changing concepts in haemochromatosis. Postgraduate Medical Journal *46*, 200.

Powell, L.W., Seckington, R.C., and Deugnier, Y. (2016). Haemochromatosis. The Lancet.

Prior, I., and Rose, B. (1966). Uric acid, gout and public health in the South Pacific. The New Zealand medical journal *65*, 295.

Prior, I., Rose, B., Harvey, H., and Davidson, F. (1966). Hyperuricaemia, gout, and diabetic abnormality in Polynesian people. The Lancet *287*, 333-338.

Pritchard, J.K. (2001). Are rare variants responsible for susceptibility to complex diseases? The American Journal of Human Genetics *69*, 124-137.

Puig, J.G., and Martinez, M.A. (2008). Hyperuricemia, gout and the metabolic syndrome. Current opinion in rheumatology *20*, 187-191.

Qi, L., Kraft, P., Hunter, D.J., and Hu, F.B. (2008). The common obesity variant near MC4R gene is associated with higher intakes of total energy and dietary fat, weight change and diabetes risk in women. Human molecular genetics *17*, 3502-3508.

Rafey, M.A., Lipkowitz, M.S., Leal-Pinto, E., and Abramson, R.G. (2003). Uric acid transport. Current opinion in nephrology and hypertension *12*, 511-516.

Raffin, S.B., Woo, C.H., Roost, K.T., Price, D.C., and Schmid, R. (1974). Intestinal absorption of hemoglobin iron-heme cleavage by mucosal heme oxygenase. Journal of Clinical Investigation *54*, 1344.

Raffin-Sanson, M., De Keyzer, Y., and Bertagna, X. (2003). Proopiomelanocortin, a polypeptide precursor with multiple functions: from physiology to pathological conditions. European Journal of Endocrinology *149*, 79-90.

Ramakrishna, V., and Jailkhani, R. (2008). Oxidative stress in non-insulin-dependent diabetes mellitus (NIDDM) patients. Acta diabetologica *45*, 41-46.

Ranasinghe, P., Mathangasinghe, Y., Jayawardena, R., Hills, A., and Misra, A. (2017). Prevalence and trends of metabolic syndrome among adults in the asia-pacific region: a systematic review. BMC Public Health *17*, 101.

Rasheed, H., Hughes, K., Flynn, T.J., and Merriman, T.R. (2014). Mendelian randomization provides no evidence for a causal role of serum urate in increasing serum triglyceride levels. Circulation: Cardiovascular Genetics, CIRCGENETICS. 114.000556.

Rasheed, H., Phipps-Green, A., Topless, R., Hollis-Moffatt, J.E., Hindmarsh, J.H., Franklin, C., Dalbeth, N., Jones, P.B., White, D.H., and Stamp, L.K. (2013a). Association of the lipoprotein receptor-related protein 2 gene with gout and non-additive interaction with alcohol consumption. Arthritis research & therapy *15*, 1.

Rasheed, H., Phipps-Green, A., Topless, R., Hollis-Moffatt, J.E., Hindmarsh, J.H., Franklin, C., Dalbeth, N., Jones, P.B., White, D.H., and Stamp, L.K. (2013b). Association of the lipoprotein receptor-related protein 2 gene with gout and non-additive interaction with alcohol consumption. Arthritis research & therapy *15*, R177.

RCore, T. (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reaven, G.M. (1988). Role of insulin resistance in human disease. Diabetes *37*, 1595-1607.

Rediger, A., Piechowski, C., Yi, C., Tarnow, P., Strotmann, R., Grüters, A., Krude, H., Schöneberg, T., Tschöp, M., and Kleinau, G. (2012). Mutually opposite signal modulation by hypothalamic heterodimerization of ghrelin-and melanocortin-3 receptors. Experimental and Clinical Endocrinology & Diabetes *120*, T4.

Reich, D.E., and Lander, E.S. (2001). On the allelic spectrum of human disease. TRENDS in Genetics *17*, 502-510.

Reis, C., and de Queiroz, M.V. (2014). Prevalence of self-reported rheumatic diseases in a Portuguese population. Acta Reumatológica Portuguesa *39*.

Reizenstein, P. (1979). Hemoglobin fortification of food and prevention of iron deficiency with heme iron. Acta medica Scandinavica Supplementum *629*, 1.

Rho, Y.H., Choi, S.J., Lee, Y.H., Ji, J.D., Choi, K.M., Baik, S.H., Chung, S.-h., Kim, C.-G., Choe, J.-Y., and Lee, S.W. (2005). The prevalence of metabolic syndrome in patients with gout: a multicenter study. Journal of Korean medical science *20*, 1029-1033.

Rho, Y.H., Choi, S.J., Lee, Y.H., Ji, J.D., and Song, G.G. (2007). The association between hyperuricemia and the Trp64Arg polymorphism of the beta-3 adrenergic receptor. Rheumatology international *27*, 835.

Rho, Y.H., Lu, N., Peloquin, C.E., Man, A., Zhu, Y., Zhang, Y., and Choi, H.K. (2016). Independent impact of gout on the risk of diabetes mellitus among women and men: a population-based, BMI-matched cohort study. Annals of the rheumatic diseases *75*, 91-95.

Richardson, R. (1991). Treatment of Asymptomatic Hyperuricemia: When do you have a good reason to provide therapy? Canadian Family Physician *37*, 459.

Richet, G. (1995). The chemistry of urinary stones around 1800: a first in clinical chemistry. Kidney international *48*, 876-886.

Richette, P., Doherty, M., Pascual, E., Barskova, V., Becce, F., Castañeda-Sanabria, J., Coyfish, M., Guillo, S., Jansen, T., and Janssens, H. (2017). 2016 updated EULAR evidence-based recommendations for the management of gout. Annals of the Rheumatic Diseases *76*, 29-42.

Rivas, M.A., Beaudoin, M., Gardet, A., Stevens, C., Sharma, Y., Zhang, C.K., Boucher, G., Ripke, S., Ellinghaus, D., and Burtt, N. (2011). Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. Nature Genetics *43*, 1066-1073.

Robert, H., and Xiaole, K. (2013). Chapter 8. Iron: Effect of Overload and Deficiency. In Interrelations between Essential Metal Ions and Human Diseases: Metal Ions in Life Sciences (Springer), pp. 229-294.

Robinson, P., Choi, H., Do, R., and Merriman, T. (2016). Insight into rheumatological cause and effect through the use of Mendelian randomization. Nature Reviews Rheumatology.

Robinson, P.C., and Horsburgh, S. (2014). Gout: joints and beyond, epidemiology, clinical features, treatment and co-morbidities. Maturitas *78*, 245-251.

Robinson, P.C., Merriman, T.R., Herbison, P., and Highton, J. (2012). Hospital admissions associated with gout and their co-morbidities in New Zealand and England 1999–2009. Rheumatology, kes253.

Robinson, P.C., Taylor, W.J., and Dalbeth, N. (2015). An observational study of gout prevalence and quality of care in a national Australian general practice population. The Journal of rheumatology *42*, 1702-1707.

Roddy, E. (2016). I155 2016 Revised British Society for Rheumatology/British Health Professionals in Rheumatology Gout Management Guideline. Rheumatology *55*, i30-i30.

Roncal-Jimenez, C.A., Lanaspa, M.A., Rivard, C.J., Nakagawa, T., Sanchez-Lozada, L.G., Jalal, D., Andres-Hernando, A., Tanabe, K., Madero, M., and Li, N. (2011). Sucrose induces fatty liver and pancreatic inflammation in male breeder rats independent of excess energy intake. Metabolism *60*, 1259-1270.

Roseff, R., Wohlgethan, J., Sipe, J., and Canoso, J. (1987). The acute phase response in gout. The Journal of rheumatology *14*, 974-977.

Rouault, T.A. (2003). How mammals acquire and distribute iron needed for oxygen-based metabolism. PLoS Biology *1*.

Roughley, M.J., Belcher, J., Mallen, C.D., and Roddy, E. (2015). Gout and risk of chronic kidney disease and nephrolithiasis: meta-analysis of observational studies. Arthritis research & therapy *17*, 90.

Ruggiero, C., Cherubini, A., Ble, A., Bos, A.J., Maggio, M., Dixit, V.D., Lauretani, F., Bandinelli, S., Senin, U., and Ferrucci, L. (2006). Uric acid and inflammatory markers. European heart journal *27*, 1174-1181.

Saito, A., Takeda, T., Hama, H., Oyama, Y., Hosaka, K., Tanuma, A., Kaseda, R., Ueno, M., Nishi, S., and Ogasawara, S. (2005). Role of megalin, a proximal tubular endocytic receptor, in the pathogenesis of diabetic and metabolic syndrome- related nephropathies: protein metabolic overload hypothesis. Nephrology *10*, S26-S31.

Saito, H. (2014). Metabolism of iron stores. Nagoya journal of medical science *76*, 235-254.

Saito, H., and Hayashi, H. (2015). Transformation rate between ferritin and hemosiderin assayed by serum ferritin kinetics in patients with normal iron stores and iron overload. Nagoya journal of medical science *77*, 571.

Saito, H., Hayashi, H., Tomita, A., Ohashi, H., Maeda, H., and Naoe, T. (2013). Increasing and decreasing phases of ferritin and hemosiderin iron determined by serum ferritin kinetics. Nagoya journal of medical science *75*, 213-223.

Sajeevan, K., Lonappan, L., George, B., and Warrier, P.K. (2016). Serum Iron Parameters in Alcoholic Cirrhosis, Cryptogenic Cirrhosis, Chronic Hepatitis B and Chronic Hepatitis C. Journal of Evidence based Medicine and Healthcare *3*, 5157-5161.

Sambrook, J., and Russell, D.W. (2006). Purification of nucleic acids by extraction with phenol: chloroform. Cold Spring Harbor Protocols *2006*, pdb. prot4455.

San Martin, C.D., Garri, C., Pizarro, F., Walter, T., Theil, E.C., and Núñez, M.T. (2008). Caco-2 intestinal epithelial cells absorb soybean ferritin by μ2 (AP2)-dependent endocytosis. The Journal of nutrition *138*, 659-666.

Sánchez-Lozada, L.G., Tapia, E., Bautista-García, P., Soto, V., Ávila-Casado, C., Vega-Campos, I.P., Nakagawa, T., Zhao, L., Franco, M., and Johnson, R.J. (2008). Effects of febuxostat on metabolic and renal alterations in rats with fructose-induced metabolic syndrome. American Journal of Physiology-Renal Physiology *294*, F710-F718.

Sanna, S., Li, B., Mulas, A., Sidore, C., Kang, H.M., Jackson, A.U., Piras, M.G., Usala, G., Maninchedda, G., and Sassu, A. (2011). Fine mapping of five loci associated with low-density lipoprotein cholesterol detects variants that double the explained heritability. PLoS Genetics *7*, e1002198.

Savage, D.B., Petersen, K.F., and Shulman, G.I. (2005). Mechanisms of insulin resistance in humans and possible links with inflammation. Hypertension *45*, 828-833.

Savastano, D.M., Tanofsky-Kraff, M., Han, J.C., Ning, C., Sorg, R.A., Roza, C.A., Wolkoff, L.E., Anandalingam, K., Jefferson-George, K.S., and Figueroa, R.E. (2009). Energy intake and energy expenditure among children with polymorphisms of the melanocortin-3 receptor. The American journal of clinical nutrition *90*, 912-920.

Schiöth, H.B., Yook, P., Muceniece, R., Wikberg, J.E., and Szardenings, M. (1998). Chimeric melanocortin MC1 and MC3 receptors: identification of domains participating in binding of melanocyte-stimulating hormone peptides. Molecular pharmacology *54*, 154-161.

Schork, N.J., Murray, S.S., Frazer, K.A., and Topol, E.J. (2009). Common vs. rare allele hypotheses for complex diseases. Current opinion in genetics & development *19*, 212-219.

Schunkert, H., König, I.R., Kathiresan, S., Reilly, M.P., Assimes, T.L., Holm, H., Preuss, M., Stewart, A.F., Barbalic, M., and Gieger, C. (2011). Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nature Genetics *43*, 333-338.

Scott, F.W., Trick, K.D., Stavric, B., Braaten, J.T., and Siddiqui, Y. (1981). Uric Acid-Induced Decrease in Rat Insulin Secretion 1 2. Proceedings of the Society for Experimental Biology and Medicine *166*, 123-128.

Selya, A.S., Rose, J.S., Dierker, L.C., Hedeker, D., and Mermelstein, R.J. (2012). A practical guide to calculating Cohen's f2, a measure of local effect size, from PROC MIXED. Frontiers in psychology *3*, 111.

Sezer, M.T., Akin, H., Demir, M., Erturk, J., Aydin, Z.D., Savik, E., and Tunc, N. (2007). The effect of serum albumin level on iron-induced oxidative stress in chronic renal failure patients. Journal of nephrology *20*, 196.

Sharp, P.A. (2010). Intestinal iron absorption: regulation by dietary & systemic factors. Int J Vitam Nutr Res *80*, 231-242.

Shekelle, P.G., Newberry, S.J., FitzGerald, J.D., Motala, A., O'Hanlon, C.E., Tariq, A., Okunogbe, A., Han, D., and Shanman, R. (2017). Management of Gout: A Systematic Review in Support of an American College of Physicians Clinical Practice GuidelineSystematic Review in Support of ACP Guideline on Management of Gout. Annals of Internal Medicine.

Shi, Y., Chen, W., Jalal, D., Li, Z., Mao, H., Yang, Q., Johnson, R., and Yu, X. (2011). Clinical outcome of hyperuricemia in IgA nephropathy: a retrospective cohort study and randomized controlled trial. Kidney and Blood Pressure Research *35*, 153-160.

Shoden, A., Gabrio, B.W., and Finch, C.A. (1953). The relationship between ferritin and hemosiderin in rabbits and man. Journal of Biological Chemistry *204*, 823-830.

Shrank, W.H., Patrick, A.R., and Brookhart, M.A. (2011). Healthy user and related biases in observational studies of preventive interventions: a primer for physicians. Journal of general internal medicine *26*, 546-550.

Sicras-Mainar, A., Navarro-Artieda, R., and Ibáñez-Nolla, J. (2013). Resource use and economic impact of patients with gout: a multicenter, population-wide study. Reumatología Clínica (English Edition) *9*, 94-100.

Simmons, D., and Thompson, C.F. (2004). Prevalence of the metabolic syndrome among adult New Zealanders of Polynesian and European descent. Diabetes Care *27*, 3002-3004.

Siow, D., Sunkara, M., Morris, A., and Wattenberg, B. (2015). Regulation of de novo sphingolipid biosynthesis by the ORMDL proteins and sphingosine kinase-1. Advances in biological regulation *57*, 42-54.

Sirota, J.C., McFann, K., Targher, G., Johnson, R.J., Chonchol, M., and Jalal, D.I. (2013). Elevated serum uric acid levels are associated with non-alcoholic fatty liver disease independently of metabolic syndrome features in the United States: Liver ultrasound data from the National Health and Nutrition Examination Survey. Metabolism *62*, 392-399.

Sjöström, L., Lindroos, A.-K., Peltonen, M., Torgerson, J., Bouchard, C., Carlsson, B., Dahlgren, S., Larsson, B., Narbro, K., and Sjöström, C.D. (2004). Lifestyle, diabetes, and cardiovascular risk factors 10 years after bariatric surgery. New England Journal of Medicine *351*, 2683-2693.

Skinner, T. (1997). National Health Survey: Summary of Results, Australia, 1995 (Canberra: Australian Bureau of Statistics).

Sliem, H.A., Ahmed, S., Nemr, N., and El-Sherif, I. (2012). Metabolic syndrome in the Middle East. Indian journal of endocrinology and metabolism *16*, 67.

Sluijs, I., Holmes, M.V., van der Schouw, Y.T., Beulens, J.W., Asselbergs, F.W., Huerta, J.M., Palmer, T.M., Arriola, L., Balkau, B., and Barricarte, A. (2015). A Mendelian randomization study of circulating uric acid and type 2 diabetes. Diabetes, db140742.

Smith, E., Hoy, D., Cross, M., Merriman, T.R., Vos, T., Buchbinder, R., Woolf, A., and March, L. (2014). The global burden of gout: estimates from the Global Burden of Disease 2010 study. Annals of the rheumatic diseases *73*, 1470-1476.

Smith, E., and March, L. (2015). Global prevalence of hyperuricemia: a systematic review of population-based epidemiological studies. Arthritis & Rheumatology *67*, 2690-2692.

Smith, G.D. (2008). Assessing intrauterine influences on offspring health outcomes: can epidemiological studies yield robust findings? Basic & clinical pharmacology & toxicology *102*, 245-256.

Smith, G.D., and Ebrahim, S. (2001). Epidemiology-is it time to call it a day? International Journal of Epidemiology *30*, 1-11.

Smith, G.D., and Ebrahim, S. (2003). 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? (IEA).

Smith, G.D., and Ebrahim, S. (2005). What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? BMJ: British Medical Journal *330*, 1076.

Smith, G.D., and Ebrahim, S. (2008). Mendelian randomization: genetic variants as instruments for strengthening causal inference in observational studies.

Smith, G.D., and Hemani, G. (2014). Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Human molecular genetics *23*, R89-R98.

Snaith, M.L. (2003). Gout (eLS. Chichester: John Wiley & Sons, Ltd.).

So, A. (2008). Developments in the scientific and clinical understanding of gout. Arthritis research & therapy *10*, 221.

Solovieff, N., Cotsapas, C., Lee, P.H., Purcell, S.M., and Smoller, J.W. (2013). Pleiotropy in complex traits: challenges and strategies. Nature Reviews Genetics *14*, 483-495.

Soltani, Z., Rasheed, K., Kapusta, D.R., and Reisin, E. (2013). Potential role of uric acid in metabolic syndrome, hypertension, kidney injury, and cardiovascular diseases: is it time for reappraisal? Current hypertension reports *15*, 175-181.

Sookoian, S., and Pirola, C.J. (2011). Metabolic syndrome: from the genetics to the pathophysiology. Current hypertension reports *13*, 149-157.

Sorensen, L., and Levinson, D. (1975). Origin and extrarenal elimination of uric acid in man. Nephron *14*, 7-20.

Speliotes, E.K., Willer, C.J., Berndt, S.I., Monda, K.L., Thorleifsson, G., Jackson, A.U., Allen, H.L., Lindgren, C.M., Luan, J.a., and Mägi, R. (2010). Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nature Genetics *42*, 937-948.

Srivastava, A., Mittal, B., Prakash, J., Narain, V.S., Natu, S., and Srivastava, N. (2014). Evaluation of MC4R [RS17782313, RS17700633], AGRP [RS3412352] and POMC [RS1042571] polymorphisms with obesity in Northern India. Oman medical journal *29*, 114.

Stamp, L.K., and Chapman, P.T. (2013). Gout and its comorbidities: implications for therapy. Rheumatology *52*, 34-44.

Stonehuerner, J., Ghio, A., Kennedy, T., Skinner, K., Skinner, H., and Parks, D. (1998). Elevations in xanthine oxidase (XO) activity after silica exposure corresponds to complexed iron concentrations. Am J Respir Crit Care Med *157*, A232.

Strazzullo, P., Iacone, R., Siani, A., Cappuccio, F.P., Russo, O., Barba, G., Barbato, A., D'elia, L., Trevisan, M., and Farinaro, E. (2001). Relationship of the Trp64Arg polymorphism of the beta3-adrenoceptor gene to central adiposity and high blood pressure: interaction with age. Cross-sectional and longitudinal findings of the Olivetti Prospective Heart Study. Journal of hypertension *19*, 399-406.

Suazo, J., Hodgson, M.I., Obregón, A.M., Valladares, M., Weisstaub, G., Amador, P., and Santos, J.L. (2013). Prevalence of metabolic syndrome in obese Chilean children and association with gene variants of the leptin-melanocortin system. Journal of Pediatric Endocrinology and Metabolism *26*, 1131-1139.

Sulem, P., Gudbjartsson, D.F., Walters, G.B., Helgadottir, H.T., Helgason, A., Gudjonsson, S.A., Zanon, C., Besenbacher, S., Bjornsdottir, G., and Magnusson, O.T. (2011). Identification of low-frequency variants associated with gout and serum uric acid levels. Nature Genetics *43*, 1127-1130.

Suliman, M.E., Johnson, R.J., García-López, E., Qureshi, A.R., Molinaei, H., Carrero, J.J., Heimbürger, O., Bárány, P., Axelsson, J., and Lindholm, B. (2006). J-shaped mortality relationship for uric acid in CKD. American Journal of Kidney Diseases *48*, 761-771.

Summerhayes, G.R., Leavesley, M., Fairbairn, A., Mandui, H., Field, J., Ford, A., and Fullagar, R. (2010). Human adaptation and plant use in highland New Guinea 49,000 to 44,000 years ago. Science *330*, 78-81.

Swerdlow, D.I., Kuchenbaecker, K.B., Shah, S., Sofat, R., Holmes, M.V., White, J., Mindell, J.S., Kivimaki, M., Brunner, E.J., and Whittaker, J.C. (2016). Selecting instruments for Mendelian randomization in the wake of genome-wide association studies. International journal of epidemiology *45*, 1600-1616.

Syrjänen, J., Mustonen, J., and Pasternack, A. (2000). Hypertriglyceridaemia and hyperuricaemia are risk factors for progression of IgA nephropathy. Nephrology Dialysis Transplantation *15*, 34-42.

Tang, L., Kubota, M., Nagai, A., Mamemoto, K., and Tokuda, M. (2010). Hyperuricemia in obese children and adolescents: the relationship with metabolic syndrome. Pediatric reports *2*, 12.

Tange, O. (2011). Gnu parallel-the command-line power tool. The USENIX Magazine *36*, 42-47.

Tao, Y.-X., and Segaloff, D.L. (2004). Functional characterization of melanocortin-3 receptor variants identify a loss-of-function mutation involving an amino acid critical for G protein-coupled receptor activation. The Journal of Clinical Endocrinology & Metabolism *89*, 3936-3942.

Tapia, E., Cristóbal, M., García-Arroyo, F.E., Soto, V., Monroy-Sánchez, F., Pacheco, U., Lanaspa, M.A., Roncal-Jiménez, C.A., Cruz-Robles, D., and Ishimoto, T. (2013). Synergistic effect of uricase blockade plus physiological amounts of fructose-glucose on glomerular hypertension and oxidative stress in rats. American Journal of Physiology-Renal Physiology *304*, F727-F736.

Taupin, D., and Podolsky, D.K. (2003). Trefoil factors: initiators of mucosal healing. Nature Reviews Molecular Cell Biology *4*, 721-732.

Tayebi, N., Ke, T., Foo, J.N., Friedlander, Y., Liu, J., and Heng, C.-K. (2013). Association of single nucleotide polymorphism rs6903956 on chromosome 6p24. 1 with coronary artery disease and lipid levels in different ethnic groups of the Singaporean population. Clinical biochemistry *46*, 755-759.

Taylor, A.E., Davies, N.M., Ware, J.J., VanderWeele, T., Smith, G.D., and Munafò, M.R. (2014). Mendelian randomization in health research: using appropriate genetic variants and avoiding biased estimates. Economics & Human Biology *13*, 99-106.

Taylor, H.A.J. (2005). The Jackson Heart Study: an overview. Ethn Dis, S6-1-S6-3.

Teng, G.G., Ang, L.-W., Saag, K.G., Mimi, C.Y., Yuan, J.-M., and Koh, W.-P. (2012). Mortality due to coronary heart disease and kidney disease among middle-aged and elderly men and women with gout in the Singapore Chinese Health Study. Annals of the rheumatic diseases *71*, 924-928.

Theil, E.C. (1987). Ferritin: structure, gene regulation, and cellular function in animals, plants, and microorganisms. Annual review of biochemistry *56*, 289-315.

Thornley, S., Chan, W.C., Crengle, S., Riddell, T., Ameratunga, S., Mehta, S., Gentles, D., Wells, S., Marshall, R., and Jackson, R. (2011). Sociodemographic differences in prevalence of diagnosed coronary heart disease in New Zealand estimated from linked national health records. The New Zealand Medical Journal (Online) *124*.

Trevaskis, J.L., Gawronska- Kozak, B., Sutton, G.M., McNeil, M., Stephens, J.M., Smith, S.R., and Butler, A.A. (2007). Role of adiponectin and inflammation in insulin resistance of Mc3r and Mc4r knockout mice. Obesity *15*, 2664-2672.

Tschritter, O., Haupt, A., Preissl, H., Ketterer, C., Hennige, A.M., Sartorius, T., Machicao, F., Fritsche, A., and Häring, H.-U. (2011). An obesity risk SNP (rs17782313) near the MC4R gene is associated with cerebrocortical insulin resistance in humans. Journal of obesity *2011*.

Tsushima, Y., Nishizawa, H., Tochino, Y., Nakatsuji, H., Sekimoto, R., Nagao, H., Shirakura, T., Kato, K., Imaizumi, K., and Takahashi, H. (2013). Uric acid secretion from adipose tissue and its increase in obesity. Journal of Biological Chemistry *288*, 27138-27149.

Underwood, M. (2006). Diagnosis and management of gout. BMJ: British Medical Journal *332*, 1315.

Valette, M., Bellisle, F., Carette, C., Poitou, C., Dubern, B., Paradis, G., Hercberg, S., Muzard, L., Clément, K., and Czernichow, S. (2013). Eating behaviour in obese patients with melanocortin-4 receptor mutations: a literature review. International journal of obesity *37*, 1027-1035.

Valve, R., Heikkinen, S., Rissanen, A., Laakso, M., and Uusitupa, M. (1998). Synergistic effect of polymorphisms in uncoupling protein 1 and β3-adrenergic receptor genes on basal metabolic rate in obese Finns. Diabetologia *41*, 357-361.

Van Der Harst, P., Zhang, W., Leach, I.M., Rendon, A., Verweij, N., Sehmi, J., Paul, D.S., Elling, U., Allayee, H., and Li, X. (2012). Seventy-five genetic loci influencing the human red blood cell. Nature *492*, 369-375.

van Vliet-Ostaptchouk, J.V., Nuotio, M.-L., Slagter, S.N., Doiron, D., Fischer, K., Foco, L., Gaye, A., Gögele, M., Heier, M., and Hiekkalinna, T. (2014). The prevalence of metabolic syndrome and metabolically healthy obesity in Europe: a collaborative analysis of ten large cohort studies. BMC endocrine disorders *14*, 9.

Vitart, V., Rudan, I., Hayward, C., Gray, N.K., Floyd, J., Palmer, C.N., Knott, S.A., Kolcic, I., Polasek, O., and Graessler, J. (2008). SLC2A9 is a newly identified urate transporter influencing serum urate concentration, urate excretion and gout. Nature genetics *40*, 437-442.

Voet, D., Voet, J., and Pratt, C. (2006). Fundamentals of biochemistry: life at the molecular level 3rd edn (United States of America: John Wiley and Sons, Inc.).

Voight, B.F., Kang, H.M., Ding, J., Palmer, C.D., Sidore, C., Chines, P.S., Burtt, N.P., Fuchsberger, C., Li, Y., and Erdmann, J. (2012). The metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. PLoS Genetics *8*, e1002793.

Voisey, J., Carroll, L., and Van Daal, A. (2003). Melanocortins and their receptors and antagonists. Current drug targets *4*, 586-597.

Vuorinen-Markkola, H., and Yki-Järvinen, H. (1994). Hyperuricemia and insulin resistance. The Journal of Clinical Endocrinology & Metabolism *78*, 25-29.

Wald, A. (1940). The fitting of straight lines if both variables are subject to error. The Annals of Mathematical Statistics *11*, 284-300.

Wallace, S.L., Robinson, H., Masi, A.T., Decker, J.L., and Mccarty, D.J. (1977). Preliminary criteria for the classification of the acute arthritis of primary gout. Arthritis & Rheumatology *20*, 895-900.

Wang, B., Meng, D., Wang, J., Jia, Z., Zhoub, S., Liu, S., Chu, N., Han, L., Zhang, K., and Ma, X. (2011a). Positive correlation between Beta-3-Adrenergic Receptor (ADRB3) gene and gout in a Chinese male population. The Journal of rheumatology *38*, 738-740.

Wang, C., Hu, Y.-M., He, J.-W., Gu, J.-M., Zhang, H., Hu, W.-W., Yue, H., Gao, G., Xiao, W.-J., and Yu, J.-B. (2011b). Association between low density lipoprotein receptor-related protein 2 gene polymorphisms and bone mineral density variation in Chinese population. PloS One *6*, e28874.

Wang, F., Xu, C.-Q., He, Q., Cai, J.-P., Li, X.-C., Wang, D., Xiong, X., Liao, Y.-H., Zeng, Q.-T., and Yang, Y.-Z. (2011c). Genome-wide association identifies a susceptibility locus for coronary artery disease in the Chinese Han population. Nature Genetics *43*, 345-349.

Wang, G., Li, Q., Niu, T., Chen, C., and Xu, X. (2002). Association of GYS1 and beta (3)-AR gene with postprandial hyperglycemia and serum uric acid in type 2 diabetes mellitus. Chinese medical journal *115*, 1308-1311.

Wang, J., and Pantopoulos, K. (2011). Regulation of cellular iron metabolism. Biochemical Journal *434*, 365-381.

Wang, T.-N., Chang, W.-T., Chiu, Y.-W., Lee, C.-Y., Lin, K.-D., Cheng, Y.Y., Su, Y.-J., Chung, H.-F., and Huang, M.-C. (2013). Relationships between changes in leptin and insulin resistance levels in obese individuals following weight loss. The Kaohsiung journal of medical sciences *29*, 436-443.

Wang, W., Bhole, V.M., and Krishnan, E. (2015). Chronic kidney disease as a risk factor for incident gout among men and women: retrospective cohort study using data from the Framingham Heart Study. BMJ open *5*, e006843.

Wannaiampikul, S., Phonrat, B., Tungtrongchitr, A., Limwongse, C., Chongviriyaphan, N., Santiprabhob, J., and Tungtrongchitr, R. (2015). Genetic variant screening of MC3R and MC4R genes in early-onset obese

children and their relatives among a Thai population: family-based study. Genetics and Molecular Research *14*, 18090-18102.

Ward, L.D., and Kellis, M. (2012). HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic acids research *40*, D930-D934.

Ward, V., McGinty, J., and Church, W. (1993). Iron (III) chloride injection increases nigral uric acid in guinea-pig. Neuroreport *4*, 787-790.

Waterhouse, D., McLaughlin, A., Sheehan, F., and O'Shea, D. (2009). An examination of the prevalence of IDF-and ATPIII-defined metabolic syndrome in an Irish screening population. Irish journal of medical science *178*, 161-166.

Wei, C.-Y., Sun, C.-C., Wei, J.C.-C., Tai, H.-C., Sun, C.-A., Chung, C.-F., Chou, Y.-C., Lin, P.-L., and Yang, T. (2015). Association between hyperuricemia and metabolic syndrome: an epidemiological study of a labor force population in Taiwan. BioMed research international *2015*.

Wen, W., Cho, Y.-S., Zheng, W., Dorajoo, R., Kato, N., Qi, L., Chen, C.-H., Delahanty, R.J., Okada, Y., and Tabara, Y. (2012). Meta-analysis identifies common variants associated with body mass index in east Asians. Nature Genetics *44*, 307-311.

West, A.P., Bennett, M.J., Sellers, V.M., Andrews, N.C., Enns, C.A., and Bjorkman, P.J. (2000). Comparison of the interactions of transferrin receptor and transferrin receptor 2 with transferrin and the hereditary hemochromatosis protein HFE. Journal of Biological Chemistry *275*, 38135-38138.

West, A.R., and Oates, P.S. (2008). Mechanisms of heme iron absorption: current questions and controversies. World Journal of Gastroenterology *14*, 4101-4110.

Westra, H.-J., Peters, M.J., Esko, T., Yaghootkar, H., Schurmann, C., Kettunen, J., Christiansen, M.W., Fairfax, B.P., Schramm, K., and Powell, J.E. (2013). Systematic identification of trans eQTLs as putative drivers of known disease associations. Nature Genetics *45*, 1238-1243.

White, J., Sofat, R., Hemani, G., Shah, T., Engmann, J., Dale, C., Shah, S., Kruger, F.A., Giambartolomei, C., and Swerdlow, D.I. (2016). Plasma urate concentration and risk of coronary heart disease: a Mendelian randomisation analysis. The lancet Diabetes & endocrinology *4*, 327-336.

WHO, W.H.O. (2011). Serum ferritin concentrations for the assessment of iron status and iron deficiency in populations.

WHO, W.H.O. (2015). The global prevalence of anaemia in 2011.

Widén, E., Lehto, M., Kanninen, T., Walston, J., Shuldiner, A.R., and Groop, L.C. (1995). Association of a polymorphism in the β3-adrenergic–receptor gene with features of the insulin resistance syndrome in Finns. New England Journal of Medicine *333*, 348-352.

Wilcox, W.R., Khalaf, A., Weinberger, A., Kippen, I., and Klinenberg, J.R. (1972). Solubility of uric acid and monosodium urate. Medical and biological engineering *10*, 522-531.

Willer, C.J., Speliotes, E.K., Loos, R.J., Li, S., Lindgren, C.M., Heid, I.M., Berndt, S.I., Elliott, A.L., Jackson, A.U., and Lamina, C. (2009). Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nature Genetics *41*, 25-34.

Williams, P.T. (2008). Effects of diet, physical activity and performance, and body weight on incident gout in ostensibly healthy, vigorously active men. The American journal of clinical nutrition *87*, 1480-1487.

Wilson, P.W., D'Agostino, R.B., Parise, H., Sullivan, L., and Meigs, J.B. (2005). Metabolic syndrome as a precursor of cardiovascular disease and type 2 diabetes mellitus. Circulation *112*, 3066-3072.

Windpessl, M., Ritelli, M., Wallner, M., and Colombi, M. (2016). A novel homozygous SLC2A9 mutation associated with renal-induced hypouricemia. American journal of nephrology *43*, 245-250.

Winnard, D., Wright, C., Jackson, G., Gow, P., Kerr, A., McLachlan, A., Orr-Walker, B., and Dalbeth, N. (2013). Gout, diabetes and cardiovascular disease in the Aotearoa New Zealand adult population: co-prevalence and implications for clinical practice. The New Zealand Medical Journal (Online) *126*.

Winnard, D., Wright, C., Taylor, W.J., Jackson, G., Te Karu, L., Gow, P.J., Arroll, B., Thornley, S., Gribben, B., and Dalbeth, N. (2012). National prevalence of gout derived from administrative health data in Aotearoa New Zealand. Rheumatology *51*, 901-909.

Wnuk, M., Hlushchuk, R., Janot, M., Tuffin, G., Martiny-Baron, G., Holzer, P., Imbach-Weese, P., Djonov, V., and Huynh-Do, U. (2012). Podocyte EphB4 signaling helps recovery from glomerular injury. Kidney international *81*, 1212-1225.

Wong, J., Love, D.R., Kyle, C., Daniels, A., White, M., Stewart, A.W., Schnell, A.H., Elston, R.C., Holdaway, I.M., and Mountjoy, K.G. (2002). Melanocortin-3 receptor gene variants in a Maori kindred with obesity and early onset type 2 diabetes. Diabetes research and clinical practice *58*, 61-71.

Wong-McClure, R.A., Gregg, E.W., Barceló, A., Lee, K., Abarca-Gómez, L., Sanabria-López, L., and Tortós-Guzmán, J. (2015). Prevalence of metabolic syndrome in Central America: a cross-sectional population-based study. Revista Panamericana de Salud Pública *38*, 202-208.

Wu, M.C., Lee, S., Cai, T., Li, Y., Boehnke, M., and Lin, X. (2011). Rare-variant association testing for sequencing data with the sequence kernel association test. The American Journal of Human Genetics *89*, 82-93.

Yamanishi, H., Iyama, S., Yamaguchi, Y., Kanakura, Y., and Iwatani, Y. (2003). Total iron-binding capacity calculated from serum transferrin concentration or serum iron concentration and unsaturated iron-binding capacity. Clinical chemistry *49*, 175-178.

Yamasaki, T., and Tomita, K. (2008). Relationship between hyperuricemia and metabolic syndrome. Nihon rinsho Japanese journal of clinical medicine *66*, 766-770.

Yan, D., Wang, J., Jiang, F., Zhang, R., Wang, T., Wang, S., Peng, D., He, Z., Chen, H., and Bao, Y. (2016). A causal relationship between uric acid and diabetic macrovascular disease in Chinese type 2 diabetes patients: A Mendelian randomization analysis. International journal of cardiology *214*, 194-199.

Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A., Heath, A.C., Martin, N.G., and Montgomery, G.W. (2010a). Common SNPs explain a large proportion of the heritability for human height. Nature Genetics *42*, 565-569.

Yang, Q., Köttgen, A., Dehghan, A., Smith, A.V., Glazer, N.L., Chen, M.-H., Chasman, D.I., Aspelund, T., Eiriksdottir, G., and Harris, T.B. (2010b). Multiple Genetic Loci Influence Serum Urate Levels and Their Relationship With Gout and Cardiovascular Disease Risk FactorsClinical Perspective. Circulation: Cardiovascular Genetics *3*, 523-530.

Yehuda, S., and Mostofsky, D.I. (2010). Iron deficiency and overload: from basic biology to clinical medicine (Springer Science & Business Media).

Yiannakouris, N., Melistas, L., Kontogianni, M., Heist, K., and Mantzoros, C. (2004). The Val81 missense mutation of the melanocortin 3 receptor gene, but not the 1908c/T nucleotide polymorphism in lamin A/C gene, is associated with hyperleptinemia and hyperinsulinemia in obese Greek caucasians. Journal of endocrinological investigation *27*, 714-720.

Yilmaz, Z., Davis, C., Loxton, N.J., Kaplan, A.S., Levitan, R.D., Carter, J.C., and Kennedy, J.L. (2015). Association between MC4R rs17782313 polymorphism and overeating behaviors. International Journal of Obesity *39*, 114-120.

Yong, Y., and Lin, H. (2005). SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci. Cell research *15*, 97-98.

Yu, K.-H., Kuo, C.-F., Luo, S.-F., See, L.-C., Chou, I.-J., Chang, H.-C., and Chiou, M.-J. (2012). Risk of end-stage renal disease associated with gout: a nationwide population study. Arthritis research & therapy *14*, R83.

Yu, Y., Jin, H., Holder, D., Ozer, J.S., Villarreal, S., Shughrue, P., Shi, S., Figueroa, D.J., Clouse, H., and Su, M. (2010). Urinary biomarkers trefoil factor 3 and albumin enable early detection of kidney tubular injury. Nature biotechnology *28*, 470-477.

Zabaneh, D., and Balding, D.J. (2010). A genome-wide association study of the metabolic syndrome in Indian Asian men. PloS One *5*, e11961.

Zegers, D., Beckers, S., Mertens, I.L., Van Gaal, L.F., and Van Hul, W. (2010). Common melanocortin-3 receptor variants are not associated with obesity, although rs3746619 does influence weight in obese individuals. Endocrine *38*, 289-293.

Zhan, S., and Ho, S.C. (2005). Meta- Analysis of the Association of the Trp64Arg Polymorphism in the β3 Adrenergic Receptor with Insulin Resistance. Obesity research *13*, 1709-1719.

Zhang, D., Bobulescu, I.A., Maalouf, N.M., Adams- Huet, B., Poindexter, J., Park, S., Wei, F., Chen, C., Moe, O.W., and Sakhaee, K. (2015). Relationship between serum uric acid and bone mineral density in the general population and in rats with experimental hyperuricemia. Journal of Bone and Mineral Research *30*, 992-999.

Zhang, F., Tao, Y., Zhang, Z., Guo, X., An, P., Shen, Y., Wu, Q., Yu, Y., and Wang, F. (2012a). Metalloreductase Steap3 coordinates the regulation of iron homeostasis and inflammatory responses. Haematologica, haematol. 2012.063974.

Zhang, Y., Chen, C., Choi, H., Chaisson, C., Hunter, D., Niu, J., and Neogi, T. (2012b). Purine-rich foods intake and recurrent gout attacks. Annals of the rheumatic diseases, annrheumdis-2011-201215.

Zhang, Z., Bian, L., and Choi, Y. (2012c). Serum uric acid: a marker of metabolic syndrome and subclinical atherosclerosis in Korean men. Angiology *63*, 420-428.

Zhao, W., Wang, L., Zhang, M., Yuan, C., and Gao, C. (2012). E3 ubiquitin ligase tripartite motif 38 negatively regulates TLR-mediated immune responses by proteasomal degradation of TNF receptor-associated factor 6 in macrophages. The Journal of Immunology *188*, 2567-2574.

Zhou, H., Sehl, M.E., Sinsheimer, J.S., and Lange, K. (2010). Association screening of common and rare genetic variants by penalized regression. Bioinformatics *26*, 2375-2382.

Zhou, Z.-W., Cui, L.-L., Han, L., Wang, C., Song, Z.-J., Shen, J.-W., Li, Z.-Q., Chen, J.-H., Wen, Z.-J., and Wang, X.-M. (2015). Polymorphisms in GCKR, SLC17A1 and SLC22A12 were associated with phenotype gout in Han Chinese males: a case–control study. BMC medical genetics *16*, 66.

Zhu, Y., Pandya, B.J., and Choi, H.K. (2011). Prevalence of gout and hyperuricemia in the US general population: the National Health and Nutrition Examination Survey 2007–2008. Arthritis & Rheumatism *63*, 3136-3141.

Zhu, Y., Pandya, B.J., and Choi, H.K. (2012). Comorbidities of gout and hyperuricemia in the US general population: NHANES 2007-2008. The American journal of medicine *125*, 679-687. e671.

Zobel, D.P., Andreasen, C.H., Grarup, N., Eiberg, H., Sørensen, T.I., Sandbæk, A., Lauritzen, T., Borch-Johnsen, K., Jørgensen, T., and Pedersen, O. (2009). Variants near MC4R are associated with obesity and influence obesity-related quantitative traits in a population of middle-aged people: studies of 14,940 Danes. Diabetes *58*, 757-764.

Zoppini, G., Targher, G., Chonchol, M., Ortalda, V., Abaterusso, C., Pichiri, I., Negri, C., and Bonora, E. (2012). Serum uric acid levels and incident chronic kidney disease in patients with type 2 diabetes and preserved kidney function. Diabetes care *35*, 99-104.

Zuk, O., Schaffner, S.F., Samocha, K., Do, R., Hechter, E., Kathiresan, S., Daly, M.J., Neale, B.M., Sunyaev, S.R., and Lander, E.S. (2014). Searching for missing heritability: designing rare variant association studies. Proceedings of the National Academy of Sciences *111*, E455-E464.

Zumbrennen-Bullough, K., and Babitt, J.L. (2014). The iron cycle in chronic kidney disease (CKD): from genetics and experimental models to CKD patients. Nephrology Dialysis Transplantation *29*, 263-273.

APPENDIX A

Appendix A Table 2.1: Association of serum iron (µg dL^{-1}), Transferrin (g L^{-1}), TIBC (µmol L^{-1}) and TSAT (%) with serum urate (µmol L^{-1})

Iron profile marker	Unadjusted		Adjusted*	
	ß [95% CI]	*P*	ß [95% CI]	*P*
		NZ European		
Serum iron	-4.44e-05 [-6.58e-04 ; 5.69e-04]	0.885	1.74e-05 [-6.04e-04 ; 6.391e-04]	0.955
Transferrin	-0.013 [-0.0607 ; 0.0338]	0.571	-0.008 [-0.0564 ; 0.0396]	0.726
TIBC	-5.35e-04 [-0.0024 ; 0.0013]	0.571	-3.36e-04 [-0.0022 ; 0.0016]	0.726
TSAT	1.35e-04 [-0.0019 ; 0.0022]	0.898	2.45e-04 [-0.0018 ; 0.0026]	0.816
		NZ Polynesian		
Serum iron	-3.86e-04 [-7.66e-04 ; -5.97e-06]	0.046	-3.40e-04 [-7.05e-04 ; 2.52e-05]	0.068
Transferrin	0.029 [0.0037 ; 0.054]	0.024	0.023 [-0.0018 ; 0.048]	0.069
TIBC	0.0012 [1.51e-04 ; 2.17e-03]	0.024	9.20e-04 [-7.47e-05 ; 0.0019]	0.069
TSAT	-0.0021 [-0.0033 ; -7.16e-04]	0.002	-1.69e-03 [-0.0029 ; -3.91e-04]	0.011
		JHS (All)		
Serum iron	1.11e-04 [4.08e-05 ; 2.63e-04]	0.152	-8.91e-05 [2.21e-04 ; 4.27e-05]	0.185
TIBC	-1.82e-04 [-2.82e-04 ; -8.19e-05]	3.77e-04	3.31e-05 [-5.45e-05 ; 1.21e-04]	0.459
TSAT	5.62e-04 [1.28e-04 ; 9.96e-02]	0.011	3.35e-04 [7.19e-04 ; 4.96e-05]	0.087
		JHS (Males)		
Serum iron	2.71e-04 [4.82e-04 ; 5.91e-05]	0.012	-1.63e-04 [3.72e-04 ; 4.54e-05]	0.125
TIBC	1.29e-04 [-3.47e-05 ; -2.93e-04]	0.122	1.07e-04 [-5.17e-05 ; 2.66e-04]	0.185
TSAT	-9.57e-04 [-1.56e-03 ; 3.51e-04]	0.002	-6.32e-04 [-1.23e-03 ; 3.31e-05]	0.038
		JHS (Females)		
Serum iron	-2.15e-04 [3.95e-04 ; 3.39e-05]	0.02	-5.98e-05 [-2.28e-04 ; 1.08e-04]	0.487
TIBC	-6.06e-05 [-1.68e-04 ; 4.67e-05]	0.268	2.49e-05 [-7.89e-05 ; 1.28e-04]	0.637

TSAT	-4.75e-04 [-8.53e-03 ; 5.79e-05]	0.081	-2.31e-04 [-7.34e-04 ; 2.71e-04]	0.366
NHANES III: EUR (All)				
Serum iron	0.209 [0.149 ; 0.269]	7.78e-12	0.159 [0.109 ; 0.208]	3.33e-10
TIBC	-0.296 [-0.514 ; -0.077]	0.008	0.306 [0.121 ; 0.491]	0.001
TSAT	0.681 [0.489 ; 0.873]	3.98e-12	0.358 [0.199 ; 0.5180]	1.07e-05
NHANES III: EUR (Males)				
Serum iron	0.135 [0.055 ; 0.216]	9.22e-04	0.135 [0.058 ; 0.212]	5.76e-04
TIBC	0.931 [0.619 ; 1.242]	5.2e-09	0.548 [0.244 ; 0.852]	4.03e-04
TSAT	0.156 [-0.096 ; 0.409]	0.225	0.265 [0.025 ; 0.506]	0.030
NHANES III: EUR (Females)				
Serum iron	0.022 [-0.046 ; 0.091]	0.519	0.171 [0.108 ; 0.233]	1.12e-07
TIBC	0.042 [-0.198 ; 0.284]	0.727	0.192 [-0.031 ; 0.416]	0.0925
TSAT	8.53e-03 [-0.219 ; 0.236]	0.942	0.411 [0.204 ; 0.618]	1.01e-04
NHANES III: AA (All)				
Serum iron	0.235 [0.159 ; 0.312]	1.49e-09	0.164 [0.097 ; 0.232]	1.5e-06
TIBC	-1.015 [-1.272 ; -0.757]	1.33e-14	0.197 [-0.031 ; 0.425]	0.09
TSAT	0.950 [0.699 ; 1.202]	1.47e-13	0.378 [0.155 ; 0.601]	8.71e-04
NHANES III: AA (Males)				
Serum iron	-0.029 [-0.135 ; 0.075]	0.577	0.131 [0.029 ; 0.233]	0.011
TIBC	0.417 [0.018 ; 0.816]	0.041	0.485 [0.105 ; 0.865]	0.012
TSAT	-0.244 [-0.593 ; 0.104]	0.17	0.252 [-0.083 ; 0.587]	0.14
NHANES III: AA (Females)				
Serum iron	-0.037 [-0.135 ; 0.059]	0.447	0.163 [0.075 ; 0.252]	2.9e-04
TIBC	-0.783 [-1.078 ; -0.487;]	2.21e-07	0.165 [-0.118 ; 0.449]	0.252
TSAT	0.036 [-0.290 ; 0.362]	0.828	0.361 [0.063 ; 0.657]	0.017

NZ; New Zealand, JHS; Jackson Heart Study, NHANES III; US Third National Health and Nutrition Examination Survey, EUR; White Caucasian/European, AA; African American, TIBC; Total iron binding capacity, TSAT; Transferrin saturation. *Adjusted for age, sex, BMI and number of self-reported Polynesian grandparents for the NZ Polynesian group.

Appendix A Table 2.2: Association of serum iron (µg dL^{-1}), Transferrin (g L^{-1}), TIBC (µmol L^{-1}) and TSAT (%) with hyperuricaemia

Iron profile marker	Unadjusted		Adjusted*	
	OR [95% CI]	*P*	OR [95% CI]	*P*
		JHS (All)		
Serum iron	1.005 [0.999 ; 1.009]	0.067	1.002 [0.994 ; 1.006]	0.958
TIBC	0.984 [0.966 ; 1.003]	0.108	1.006 [0.983 ; 1.029]	0.607
TSAT	1.016 [1.001 ; 1.031]	0.034	0.996 [0.979 ; 1.014]	0.735
		JHS (Males)		
Serum iron	0.997 [0.990 ; 1.003]	0.379	0.998 [0.991 ; 1.005]	0.713
TIBC	1.013 [0.987 ; 1.041]	0.330	1.012 [0.985 ; 1.040]	0.396
TSAT	0.987 [0.968 ; 1.006]	0.186	0.992 [0.972 ; 1.012]	0.432
		JHS (Females)		
Serum iron	0.999 [0.986 ;1.011]	0.976	1.004 [9.88e-01 ; 1.016]	0.577
TIBC	0.988 [0.948 ; 1.029]	0.587	1.001 [9.55e-01 ; 1.046]	0.981
TSAT	1.001 [0.962 ; 1.036]	0.953	1.007 [9.645e-01 ; 1.048]	0.714
		NHANES III: EUR (All)		
Serum iron	1.003 [1.002 ; 1.004]	5.05e-06	1.002 [1.001 ; 1.004]	6.4E-04
TIBC	0.994 [0.989 ; 0.999]	0.026	1.013 [1.008 ; 1.019]	1.07e-06
TSAT	1.009 [1.005 ; 1.013]	9.4e-06	1.002 [0.997 ; 1.007]	0.302
		NHANES III: EUR (Males)		
Serum iron	1.001 [0.999 ; 1.002]	0.252	1.002 [1.001 ; 1.004]	0.00646
TIBC	1.015 [1.008 ; 1.021]	4.51e-06	1.019 [1.012 ; 1.026]	2.71e-08
TSAT	0.998 [0.993 ; 1.003]	0.526	1.001 [0.996 ; 1.006]	0.607
		NHANES III: EUR (Females)		
Serum iron	0.996 [0.993 ; 0.999]	0.016	1.001 [0.998 ; 1.005]	0.413
TIBC	0.986 [0.977 ; 0.995]	0.00345	1.006 [9.9e-01 ; 1.016]	0.183
TSAT	0.992 [0.983 ; 1.002]	0.147	1.001 [0.989 ; 1.011]	0.899
		NHANES III: AA (All)		
Serum iron	1.003 [1.001 ; 1.005]	4.44e-04	1.005 [1.002 ; 1.006]	1.29e-04
TIBC	0.977 [0.971 ; 0.985]	1.12e-09	1.005 [0.996 ; 1.013]	0.241
TSAT	1.016 [1.009 ; 1.023]	6.63e-07	1.012 [1.004 ; 1.019]	2.77e-03

NHANES III: AA (Males)				
Serum iron	0.999 [0.996 ; 1.002]	0.6	1.004 [1.001 ; 1.007]	0.003
TIBC	0.992 [0.983 ; 1.002]	0.1509	1.004 [0.993 ; 1.015]	0.407
TSAT	0.999 [0.991 ; 1.007]	0.908	1.012 [1.002 ; 1.021]	0.013
NHANES III: AA (Females)				
Serum iron	0.999 [0.995 ; 1.003]	0.724	1.005 [1.001 ; 1.009]	0.024
TIBC	0.974 [0.961 ; 0.987]	9.3e-05	1.009 [0.995 ; 1.024]	0.176
TSAT	1.004 [0.991 ; 1.017]	0.503	1.011 [0.994 ; 1.025]	0.19

JHS; Jackson Heart Study, NHANES III; US Third National Health and Nutrition Examination Survey, EUR; White Caucasian/European, AA; African American, TIBC; Total iron binding capacity, TSAT; Transferrin saturation. *Adjusted for age, sex, BMI and number of self-reported Polynesian grandparents for the NZ Polynesian group.

Appendix A Table 2.3: Association of serum iron (µg dL^{-1}), Transferrin (g L^{-1}), TIBC (µmol L^{-1}) and TSAT (%) with gout

Iron profile marker	Unadjusted		Adjusted*	
	OR [95% CI]	***P***	**OR [95% CI]**	***P***
NZ European				
Serum iron	0.991 [0.981 ; 1.002]	0.115	0.994 [0.981 ; 1.008]	0.459
Transferrin	1.393 [0.606 ; 3.307]	0.440	1.318 [0.045 ; 3.801]	0.602
TIBC	1.013 [0.981 ; 1.048]	0.440	1.011 [0.097 ; 1.054]	0.602
TSAT	0.968 [0.934 ; 1.003]	0.077	0.096 [0.092 ; 1.013]	0.171
NZ Polynesian				
Serum iron	0.995 [0.984 ; 1.005]	0.354	0.993 [0.981 ; 1.005]	0.263
Transferrin	1.708 [0.934 ; 3.252]	0.090	1.663 [0.839 ; 3.476]	0.156
TIBC	1.021 [0.997 ; 1.048]	0.090	1.020 [0.099 ; 1.051]	0.156
TSAT	0.964 [0.925 ; 1.003]	0.074	0.960 [0.916 ; 1.004]	0.082

NZ; New Zealand, TIBC; Total iron binding capacity, TSAT; Transferrin saturation. *Adjusted for age, sex, BMI and number of self-reported Polynesian grandparents for the NZ Polynesian group.

Appendix A Table 2.4: Comparison of average values of iron profile markers in gout case-control groups

Population	*P*	[95% CI] for difference
NZ European		
Serum iron (μg dL^{-1})	0.114	[-1.96 ; 18.06]
Serum ferritin (ng mL^{-1})	0.213	[-97.46 ; 22.02]
Transferrin (g L^{-1})	0.446	[-0.17 ; 0.07]
TIBC (μmol L^{-1})	0.446	[-4.42 ; 1.96]
TSAT (%)	0.074	[-0.27 ; 5.64]
NZ Polynesian		
Serum iron (μg dL^{-1})	0.342	[-5.09 ; 14.57]
Serum ferritin (ng mL^{-1})	2.29E-04	[-211.88 ; -66.31]
Transferrin (g L^{-1})	0.073	[-0.32 ; 0.014]
TIBC (μmol L^{-1})	0.073	[-8.05 ; 0.36]
TSAT (%)	0.071	[-0.21 ; 5.01]
US		
Serum ferritin (ng mL^{-1})	6.60E-17	[-167.19 to -107.16]

NZ; New Zealand, TIBC; Total iron binding capacity, TSAT; Transferrin saturation.

Appendix A Table 2.5 Association of serum ferritin with gout (Transformed data)

Population	OR [95% CI]	*P*	OR [95% CI]*	*P**
NZ European	1.19 [0.81 ; 1.76]	0.37	0.88 [0.49 ; 1.52]	0.64
NZ Polynesian	1.76 [1.12 ; 2.84]	1.63E-02	2.24 [1.27 ; 4.17]	7.41E-03
US	2.42 [1.79 ; 3.39]	7.02E-08	2.31 [1.65 ; 3.31]	2.26E-06

NZ; New Zealand, US; United States of America. *Adjusted for age, sex, BMI, C-reactive protein and number of self-reported Polynesian grandparents in the NZ Polynesian analyses. All odds ratios represent a change per unit log of ferritin.

Appendix A Table 2.6 Association of serum ferritin with gout flares/year (Transformed data)

Population	ß [95% CI]	*P*	ß [95% CI]*	*P**
NZ European	1.74 [0.17 ; 3.32]	0.030	1.94 [0.24 ; 3.63]	0.025
NZ Polynesian	-4.42 [-9.750 ; 0.90]	0.10	-4.68 [-10.11 ; 0.75]	0.090
US	0.38 [0.11 ; 0.64]	0.0063	0.37 [0.10 ; 0.63]	0.0066

NZ; New Zealand, US; United States of America. *Adjusted for age, sex, BMI and number of self-reported Polynesian grandparents in the NZ Polynesian analyses. All ß-estimates represent a change per unit log of ferritin.

Appendix A Table 2.7 Association of serum ferritin with gout flares/year

Population	^Unadjusted		*Adjusted		†Adjusted	
	ß [95% CI]	*P*	ß [95% CI]	*P*	ß [95% CI]	*P*
NZ European	0.08 [0.001 ; 0.16]	0.045	0.09 [0.003 ; 0.17]	0.042	0.069 [-0.005 ; 0.144]	0.067
NZ Polynesian	-0.11 [-0.24 ; 0.03]	0.13	-0.11 [-0.24 ; 0.03]	0.14	-0.091 [-0.22 ; 0.034]	0.15
US	0.02 [0.01 ; 0.04]	0.003	0.02 [0.01 ; 0.04]	0.002	0.018 [0.001 ; 0.036]	0.043

^Unadjusted; *Adjusted for age, BMI and number of self-reported Polynesian grandparents in the NZ Polynesian analyses; †Adjusted for age, BMI, allopurinol usage and number of self-reported Polynesian grandparents in the NZ Polynesian analyses.

Appendix A Table 2.8 Association of serum ferritin (ng mL^{-1}) with gout excluding those 'not taking allopurinol'

Population	All				Taking allopurinol			
	OR [95% CI]	*P*	OR [95% CI]*	*P**	OR [95% CI]	*P*	OR [95% CI]*	*P**
NZ European	1.011 [0.993 ; 1.031]	0.22	0.997 [0.971 ; 1.023]	0.84	1.009 [0.987 ; 1.033]	0.42	0.993 [0.961 ; 1.025]	0.67
NZ Polynesian	1.026 [1.011 ; 1.045]	1.71E-03	1.032 [1.013 ; 1.055]	1.76E-03	1.039 [1.013 ; 1.0725]	6.83E-03	1.061 [1.024 ; 1.112]	3.93E-03
US	1.115 [1.071 ; 1.170]	1.47E-06	1.112 [1.066 ; 1.170]	7.41E-06	1.149 [1.086 ; 1.233]	1.57E-05	1.270 [1.148 ; 1.469]	8.89E-05

^Unadjusted; *Adjusted for age, BMI, C-reactive protein and number of self-reported Polynesian grandparents in the NZ Polynesian analyses. All OR values represent change in risk for per 10 ng mL^{-1} increase in serum ferritin.

Appendix A Table 2.7 Association of ferritin (ng mL^{-1}) with serum urate (µmol L^{-1}) (Transformed data)

Population	ß [95% CI]	*P*	ß [95% CI]*	*P**
NZ European	21.97 [-2.452 ; 46.389]	0.076	20.027 [-6.174 ; 46.229]	0.13
NZ Polynesian	22.031 [7.465 ; 37.151]	3.35E-03	22.824 [8.098 ; 37.551]	2.5E-03
US	22.98 [-3.882 ; 49.836]	0.092	19.326 [-9.281 ; 47.933]	0.18
JHS (males)	10.921 [2.642 ; 19.200]	9.82E-03	9.476 [1.457 ; 17.495]	2.06E-02
JHS (females)	17.48 [11.863 ; 23.093]	1.64E-09	9.527 [3.765 ; 15.288]	1.22E-03
JHS (combined)	31.470 [27.021 ; 35.919]	7.16E-41	10.850 [6.249 ; 15.452]	4.11E-06
NHANES III EUR (Males)	8.663 [5.215 ; 12.111]	8.91E-07	5.118 [1.816 ; 8.419]	2.39E-03
NHANES III EUR (Females)	17.814 [15.342 ; 20.286]	9.27E-44	11.571 [9.074 ; 14.068]	2.07E-19
NHANES III EUR (Combined)	28.779 [26.779 ; 30.777]	6.12E-163	10.645 [8.655 ; 12.634]	1.76E-25
NHANES III AA (Males)	13.809 [9.425 ; 18.193]	7.93E-10	7.389 [3.135 ; 11.643]	6.70E-04
NHANES III AA (Females)	22.585 [19.856 ; 25.313]	2.62E-56	12.526 [9.658 ; 15.394]	1.91E-17
NHANES III AA (Combined)	31.375 [29.188 ; 33.562]	4.51E-160	12.347 [9.994 ; 14.701]	1.53E-24

NZ; New Zealand, JHS; Jackson Heart Study, NHANES III; US Third National Health and Nutrition Examination Survey, EUR; White Caucasian/European, AA; African American. *Adjusted for age, sex, BMI, C-reactive protein and number of self-reported Polynesian grandparents in the NZ Polynesian analyses. All ß-estimates represent a change per unit log of ferritin.

Appendix A Table 2.8 Association of serum ferritin with hyperuricaemia (Transformed data)

Population	OR [95% CI]	*P*	OR [95% CI]*	*P**
JHS (Males)	1.31 [1.01 ; 1.72]	4.22E-02	1.30 [1.00 ; 1.70]	5.74E-02
JHS (Females)	2.40 [1.55 ; 3.84]	1.41E-04	2.12 [1.31 ; 3.61]	3.029E-03
JHS (Combined)	2.17 [1.78 ; 2.68]	7.22E-14	1.52 [1.20 ; 1.92]	4.65E-04
NHANES III EUR (Males)	1.40 [1.31 ; 1.51]	1.22E-20	1.18 [1.09 ; 1.27]	2.31E-05
NHANES III EUR (Females)	2.21 [1.99 ; 2.46]	2.81E-48	1.52 [1.36 ; 1.71]	2.1E-12
NHANES III EUR (Combined)	1.97 [1.86 ; 2.08]	3.89E-126	1.35 [1.27 ; 1.44]	6.00E-21
NHANES III AA (Males)	1.65 [1.49 ; 1.84]	3.82E-20	1.31 [1.16 ; 1.47]	9.14E-06
NHANES III AA (Females)	2.36 [2.05 ; 2.72]	2.15E-32	1.74 [1.48 ; 2.05]	3.35E-11
NHANES III AA (Combined)	2.21 [2.04 ; 2.39]	7.91E-83	1.50 [1.37 ; 1.65]	2.41E-17

NZ; New Zealand, JHS; Jackson Heart Study, NHANES III; US Third National Health and Nutrition Examination Survey, EUR; White Caucasian/European, AA; African American. *Adjusted for age, sex, BMI, C-reactive protein and number of self-reported Polynesian grandparents in the NZ Polynesian analyses. All odds ratios represent a change per unit log of ferritin.

Appendix A Table 3.1 Details of study cohorts included in GWAS used in MR-analysis

Köttgen *et al.* (2013)		Benyamin *et al.* (2014) data	
Name of the study	**Number of individuals**	**Name of the study**	**Number of individuals**
AGES Reyjavik Study	3219	Australia-Adult	9148
Amish	1193	Autralia-Adolescent	2544
ARIC	9049	BHS	877
ASPS	845	Cambridge	2419
AUSTWIN	11520	CoLAUS	5409
BLSA	521	ERF/Rotterdam	871
BRIGHT	1743	Estonian Biobank (Original)	893
CARDIA	1713	Estonian Biobank (Replication)	1017
CHS	3252	FENLAND	1402
CoLaus	5409	InCHIANTI	1206
CROATIA-CORCULA	895	INTERACT	9294
CROATIS-VIS	490	KORA F3	1634
CROATSIS-SPLIT	912	KORA F4	1809
DESIR	716	Micros/EURAC	1218
ERF	1835	NBS	1791
EPIC-Norfolk cohort	889	PREVEND	3644
Estonian Biobank (Original)	931	SardiNIA	4694
Family Heart Study (FamHS)	7837	Val Borbera	1659
FHS	7699	Total number	51529
Health 2000	2069		
InCHIANTI	1205		
INCIPE	940		
INGI-Carlantiono	432		
INGI-CLIENTO	859		
INGI-FVG	1018		
INGI-Val Borbera	1658		
KORA F3	1643		
KORA F4	1814		
LBC1936	769		
LifeLines	3343		
LOLIPOP_EW_A	587		
LOLIPOP_EW_P	650		
LOLIPOP_EW610	924		
LURIC	963		
MICROS	1236		
NESDA	1731		
NSPHS	655		
ORCADES	888		
PREVEND	3785		
PROCARDIS	3742		
RS-I	4274		
RS-II	2123		
SardiNIA	4694		
SHIP	4067		
SOCCS	1105		
Sorbs	896		
TwinsUK	3640		
WGHS	NA		
Young Finns Study	2023		
Total	114401		

Appendix A Table 5.1: Sixty-seven Genomic control SNPs genotyped to calculate estimates of individual Eastern Polynesian ancestry

TaqMan Genotyped	Sequenom MassArray Genotyped	PCR-RFLP
rs1183201	*rs10025373*	*rs7725*
rs9358890	*rs1143634*	*rs573816*
rs3799344	*rs11536879*	
rs12664474	*rs1205*	
rs2075876	*rs2812378*	
rs1816532	*rs3014875*	
rs13419122	*rs344542*	
rs12401573	*rs40401*	
rs6945435	*rs452204*	
rs743777	*rs4780884*	
rs10511216	*rs4781011*	
rs12745968	*rs4804221*	
rs1539438	*rs4845622*	
rs729749	*rs4889640*	
rs3738919	*rs507879*	
rs1130214	*rs6005863*	
rs755622	*rs6819597*	
rs7901695	*rs6835636*	
rs7578597	*rs7811892*	
rs2043211	*rs7842*	
rs10733113	*rs795467*	
rs900865	*rs8075846*	
rs2059606	*rs8122*	
rs4129148	*rs9639436*	
rs831628	*rs9882205*	
rs1929480	*rs11078855*	
rs12917707	*rs11119568*	
	rs1184835	
	rs12535365	
	rs12877336	
	rs2683764	
	rs35958249	
	rs4256629	
	rs4571803	
	rs493430	
	rs615204	
	rs693916	
	rs7108425	
	rs7118682	
	rs730275	
	rs9294168	
	rs9690688	

PCR-RFLP: Polymerase Chain Reaction-Restriction Fragment Length Polymorphism

Appendix A Table 5.2: *Brief summary of the 19 renal cell samples used to identify DNase hypersensitivity regions for sequencing

	Sample			Analysis	
GEO accession	**Sex**	**Age**	**Tissue Type**	**Marker Type**	**Protocol**
GSM621648	Male	67 (Years)	Adult Kidney	Histone Methylation	ChIP-Seq
GSM701494	Male	15 (Weeks)	Foetal Renal Cortex	DNase-Hypersensitivity	DNase-Seq
GSM701502	Male	15 (Weeks)	Foetal Renal Cortex	DNase-Hypersensitivity	DNase-Seq
GSM701511	Male	16 (Weeks)	Foetal Renal Cortex	DNase-Hypersensitivity	DNase-Seq
GSM701517	Male	15 (Weeks)	Foetal Renal Cortex	DNase-Hypersensitivity	DNase-Seq
GSM701519	Male	15 (Weeks)	Foetal Renal Cortex	DNase-Hypersensitivity	DNase-Seq
GSM701529	Male	13 (Weeks)	Foetal Renal Cortex	DNase-Hypersensitivity	DNase-Seq
GSM701532	Female	17 (Weeks)	Foetal Renal Cortex	DNase-Hypersensitivity	DNase-Seq
GSM772811	Male	67 (Years)	Adult Kidney	Histone Acetylation	Chip-Seq
GSM773005	Male	50 (years)	Adult Kidney	Histone Methylation	ChIP-Seq
GSM773006	Male	50 (years)	Adult Kidney	Histone Acetylation	ChIP-Seq
GSM817176	Male	18 (Weeks)	Foetal Renal Cortex	DNase-Hypersensitivity	DNase-Seq
GSM817190	Male	15 (Weeks)	Foetal Renal Cortex	DNase-Hypersensitivity	DNase-Seq
GSM817202	Male	17 (Weeks)	Foetal Renal Cortex	DNase-Hypersensitivity	DNase-Seq
GSM817203	Male	17 (Weeks)	Foetal Renal Cortex	DNase-Hypersensitivity	DNase-Seq
GSM817210	Male	15 (Weeks)	Foetal Renal Cortex	DNase-Hypersensitivity	DNase-Seq
GSM817211	Male	15 (Weeks)	Foetal Renal Cortex	DNase-Hypersensitivity	DNase-Seq
GSM878629	Male	14 (Weeks)	Foetal Renal Cortex	DNase-Hypersensitivity	DNase-Seq
GSM878667	Female	13 (Weeks)	Foetal Renal Cortex	DNase-Hypersensitivity	DNase-Seq

*Modified from Flynn (2016)

Appendix A Table 5.3: Meta-analysis results of single variant association analysis within *LRP2* and *A1CF* for hyperuricaemia

Exon	Chr: location	rs ID/var number	P_{OR}	P_{Het}	OR [95% CI]
		LRP2			
2	2:170177382	*rs144829356*	0.198	1.000	1.27 [0.88; 1.83]
3	2:170175334	*rs2229263*	0.841	0.577	0.97 [0.73; 1.29]
4	2:170163816	*rs34104660*	0.191	0.984	0.62 [0.30; 1.27]
6	2:170150671	*rs2229266*	0.401	0.481	0.90 [0.70; 1.16]
8	2:170147502	*rs34693334*	0.094	0.096	1.76 [0.91; 3.39]
13	2:170131548	*rs111360923*	0.019	1.000	0.54 [0.32; 0.90]
14	2:170129474	*rs145709922*	0.850	1.000	0.76 [0.05; 12.42]
14	2:170129547	*rs34291900*	0.583	0.318	1.30 [0.51; 3.36]
14	2:170129528	*rs830994*	0.151	0.051	0.82 [0.62; 1.08]
15	2:170115588	*rs2241190*	0.735	0.479	1.04 [0.82; 1.32]
15	2:170115672	*rs33954745*	0.625	0.703	0.85 [0.45; 1.61]
15	2:170115626	var170115626	0.029	1.000	0.56 [0.34; 0.94]
16	2:170127559	*rs141180155*	0.298	0.982	2.45 [0.45; 13.23]
19	2:170103471	*rs17848143*	0.582	1.000	0.67 [0.16; 2.82]
19	2:170103351	*rs2075249*	0.314	0.614	1.14 [0.88; 1.47]
19	2:170103336	*rs831043*	0.664	0.355	1.05 [0.83; 1.34]
21	2:170100011	*rs150552608*	0.429	1.000	0.38 [0.03; 4.23]
22	2:170099473	*rs831042*	0.605	0.301	1.07 [0.84; 1.35]
23	2:170097707	*rs17848149*	0.693	0.755	1.11 [0.65; 1.91]
26	2:170096095	*rs34915742*	0.429	0.979	0.38 [0.03; 4.23]
27	2:170094756	*rs146289506*	0.850	1.000	0.76 [0.05; 12.42]
28	2:170093726	var170093726	0.249	1.000	0.35 [0.06; 2.07]
29	2:170092439	*rs151079411*	0.668	1.000	1.97 [0.09; 43.53]
29	2:170092395	*rs2229267*	0.910	0.612	0.99 [0.76; 1.28]
29	2:170092504	var170092504	0.345	1.000	1.78 [0.54; 5.86]
30	2:170089934	*rs145384264*	0.943	0.361	1.05 [0.26; 4.31]
31	2:170088351	*rs2302694*	0.001	0.291	0.54 [0.37; 0.79]
32	2:170082936	*rs138070797*	0.725	0.000	1.54 [0.14; 17.36]
36	2:170070348	*rs11886219*	0.565	0.948	0.80 [0.37; 1.73]
36	2:170070172	*rs4667596*	0.033	0.308	0.36 [0.14; 0.92]
39	2:170063263	*rs149367019*	0.850	0.982	0.76 [0.05; 12.42]
39	2:170063471	*rs35114151*	0.623	0.948	1.14 [0.67; 1.94]
39	2:170062977	*rs61995915*	0.387	0.982	2.00 [0.42; 9.62]

41	2:170062078	*rs13397109*	0.083	0.152	0.71 [0.49; 1.05]
42	2:170060603	*rs17848169*	0.854	0.398	1.09 [0.44; 2.72]
44	2:170058345	var170058345	0.250	1.000	0.41 [0.09; 1.88]
46	2:170053505	*rs2228171*	0.234	0.368	0.86 [0.68; 1.10]
48	2:170048482	*rs149148763*	0.510	0.000	0.45 [0.04; 4.79]
50	2:170042245	*rs35734447*	0.668	0.000	1.97 [0.09; 43.53]
54	2:170032989	*rs2229265*	0.027	0.809	1.32 [1.03; 1.69]
55	2:170031824	*rs17848184*	0.399	0.000	1.17 [0.81; 1.69]
57	2:170029657	*rs34355135*	0.543	0.000	1.98 [0.22; 17.85]
58	2:170028529	*rs199528723*	0.375	1.000	0.56 [0.15; 2.04]
60	2:170026248	var170026248	0.668	1.000	1.97 [0.09; 43.53]
61	2:170025083	*rs2229268*	0.336	0.408	1.18 [0.84; 1.67]
64	2:170013904	*rs79723119*	0.482	0.981	1.71 [0.38; 7.60]
66	2:170010985	*rs2075252*	0.028	0.415	0.75 [0.58; 0.97]
67	2:170009390	*rs142934522*	0.429	1.000	0.38 [0.03; 4.23]
69	2:170003432	*rs4667591*	0.012	0.430	0.72 [0.56; 0.93]
70	2:170002412	var170002412	0.607	1.000	0.48 [0.03; 7.79]
72	2:169997051	*rs990626*	0.201	0.470	0.85 [0.67; 1.09]
73	2:169996070	*rs41268685*	0.045	0.980	3.85 [1.03; 14.38]
77	2:169989127	*rs142245618*	0.682	1.000	0.67 [0.10; 4.45]
79	2:169985338	*rs34564141*	0.718	1.000	1.38 [0.24; 8.11]
		A1CF			
5'UTR	10:52645424	*rs10994860*	0.272	0.487	1.22 [0.86; 1.73]
5'UTR	10:52645409	var_52645409	0.362	1.000	0.62 [0.22; 1.74]
3	10:52603874	*rs142969066*	0.576	0.982	0.65 [0.14; 2.99]
3	10:52603754	*rs35967725*	0.857	0.534	0.86 [0.16; 4.51]
6	10:52587964	*rs143315865*	0.850	1.000	0.76 [0.05; 12.42]
7	10:52580318	*rs372408821*	0.237	0.426	3.31 [0.45; 24.17]
7	10:52576068	*rs4245008*	0.263	0.980	3.40 [0.40; 29.06]
8	10:52576025	*rs142026324*	0.935	1.000	0.90 [0.08; 10.72]
9	10:52573772	*rs41274050*	0.317	0.355	1.18 [0.86; 1.61]
3'UTR	10:52566049	*rs10821846*	0.529	0.833	1.08 [0.85; 1.37]
3'UTR	10:52559634	*rs10994507*	0.701	0.715	0.91 [0.56; 1.47]
3'UTR	10:52560658	*rs10994521*	0.850	1.000	0.76 [0.05; 12.42]
3'UTR	10:52563898	*rs112824128*	0.885	1.000	1.17 [0.14; 9.71]
3'UTR	10:52561919	*rs12571156*	0.543	0.994	1.11 [0.79; 1.57]
3'UTR	10:52563981	*rs150545950*	0.707	0.980	0.61 [0.05; 7.93]
3'UTR	10:52563196	*rs16751*	0.529	0.833	1.08 [0.85; 1.37]

3'UTR	10:52565132	*rs185182715*	0.068	1.000	0.72 [0.50; 1.03]
3'UTR	10:52563051	rs4282939	0.317	0.355	1.18 [0.86; 1.61]
3'UTR	10:52564065	*rs4619096*	0.093	0.547	0.82 [0.65; 1.03]
3'UTR	10:52565951	*rs4619097*	0.364	0.293	1.16 [0.84; 1.59]
3'UTR	10:52561643	*rs59030524*	0.429	1.000	0.38 [0.03; 4.23]
3'UTR	10:52560557	*rs6479731*	0.745	0.990	1.46 [0.15; 14.23]
3'UTR	10:52561829	*rs7072584*	0.364	0.293	1.16 [0.84; 1.59]
3'UTR	10:52559843	*rs7084132*	0.364	0.293	1.16 [0.84; 1.59]
3'UTR	10:52561803	*rs74352101*	0.364	0.293	1.16 [0.84; 1.59]
3'UTR	10:52564700	*rs74874346*	0.723	0.695	0.92 [0.57; 1.49]
3'UTR	10:52565377	*rs75583477*	0.830	0.598	0.95 [0.58; 1.54]
3'UTR	10:52564421	*rs80080606*	0.543	0.994	1.11 [0.79; 1.57]
3'UTR	10:52559853	var52559853	0.850	1.000	0.76 [0.05; 12.42]
3'UTR	10:52561920	var52561920	0.357	1.000	2.07 [0.44; 9.70]
3'UTR	10:52562642	var52562642	0.668	1.000	1.97 [0.09; 43.53]
3'UTR	10:52563862	var52563898	0.850	1.000	0.76 [0.05; 12.42]
3'UTR	10:52564524	var52564700	0.530	1.000	0.47 [0.04; 5.00]
3'UTR	10:52565940	var52565951	0.850	1.000	0.76 [0.05; 12.42]
3'UTR	10:52566333	var52566417	0.207	1.000	0.64 [0.32; 1.28]

Exon: Exon number, UTR: Un-translated region, rs ID: Reference SNP cluster ID, Chr: Chromosome, var: Variant, Ref: Reference, Alt: Alternate. Normo: Normouricaemic, Hyper: Hyperuricaemic, OR: Odds ratio for the alternate allele, 95% CI: 95% confidence interval, P_{OR}: *p*-value for odds ratio. P_{Het}: *p*-value for intra population heterogeneity. *Odds ratios are adjusted for sex, age, BMI and estimates of grand-parental ancestry for Polynesian dataset. Exons numbers and variant positions are sourced from the Genome Reference Consortium human genome build 37 (GRCh37) on Ensembl (http://grch37.ensembl.org/).

Appendix A Table 5.4: Summary of single variant association analysis within *LRP2* for eGFR in NZ Polynesians

Exon	rs ID/var number	Allele		Altered allele Frequency				Beta	95% CI		P_{Beta}
		Ref	Alt	Normouricaemic		Hyperuricaemic					
				n	f	n	f		Lower	Upper	
2	*rs144829356*	G	A	70	0.164	1	0.211	-2.169	-5.84	1.50	0.247
3	*rs2229263*	T	C	66	0.155	187	0.143	0.774	-3.42	4.97	0.718
4	*rs34104660*	G	T	3	0.007	5	0.000	-6.026	-30.86	18.81	0.635
4	var170163808	G	T	0	0.000	1	0.002	-11.914	-54.92	31.09	0.587
6	*rs2229266*	G	A	154	0.362	362	0.363	0.516	-2.67	3.70	0.751
7	var170148871	T	C	0	0.000	1	0.002	9.949	-33.00	52.90	0.650
8	*rs34693334*	C	G	4	0.009	0	0.000	3.953	-11.37	19.28	0.613
8	var170147408	A	G	0	0.000	0	0.000	-3.109	-46.07	39.85	0.887
13	*rs111360923*	T	C	47	0.110	0	0.000	-3.600	-8.71	1.51	0.168
14	*rs34291900*	C	T	2	0.005	0	0.000	-2.618	-33.06	27.82	0.866
14	*rs830994*	A	G	95	0.223	0	0.000	-1.845	-5.58	1.88	0.333
15	*rs2241190*	T	C	169	0.397	39	0.388	0.367	-2.65	3.39	0.812
15	*rs33954745*	A	G	7	0.016	43	0.011	-5.523	-18.66	7.61	0.410
15	var170115626	T	C	46	0.108	0	0.000	-3.473	-8.61	1.66	0.186
16	*rs141180155*	G	A	0	0.000	0	0.000	-16.177	-59.54	27.19	0.465
16	*rs147621120*	T	A	0	0.000	0	0.000	0.909	-42.05	43.87	0.967
16	var170127497	A	G	0	0.000	0	0.000	-7.127	-50.08	35.83	0.745
18	var170104017	A	C	2	0.005	0	0.000	-3.404	-33.80	26.99	0.826
19	*rs17848143*	C	T	5	0.012	0	0.000	9.363	-5.10	23.82	0.205
19	*rs2075249*	G	T	101	0.237	0	0.000	1.487	-1.92	4.90	0.393
19	*rs831043*	T	C	169	0.397	175	0.385	0.220	-2.81	3.25	0.887
19	var170103219	T	A	1	0.002	0	0.000	-0.095	-43.05	42.86	0.997
22	*rs144723964*	G	A	0	0.000	0	0.000	0.909	-42.05	43.87	0.967
22	*rs831042*	T	C	169	0.397	76	0.385	0.220	-2.81	3.25	0.887
23	*rs138030034*	G	A	0	0.000	55	0.002	0.909	-42.05	43.87	0.967
23	*rs17848149*	T	G	26	0.061	4	0.064	0.381	-5.62	6.38	0.901
26	*rs34915742*	C	G	0	0.000	4	0.002	0.909	-42.05	43.87	0.967
28	var170093726	T	G	5	0.012	10	0.004	-1.066	-17.42	15.29	0.898
29	*rs2229267*	A	G	128	0.300	165	0.284	-1.222	-4.53	2.08	0.469
29	var170092504	C	T	5	0.012	0	0.000	3.401	-7.88	14.68	0.555
30	*rs145384264*	C	T	2	0.005	0	0.000	-11.735	-36.55	13.08	0.355
31	*rs144054579*	C	G	0	0.000	0	0.000	0.909	-42.05	43.87	0.967
31	*rs149469954*	G	A	0	0.000	0	0.000	-7.997	-51.00	35.01	0.716

31	*rs2302694*	G	A	64	0.150	176	0.095	-2.902	-7.47	1.66	0.213
36	*rs11886219*	T	C	7	0.016	0	0.000	-5.665	-18.26	6.93	0.378
36	*rs4667596*	C	T	10	0.023	155	0.009	4.447	-7.66	16.56	0.472
39	*rs149367019*	G	A	1	0.002	0	0.000	-44.147	-86.96	-1.34	0.044
39	*rs35114151*	A	G	26	0.061	1	0.062	-0.687	-6.73	5.36	0.824
39	*rs35413340*	T	C	0	0.000	28	0.002	28.626	-14.37	71.63	0.193
39	*rs61995915*	T	C	0	0.000	4	0.002	-7.127	-50.08	35.83	0.745
41	*rs13397109*	G	C	72	0.169	5	0.121	-1.675	-5.91	2.56	0.438
42	*rs17848169*	T	C	2	0.005	29	0.002	-2.618	-33.06	27.82	0.866
42	*rs199593393*	C	T	0	0.000	5	0.002	-12.078	-55.45	31.30	0.586
44	var170058345	C	A	5	0.012	0	0.000	3.756	-11.56	19.07	0.631
45	var170055385	T	C	1	0.002	0	0.000	11.254	-31.75	54.26	0.608
46	*rs2228171*	T	C	182	0.427	131	0.412	0.305	-2.65	3.26	0.839
51	var170038806	C	T	0	0.000	0	0.000	4.318	-39.07	47.71	0.845
54	*rs2229265*	T	C	320	0.751	65	0.797	0.292	-3.16	3.75	0.868
55	*rs17848184*	C	T	85	0.200	1	0.218	1.626	-2.06	5.31	0.387
56	*rs142549310*	C	T	0	0.000	0	0.000	13.541	-29.83	56.91	0.541
56	var170030556	G	A	0	0.000	0	0.000	-7.114	-26.35	12.12	0.469
58	*rs199528723*	G	A	5	0.012	99	0.011	4.400	-9.36	18.17	0.531
61	*rs2229268*	A	G	28	0.066	129	0.086	0.673	-4.93	6.28	0.814
62	var170022511	T	C	1	0.002	1	0.000	11.254	-31.75	54.26	0.608
64	*rs79723119*	A	C	1	0.002	0	0.000	-18.965	-61.94	24.02	0.388
64	var170013979	A	G	0	0.000	0	0.000	1.244	-42.15	44.63	0.955
66	*rs2075252*	T	C	139	0.326	120	0.289	0.124	-3.06	3.31	0.939
69	*rs4667591*	T	G	170	0.399	0	0.000	0.887	-2.15	3.92	0.567
70	var170002412	T	C	1	0.002	0	0.000	-12.650	-43.05	17.75	0.415
72	*rs990626*	G	A	180	0.423	175	0.394	0.717	-2.26	3.69	0.637
74	*rs370978040*	G	A	0	0.000	0	0.000	13.320	-29.82	56.46	0.545

rs ID: Reference SNP cluster ID, Chr: Chromosome, var: Variant, Ref: Reference, Alt: Alternate.n: Total number of altered alleles, f: Altered allele frequency, Beta: Effect estimates (mL/min/1.73m^2) for the alternate allele, 95% CI: 95% confidence interval, P_{Beta}: *p*-value for effect estimates. *Values are adjusted for sex, age, BMI and estimates of grand-parental ancestry. Exons numbers and variant positions are sourced from the Genome Reference Consortium human genome build 37 (GRCh37) on Ensembl (http://grch37.ensembl.org/). Note: eGFR was calculated using CKD-EPI formula (mL/min/1.73m2) (https://www.qxmd.com/calculate/calculator_251/egfr-using-ckd-epi).

Appendix A Table 5.5: List of the variants selected for rare-variant burden analyses for *LRP2* and *A1CF*

LRP2				*A1CF*	
rs ID/var position	**CAF**	**rs ID/var position**	**CAF**	**rs ID/var position**	**CAF**
rs114460450	0.003	*rs34564141*	0.009	*rs10994521*	0.003
rs114842875	0.003	*rs34592807*	0.001	*rs112824128*	0.007
rs116332504	0.001	*rs34693334*	0.034	*rs141891504*	0.001
rs11886219	0.028	*rs34915742*	0.002	*rs142026324*	0.007
rs138030034	0.001	*rs35114151*	0.047	*rs142968717*	0.001
rs138070797	0.005	*rs35413340*	0.003	*rs142969066*	0.006
rs138269726	0.001	*rs35734447*	0.003	*rs143123872*	0.003
rs139514301	0.003	*rs35942532*	0.003	*rs143315865*	0.003
rs140272085	0.003	*rs370978040*	0.001	*rs146662131*	0.001
rs140586887	0.001	*rs371966515*	0.001	*rs150441974*	0.001
rs140918583	0.003	*rs374368151*	0.001	*rs150545950*	0.003
rs141068435	0.001	*rs375394006*	0.001	*rs16909156*	0.001
rs141180155	0.006	*rs41268685*	0.011	*rs181769526*	0.003
rs141305635	0.001	*rs4667596*	0.017	*rs183260900*	0.001
rs142093111	0.001	*rs61995915*	0.007	*rs185182715*	0.001
rs142221587	0.001	*rs7598640*	0.001	*rs34190540*	0.004
rs142245618	0.007	*rs79723119*	0.009	*rs35967725*	0.006
rs142549310	0.002	var169995201	0.001	*rs372408821*	0.001
rs142934522	0.004	var169995769	0.001	*rs41274050*	0.004
rs143150497	0.001	var169996058	0.001	*rs4619097*	0.004
rs143893803	0.001	var170002291	0.001	*rs59030524*	0.001
rs144054579	0.001	var170002412	0.002	*rs61742973*	0.001
rs144451000	0.001	var170012798	0.001	*rs61856570*	0.007
rs144723964	0.001	var170013979	0.001	*rs75583477*	0.001
rs145201961	0.003	var170022511	0.001	*rs80080606*	0.001
rs145365776	0.001	var170026248	0.003	*rs9073*	0.001
rs145384264	0.009	var170029656	0.001	var52559596	0.003
rs145709922	0.004	var170030458	0.001	var52559853	0.001
rs146149181	0.003	var170031714	0.001	var52559869	0.001
rs146289506	0.003	var170038795	0.001	var52559874	0.001
rs146783211	0.001	var170038806	0.001	var52560181	0.001
rs147267007	0.001	var170042056	0.001	var52561178	0.001
rs147287428	0.001	var170050399	0.001	var52561212	0.001
rs147621120	0.001	var170055385	0.001	var52561680	0.010
rs148356370	0.005	var170058345	0.009	var52561920	0.001
rs148503556	0.001	var170060706	0.001	var52562339	0.001
rs149148763	0.007	var170062881	0.001	var52562434	0.003
rs149367019	0.004	var170063080	0.001	var52562642	0.001
rs149469954	0.001	var170063503	0.001	var52563248	0.001
rs149558767	0.001	var170066058	0.001	var52563513	0.001
rs149853330	0.001	var170068592	0.001	var52563650	0.001
rs150552608	0.004	var170092504	0.017	var52563845	0.001
rs150752263	0.001	var170093726	0.008	var52563862	0.003
rs151079411	0.003	var170096262	0.001	var52563898	0.001
rs17848143	0.01	var170101294	0.001	var52563981	0.006
rs17848149	0.046	var170103219	0.001	var52564700	0.001
rs17848169	0.02	var170103416	0.001	var52565100	0.001
rs199528723	0.011	var170103488	0.001	var52565132	0.004
rs199593393	0.001	var170104017	0.002	var52566417	0.001
rs200475391	0.004	var170115652	0.001	var52595864	0.001
rs200587303	0.001	var170127497	0.001	var52619722	0.001
rs201490492	0.001	var170135946	0.001	var52623804	0.019
rs202154723	0.001	var170137024	0.001	var52645409	0.001
rs33954745	0.041	var170147408	0.001		
rs34104660	0.043	var170148871	0.001		
rs34291900	0.018	var170163808	0.001		
rs34355135	0.008				

CAF: Combined frequency of altered allele in NZ Polynesian and European populations, rs ID: Reference SNP cluster ID, var; Variant.

Appendix A Table 5.6: List of all exonic variants/ SNPs for *LRP2* and *A1CF* exported from Ensembl Variant Effect Predictor (V*e*!P) and their functional consequences with annotation scores

Sr. #	Exon	Location	Conseq	rs/var number	CADD	PolyPhen2		SIFT		POVEAN		Protein	
					Score	Score	Prediction	Score	Prediction	Score	Prediction	Pos	Ref|Alt
LRP2													
1	2	2:170177325	MV	*rs114460450*	20.4	Not scored	Not scored	Not scored	Not scored	Not scored	Not scored	50	S|T
2	2	2:170177382	MV	*rs144829356*	Not scored	Not scored	Not scored	Not scored	Not scored	Not scored	Not scored	31	G|E
3	3	2:170175334	MV	*rs2229263*	0.001	Not scored	Not scored	Not scored	Not scored	Not scored	Not scored	83	I|T
4	4	2:170163816	MV	*rs34104660*	2.735	-	-	-	-	-	-	134	P
5	4	2:170163827	SL	*rs34592807*	0.245	-	-	-	-	-	-	131	*|R
6	4	2:170163808	SV	var170163808	Not scored	Not scored	Not scored	Not scored	Not scored	Not scored	Not scored	137	G|V
7	6	2:170150671	SV	*rs2229266*	12.22	-	-	-	-	-	-	213	E
8	7	2:170148871	MV	var170148871	Not scored	Not scored	Not scored	Not scored	Not scored	Not scored	Not scored	221	S|P
9	8	2:170147502	MV	*rs34693334*	10.32	Not scored	Not scored	Not scored	Not scored	Not scored	Not scored	259	R|G
10	8	2:170147368	SReV	*rs375394006*	7.89	-	-	-	-	-	-	303	T
11	8	2:170147408	SV	var170147408	-	-	-	-	-	-	-	290	*
12	11	2:170136871	SG	*rs148503556*	7.379	-	-	-	-	-	-	444	L
13	11	2:170137024	SV	var170137024	0.021	-	-	-	-	-	-	393	Q|*
14	12	2:170136059	MV	*rs202154723*	9.236	Not scored	Not scored	Not scored	Not scored	Not scored	Not scored	463	I|T
15	12	2:170135946	MV	var170135946	26.7	Not scored	Not scored	Not scored	Not scored	Not scored	Not scored	501	G|R
16	13	2:170131548	MV	*rs111360923*	Not scored	Not scored	Not scored	Not scored	Not scored	Not scored	Not scored	589	F|S
17	13	2:170134324	MV	*rs141305635*	13.99	-	-	-	-	-	-	568	W|*
18	13	2:170134385	MV	*rs201490492*	31	3.583	Deleterious	0.02	Deleterious	-4.29	Deleterious	548	G|S
19	13	2:170131729	SG	*rs374368151*	24.8	2.077	Deleterious	0.5	Neutral	-3.44	Deleterious	529	R|C
20	14	2:170129529	MV	*rs116332504*	24	1.423	Deleterious	0	Deleterious	-3.8	Deleterious	606	D|G
21	14	2:170129474	MV	*rs145709922*	8.443	-	-	-	-	-	-	624	L
22	14	2:170129547	SV	*rs34291900*	33	2.385	Deleterious	0.07	Neutral	-6.41	Deleterious	600	A|V
23	14	2:170129528	SV	*rs830994*	-	-	-	-	-	-	-	606	V
24	15	2:170115588	MV	*rs2241190*	0.34	-	-	-	-	-	-	683	T
25	15	2:170115672	MV	*rs33954745*	8.978	-	-	-	-	-	-	655	E
26	15	2:170115706	MV	*rs371966515*	7.774	0.647	Neutral	0.9	Neutral	-1.05	Neutral	644	I|T
27	15	2:170115626	SV	var170115626	Not scored	-0.12	Neutral	0.33	Neutral	-0.44	Neutral	671	F|L
28	15	2:170115652	SV	var170115652	28.7	3.458	Deleterious	0	Deleterious	-8.28	Deleterious	662	F|S
29	16	2:170127559	MV	*rs141180155*	12.92	-	-	-	-	-	-	725	T
30	16	2:170127556	MV	*rs144451000*	12.96	0.879	Neutral	0.08	Neutral	-2.31	Neutral	726	F|L
31	16	2:170113671	MV	*rs147621120*	Not scored	0.386	Neutral	0.25	Neutral	-2.4	Neutral	731	S|T
32	16	2:170113670	MV	*rs150752263*	19.02	0.386	Neutral	0.25	Neutral	-2.4	Neutral	731	S|T
33	16	2:170127497	SV	var170127497	Not scored	1.757	Deleterious	0.01	Deleterious	-3.39	Deleterious	746	N|S
34	18	2:170104017	MV	var170104017	Not scored	1.583	Deleterious	0	Deleterious	-4.57	Deleterious	790	I|L
35	19	2:170103472	MV	*rs114842875*	15.16	2.966	Deleterious	0.18	Neutral	-2.35	Neutral	841	R|K
36	19	2:170103471	MV	*rs17848143*	-	-	-	-	-	-	-	841	T

37	19	2:170103351	SV	*rs2075249*	0.593	-	-	-	-	-	-	881	T
38	19	2:170103336	SV	*rs831043*	0.126	-	-	-	-	-	-	886	T
39	19	2:170103219	SV	var170103219	-	-	-	-	-	-	-	925	T
40	19	2:170103416	SV	var170103416	27.2	-	-	-	-	-	-	860	L
41	19	2:170103488	SV	var170103488	13.07	1.462	Deleterious	0.65	Neutral	-0.42	Neutral	836	P|S
42	20	2:170101269	MV	*rs149853330*	3.687	0.171	Neutral	0.11	Neutral	-0.59	Neutral	985	F|L
43	20	2:170101294	SV	var170101294	0.008	-	-	-	-	-	-	976	H
44	21	2:170100011	MV	*rs150552608*	10.62	1.605	Neutral	0.5	Neutral	-3.42	Deleterious	1014	R|H
45	22	2:170099474	MV	*rs144723964*	Not scored	0.464	Neutral	0.32	Neutral	-0.73	Neutral	1083	G|E
46	22	2:170099473	SV	*rs831042*	10.6	-	-	-	-	-	-	1083	A
47	23	2:170097655	MV	*rs138030034*	-	-	-	-	-	-	-	1159	W|*
48	23	2:170097707	SG	*rs17848149*	0.758	0.684	Neutral	0.74	Neutral	-3.97	Deleterious	1142	V|G
49	23	2:170097508	SV	*rs200587303*	10.16	-	-	-	-	-	-	1208	S
50	26	2:170096095	MV	*rs34915742*	11.14	-	-	-	-	-	-	1412	R
51	26	2:170096262	SV	var170096262	24.6	1.634	Deleterious	0.37	Neutral	-3.03	Deleterious	1357	Y|H
52	27	2:170094756	MV	*rs146289506*	12.52	1.642	Neutral	0.04	Deleterious	-1.63	Neutral	1451	L|I
53	28	2:170093726	SV	var170093726	-	-	-	-	-	-	-	1526	T
54	29	2:170092467	MV	*rs141068435*	11.62	-	-	-	-	-	-	1601	W|*
55	29	2:170092439	MV	*rs151079411*	11.52	Not scored	Not scored	Not scored	Not scored	Not scored	Not scored	1611	V|M
56	29	2:170092395	SG	*rs2229267*	6.06	-	-	-	-	-	-	1625	*|W
57	29	2:170092504	SL	var170092504	Not scored	2.726	Deleterious	0	Deleterious	-4.85	Deleterious	1589	P|L
58	30	2:170089947	MV	*rs145365776*	13.38	0.476	Neutral	1	Neutral	0.15	Neutral	1691	G|E
59	30	2:170089934	SV	*rs145384264*	14.25	-	-	-	-	-	-	1695	S
60	31	2:170088328	MV	*rs142221587*	18.15	0.471	Neutral	0.22	Neutral	-0.68	Neutral	1708	P|L
61	31	2:170088296	MV	*rs144054579*	Not scored	1.042	Neutral	0.54	Neutral	-0.86	Neutral	1719	R|G
62	31	2:170088242	MV	*rs149469954*	11.88	1.621	Neutral	0.06	Neutral	-0.01	Neutral	1737	V|I
63	31	2:170088351	SRV	*rs2302694*	3.454	-	-	-	-	-	-	1700	S
64	32	2:170082936	MV	*rs138070797*	9.969	1.557	Deleterious	0.77	Neutral	-1.53	Neutral	1797	I|T
65	36	2:170070348	MV	*rs11886219*	4.612	-	-	-	-	-	-	1953	R
66	36	2:170070172	SV	*rs4667596*	17.88	0.496	Neutral	0.99	Neutral	-0.13	Neutral	2012	T|I
67	37	2:170068598	MV	*rs138269726*	23.7	1.943	Deleterious	0.62	Neutral	-1.85	Neutral	2054	H|Y
68	37	2:170068502	MV	*rs146149181*	23	-0.021	Neutral	0.07	Neutral	0.64	Neutral	2086	S|T
69	37	2:170068592	MV	var170068592	34	3.621	Deleterious	0.02	Deleterious	-3.34	Deleterious	2056	G|R
70	38	2:170066058	MV	var170066058	23.1	2.726	Deleterious	0.54	Neutral	-1.05	Neutral	2125	P|L
71	39	2:170063372	MV	*rs140918583*	15.51	0.49	Neutral	0.26	Neutral	-3.83	Deleterious	2286	L|F
72	39	2:170063339	MV	*rs143893803*	25	2.277	Deleterious	0.24	Neutral	-3.53	Deleterious	2297	F|L
73	39	2:170063223	MV	*rs147287428*	11.29	1.72	Deleterious	0.69	Neutral	-0.32	Neutral	2336	P|L
74	39	2:170063263	MV	*rs149367019*	6.299	Not scored	Not scored	Not scored	Not scored	Not scored	Not scored	2323	V|I
75	39	2:170063471	MV	*rs35114151*	1.78	-	-	-	-	-	-	2253	E
76	39	2:170063380	MV	*rs35413340*	5.194	0.548	Neutral	1	Neutral	0.06	Neutral	2284	S|P
77	39	2:170062977	MV	*rs61995915*	0.016	0.053	Neutral	0.54	Neutral	-0.96	Neutral	2418	V|A
78	39	2:170062881	MV	var170062881	9.935	1.599	Neutral	0.14	Neutral	-2.45	Neutral	2450	S|N
79	39	2:170063080	SG	var170063080	3.326	0.043	Neutral	0.67	Neutral	-0.02	Neutral	2384	F|L

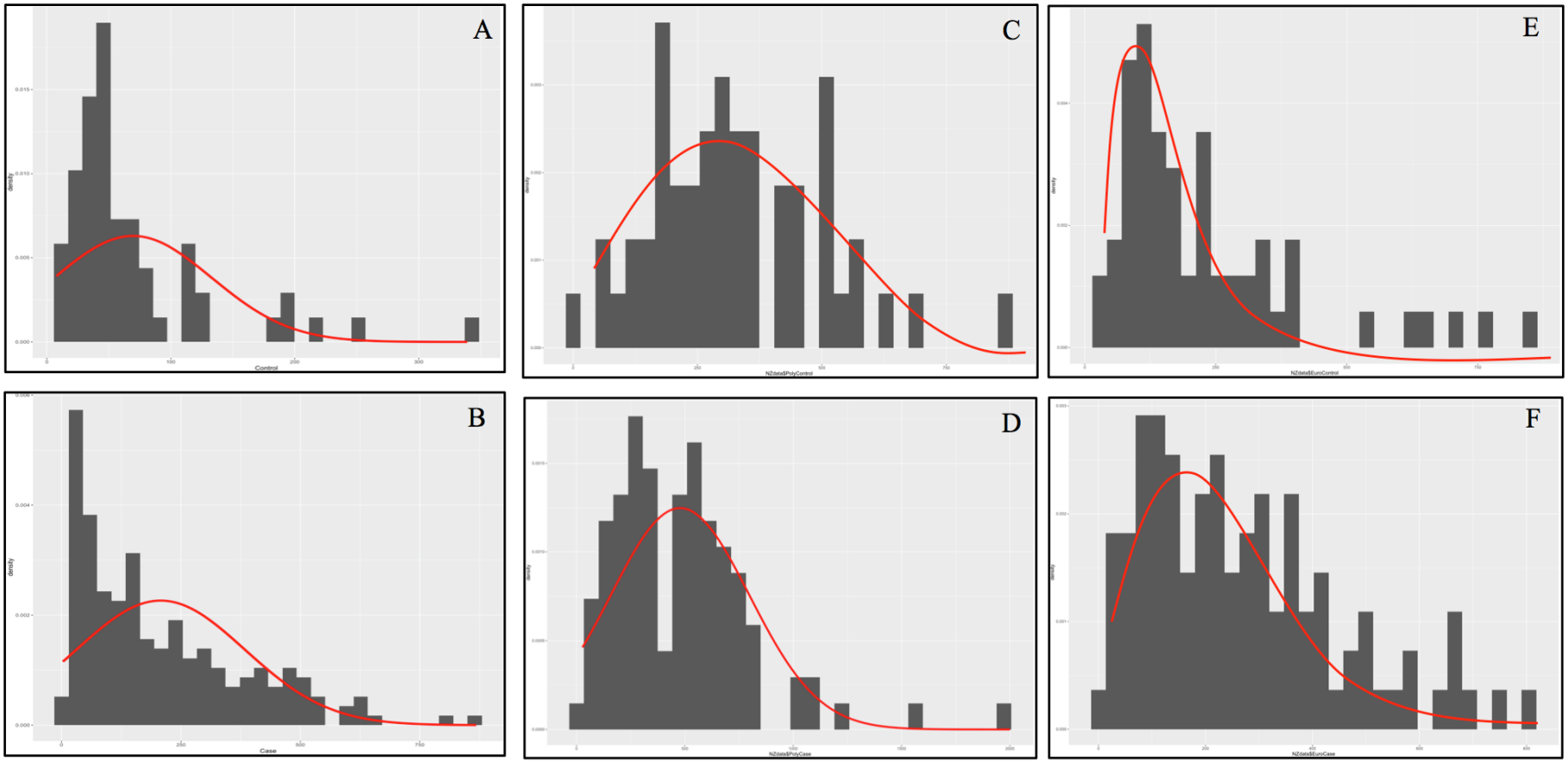
80	39	2:170063503	SV	var170063503	7.804	-	-	-	-	-	-	2243	G|*
81	41	2:170062078	SV	*rs13397109*	9.831	-	-	-	-	-	-	2542	R
82	42	2:170060631	MV	*rs139514301*	11.51	-	-	-	-	-	-	2622	E
83	42	2:170060750	SV	*rs145201961*	11.03	-	-	-	-	-	-	2583	L
84	42	2:170060603	SV	*rs17848169*	0.307	1.407	Deleterious	0.87	Neutral	-1.73	Neutral	2632	Y|H
85	42	2:170060619	SV	*rs199593393*	-	-	-	-	-	-	-	2626	H
86	42	2:170060706	SV	var170060706	10.98	-	-	-	-	-	-	2597	G
87	44	2:170058345	MV	var170058345	Not scored	1.267	Neutral	0.83	Neutral	0.2	Neutral	2749	P|T
88	45	2:170055416	MV	*rs140586887*	24.6	1.637	Deleterious	0.48	Neutral	-1.67	Neutral	2820	G|C
89	45	2:170055385	MV	var170055385	Not scored	-0.216	Neutral	0.41	Neutral	0.02	Neutral	2830	L|P
90	46	2:170053505	MV	*rs2228171*	0.003	-1.003	Neutral	-	Not scored	-	-	2872	S|P
91	47	2:170050399	MV	var170050399	0.001	0.143	Neutral	0.37	Neutral	-1.19	Neutral	2901	L|P
92	48	2:170048482	SV	*rs149148763*	13.31	-	-	-	-	-	-	2964	S
93	49	2:170044768	MV	*rs142093111*	23.9	0.848	Neutral	0.13	Neutral	-3.9	Deleterious	3014	G|R
94	50	2:170042208	MV	*rs143150497*	10.43	2.151	Deleterious	0.06	Neutral	-1.54	Neutral	3217	F|S
95	50	2:170042495	MV	*rs146783211*	7.817	-	-	-	-	-	-	3121	G
96	50	2:170042245	SL	*rs35734447*	27.8	1.634	Deleterious	0.01	Deleterious	-4.96	Deleterious	3205	Y|H
97	50	2:170042056	SV	var170042056	6.843	-	-	-	-	-	-	3268	*|R
98	51	2:170038694	MV	*rs149558767*	2.376	-	-	-	-	-	-	3327	A
99	51	2:170038795	MV	var170038795	25	2.199	Deleterious	0.01	Deleterious	-3.22	Deleterious	3294	V|I
100	51	2:170038806	SV	var170038806	Not scored	-0.937	Neutral	0.6	Neutral	2.84	Neutral	3290	P|L
101	52	2:170037962	SV	*rs140272085*	23.5	-	-	-	-	-	-	3389	L
102	53	2:170034424	MV	*rs147267007*	26.6	1.765	Deleterious	0.06	Neutral	-5.06	Deleterious	3428	W|R
103	54	2:170032989	SV	*rs2229265*	2.762	-	-	-	-	-	-	3501	H
104	55	2:170031824	MV	*rs17848184*	-	-	-	-	-	-	-	3549	P
105	55	2:170031714	SV	var170031714	12.34	-0.087	Neutral	0.27	Neutral	0.4	Neutral	3586	P|L
106	56	2:170030506	MV	*rs142549310*	19.26	0.514	Neutral	0.12	Neutral	-3.05	Deleterious	3646	P|L
107	56	2:170030458	MV	var170030458	28.1	2.023	Deleterious	0	Deleterious	-7.67	Deleterious	3662	V|G
108	56	2:170030556	SV	var170030556	-	-	-	-	-	-	-	3629	Q
109	57	2:170029657	MV	*rs34355135*	13.06	-	-	-	-	-	-	3698	L
110	57	2:170029656	SV	var170029656	13.97	1.423	Deleterious	0.51	Neutral	-2.02	Neutral	3698	E|G
111	58	2:170028529	SV	*rs199528723*	-	-	-	-	-	-	-	3753	K
112	59	2:170027153	MV	*rs200475391*	32	2.148	Deleterious	0	Deleterious	-6.94	Deleterious	3763	V|E
113	59	2:170027095	SV	*rs7598640*	12.18	-	-	-	-	-	-	3782	K
114	60	2:170026248	MV	var170026248	23	1.657	Deleterious	0.58	Neutral	-1.59	Neutral	3821	P|S
115	61	2:170025083	SL	*rs2229268*	4.949	-	-	-	-	-	-	3867	*|W
116	62	2:170022511	MV	var170022511	Not scored	0.186	Neutral	0.44	Neutral	-0.64	Neutral	3897	F|L
117	64	2:170013904	MV	*rs79723119*	0.005	0.759	Neutral	0.42	Neutral	-0.34	Neutral	3999	D|A
118	64	2:170013979	MV	var170013979	Not scored	1.665	Deleterious	0.62	Neutral	-0.53	Neutral	3974	K|R
119	65	2:170012798	MV	var170012798	22.4	0.253	Neutral	0.59	Neutral	-0.15	Neutral	4046	P|L
120	66	2:170010985	SL	*rs2075252*	16.49	-	-	-	-	-	-	4094	*|Q
121	67	2:170009390	MV	*rs142934522*	11.96	2.116	Deleterious	0.22	Neutral	0.78	Neutral	4127	P|L
122	67	2:170009391	MV	*rs148356370*	5.278	-0.329	Neutral	0.8	Neutral	0.58	Neutral	4127	G|C

123	69	2:170003335	MV	*rs35942532*	22.1	1.609	Deleterious	0.51	Neutral	-3.63	Deleterious	4242	V|A
124	69	2:170003432	MV	*rs4667591*	31	0.922	Deleterious	0.03	Deleterious	-1.91	Neutral	4210	F|V
125	70	2:170002291	MV	var170002291	5.035	-	-	-	-	-	-	4318	V
126	70	2:170002412	SV	var170002412	Not scored	1.853	Deleterious	0.01	Deleterious	-6.15	Deleterious	4278	V|A
127	72	2:169997051	MV	*rs990626*	14.46	Not scored	Not scored	Not scored	Not scored	Not scored	Not scored	4371	M|I
128	73	2:169996070	MV	*rs41268685*	23	2.095	Deleterious	0.11	Neutral	-2.58	Deleterious	4417	A|V
129	73	2:169996058	MV	var169996058	29.7	2.1	Deleterious	0.03	Deleterious	-3.15	Deleterious	4421	A|V
130	74	2:169995880	MV	*rs370978040*	-	-	-	-	-	-	-	4423	T
131	74	2:169995769	SRV	var169995769	25.3	1.991	Deleterious	0.01	Deleterious	-2.13	Neutral	4460	N|K
132	75	2:169995201	SV	var169995201	9.565	-	-	-	-	-	-	4468	V
133	77	2:169989127	MV	*rs142245618*	6.369	0.076	Neutral	0.79	Neutral	-1.31	Neutral	4562	D|G
134	79	2:169985338	SRV	*rs34564141*	22.9	-	-	-	-	-	-	4601	I
						A1CF							
1	5'	10:52645409	5' UTR	var52645409									
2	5'	10:52645424	5' UTR	*rs10994860*	-	-	-	-	-	-	-	-	-
3	5'	10:52619722	5' UTR	var52619722	-	-	-	-	-	-	-	-	-
4	2	10:52623804	NCT	var52623804	-	-	-	-	-	-	-	-	-
5	3	10:52603754	SL	*rs35967725*	-	-	-	-	-	-	-	76	*|W
6	3	10:52603874	SV	*rs142969066*	-	-	-	-	-	-	-	36	G
7	4	10:52601702	SV	*rs141891504*	-	-	-	-	-	-	-	95	K
8	5	10:52595853	SV	*rs61742973*	-	-	-	-	-	-	-	195	A
9	5	10:52595864	MV	var52595864	Not scored	Not scored	Not scored	0	Deleterious	-1.99	Neutral	192	P|S
10	5	10:52595978	MV	*rs143123872*	10.73	1.545	Deleterious	0.03	Deleterious	-1.99	Deleterious	154	T|A
11	6	10:52587964	MV	*rs143315865*	12.64	2.381	Deleterious	0.77	Tolerated	-1.99	Neutral	232	M|I
12	6	10:52588045	SV	*rs146662131*	-	-	-	-	-	-	-	205	H
13	7	10:52576068	NCT	*rs4245008*	-	-	-	-	-	-	-	-	-
14	7	10:52580318	SV	*rs372408821*	-	-	-	-	-	-	-	287	K
15	8	10:52576025	SV	*rs142026324*	-	-	-	-	-	-	-	294	S
16	9	10:52573772	MV	*rs41274050*	23.8	1.792	Deleterious	0.24	Tolerated	-1.99	Neutral	390	R|C
17	12	10:52566594	SV	*rs34190540*	-	-	-	-	-	-	-	552	L
18	12	10:52566611	SV	*rs9073*	-	-	-	-	-	-	-	547	L
19	3'	10:52559291	3' UTR	*rs61856570*	-	-	-	-	-	-	-	-	-
20	3'	10:52559596	3' UTR	var52559596	-	-	-	-	-	-	-	-	-
21	3'	10:52559634	3' UTR	*rs10994507*	-	-	-	-	-	-	-	-	-
22	3'	10:52559843	3' UTR	*rs7084132*	-	-	-	-	-	-	-	-	-
23	3'	10:52559853	3' UTR	var52559853	-	-	-	-	-	-	-	-	-
24	3'	10:52559869	3' UTR	var52559869	-	-	-	-	-	-	-	-	-
25	3'	10:52559874	3' UTR	var52559874	-	-	-	-	-	-	-	-	-
26	3'	10:52560181	3' UTR	var52560181	-	-	-	-	-	-	-	-	-
27	3'	10:52560476	3' UTR	*rs75907017*	-	-	-	-	-	-	-	-	-
28	3'	10:52560557	3' UTR	*rs6479731*	-	-	-	-	-	-	-	-	-
29	3'	10:52560658	3' UTR	*rs10994521*	-	-	-	-	-	-	-	-	-

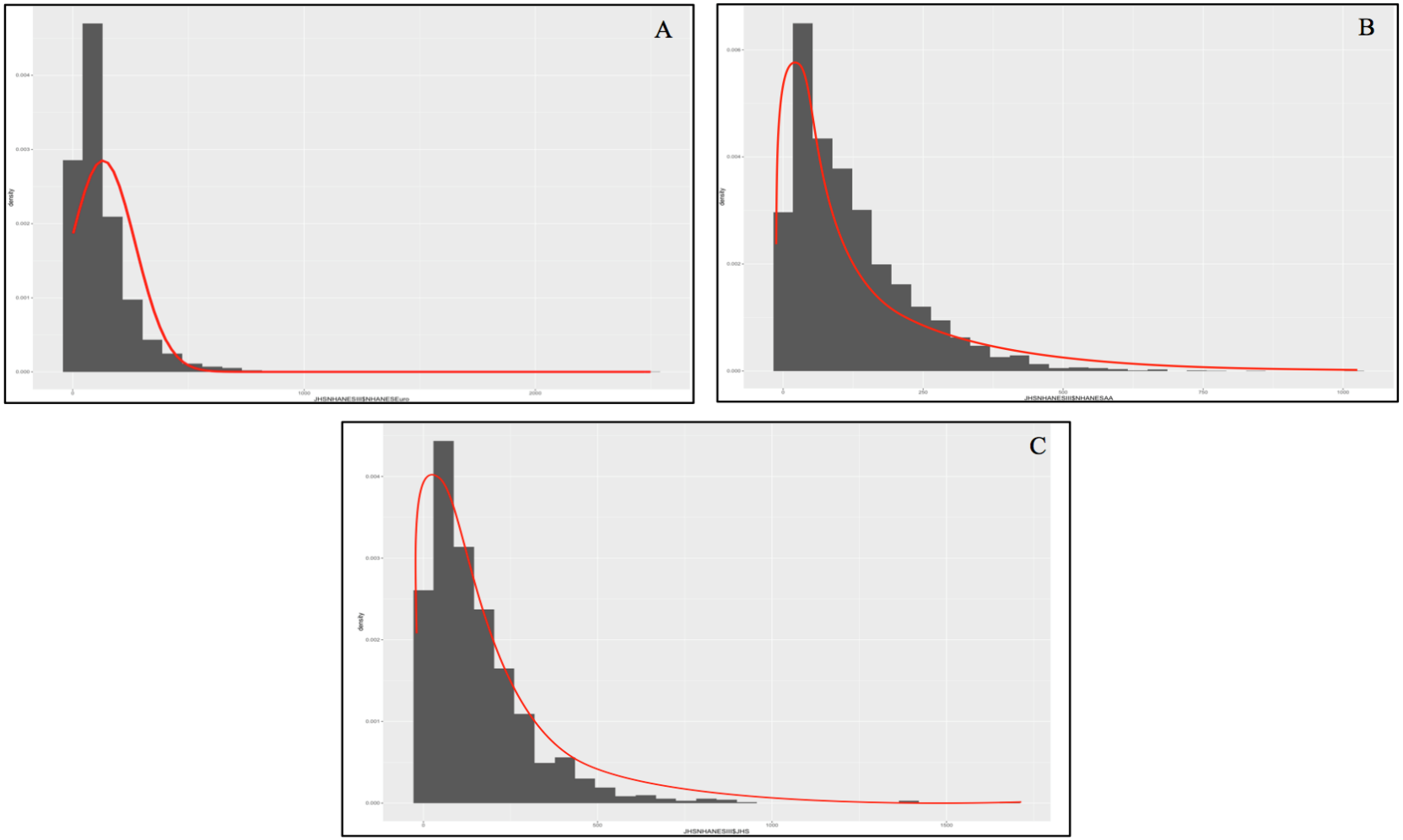
30	3'	10:52560693	3' UTR	*rs142968717*	-	-	-	-	-	-	-	-	-
31	3'	10:52561178	3' UTR	var52561178	-	-	-	-	-	-	-	-	-
32	3'	10:52561212	3' UTR	var52561212	-	-	-	-	-	-	-	-	-
33	3'	10:52561643	3' UTR	*rs59030524*	-	-	-	-	-	-	-	-	-
34	3'	10:52561680	3' UTR	var52561680	-	-	-	-	-	-	-	-	-
35	3'	10:52561803	3' UTR	*rs74352101*	-	-	-	-	-	-	-	-	-
36	3'	10:52561829	3' UTR	*rs7072584*	-	-	-	-	-	-	-	-	-
37	3'	10:52561919	3' UTR	*rs12571156*	-	-	-	-	-	-	-	-	-
38	3'	10:52561920	3' UTR	var52561920	-	-	-	-	-	-	-	-	-
39	3'	10:52562099	3' UTR	*rs150441974*	-	-	-	-	-	-	-	-	-
40	3'	10:52562339	3' UTR	var52562339	-	-	-	-	-	-	-	-	-
41	3'	10:52562434	3' UTR	var52562434	-	-	-	-	-	-	-	-	-
42	3'	10:52562642	3' UTR	var52562642	-	-	-	-	-	-	-	-	-
43	3'	10:52563051	3' UTR	*rs4282939*	-	-	-	-	-	-	-	-	-
44	3'	10:52563196	3' UTR	*rs16751*	-	-	-	-	-	-	-	-	-
45	3'	10:52563248	3' UTR	var52563248	-	-	-	-	-	-	-	-	-
46	3'	10:52563345	3' UTR	*rs16909156*	-	-	-	-	-	-	-	-	-
47	3'	10:52563513	3' UTR	var52563513	-	-	-	-	-	-	-	-	-
48	3'	10:52563650	3' UTR	var52563650	-	-	-	-	-	-	-	-	-
49	3'	10:52563707	3' UTR	*rs184644838*	-	-	-	-	-	-	-	-	-
50	3'	10:52563725	3' UTR	var52563845	-	-	-	-	-	-	-	-	-
51	3'	10:52563845	3' UTR	var52563862	-	-	-	-	-	-	-	-	-
52	3'	10:52563862	3' UTR	var52563898	-	-	-	-	-	-	-	-	-
53	3'	10:52563898	3' UTR	*rs112824128*	-	-	-	-	-	-	-	-	-
54	3'	10:52563904	3' UTR	var52563981	-	-	-	-	-	-	-	-	-
55	3'	10:52563981	3' UTR	*rs150545950*	-	-	-	-	-	-	-	-	-
56	3'	10:52564065	3' UTR	*rs4619096*	-	-	-	-	-	-	-	-	-
57	3'	10:52564421	3' UTR	*rs80080606*	-	-	-	-	-	-	-	-	-
58	3'	10:52564524	3' UTR	var52564700	-	-	-	-	-	-	-	-	-
59	3'	10:52564700	3' UTR	*rs74874346*	-	-	-	-	-	-	-	-	-
60	3'	10:52564768	3' UTR	var52565100	-	-	-	-	-	-	-	-	-
61	3'	10:52565100	3' UTR	var52565132	-	-	-	-	-	-	-	-	-
62	3'	10:52565132	3' UTR	*rs185182715*	-	-	-	-	-	-	-	-	-
63	3'	10:52565377	3' UTR	*rs75583477*	-	-	-	-	-	-	-	-	-
64	3'	10:52565903	3' UTR	*rs4078160*	-	-	-	-	-	-	-	-	-
65	3'	10:52565940	3' UTR	var52565951	-	-	-	-	-	-	-	-	-
66	3'	10:52565951	3' UTR	*rs4619097*	-	-	-	-	-	-	-	-	-
67	3'	10:52566049	3' UTR	*rs10821846*	-	-	-	-	-	-	-	-	-
68	3'	10:52566057	3' UTR	*rs183260900*	-	-	-	-	-	-	-	-	-
69	3'	10:52566333	3' UTR	var52566417	-	-	-	-	-	-	-	-	-

MS: Missense variant, NS: Non-synonymous variant, SRV: Splice region variant, SReV: Splice retained variant, *: Stop, SL/G: Stop lost/gained, NCT: Non-coding transcript, Ref|Alt: Reference|Alternate amino acid, Conseq: Functional consequence, Pos: Position, UTR: Untranslated region.

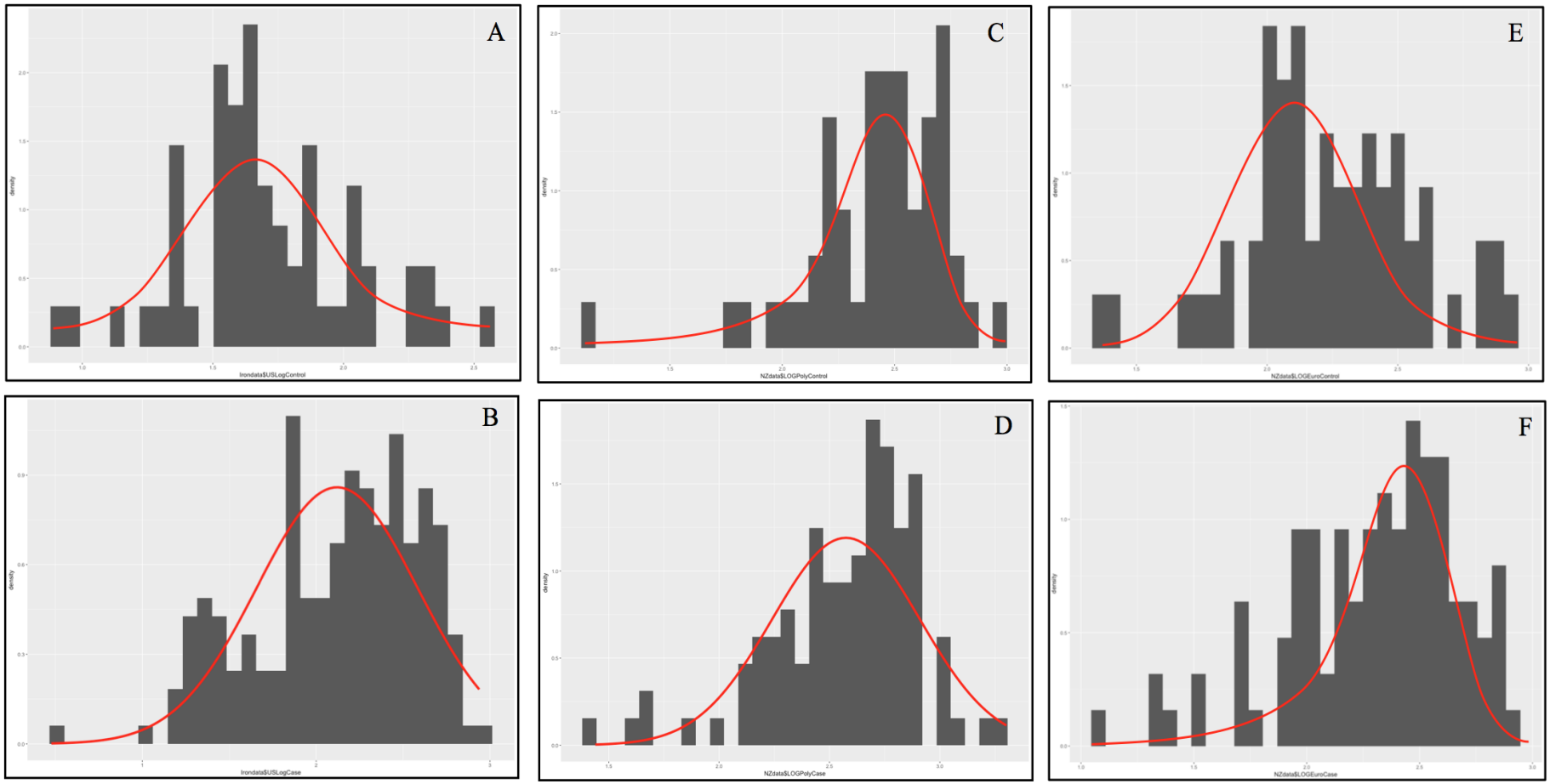
APPENDIX B



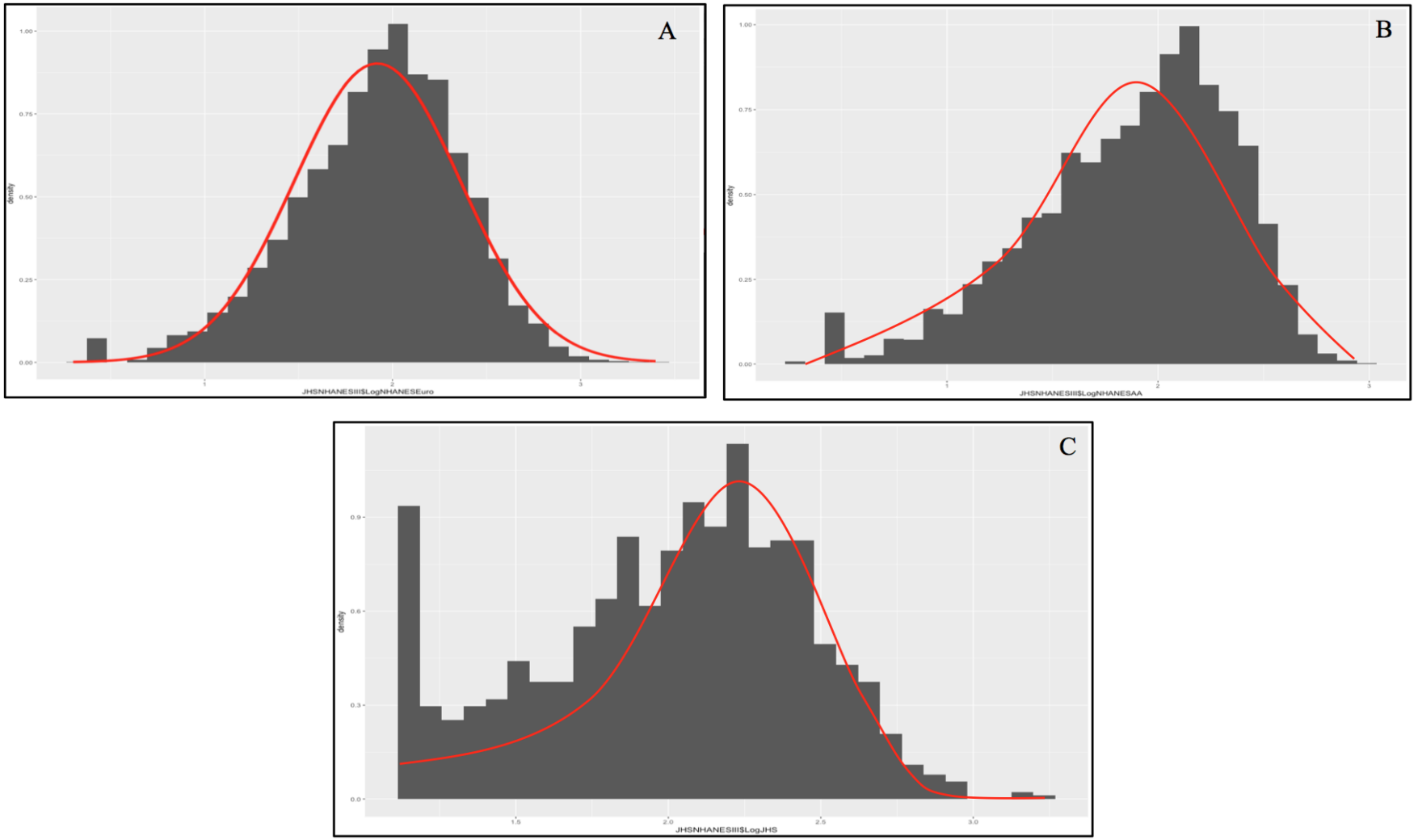
Appendix B Figure 2.1: Histograms indicating non-linear distribution of ferritin data in US (A) Control and (B) Gout, NZ Polynesian (C) Control and (D) Gout and NZ European (E) Control and (F) Gout case groups.



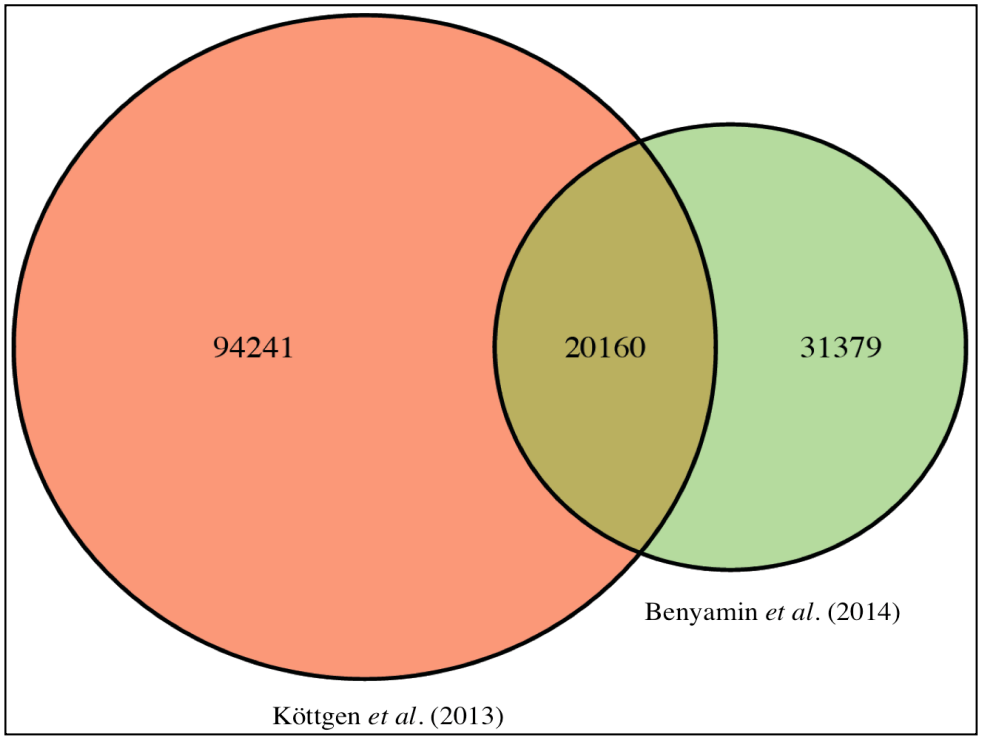
Appendix B Figure 2.2: Histograms indicating non-linear distribution of ferritin data in (A) NHANESIII European, (B) NHANESIII African American and (C) JHS African American non-gout (control) groups.



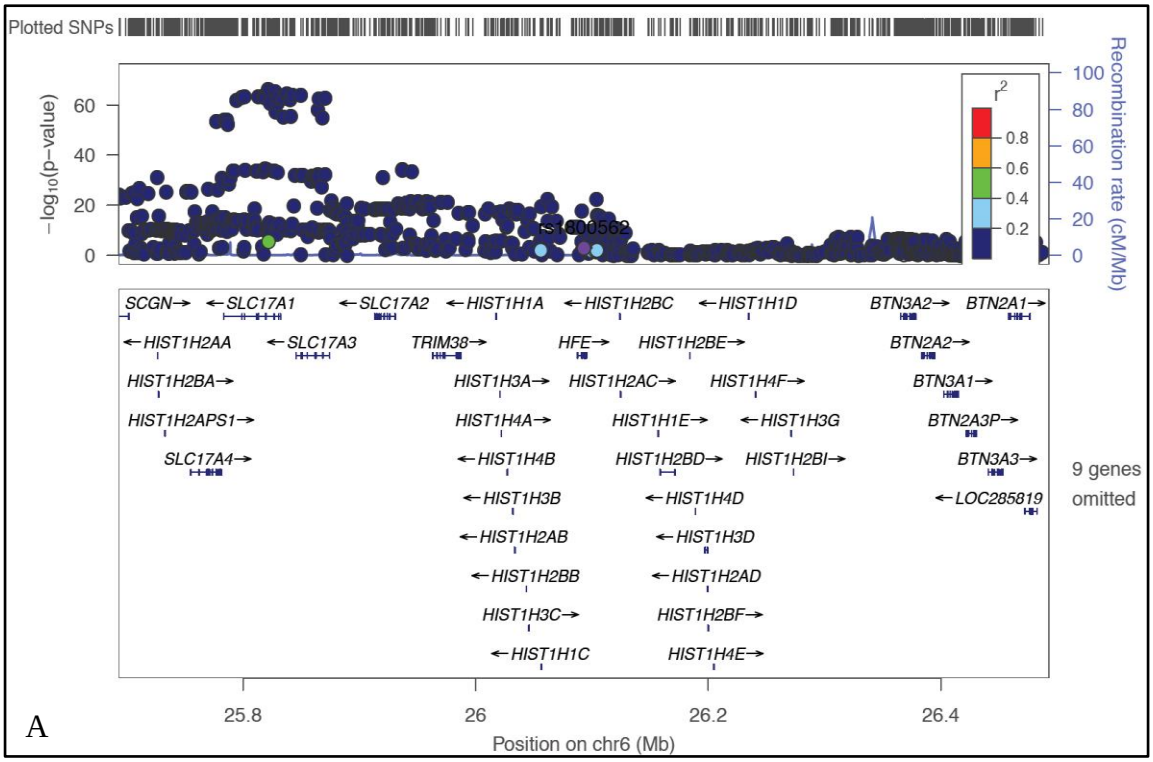
Appendix B Figure 2.3: Histograms indicating log-transformed distribution of ferritin data in US (A) Control and (B) Gout, NZ Polynesian (C) Control and (D) Gout and NZ European (E) Control and (F) Gout case groups.

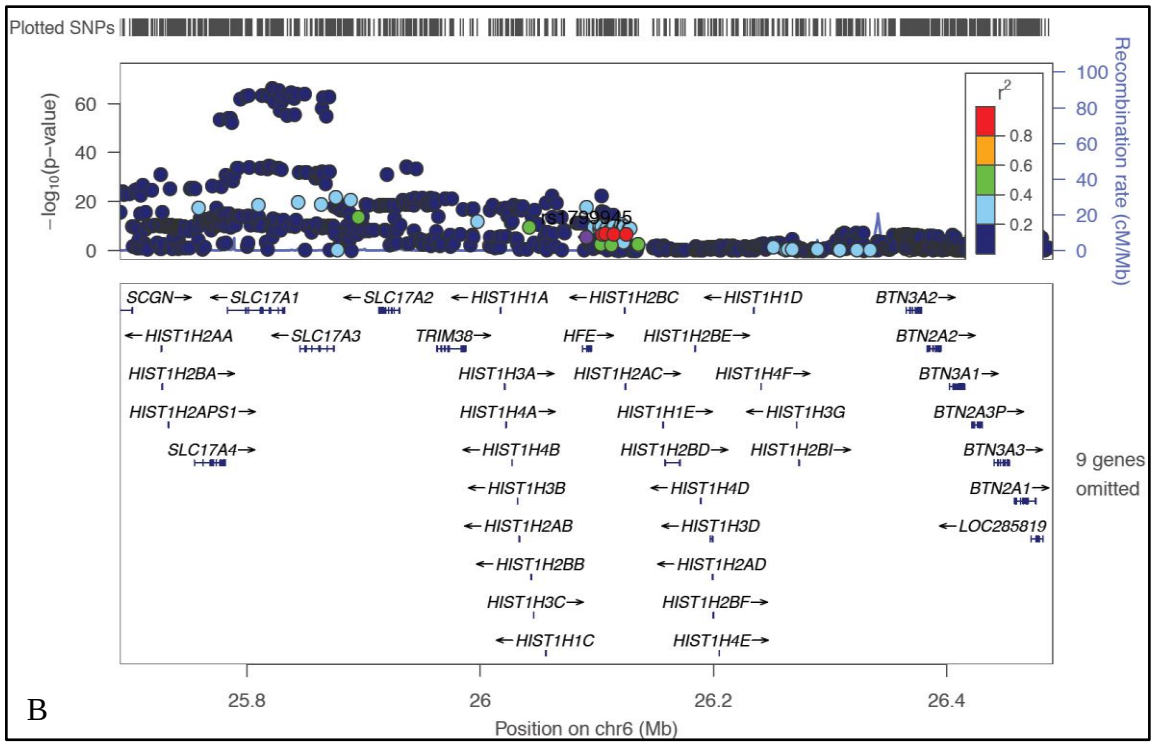


Appendix B Figure 2.4: Histograms indicating log-transformed distribution of ferritin data in (A) NHANESIII European, (B) NHANESIII African American and (C) JHS African American non-gout (control) groups.

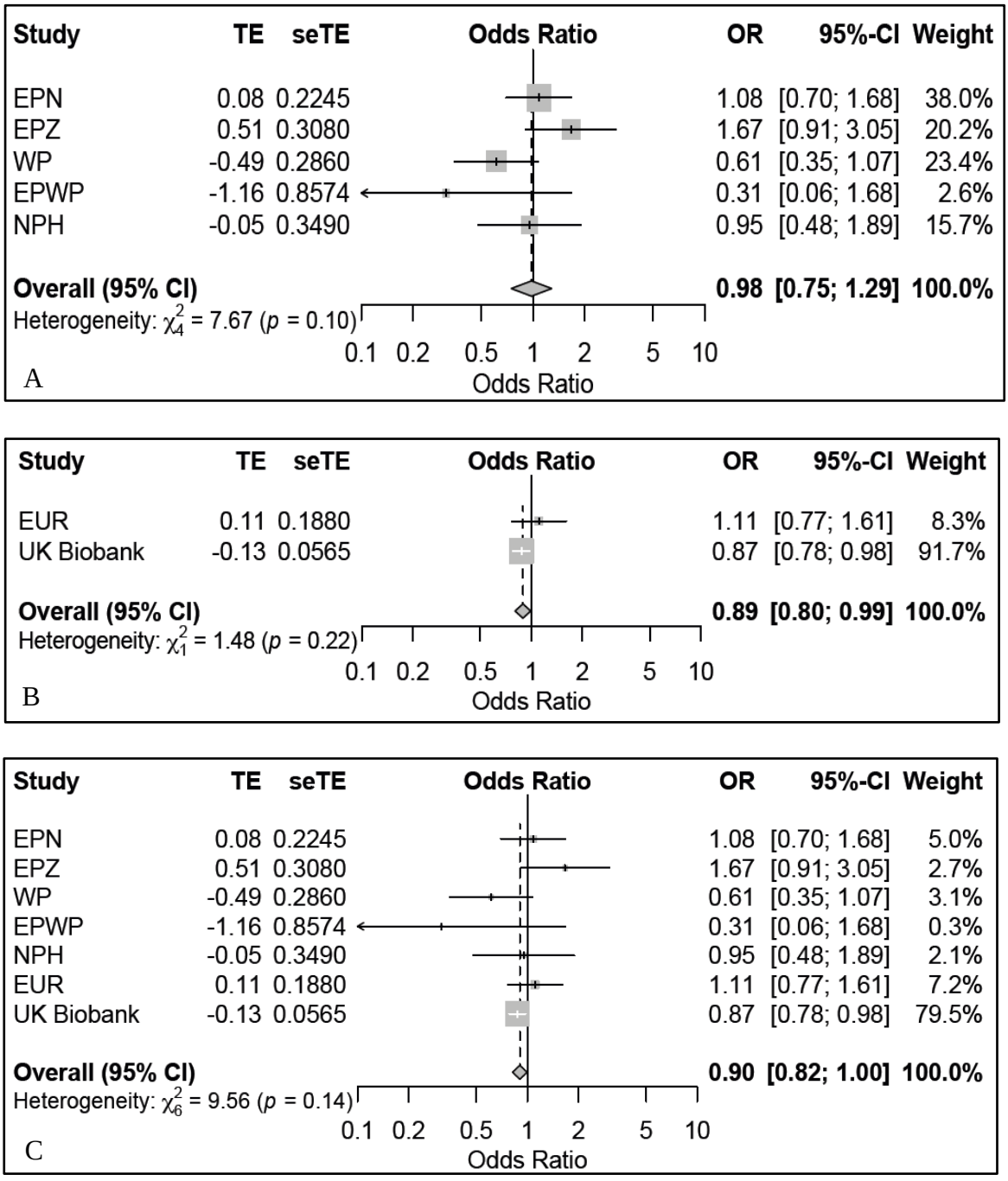


Appendix B Figure 3.1: Venn diagram representing total sample overlap between the two GWAS (Benyamin *et al.*, 2014; Köttgen *et al.*, 2013) selected for two-sample Mendelian randomisation analysis.

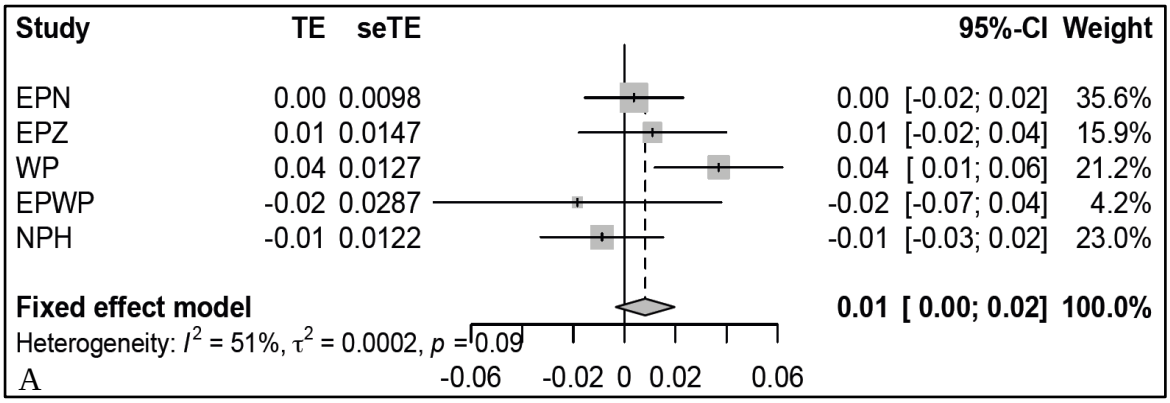


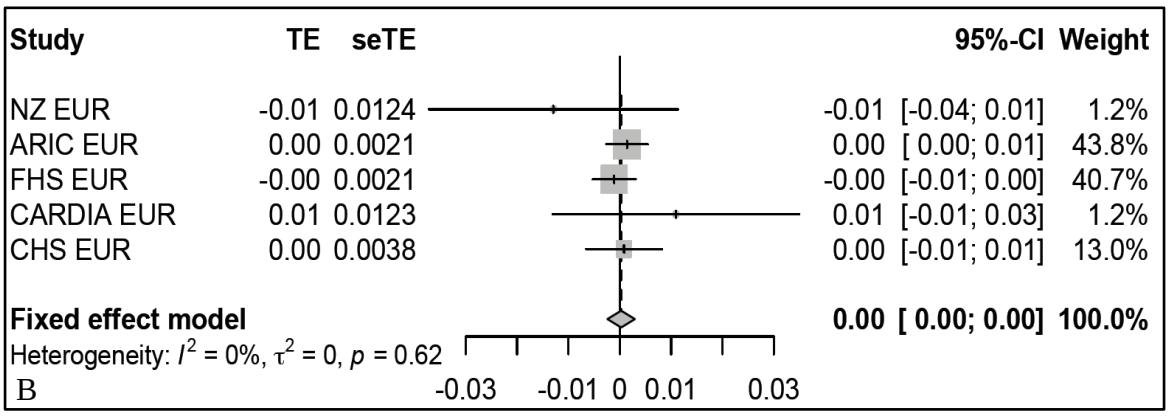


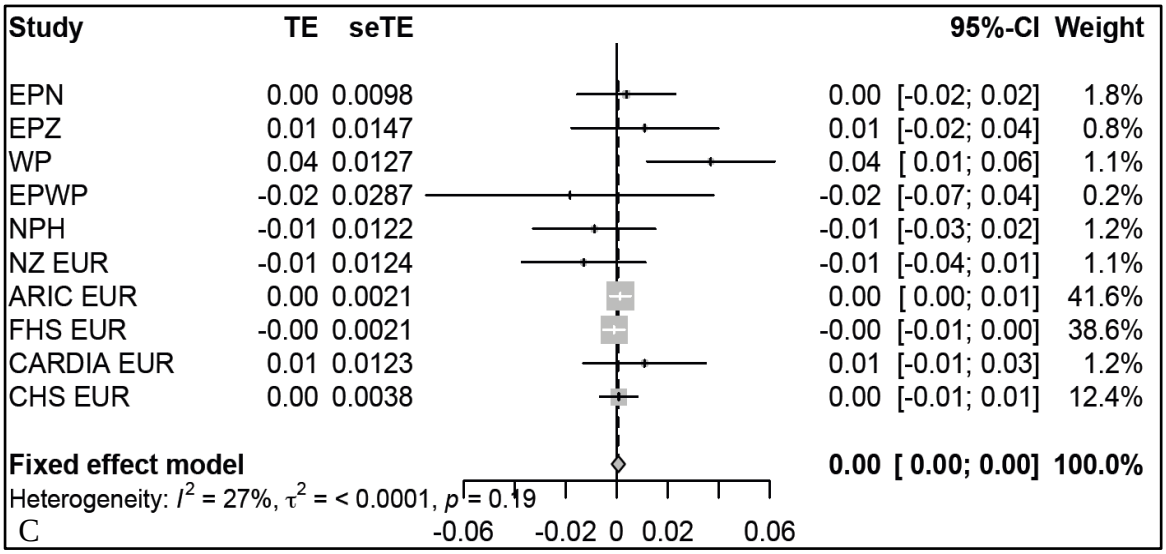
Appendix B Figure 3.2: Locus zoom plot indicating two *HFE* gene variants (A) *rs1800562* (*P* = 0.001) and (B) *rs1799945* (*P* = 2.94E-06) having an LD with variants within/near *SLC17A1-3* loci in a recent GWAS (Köttgen *et al*., 2013). Plots were created online (http://locuszoom.sph.umich.edu/).



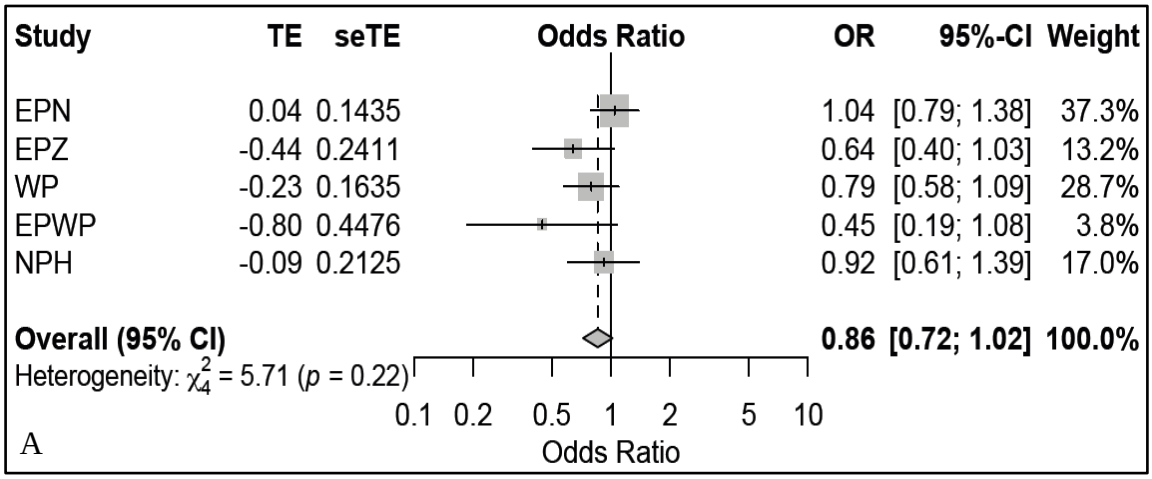
Appendix B Figure 4.1: Meta-analysis of (A) NZ Polynesian, (B) European and (C) All sample sets combined for association of *rs4994* (G allele) with gout.

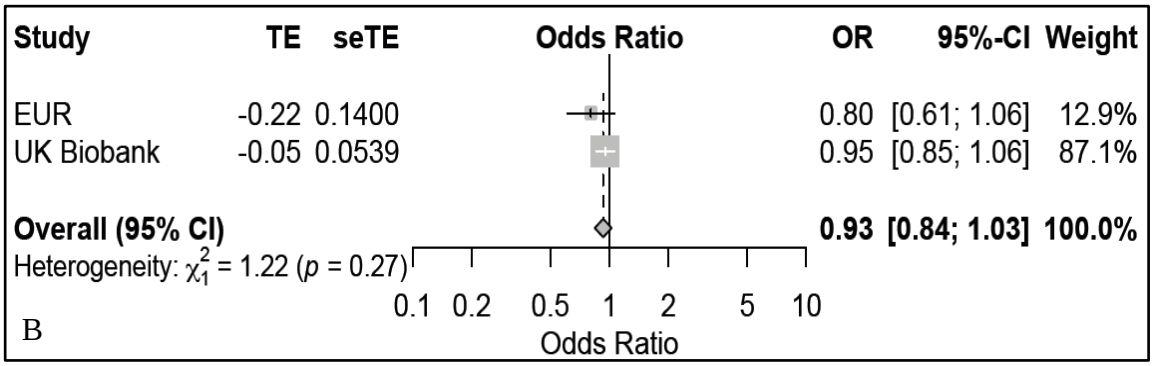


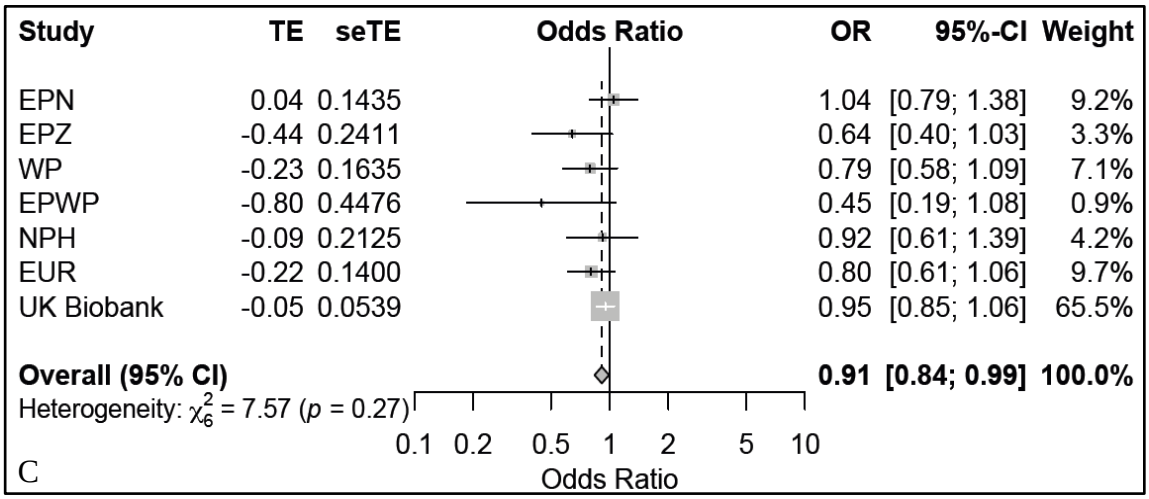




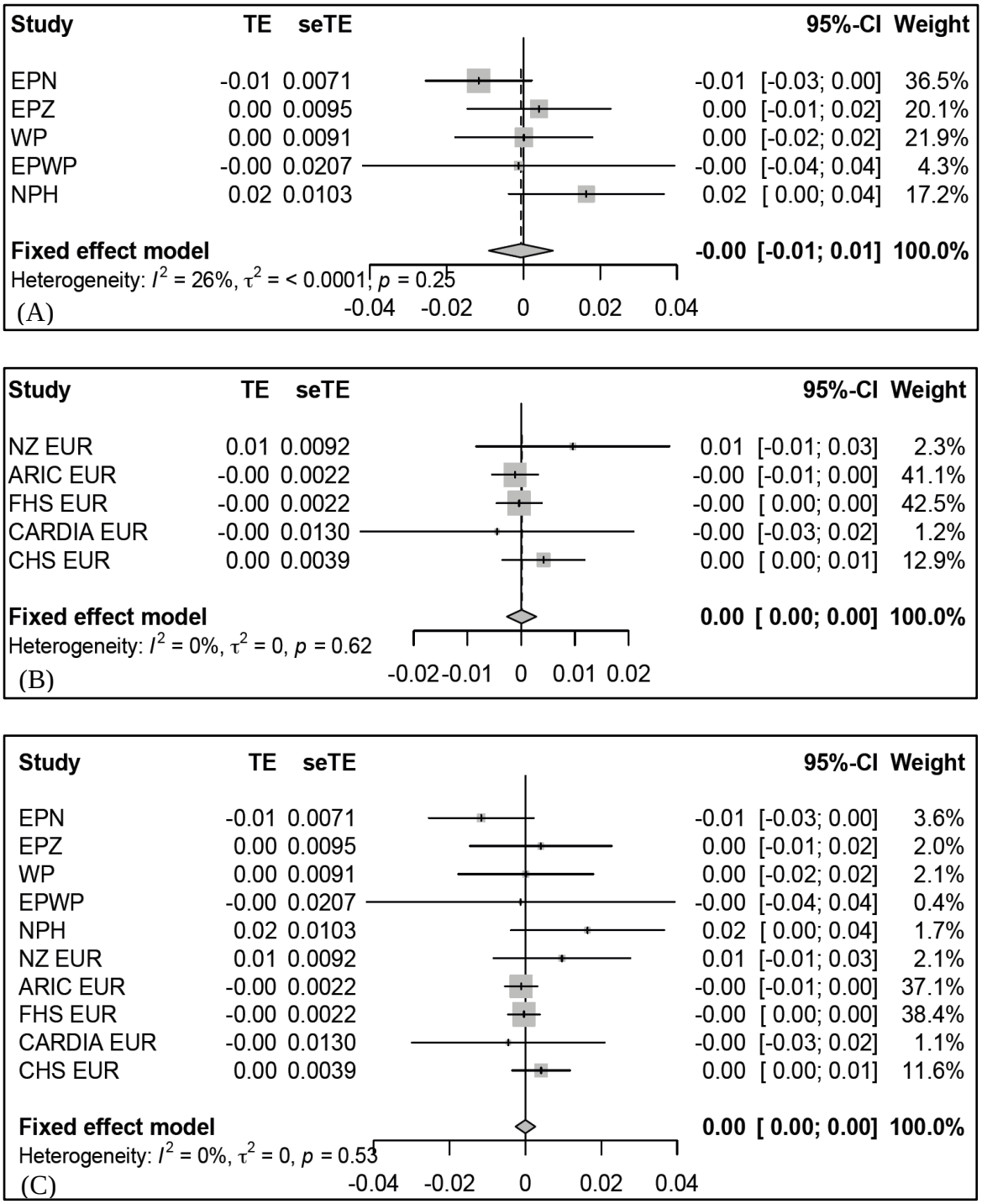
Appendix B Figure 4.2: Meta-analysis of (A) NZ Polynesian, (B) European and (C) All sample sets combined for association of *rs4994* (G allele) with serum urate. Values on x-axis represent effect size (ß) for serum urate in mmol L^{-1}.



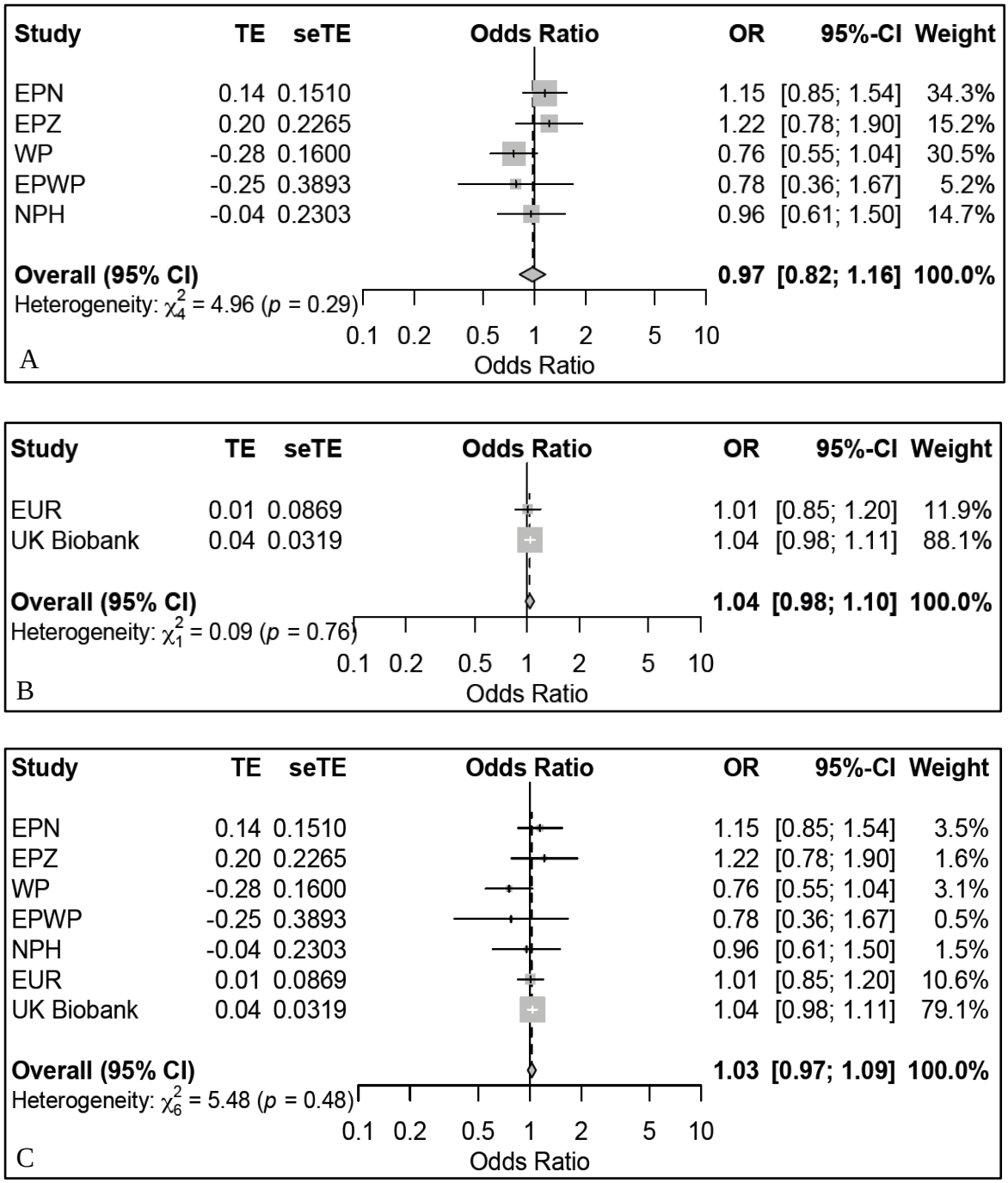




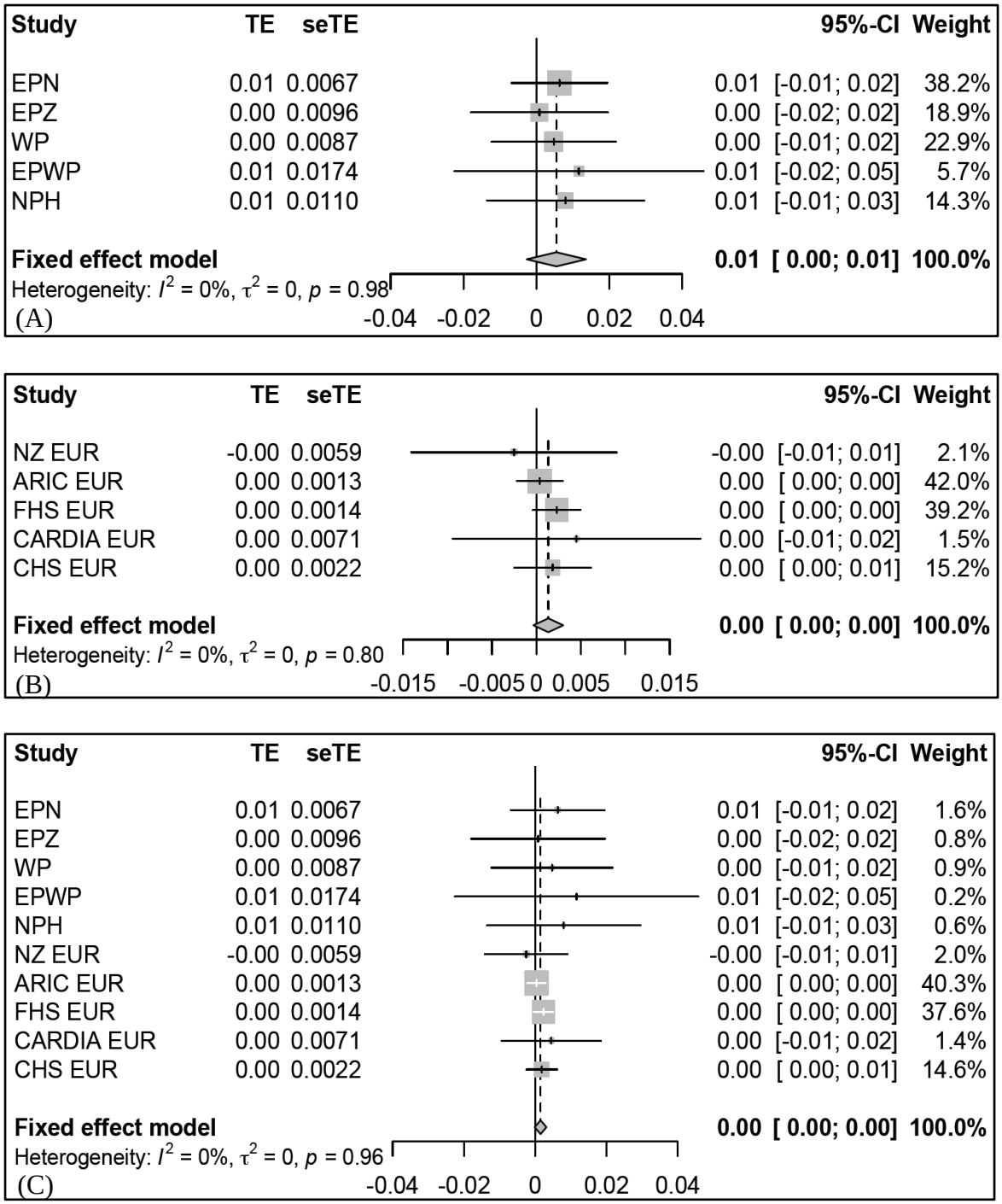
Appendix B Figure 4.3: Meta-analysis of (A) NZ Polynesian, (B) European and (C) All sample sets combined for association of *rs3827103* (A allele) with gout.



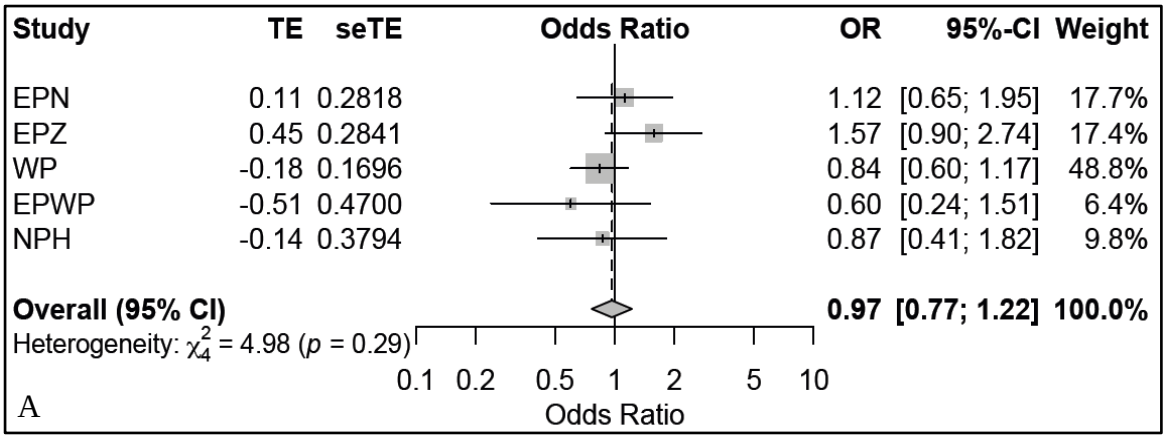
Appendix B Figure 4.4: Meta-analysis of (A) NZ Polynesian, (B) European and (C) All sample sets combined for association of *rs3827103* (A allele) with serum urate. Values on x-axis represent effect size (ß) for serum urate in mmol L^{-1}.

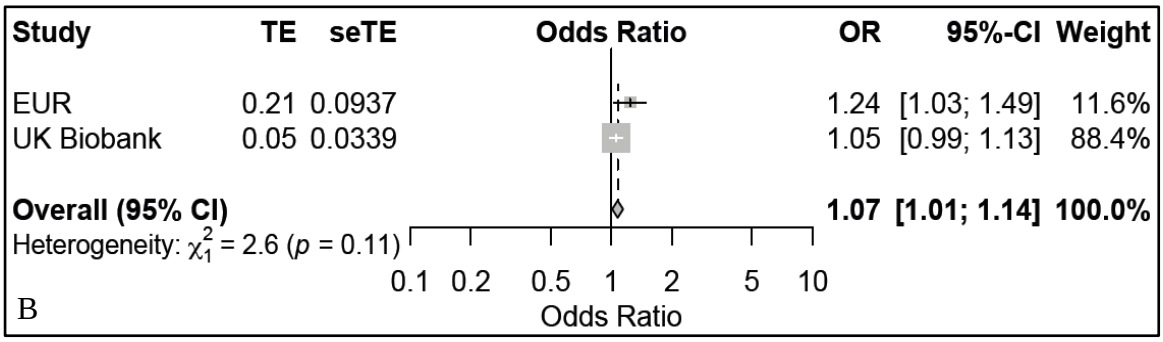


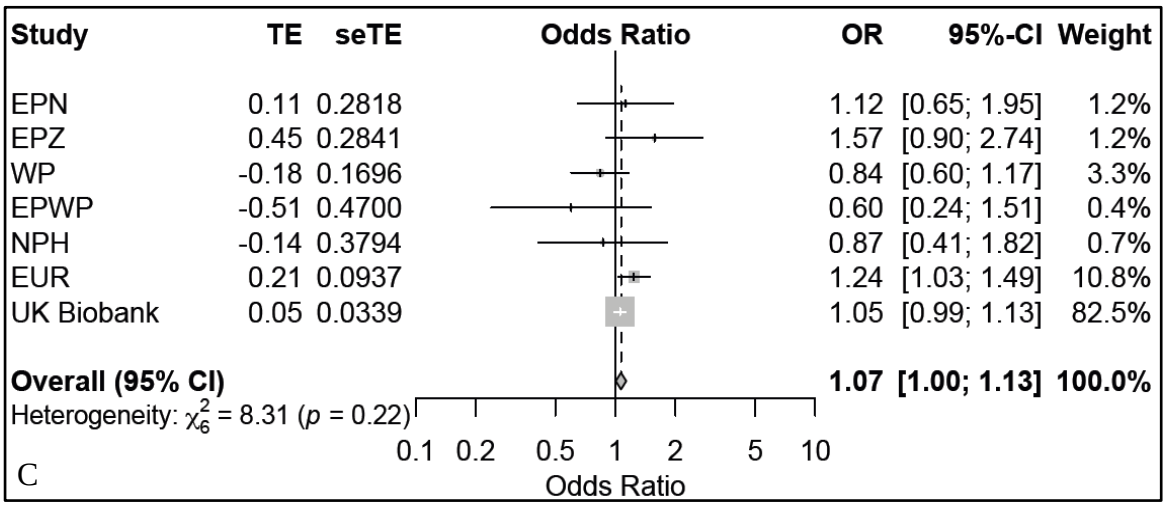
Appendix B Figure 4.5: Meta-analysis of (A) NZ Polynesian, (B) European and (C) All sample sets combined for association of *rs17700633* (A allele) with gout.



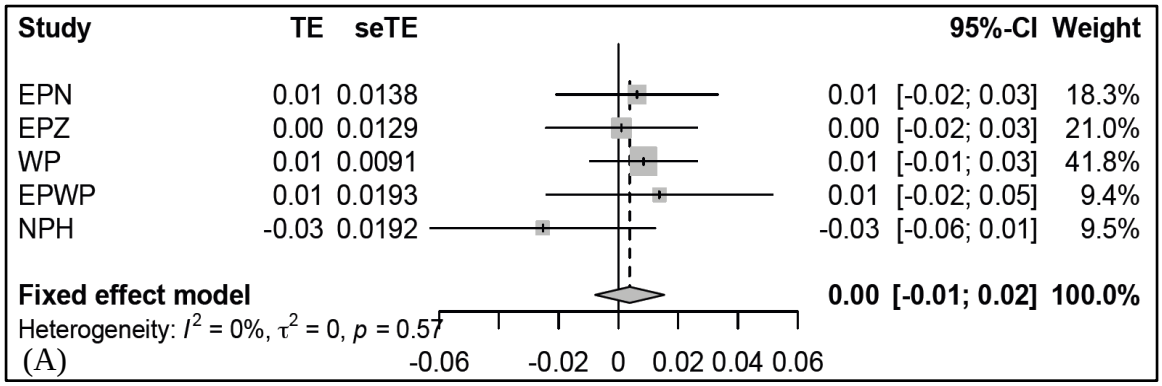
Appendix B Figure 4.6: Meta-analysis of (A) NZ Polynesian, (B) European and (C) All sample sets combined for association of *rs17700633* (A allele) with serum urate. Values on x-axis represent effect size (ß) for serum urate in mmol L^{-1}.

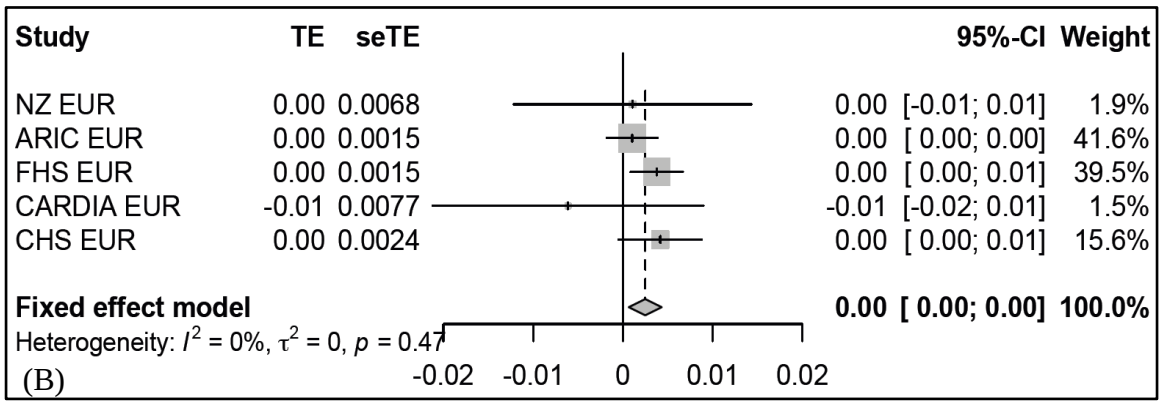


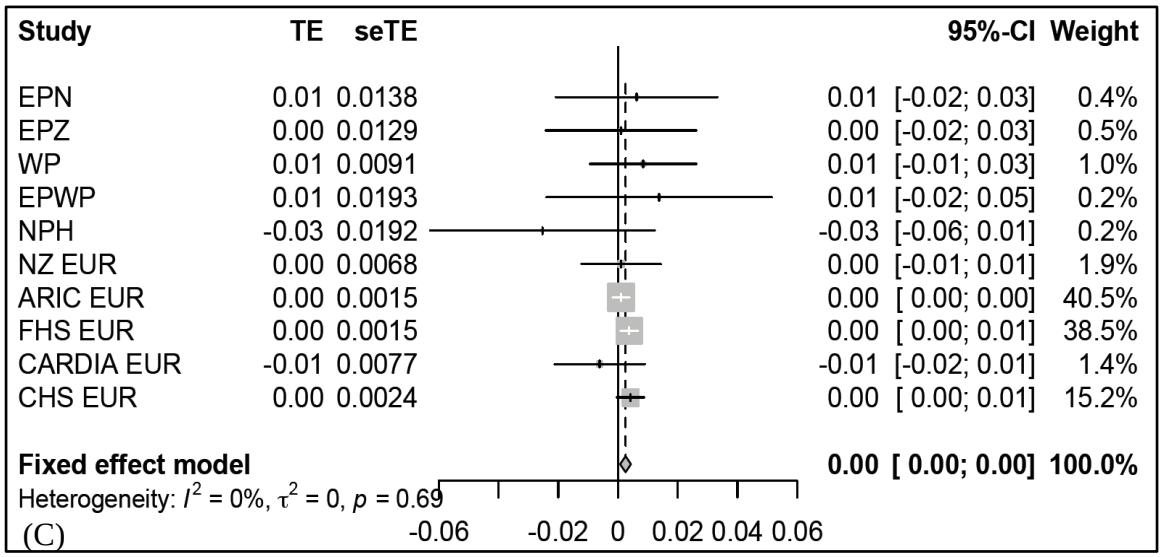




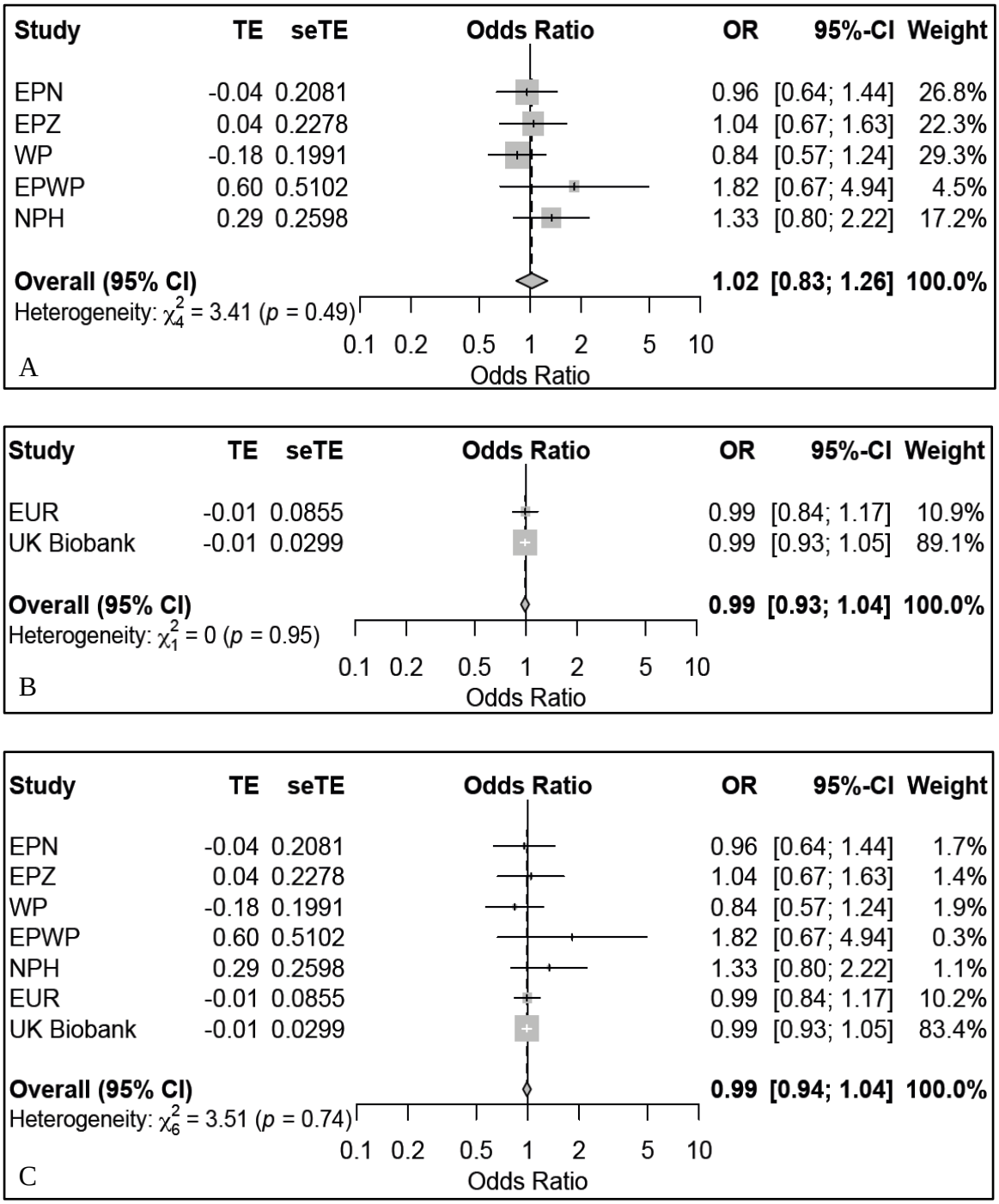
Appendix B Figure 4.7: Meta-analysis of (A) NZ Polynesian, (B) European and (C) All sample sets combined for association of *rs17782313* (C allele) with gout.



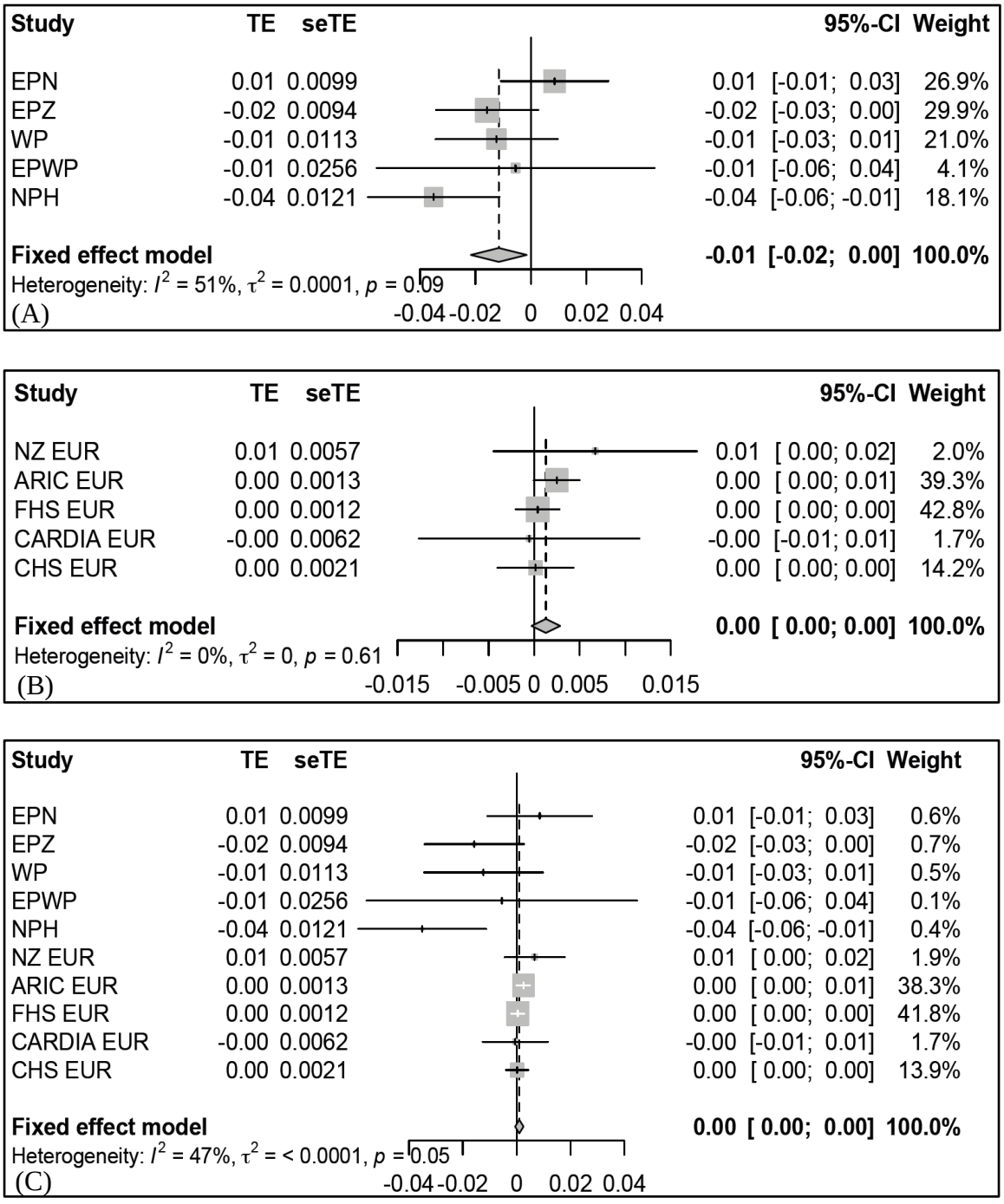




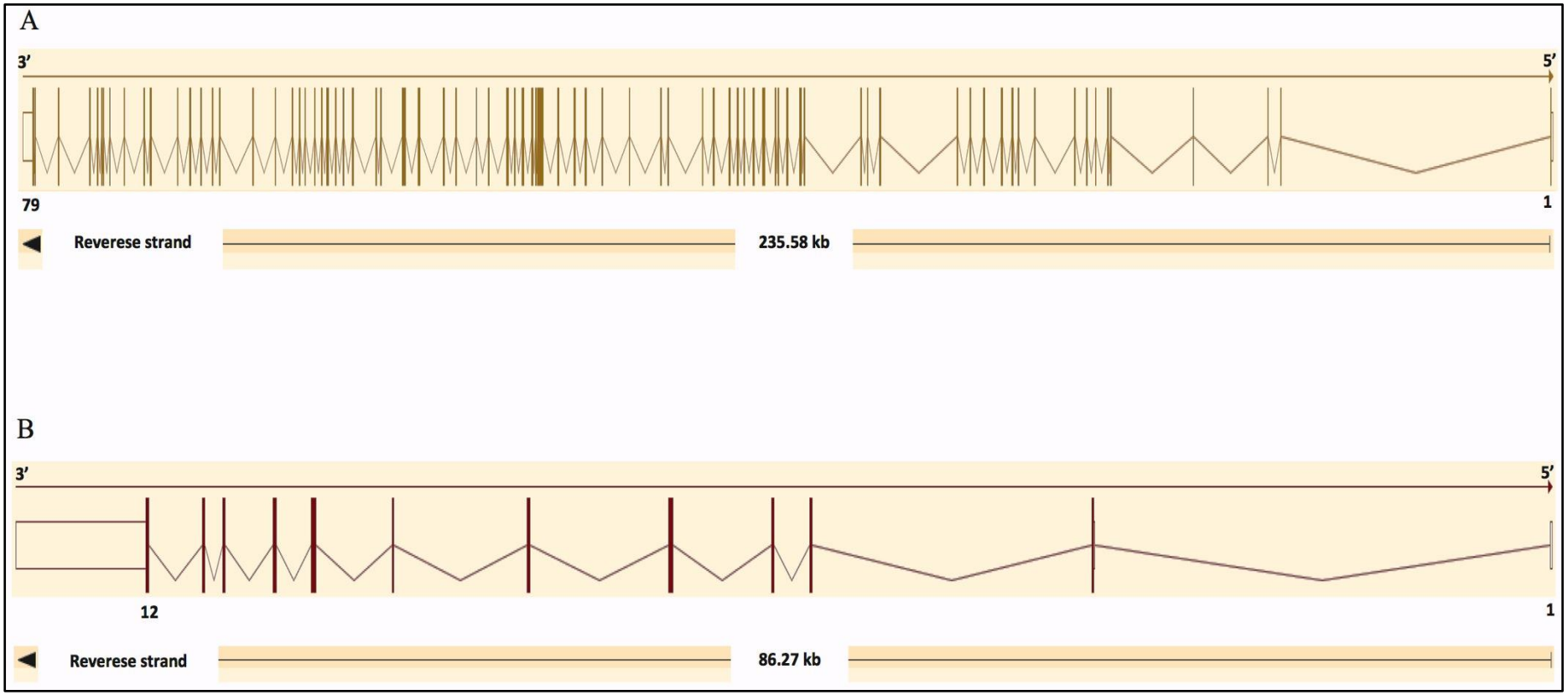
Appendix B Figure 4.8: Meta-analysis of (A) NZ Polynesian, (B) European and (C) All sample sets combined for association of *rs17782313* (C allele) with serum urate. Values on x-axis represent effect size (ß) for serum urate in mmol L^{-1}.



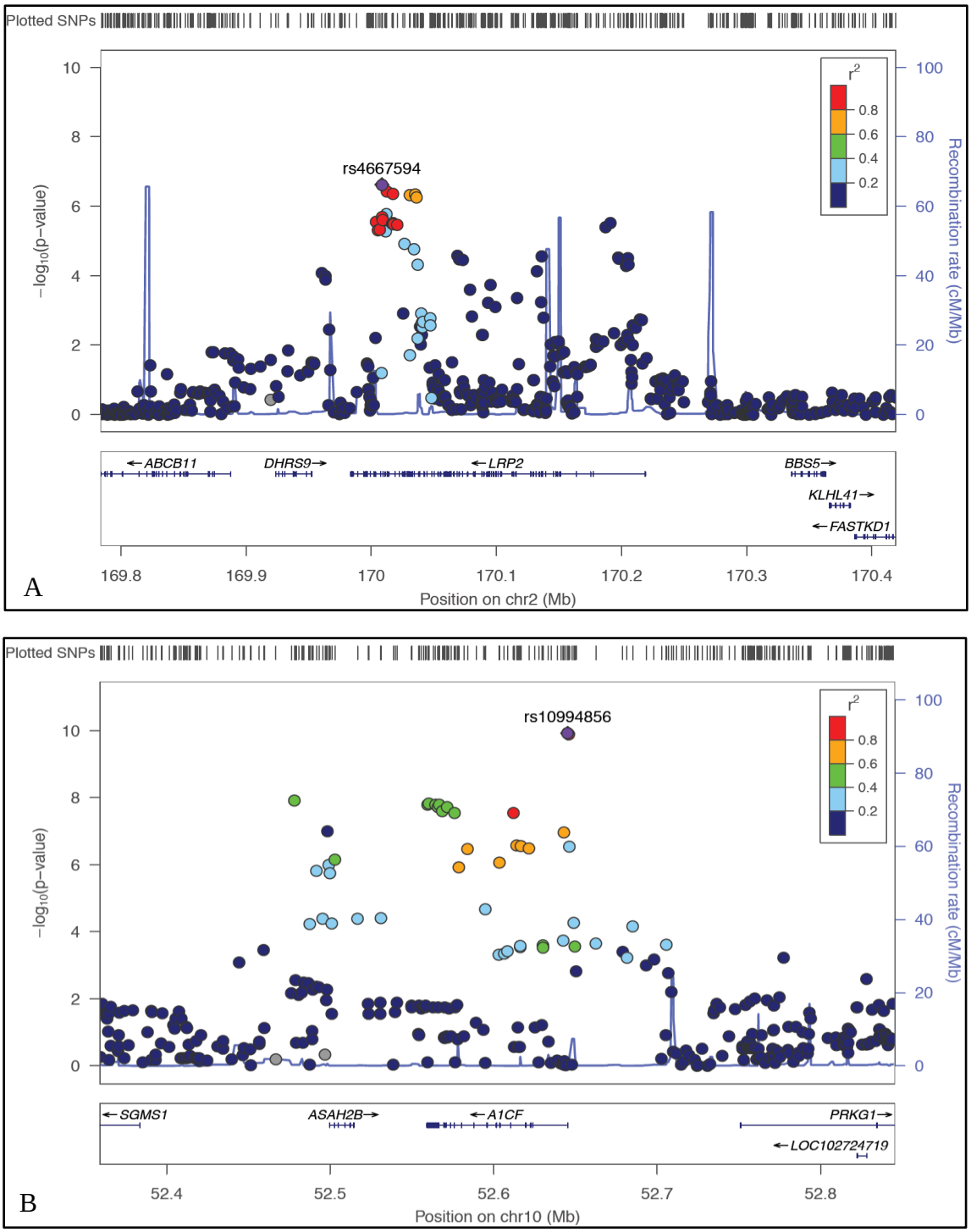
Appendix B Figure 4.9: Meta-analysis of (A) NZ Polynesian, (B) European and (C) All sample sets combined for association of *rs6903956* (A allele) with gout.



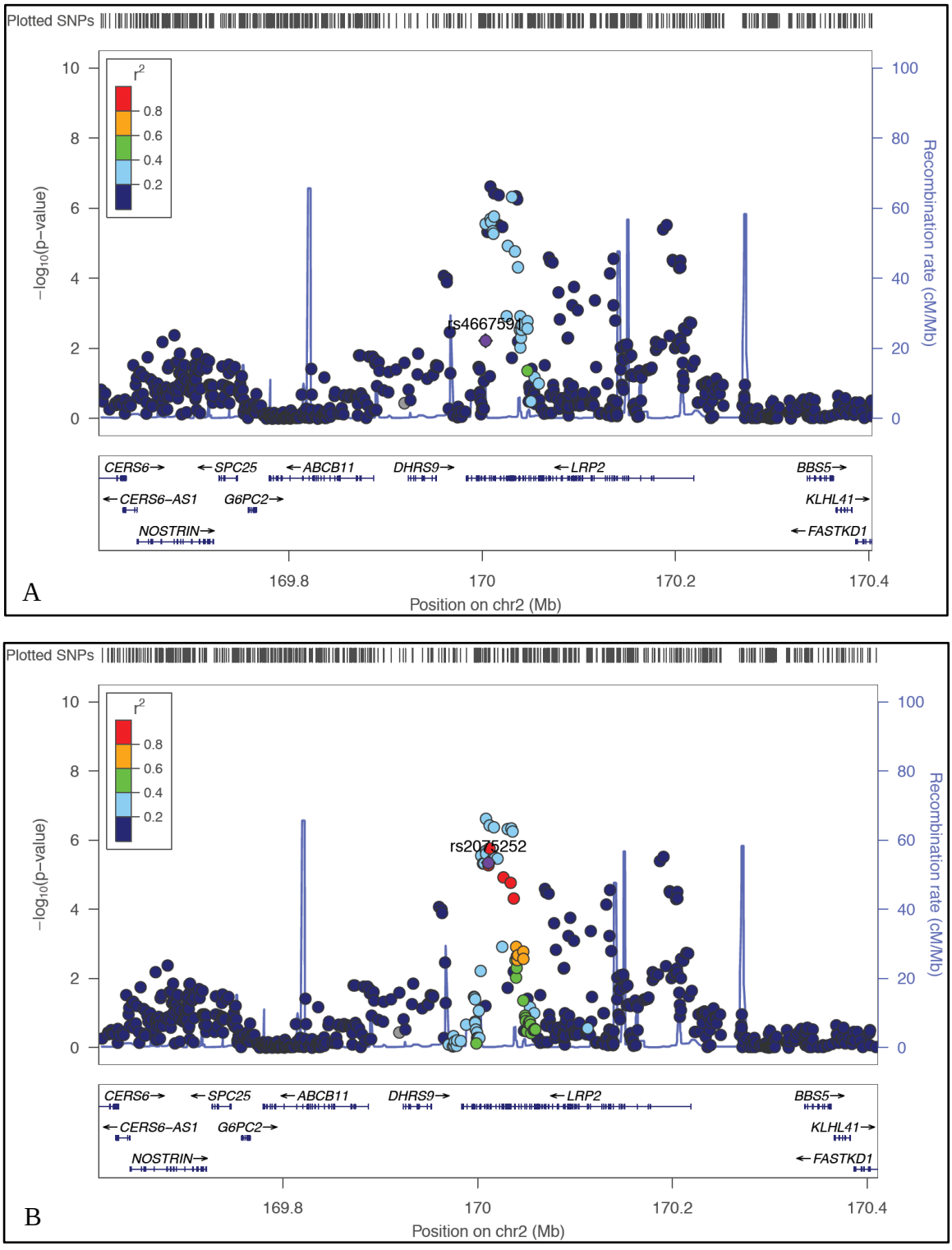
Appendix B Figure 4.10: Meta-analysis of (A) NZ Polynesian, (B) European and (C) All sample sets combined for association of *rs6903956* (A allele) with serum urate. Values on x-axis represent effect size (ß) for serum urate in mmol L^{-1}.



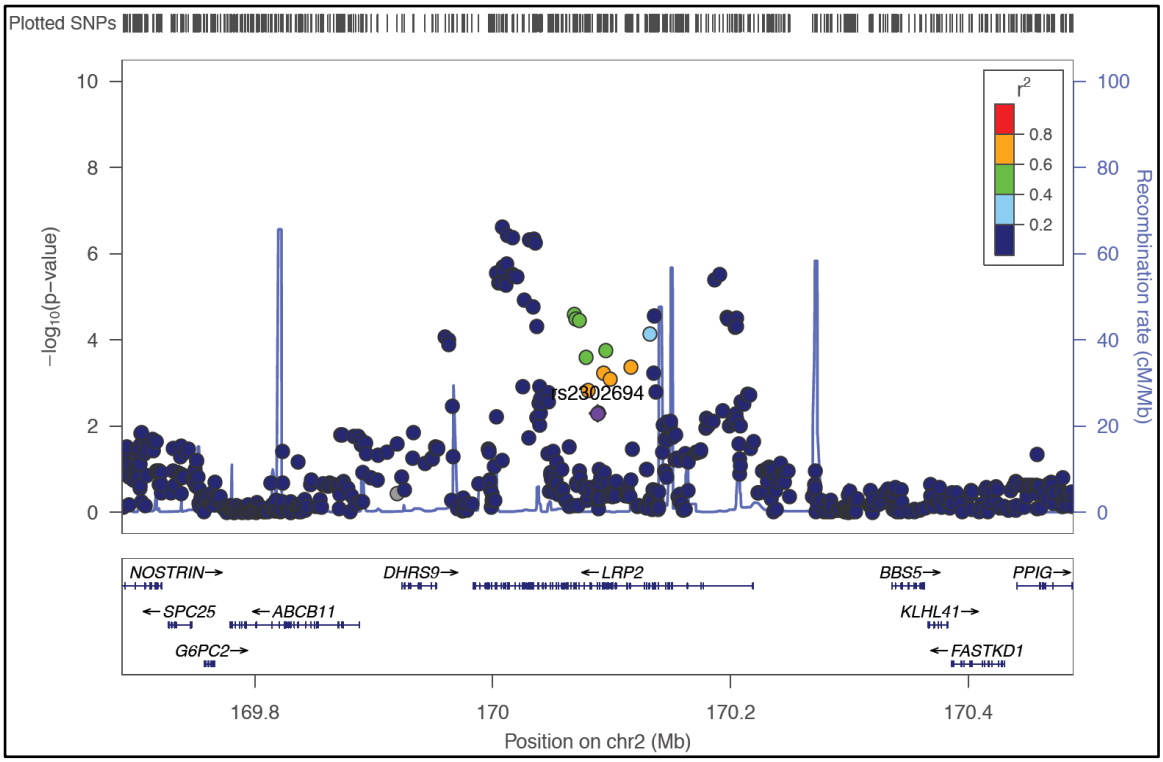
Appendix B Figure 5.1: An overview of coding (exonic) and non-coding (intronic) regions of (A) *LRP2* gene: total length – 235.58 kb, exons – 79, coding exons – 79, transcript length – 15,808 bps, introns – 78, and (B) *A1CF* gene: total length – 86.27 kb, exons – 12, coding exons – 11, transcript length – 9,221 bps, introns – 11. Sourced from Genome Reference consortium human genome build 37 (GRCh37) on Ensembl (http://grch37.ensembl.org/).



Appendix B Figure 5.2: Locus zoom plot indicating association signals for two lipid-related genes (A) *LRP2* and (B) *A1CF* with eGFR (mL/min/1.73m^2) in a recent GWAS (Pattaro *et al.*, 2016). Plots were created online (http://locuszoom.sph.umich.edu/).



Appendix B Figure 5.3: Locus zoom plot indicating association signals for two *LRP2* variants (A) *rs4667591* (*P* = 0.006) and (B) *rs2075252* (*P* = 0.004) with eGFR (mL/min/1.73m^2) in a recent GWAS (Pattaro *et al.*, 2016). Plots were created online (http://locuszoom.sph.umich.edu/).



Appendix B Figure 5.4: Locus zoom plot indicating association signals for two *LRP2* variant *rs2302694* with eGFR (mL/min/1.73m^2) in a recent GWAS (*P* = 0.005) (Pattaro *et al.*, 2016). Plots were created online (http://locuszoom.sph.umich.edu/).

APPENDIX C

Rheumatol Int (2016) 36:255–261
DOI 10.1007/s00296-015-3370-6

Rheumatology INTERNATIONAL

ORIGINAL ARTICLE - GENES AND DISEASE

Association analysis of the beta-3 adrenergic receptor Trp64Arg (*rs4994*) polymorphism with urate and gout

Tahzeeb Fatima[1] · Sara Altaf[1] · Amanda Phipps-Green[1] · Ruth Topless[1] · Tanya J. Flynn[1] · Lisa K. Stamp[2] · Nicola Dalbeth[3] · Tony R. Merriman[1]

Received: 18 July 2015 / Accepted: 18 September 2015 / Published online: 26 September 2015

Abstract The Arg64 allele of variant *rs4994* (Trp64Arg) in the β3-adrenergic receptor gene has been associated with increased serum urate and risk of gout. Our objective was to investigate the relationship of *rs4994* with serum urate and gout in New Zealand European, Māori and Pacific subjects. A total of 1730 clinically ascertained gout cases and 2145 controls were genotyped for *rs4994* by Taqman®. Māori and Pacific subjects were subdivided into Eastern Polynesian (EP) and Western Polynesian (WP) sample sets. Publicly available genotype data from the Atherosclerosis Risk in Communities Study and the Framingham Heart Study were utilized for serum urate association analysis. Multivariate logistic and linear regression adjusted for potential confounders was carried out using R version 2.15.2. No significant association of the minor Arg64 (G) allele of *rs4994* with gout was found in the combined Polynesian cohorts (OR = 0.98, $P = 0.88$), although there was evidence, after adjustment for renal disease, for association in both the WP (OR = 0.53, $P = 0.03$) and the lower Polynesian ancestry EP sample sets (OR = 1.86, $P = 0.05$). There was no evidence for association with gout in the European sample set (OR = 1.11, $P = 0.57$). However, the Arg64 allele was positively associated with urate in the WP data set ($\beta = 0.036$, $P = 0.004$, $P_{Corrected} = 0.032$). Association of the Arg64 variant with increased urate in the WP sample set was consistent with the previous literature, although the protective effect of this variant with gout in WP was inconsistent. This association provides an etiological link between metabolic syndrome components and urate homeostasis.

Keywords Gout · Urate · ADRB3 · Genetic · Association

Electronic supplementary material The online version of this article (doi:10.1007/s00296-015-3370-6) contains supplementary material, which is available to authorized users.

✉ Tony R. Merriman
tony.merriman@otago.ac.nz

[1] Department of Biochemistry, University of Otago, Dunedin, New Zealand

[2] Department of Medicine, University of Otago, Christchurch, New Zealand

[3] Department of Medicine, University of Auckland, Auckland, New Zealand

Introduction

Urate is the end product of endogenous and dietary purine metabolism in humans. The most important regulator of serum urate levels is reduced excretion of uric acid in the urine [1]. When supersaturation of urate is reached, monosodium urate (MSU) crystals are able to form within synovial fluid. Gout is the extremely painful innate immune response to these crystals. Individual variation in serum urate concentrations is partly explained by genetic factors, with 28 loci associated with serum urate in Europeans at a genome-wide level of significance [2]. The strongest genetic effects are with uric acid transporter genes, which are also associated with gout [2, 3].

Hyperuricemia and gout are closely related to metabolic diatheses like obesity, dyslipidemia, glucose intolerance and hypertension [4, 5]. Renal clearance of uric acid is inversely related to insulin resistance [6], and the evidence that gout is associated with the metabolic syndrome has led to the hypothesis that gout and hyperuricemia may also have a causal relationship with insulin resistance and obesity. This is consistent with Mendelian randomization studies that have demonstrated increased triglyceride levels and

body mass index (BMI) as causal of increased urate [7, 8]. Therefore, genes influencing insulin resistance could also contribute to the development of hyperuricemia.

The beta-3 adrenergic receptor (ADRB3) is part of the adrenergic system. It is expressed mainly in adipose tissue in humans and is involved in the regulation of lipid metabolism and glucose homeostasis [9]. The Trp64Arg (*rs4994*) polymorphism in the first transmembrane domain of ADRB3 has been associated with BMI in meta-analysis of 44,833 individuals with Arg64 carriers having increased BMI and a stronger effect in East Asian sample sets [10]. This is consistent with reports of association of type 2 diabetes mellitus and insulin resistance with Arg64 [11, 12] and association of the Arg64 allele with increased adiposity measures, serum urate and blood pressure in older men [13].

The Trp64Arg polymorphism has previously been tested for association with hyperuricemia and gout under the hypothesis that genetic variants contributing to insulin resistance could also contribute to hyperuricemia. A combination of increased BMI and Arg64 increased the risk of developing hyperuricemia by fourfold in a postprandial diabetic group drawn from the Chinese population [14]. Rho et al. [15] and Huang et al. [16] reported association of the Arg64 allele with hyperuricemia in Korean and Chinese sample sets. A similar association was reported by Morcillo et al. [17] who reported that the Arg64 allele predicts the risk of developing hyperuricemia in a prospective study in a population from southern Spain. Wang et al. [18] reported the Arg64 allele was associated with gout in a male Chinese population.

The aim of the study reported here was to further investigate the relationship of the Trp64Arg (*rs4994*) variant of *ADRB3* with hyperuricemia and gout in European and New Zealand Māori and Pacific Island (Polynesian) sample sets. The New Zealand Polynesian population exhibits the highest rate of gout in the world, with considerable comorbidity with type 2 diabetes and cardiovascular disease [19, 20].

Materials and methods

Study participants

All New Zealand (NZ) gout cases and controls included in this study were recruited during the years 2006–2013. Gout cases fulfilled the American Rheumatology Association criteria for gout by clinical examination [21], while controls self-reported no history of gouty arthritis. Except for the biochemical measurements and BMI, all other variables were self-reported. Written informed consent was obtained from all subjects for collection of samples and subsequent analyses. Publicly available genotype and phenotype data from the Framingham Heart Study (FHS) and Atherosclerosis Risk in Communities (ARIC) cohorts were accessed from the ARIC and FHS studies under the project name "The Genetic Basis of Gout" and approval number 834. Table S1 reports the demographic and clinical data for all study groups.

The NZ gout case–control sample set was divided into four ancestral groups [22]: NZ European (648 cases and 877 controls), Eastern Polynesian (EP; Cook Islands and NZ Māori; 491 cases and 696 controls), Western Polynesian (WP; Samoa, Tonga, Tuvalu, Niue and Tokelau; 367 cases and 310 controls) and mixed Eastern and Western Polynesian (EP/WP; 29 cases and 70 controls). Eastern Polynesian participants were further subdivided into EPN (subjects with high EP ancestry; 334 cases and 392 controls) and EPZ (subjects with low EP ancestry; 157 cases and 311 controls) [22]. A separate Māori sample set (NPH; 195 cases and 192 controls) was also included in the study, ascertained with criteria described above. These participants were recruited in collaboration with Ngāti Porou Hauora Charitable Trust (NPHCT) from the Ngāti Porou *rohe* (tribal territory) located in the East Coast (Tairawhiti) region of the North Island of New Zealand. The procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation and with the Helsinki Declaration of 1975, as revised in 1983. The NPHCT study was approved by the Ngāti Porou Hauora Board. The Lower South Ethics Committee (OTA/99/11/098) and New Zealand Multi-region Ethics Committee (MEC/05/10/130) granted ethical approval for this study. The Northern Y Region Health Research Ethics Committee granted ethical approval for the NPHCT study (NTY07/07/074).

Control (non-gout) subjects were recruited from NZ and sourced from ARIC (http://www2.cscc.unc.edu/aric/) and FHS (Offspring and Generation 3) (http://www.framinghamheartstudy.org/) cohorts for the purpose of evaluating the association of *rs4994* with serum urate. The ARIC data set consisted of 4144 subjects and the FHS of 5109 subjects. Subjects who self-reported as taking diuretic medication, or had renal failure, gout, or had first-degree relatives with gout or were not of European ancestry were excluded.

Data collection and genotyping

The uricase oxidation method was applied to measure serum urate levels for NZ subjects using a Roche chemistry modular P/D analyzer. Data for serum urate were obtained from visit 1 (1987–1989) for ARIC and examination 1 (Offspring: 1971–1975 and Generation 3: 2002–2005) for FHS cohorts. Genotyping of NZ samples for *rs4994* was done using a TaqMan® assay (C_2215549_20; Applied Biosystems, Foster City, USA) using a Lightcycler® 480

Table 1 Association analysis of *rs4994* with gout

	AA	AG	GG	G	OR (95 % CI) (unadjusted)	$P_{unadjusted}$	OR (95 % CI) (adjusted)[a]	$P^{a}_{adjusted}$	HWE
EPN									
Cases	266 (0.796)	66 (0.197)	2 (0.006)	70 (0.105)	0.89 (0.64–1.25)	0.53	1.08 (0.69–1.68)	0.72	0.33
Control	306 (0.780)	82 (0.209)	4 (0.011)	90 (0.114)					0.56
EPZ									
Cases	122 (0.777)	33 (0.210)	2 (0.012)	37 (0.117)	1.49 (0.95–2.35)	0.07	1.66 (0.91–3.08)	0.09	0.89
Control	262 (0.842)	47 (0.151)	2 (0.006)	51 (0.081)					0.94
WP									
Cases	336 (0.915)	31 (0.084)	0 (0.000)	31 (0.042)	0.62 (0.38–0.99)	0.04	0.61 (0.34–1.06)	0.08	0.39
Control	272 (0.877)	34 (0.109)	4 (0.012)	42 (0.067)					0.02
EP/WP									
Cases	27 (0.931)	2 (0.069)	0 (0.000)	2 (0.034)	0.31 (0.04–1.21)	0.13	0.31 (0.04–1.44)	0.17	0.84
Control	57 (0.814)	13 (0.185)	0 (0.000)	13 (0.092)					0.39
NPH									
Cases	156 (0.800)	39 (0.200)	0 (0.000)	39 (0.100)	1.07 (0.66–1.73)	0.77	0.95 (0.47–1.90)	0.89	0.11
Control	160 (0.833)	28 (0.144)	4 (0.020)	36 (0.093)					0.04
NZ European									
Cases	543 (0.838)	102 (0.157)	3 (0.004)	108 (0.083)	1.00 (0.76–1.30)	0.99	1.11 (0.77–1.61)	0.57	0.44
Control	735 (0.838)	138 (0.157)	4 (0.004)	146 (0.083)					0.35
Han Chinese [18]									
Cases	298 (0.793)	104 (0.252)	10 (0.024)	124 (0.150)	1.50 (1.09-2.06)	0.01	1.95 (1.22–3.13)	0.02	0.79
Control	248 (0.795)	62 (0.199)	2 (0.006)	66 (0.106)					0.37

[a] All values are adjusted against sex, age and BMI and self-reported grandparental ancestry for Polynesian data sets

Real-Time Polymerase Chain Reaction (RT-PCR) System (Roche Applied Science, Indianapolis, IN, USA). For the FHS and ARIC cohorts, *rs4994* genotype was imputed using all 1000 Genomes haplotype data (phase 1; 2012) as a reference panel using IMPUTE2 [23].

Statistical analysis

Logistic and linear regression analysis was done using statistical software R (v 2.15.2) [24] to test for an association of *rs4994* (explanatory variable) with gout (binary response variable) and serum urate (continuous response variable), respectively. Any individual with missing data for any variable was excluded from the various analyses. The adjusted odds ratio (OR) and β-coefficients were obtained by including age, sex and BMI as covariates in the regression model. In additional analyses, the presence or absence of type 2 diabetes, hypertension and renal disease was included as an adjustor. Self-reported grandparental ancestry was also included as a covariate in Polynesian data sets. Adherence to Hardy–Weinberg equilibrium (HWE) was calculated using the SHEsis package (http://analysis2.bio-x.cn/myAnalysis.php) with a significant deviation from HWE if $P < 0.004$ (0.05 divided by 12—the number of data sets tested in Table 1). Meta-analysis was done using the meta package (http://CRAN.R-project.org/package=meta, 2014) within R using a fixed-effect model. For analyses showing heterogeneity ($P_{Het} < 0.05$), the fixed-effect model was replaced with a random-effect model. A threshold of $P \leq 0.05$ was used to indicate nominal statistical significance between response and explanatory variables. $P_{Corrected}$ was calculated by dividing the P by the number of tests performed (eight), and a $P_{Corrected} < 0.05$ indicated significance.

Results

There was nominally significant association between the *ADRB3* variant *rs4994* and gout only in the WP sample set where the minor Arg64 (G) allele was associated with reduced risk of gout (OR = 0.62, $P = 0.04$) although this association became nonsignificant when adjusted for confounding variables (OR = 0.61, $P = 0.08$) (Table 1). A strong trend of association of the Arg64 allele toward increased risk of gout was observed in EPZ (OR = 1.49, $P = 0.07$) (Table 1). To increase the power of the analysis, all NZ case–control groups were combined with the

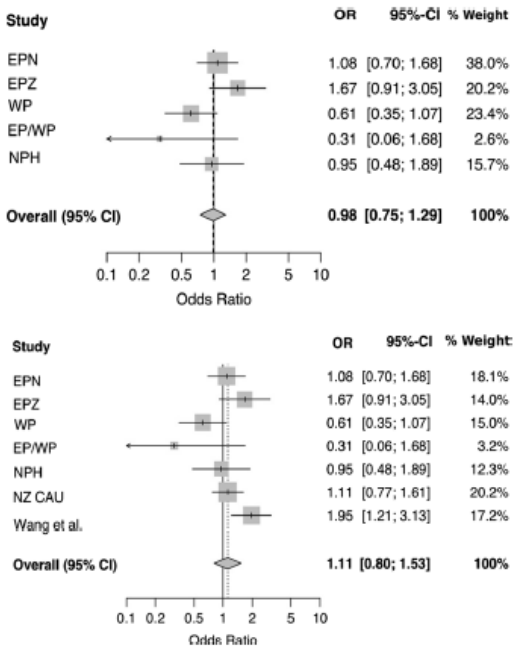


Fig. 1 Meta-analysis of NZ Polynesian sample sets for association of *rs4994* with gout (*top*; $P_{OR} = 0.88$ and $P_{Het.} = 0.10$) and meta-analysis of all available sample sets for association of *rs4994* with gout (*bottom*; $P_{OR} = 0.52$ and $P_{Het.} = 0.03$)

Wang et al. [18] data set by meta-analysis. No evidence of association was found (Fig. 1; OR = 1.11, $P = 0.52$), although P_{Het} was 0.03 which suggested some heterogeneity. Considering ancestry as a possible source of heterogeneity, Polynesian data sets only were combined in meta-analysis. Again the association was not significant (Fig. 1; OR = 0.98, $P = 0.88$, $P_{Het} = 0.10$). Significant protective association of the *rs4994* Arg64 allele was still observed in the WP sample set when adjusted for diabetes (OR = 0.57, $P = 0.05$), hypertension (OR = 0.56, $P = 0.05$) and renal dysfunction (OR = 0.53, $P = 0.03$), whereas there was evidence for a risk effect of Arg64 on gout in EPZ after adjustment for renal dysfunction (OR = 1.86, $P = 0.05$) (Table 2).

Linear regression analysis was performed to test for association of *rs4994* with urate in the NZ and ARIC/FHS data sets. A significant association of the Arg64 allele with urate was found in the WP sample set ($\beta = 0.036$, $P = 0.004$), which retained significance after correction for the eight sample sets examined ($P_{Corrected} = 0.03$). This indicates that each copy of the Arg64 allele increases serum urate by 0.036 mmol/L (Table 3). To increase the power of the analysis, the various Polynesian and European (including ARIC and FHS) data sets were combined separately by meta-analysis. No significant association was found in either analysis (Arg64 allele: $\beta = 0.01$, $P = 0.16$ and $\beta = 2.99 \times 10^{-5}$, $P = 0.98$, respectively) (Fig. 2).

Discussion

Our study reports association of the Arg64 allele of the *rs4994* polymorphism of *ADRB3* with increased serum urate in WP individuals ($\beta = 0.04$, $P = 0.004$, $P_{Corrected} = 0.03$). This is consistent with the previously reported findings of Morcillo et al. [17] who demonstrated that the Arg64 allele predicts the development of hyperuricemia in the population of southern Spain and with studies in Chinese and Korean sample sets that also associated the Arg64 allele with increased serum urate and risk of gout in Asian subjects [14–16, 18]. Collectively our and previous studies increase the support for a causal role of *ADRB3* and the adrenergic system in urate control. We could not meta-analyze our serum urate findings with previously published studies as they were described as a secondary finding in conjunction with other metabolic conditions, or from a population subgroup, or as a binary outcome [13–15]. It is important to note that the genome-wide association study

Table 2 Association analysis of *rs4994* with gout adjusted for comorbidities

	Baseline adjustment[a]		Diabetes (T2D)[a]		Hypertension[a]		Renal dysfunction[a]	
	OR (95 % CI)	*P*	OR (95 % CI)	*P*	OR (95 % CI)	*P*	OR (95 % CI)	*P*
EPN	1.08 (0.69–1.68)	0.72	1.03 (0.66–1.62)	0.86	1.08 (0.69–1.70)	0.72	0.93 (0.58–1.49)	0.77
EPZ	1.66 (0.91–3.08)	0.09	1.78 (0.96–3.34)	0.06	1.56 (0.80–3.05)	0.18	1.86 (0.98–3.60)	0.05
WP	0.61 (0.34–1.06)	0.08	0.57 (0.32–1.01)	0.05	0.56 (0.31–1.01)	0.05	0.53 (0.29–0.95)	0.03
EP/WP	0.31 (0.04–1.44)	0.17	0.31 (0.04–1.45)	0.17	0.35 (0.04–1.76)	0.24	0.41 (0.05–2.02)	0.31
NPH	0.95 (0.47–1.90)	0.89	0.93 (0.46–1.89)	0.85	0.92 (0.45–1.90)	0.83	0.98 (0.48–1.98)	0.96
NZ European	1.11 (0.77–1.61)	0.57	1.12 (0.77–1.63)	0.55	1.12 (0.76–1.65)	0.56	1.11 (0.76–1.63)	0.58

[a] All values are adjusted against sex, age and BMI and self-reported grandparental ancestry for Polynesian data sets

Table 3 Association analysis of *rs4994* with serum urate (mmol L^{-1})

	Unadjusted		Adjusted[a]	
	β-Coef. (95 % CI)	*P*	β-Coef. (95 % CI)	*P*
EPN	0.004 (−0.016–0.026)	0.66	0.003 (−0.015–0.023)	0.70
EPZ	0.017 (−0.018–0.054)	0.33	0.011 (−0.018–0.040)	0.45
WP	0.036 (0.009–0.063)	0.01	0.036 (0.011–0.062)	0.01
EP/WP	−0.021 (−0.088–0.044)	0.51	−0.018 (−0.076–0.039)	0.53
NPH	−0.006 (−0.041–0.026)	0.68	−0.021 (−0.051–0.008)	0.16
NZ CAU	−0.025 (−0.054–0.004)	0.09	−0.012 (−0.037–0.012)	0.29
ARIC CAU	0.001 (−0.003–0.006)	0.56	0.001 (−0.002–0.005)	0.49
FHS CAU	0.001 (−0.005–0.006)	0.90	−0.001 (−0.005–0.003)	0.61

[a] All values are adjusted against sex, age and BMI and self-reported grandparental ancestry for Polynesian data sets

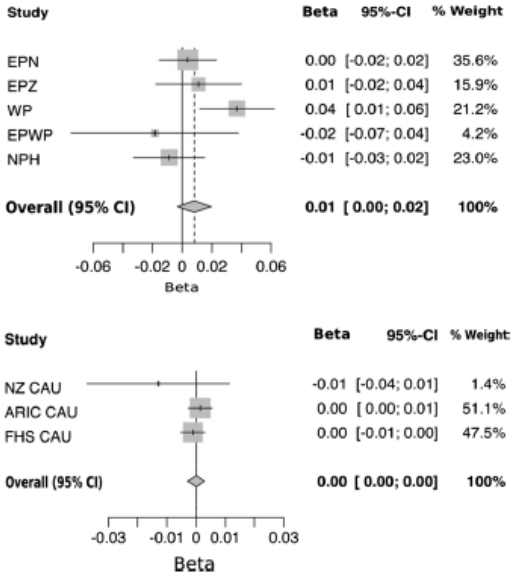


Fig. 2 Meta-analysis of NZ Polynesian sample sets for association of *rs4994* with serum urate (*top*; $P_{Beta} = 0.16$ and $P_{Het.} = 0.08$) and meta-analysis of European sample sets for association of *rs4994* with serum urate (*bottom*; $P_{Beta} = 0.98$ and $P_{Het.} = 0.41$)

of Köttgen et al. [2] using ~140,000 individuals would not have been able to test variants in the ADRB3 region for association with urate because of the absence of any HapMap2 data including *rs4994* that could be used for imputation (http://hapmap.ncbi.nlm.nih.gov). Using more recently available data, it should be possible to impute this region in the Kottgen et al. data using 1000 Genomes haplotype data in which the *ADRB3* region is adequately covered by common variants and to test for association with urate.

The direction of association of the Arg64 allele with urate opposed that observed in gout in WP, where it was associated with reduced risk of gout. This direction of association also conflicted with the Arg64-mediated increased risk of gout reported by Wang et al. [18] and observed by us in the EPZ sample set (Table 2). Adjusting for the effect of comorbidities type 2 diabetes, hypertension and renal dysfunction did not influence the protective effect of Arg64 with gout in WP (Table 2). Acknowledging that this could be a false-positive finding, the opposing direction of association of Arg64 with gout and hyperuricemia could represent a pleiotropic effect of this allele in the WP population, perhaps having a role both in determining hyperuricemia and in the inflammatory processes leading to gout. The opposing direction to EPZ may reflect different ancestral haplotypes—we have previously observed differential effects at *ABCG2* between Eastern and Western Polynesian sample sets [25]. Most importantly, however, the association of *rs4994* with gout needs to be studied in larger sample sets of diverse ancestries.

The *ADRB3* gene is expressed mainly in adipose tissue and encodes for the beta-3 adrenergic receptor. The activation of this receptor induces lipolysis in adipose tissue and thermogenesis in skeletal muscles. It is also responsible for the delivery of free fatty acids into the portal vein. These free fatty acids and other products of lipolysis disrupt the insulin receptor signaling pathway, thereby leading to insulin resistance [26]. The Arg64 allele of *rs4994* has been reported to be associated with the development of obesity, increased BMI and insulin resistance [10, 27, 28]. Decreased activity of the Arg64 variant receptor could lead to a decline in lipolysis and increased deposition of adipose tissue. Furthermore, ADRB3 is proposed to be a part of the "leptin-sympathetic-leptin feedback loop," whereby decreased activity of this receptor causes an increase in leptin secretion from the adipose tissue [29]. The observation of elevated leptin levels in hyperuricemic patients [30] is

consistent with the association of Arg64 with hyperuricemia. Insulin resistance is a possible link between hyperuricemia and obesity [31, 32], which supports the hypothesis that the Arg64 allele could promote hyperuricemia in the obese.

Acknowledgments This work was supported by the Health Research Council of New Zealand, Arthritis New Zealand, New Zealand Lottery Health and the University of Otago. The authors would like to thank Jill Drake (Canterbury District Health Board), Roddi Laurence, Chris Franklin, Meaghen House (all University of Auckland) and Gabrielle Sexton (University of Otago) for recruitment. We also thank Ria Akuhata and Nancy Aupouri of NPHCT for recruitment, coordinated by Dr. Jennie Harré Hindmarsh. The Atherosclerosis Risk in Communities and Framingham Heart study analyses (Project #834) were approved by the relevant Database of Genotype and Phenotype (dbGaP; www.ncbi.nim.nih/gov/dbgap) Data Access Committees. The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, N01-HC-55022, R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant No. UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research. The Framingham Heart Study and the Framingham SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with Boston University. The Framingham SHARe data used for the analyses described in this manuscript were obtained through dbGaP. This manuscript was not prepared in collaboration with investigators of the Framingham Heart Study and does not necessarily reflect the opinions or views of the Framingham Heart Study, Boston University, or the NHLBI.

Funding This study was funded by Grant 11/1075 from the Health Research Council of New Zealand.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Dalbeth N, Merriman T (2009) Crystal ball gazing: new therapeutic targets for hyperuricaemia and gout. Rheumatology 48:222–226
2. Kottgen A, Albrecht E, Teumer A, Vitart V, Krumsiek J, Hundertmark C et al (2013) Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. Nat Genet 45:145–154
3. Phipps-Green A, Merriman M, Topless R, Altaf S, Montgomery G, Franklin C et al (2014) Twenty-eight loci that influence serum urate levels: analysis of association with gout. Ann Rheum Dis. doi:10.1136/annrheumdis-2014-205877
4. Choi HK, Ford ES (2007) Prevalence of the metabolic syndrome in individuals with hyperuricemia. Am J Med 120:442–447
5. Choi HK, Ford ES, Li C, Curhan G (2007) Prevalence of the metabolic syndrome in patients with gout: the Third National Health and Nutrition Examination Survey. Arthritis Care Res 57:109–115
6. Facchini F, Chen Y-DI, Hollenbeck CB, Reaven GM (1991) Relationship between resistance to insulin-mediated glucose uptake, urinary uric acid clearance, and plasma uric acid concentration. JAMA 266:3008–3011
7. Lyngdoh T, Vuistiner P, Marques-Vidal P, Rousson V, Waeber G, Vollenweider P et al (2012) Serum uric acid and adiposity: deciphering causality using a bidirectional Mendelian randomization approach. PLoS One 7:e39321
8. Rasheed H, Hughes K, Flynn TJ, Merriman TR (2014) Mendelian randomisation provides no evidence for a causal role of serum urate in increasing serum triglyceride levels. Circ Cardiovasc Genet 7:830–837
9. Krief S, Lönnqvist F, Raimbault S, Baude B, Van Spronsen A, Arner P et al (1993) Tissue distribution of beta 3-adrenergic receptor mRNA in man. J Clin Invest 91:344–349
10. Kurokawa N, Young EH, Oka Y, Satoh H, Wareham NJ, Sandhu M et al (2008) The ADRB3 Trp64Arg variant and BMI: a meta-analysis of 44,833 individuals. Int J Obes 32:1240–1249
11. Gjesing A, Andersen G, Borch-Johnsen K, Jørgensen T, Hansen T, Pedersen O (2008) Association of the β3-adrenergic receptor Trp64Arg polymorphism with common metabolic traits: studies of 7605 middle-aged white people. Mol Genet Metab 94:90–97
12. Widen E, Lehto M, Kanninen T, Walston J, Shuldiner AR, Groop LC (1995) Association of a polymorphism in the β3-adrenergic–receptor gene with features of the insulin resistance syndrome in Finns. N Engl J Med 333:348–352
13. Strazzullo P, Iacone R, Siani A, Cappuccio FP, Russo O, Barba G et al (2001) Relationship of the Trp64Arg polymorphism of the beta3-adrenoceptor gene to central adiposity and high blood pressure: interaction with age. Cross-sectional and longitudinal findings of the Olivetti Prospective Heart Study. J Hypertens 19:399–406
14. Wang G, Li Q, Niu T, Chen C, Xu X (2002) Association of GYS1 and beta (3)-AR gene with postprandial hyperglycemia and serum uric acid in type 2 diabetes mellitus. Chin Med J 115:1308–1311
15. Rho YH, Choi SJ, Lee YH, Ji JD, Song GG (2007) The association between hyperuricemia and the Trp64Arg polymorphism of the beta-3 adrenergic receptor. Rheumatol Int 27:835–839
16. Huang Q, Zhang L-F, Cheng Y, Zhao Y-C, Si L, Gao Y et al (2013) Trp64Arg (rs4994) polymorphism of β3-adrenergic receptor gene is associated with hyperuricemia in a Chinese male population. Clin Chem Lab Med 51:1755–1760
17. Morcillo S, Rojo-Martínez G, Martín-Núñez GM, Gómez-Zumaquero JM, García-Fuentes E, de Adana MR et al (2010) Trp64Arg polymorphism of the ADRB3 gene predicts hyperuricemia risk in a population from southern Spain. J Rheumatol 37:417–421
18. Wang B, Meng D, Wang J, Jia Z, Zhoub S, Liu S et al (2011) Positive correlation between Beta-3-Adrenergic Receptor (ADRB3) gene and gout in a Chinese male population. J Rheumatol 38:738–740
19. Winnard D, Wright C, Jackson G, Gow P, Kerr A, McLachlan A et al (2013) Gout, diabetes and cardiovascular disease in the Aotearoa New Zealand adult population: co-prevalence and implications for clinical practice. NZ Med J. 126:53–64
20. Winnard D, Wright C, Taylor WJ, Jackson G, Te Karu L, Gow PJ et al (2012) National prevalence of gout derived from administrative health data in Aotearoa New Zealand. Rheumatol 51:901–909
21. Wallace SL, Robinson H, Masi AT, Decker JL, McCarty DJ, Yu TF (1977) Preliminary criteria for the classification of the acute arthritis of primary gout. Arthritis Rheum 20:895–900
22. Hollis-Moffatt JE, Phipps-Green AJ, Chapman B, Jones GT, van Rij A, Gow PJ et al (2012) The renal urate transporter SLC17A1

locus: confirmation of association with gout. Arthritis Res Ther 14:R92
23. Howie B, Marchini J, Stephens M (2011) Genotype imputation with thousands of genomes. G3: genes, genomes. Genetics 1:457–470
24. Team RC (2012) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. ISBN 3-900051-07-0
25. Phipps-Green AJ, Hollis-Moffatt J, Dalbeth N, Merriman ME, Topless R, Gow PJ et al (2010) A strong role for the ABCG2 gene in susceptibility to gout in New Zealand Pacific Island and Caucasian, but not Māori, case and control sample sets. Hum Mol Genet 19:4813–4819
26. Savage DB, Petersen KF, Shulman GI (2005) Mechanisms of insulin resistance in humans and possible links with inflammation. Hypertension 45:828–833
27. Park HS, Shin ES, Lee JE (2008) Genotypes and haplotypes of β2-adrenergic receptor and parameters of the metabolic syndrome in Korean adolescents. Metabolism 57:1064–1070
28. Zhan S, Ho SC (2005) Meta-analysis of the association of the Trp64Arg polymorphism in the β3 adrenergic receptor with insulin resistance. Obes Res 13:1709–1719
29. Mark A, Rahmouni K, Correia M, Haynes W (2003) A leptin-sympathetic-leptin feedback loop: potential implications for regulation of arterial pressure and body fat. Acta Physiol Scand 177:345–349
30. Bedir A, Topbas M, Tanyeri F, Alvur M, Arik N (2003) Leptin might be a regulator of serum uric acid concentrations in humans. Jpn Heart J 44:527–536
31. Modan M, Halkin H, Karasik A, Lusky A (1987) Elevated serum uric acid—a facet of hyperinsulinaemia. Diabetologia 30:713–718
32. Quiñones-Galvan A, Ferrannini E (1996) Renal effects of insulin in man. J Nephrol 10:188–191