Investigation of the mechanisms mediating genetic associations with atrial fibrillation

Ruairidh Iain Russell Martin

Thesis submitted in partial fulfilment of the requirements of the degree of Doctor of Medicine

Newcastle University

Faculty of Medical Sciences

Institute of Genetic Medicine

March 2015

Abstract

Genome wide association studies (GWAS) have identified multiple loci which are associated with increased risk of atrial fibrillation (AF). The mechanisms underlying these associations are not understood. None of the variants identified result in alteration of coding sequences; they are therefore likely to act by altering gene expression. Identification of intermediate gene expression phenotypes which are associated with the risk variants could provide important insights into disease pathogenesis which could in turn lead to development of therapeutic targets.

One of the genetic variants identified by GWAS is situated within an intron of the gene *HCN4*, which encodes the major component of the I_f pacemaker current. This raised the question whether lower activity of *HCN4* could be a risk factor for AF. By performing meta-analysis of randomised controlled trials of ivabradine, an I_f inhibitor, I demonstrated an increase in relative risk of incident AF of 15% in patients treated with ivabradine vs. controls, supporting the role of *HCN4* in AF susceptibility.

Analysis of total expression and allelic expression ratios of candidate genes in the GWAS hit regions in whole blood identified associations between AF risk variants and increased expression of *KCNN3* and *SYNE2* and decreased expression of *CAV1*.

Analysis of total expression and allelic expression ratio in right atrial appendage tissue identified further associations between AF risk variants and increased expression of *PITX2a/b* and decreased expression of *MYOZ1*, *CAV1*, *C9orf3* and *FANCC* in right atrial tissue. Furthermore, although the experiments were not designed to detect difference between AF cases and controls, I have shown that AF is associated with reduced expression of *SYNE2*, *HCN4* and *CAV1*.

Acknowledgements

I would like to thank the many people who have contributed to my research. Thanks must go first to my supervisors, Professor Bernard Keavney and Dr Mauro Santibáñez Koref. Without their support, help, advice and encouragement, this research would not have been possible.

I must also thank the British Heart Foundation, for funding the project and the NHS National Institute of Health Research, for including the project on the Clinical Research Network Portfolio.

Thanks must also go to my collaborators, without whom the recruitment of so many patients to the research study would not have possible. Mr Andrew Owens in the James Cook University Hospital, Middlesbrough; Professor Azfar Zaman at the Freeman Hospital, Newcastle; Dr Tim Chico at the Northern General Hospital, Sheffield; Professor Daniel Keenan at Manchester Royal Infirmary, Manchester and Professor Nizar Yonan at the Wythenshawe Hospital, Manchester. I would also like to thank the clinical research teams at each of the above hospitals who assisted with recruitment. I also was very fortunate to be given access to genetic material from patients recruited in South Africa by Dr Mike Cunnington and Professor Bongani Mayosi. Sequenom genotyping and allelic expression analysis was performed by Dr Christine Blancher at the Wellcome Trust Centre for Human Genetics, University of Oxford.

The project was also assisted by a number of people in the Institute of Genetic Medicine. Dr Darroch Hall provided me with training in laboratory techniques and was invaluable in providing advice on experimental design. Dr Mauro Santibáñez Koref assisted with the custom programmes that were used for simulations and allelic expression analyses. Mr Rafiqul Hussain and Mr Mzwandile Mbele assisted with DNA extraction, Dr Elise Glen and Ms Mahsa Sheikhali Babaei assisted with eQTL experiments. Dr Oksana Pogoryelova translated papers from Russian.

Finally, I must thank my family, who have shown me endless love and kindness.

The work in this thesis, except where otherwise specified, is entirely my own, and I have not submitted it previously for a degree in this or any other institution.

Publications

Martin RIR, Pogoryelova O, Koref MS, Bourke JP, Teare MD, Keavney BD. Atrial fibrillation associated with ivabradine treatment: meta-analysis of randomised controlled trials. Heart 2014;100:1506-10.

Martin RIR, Owens WA, Cunnington MS, Mayosi BM, Santibanez Koref M, Keavney BD. Chromosome 16q22 variants in a region associated with cardiovascular phenotypes correlate with ZFHX3 expression in a transcript-specific manner. BMC Genetics. Accepted pending publication.

Martin RIR, Keavney BD. Genetics of Atrial Fibrillation. In *Recent Advances in Cardiology 16*. Rowlands D, Clarke B. 2014; JP Medical Ltd, London.

Publications in submission

Martin RIR, Sheikhali Babaei M, Dark JD, Chico T, Keenan DJM, Yonan N, Owens A, Santibanez Koref MS, Keavney BD. Genetic Risk Markers for Atrial Fibrillation Influence Allelic Expression of Candidate Genes in Blood and Atrial Tissue. In preparation

Presentations at national and international meetings

Martin RIR, Owens WA, Santibanez Koref M, Keavney B. Genetic variants on chromosome 4q25 associated with atrial fibrillation upregulate PITX2 expression. Abstract P5841

ESC Congress 2014 (moderated poster)

Martin RIR, Owens WA, Santibanez Koref M, Keavney BD. Heart. Genetic Risk Markers for Atrial Fibrillation Influence Allelic Expression of Nearby Candidate Genes. **YIA4** Jun 2014;100 Suppl 3:A123-4. doi: 10.1136/heartjnl-2014-306118.227.

British Cardiac Society 2014. (BAS/BSCR Young Investigator's Award runner up)

Martin RIR, Glen E, Hall DH, Owens WA, Santibanez Koref M, Keavney BD. Genetic variation in two chromosomal regions associated with AF influences local gene expression. Europace. 2013: 15 suppl 2; ii119. doi:10.1093/europace/eut199

EHRA Europace 2013 (moderated poster)

Martin RIR, Santibanez Koref M, Owens WA, Keavney BD. Genetic Variation Associated With Low-Density Lipoprotein Cholesterol Levels Influences Zfhx3 Expression. Heart. May 2013: 99 Suppl S2: A102-3. doi:10.1136/heartjnl-2013-304019.180

British Cardiac Society 2013 (moderated poster)

Table of contents

List of figures

List of tables

List of abbreviations

ABCC9	ATP-binding cassette, sub-family C member 9 gene
ACE	Angiotensin converting enzyme
ACTB	Actin, Beta gene
AD	Autosomal dominant
AEI	Allelic expression imbalance
AER	Allelic expression ratio
aerQTL	Allelic expression ratio quantitative trait locus
AF	Atrial Fibrillation
ANRIL	antisense non-coding RNA in the INK4 locus gene (also known as CDKN2BAS)
APD	Action potential duration
APO	Aminopeptidase O gene
AR	Autosomal recessive
ASD	Atrial septal defect
ATFB1	Atrial Fibrillation, Familial 1 gene
ATP	Adenosine triphosphate
AV	Atrio-ventricular
bpm	Beats per minute
bp	Base pairs
CAGE	Cap analysis gene expression
CDKN2A	Cyclin-dependent kinase inhibitor 2A gene
CDKN2B	Cyclin-dependent kinase inhibitor 2B gene
CAV1	Caveolin 1 gene
CAV2	Caveolin 2 gene
cDNA	Complimentary DNA
CEU	Utah residents with Northern and Western European ancestry

CH domain	Calponin homology domain
CI	Confidence interval
CNV	Copy number variation
C9orf3	Chromosome 9 open reading frame 3 gene
DEAD box	Protein motif containing the amino acids D-E-A-D
DEAH box	Protein motif containing the amino acids D-E-A-H
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
ECG	Electrocardiogram
EDMD	Emery Dreifuss muscular dystrophy
EDTA	Ethylenediaminetetracetic acid
EMeA	European Medicines Agency
ERP	Effective refractory period
eQTL	Expression quantitative trait locus
FA	Fanconi anaemia
FANCC	Fanconi anaemia group C gene
FTO	Fat mass and obesity-associated gene
FWER	Family-wise error rate
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase gene
GATA4	Transcription factor GATA-4 gene
GATA6	Transcription factor GATA-6 gene
GJA1	Gap junction alpha-1 gene
GJA5	Gap junction alpha-5 gene
GWAS	Genome wide association study
HCN	Hyperpolarisation-activated cyclic nucleotide-gated channels
HCN2	Hyperpolarisation-activated cyclic nucleotide-gated ion channel 2 gene
HCN4	Hyperpolarisation-activated cyclic nucleotide-gated ion channel 4 gene
HPRT1	Hypoxanthine-guanine phosphoribosyltransferase gene

IMA	Internal mammary artery
IRX1	Iroquois-class homeodomain gene
IST	Inappropriate sinus tachycardia
JPH2	Junctophilin 2 gene
JPX	Just Proximal to XIST gene
KASH	Klarsicht/ANC-1/Syne homology transmembrane domain
kb	kilobase
KCND3	Potassium voltage-gated channel subfamily D member 3 gene
KCNE1	Potassium voltage-gated channel subfamily E member 1 gene
KCNE1L	Potassium voltage-gated channel subfamily E member 1 like gene
KCNE2	Potassium voltage-gated channel subfamily E member 2 gene
KCNE3	Potassium voltage-gated channel subfamily E member 3 gene
KCNE4	Potassium voltage-gated channel subfamily E member 4 gene
KCNH2	Potassium voltage-gated channel subfamily H member 2 gene (hERG)
KCNJ2	Potassium voltage-gated channel subfamily J member 2 gene
KCNJ8	Potassium voltage-gated channel subfamily J member 8 gene
KCNN3	Potassium voltage-gated channel subfamily N member 3 gene
KCNQ1	Potassium voltage-gated channel subfamily Q member 1 gene
LAA	Left atrial appendage
LCL	Lymphoblastoid cell line
LD	Linkage disequilibrium
LIMA	Left internal mammary artery
LMNA	Lamin A gene
LOD	Logarithm (base 10) of odds
LVH	Left ventricular hypertrophy
LVSD	Left ventricular systolic dysfunction
MAF	Minor allele frequency
MALDI-TOF	Matrix-assisted laser desorption/ionisation time of flight mass spectrometry

Mb	Megabase
MIQE	Minimum information for publication of quantitative real-time PCR experiments
mRNA	Messenger RNA
MuTHER	Multiple tissue human expression resource
MYOZ1	Myozenin-1 gene
NE cohort	Cohort recruited in the North of England
NKX2-5	NK2 homeobox 5 gene
NNH	Number needed to harm
NNT	Number needed to treat
NSTE-ACS	Non-ST elevation acute coronary syndrome
NPPA	Natriuretic peptide A gene
NUP155	Nuclear pore complex 155kDa gene
OMIM	Online Mendelian inheritance in man database
OOSS	Overall oral safety set (of the EMeA assessment of ivabradine)
OR	Odds ratio
PCR	Polymerase chain reaction
PDGF	Platelet-derived growth factor
PDGFRB	Platelet-derived growth factor receptor beta gene
PITX2	Paired-like homeodomain transcription factor 2 gene
PPARγ	Peroxisome proliferator-activated receptor gamma
PRRX1	Paired related homeobox 1 gene
qPCR	quantitative PCR
RAA	Right atrial appendage
RCT	Randomised controlled trial
RNA	Ribonucleic acid
RNAseq	RNA sequencing
RN18S1	18S ribosomal RNA gene

RYR2	Ryanodine receptor 2 gene
SA cohort	Cohort recruited in South Africa
SCN1B	Sodium channel subunit beta-1 gene
SCN2B	Sodium channel subunit beta-2 gene
SCN3B	Sodium channel subunit beta-3 gene
SCN5A	Sodium channel subunit alpha-5 gene
SK2	Small conductance calcium-activated potassium channel 2
SK3	Small conductance calcium-activated potassium channel 3
SNP	Single nucleotide polymorphism
STEMI	ST elevation myocardial infarction
SYNE2	Nesprin-2 gene
SYNPO2L	Synaptopodin 2-like gene
TBX5	T-box transcription factor gene
TNF	Tissue necrosis factor
TGF-β1	Transforming growth factor beta 1
UCSC	University of California, Santa Cruz
UTR	Untranslated region
VSD	Ventricular septal defect
Xist	X-inactive specific transcript gene

Chapter 1

Introduction

1 Introduction

1.1 Introduction

AF is the most common sustained cardiac arrhythmia, with a prevalence of 0.5% in the fifth decade of life, rising to 5-15% over the subsequent 20 years[1]. AF is a rhythm which is characterised by apparently chaotic electrical activity in the atria. A number of potential underlying electrophysiological mechanisms have been proposed and it is likely that all of them play some role in the initiation and maintenance of AF. All of the proposed mechanisms have a common appearance on the electrocardiogram (ECG). In AF the electrical activity of the atrium can be detected as fine, rapid deflections of a variable rate and morphology at a rate of between 300 and 600 beats per minute (bpm), called f waves. The ventricular response is irregularly irregular and, in the context of normal atrioventricular (AV) nodal conduction, is usually between 100 and 160 bpm[2].

The three most commonly proposed mechanisms which underlie AF are focal ectopic firing, a single localised re-entry circuit or multiple variable functional re-entry circuits[3]. It is likely that the importance of each of these mechanisms varies with the time point in the natural history of the disease. Usually AF presents as the paroxysmal form, where episodes of AF are short lived and resolve spontaneously. There is often progression to persistent AF, defined as an episode which lasts longer than seven days, and usually requires chemical or electrical cardioversion to restore normal sinus rhythm. The final stage, permanent AF, is defined as the situation where restoration of sinus rhythm is no longer possible or considered advisable by the treating physician. Paroxysmal AF is largely driven by focal ectopic activity, often from the pulmonary vein myocardium[4]. Permanent AF, on the other hand is characterised by multiple re-entrant circuits involving the entire atrial myocardium, whereas the intermediate persistent stage may be characterised by one or more relatively stable rotors of re-entry[5]. AF results in electrical and structural remodelling of the atria, which in turn provides a substrate for further AF, leading to the natural progression of the disease from paroxysmal to persistent and finally to permanent AF[3].

AF is not a benign arrhythmia: it is associated with a significant reduction in quality of life and exercise capacity. More importantly; AF is associated with a 5-fold increase in the risk of thromboembolic stroke and a 2-fold increase in all-cause mortality[6].

There are many well recognised environmental factors and disease states which contribute to AF risk; including hypertension, heart failure, valvular heart disease, obesity, diabetes mellitus, chronic obstructive pulmonary disease, sleep apnoea, thyroid disease and renal failure[6]. A significant number of individuals, however, develop AF in the absence of these factors, so-called 'lone AF'[7]. The first case of familial AF was identified in 1943, in three brothers who developed AF in childhood[8]. However it was not until the late 1990s that the importance of genetic risk in AF became increasingly appreciated following the first published report of AF segregating as a Mendelian trait[9].

Management strategies for AF can be broadly divided into two: rate control strategies, where the AF is presumed to be permanent and the focus is on controlling the ventricular rate; and rhythm control strategies where attempts are made to restore sinus rhythm. Where a rate control strategy is adopted, the ventricular rate is limited by pharmacological inhibition of conduction through the AV node with Beta adrenoceptor antagonists, rate-limiting calcium channel antagonists or digoxin, or occasionally by radiofrequency ablation of the AV node (which would necessitate implantation of a permanent pacemaker). The most commonly used pharmacological agents used for rhythm control are Vaughn Williams class Ic agents, which block sodium channels, such as flecainide or propafenone or class III agents which block potassium channels, such as amiodarone or dronedarone[10]. A more effective approach to rhythm control is to prevent focal atrial activity arising in the pulmonary vein myocardium form initiating AF by electrically isolating the pulmonary veins using radiofrequency energy, cryotherapy or laser energy[11].

Current pharmacological strategies have a very low success rate in maintaining sinus rhythm in patients with paroxysmal or persistent AF, and while pulmonary vein isolation is more effective, 25% of patients have recurrent symptomatic AF despite multiple ablation procedures[12]. No rhythm control strategy has been shown to reduce mortality or stroke risk, probably due to the relatively high incidence of asymptomatic

AF. It is therefore important to improve our understanding of the mechanisms which underlie AF at a molecular level in order to develop new therapies which can improve our ability to maintain sinus rhythm.

1.1.1 Evidence for genetic susceptibility to AF

AF has been shown to exhibit a significant degree of heritability. A Danish twin study of 1137 single-sex twin pairs identified increased concordance in monozygotic vs. dizygotic twins (22.0% versus 11.6%, *P*<0.0001) and estimated the heritability of AF to be 62% (95% CI 55-68%)[13]. By comparison the heritability of coronary artery disease is estimated to be around 40%[14], and that of breast cancer to be 30%[15]. Although this study provides strong evidence for a genetic component to atrial fibrillation, twin and family studies in selected populations with lone AF may overestimate the genetic contribution to risk in a population with multiple risk factors for AF. Analysis of participants in the Framingham study, however, identified a significantly increased rate of AF in the children of patients with AF with an odds ratio (OR) 1.85 [95% confidence interval (CI) 1.12–3.06, P=0.02][16]; demonstrating that AF displays a significant degree of heritability even in a more general population. In keeping with there being a stronger heritable component in lone AF, the association was stronger when the analysis was limited to parents and offspring developing AF under the age of 75 (OR 3.23, 95% CI 1.87–5.58, P<0.001). These odds ratios were calculated after adjustment for multiple variables, including hypertension, diabetes mellitus, clinically overt heart disease, age and sex. The unadjusted odds ratios were somewhat higher, highlighting the importance of these risk factors. It is noteworthy that some of the recognised risk factors for AF, including diabetes and hypertension, exhibit significant heritability[17,18], and this will indirectly contribute to the overall heritability of AF.

1.1.2 The genetic architecture of AF susceptibility

Although there is good evidence that there is an important genetic contribution to AF susceptibility, establishing the genetic elements that are responsible for the heritable component is not straightforward. AF is the result of complex electrical and structural

changes within the atria[19]. Multiple genetic loci have been identified as conferring AF risk. Some of these were identified in linkage studies of families with early-onset AF. This type of study has identified AF-causing mutations in multiple genes, including those coding for sodium channels (*SCN5A*)[20], Potassium channels (*KCNQ1*)[21], atrial natriuretic peptide (*NPPA*)[22] and atrial gap junction connexins (*GJA5*)[23]. These mutations have a large effect size with a high incidence of early onset AF in carriers. These mutations give insight into genes that may be involved in atrial electrical signalling; but they are rare and they are unlikely to explain the overall heritability of AF susceptibility in the general population; candidate gene studies, including some of those described in Table 1.1, typically find mutations which are likely to be causative of AF in less than 5% of lone AF patients screened[13,24-27]. Conversely, genome-wide association studies (GWAS) have identified nine genetic regions that are associated with AF and several further regions associated with alterations in heart rate[28,29] and PR interval, which is considered a marker for AF susceptibility[30]. Typical effect sizes for loci identified by GWAS are, unsurprisingly, smaller than those seen in families with inherited lone AF (median OR 1.33, inter-quartile range (IQR) 1.20– 1.61)[31].

1.1.3 Common variation in human populations

The Human Genome Project was a major collaborative publicly-funded project to sequence the whole human genome using the then state-of-the-art Sanger sequencing technology. It reported the human genome as a single haploid sequence in 2001[32]. The human sequence, of course, is diploid and contains a significant amount of genetic variation, both within and between human populations. Genetic variation can be divided into two major groups, single nucleotide variants and structural variants (Figure 1.1)[33]. Common single nucleotide variants can also be referred to as single nucleotide polymorphisms (SNPs), a polymorphism being an occurrence in a population of more than one allele where the variant allele has a frequency of greater than 1%.

Before the invention of sequencing technology, microscopic, *i.e.* large enough to see with a microscope, structural abnormalities had been observed in chromosomes. These included rearrangements, aneuploidies (an abnormal number of chromosomes),

heteromorphisms (visible changes in the size, morphology or staining of a chromosome) and fragile sites (small breaks or constriction of a chromosome, visible under special cell culture conditions)[34]. The advent of sequencing technologies has allowed the identification of sub-microscopic variation, relatively short repetitive DNA sequences such as mini- and microsatellites and the sub-microscopic structural variants described below[35].

Structural variations involve more significant sequence alterations than single nucleotide variants. Insertion/deletion (indel) mutations are segments of DNA that are present in some chromosomes but not others; they are mostly a few bases in length but can be up to 1 kb long[36]. Longer segments of DNA (*i.e.* > 1000bp), that are present in some chromosomes but not others, are called copy number variants (CNVs). CNVs can be surprisingly large without being harmful, for example, the largest CNV in the Venter genome (the genome of Craig Venter, who was one of the first individuals to be sequenced) was almost 2 megabases long[36]. Block substitutions are similar to SNPs but several consecutive base pairs vary between genomes. Inversions occur where segments of DNA are present in the reverse order in some chromosomes; for example there is a 900kb segment of chromosome 17 that is inverted in 20% of Europeans[37].

Figure 1.1 Common forms of genetic variation

Variants are highlighted in red. Figure adapted from Frazer *et al*[33].

1.1.4 Single nucleotide polymorphisms

Of the variants above, the easiest to determine and analyse are SNPs. For this reason, SNPs have been more widely studied than other variants. Each individual has approximately 3.5 million single nucleotide variants with respect to the reference genome[38], over 62 million are reported in build 138 of the single nucleotide database (dbSNP)[39], and there are estimated to be around 11 million with a minor allele frequency (MAF) of >1%[40], *i.e.* more than 1:100 chromosomes carry the variant allele. Interestingly, a large proportion of the single nucleotide variants carried by any given individual are common, suggesting that most variation between two individuals is the result of common, rather than rare variants.

1.1.5 Linkage disequilibrium and haplotypes

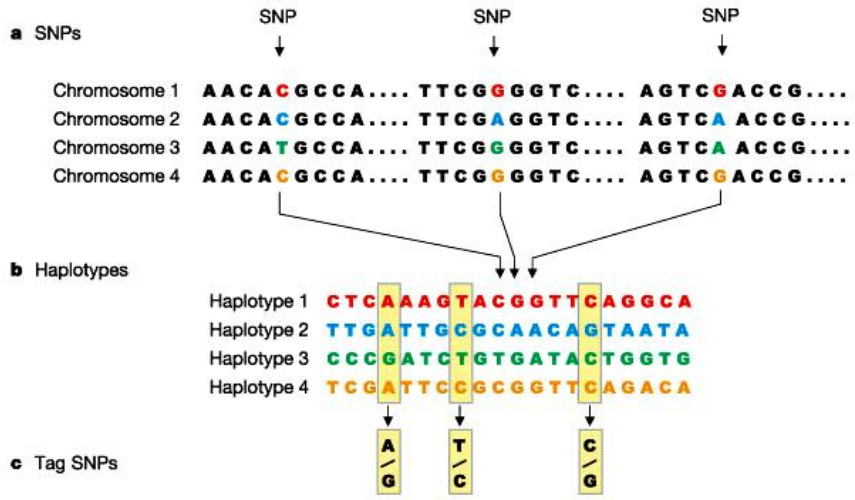
The diploid human genome consists of 46 chromosomes, 22 pairs of autosomes and two sex chromosomes. During gametogenesis, the formation of egg and sperm cells, the genome is reduced from diploid to haploid (22 autosomes and a single sex chromosome) through meiosis, consisting of two cell divisions, meiosis I and II. Prior to division the chromosomes divide to form two chromatids, joined at the centromere.

The maternally and paternally derived chromatids align to form a four strand structure called a tetrad. When the chromatids are aligned they can exchange material via breakage and recombination of the DNA strand at points called chiasmata. The chromatids are then separated and the cell divides to form the gametes. The consequence of this recombination is that the chromosomes within each gamete contain genetic material from both the maternally and paternally derived chromosomes[41]. Alleles of marker polymorphisms on the same chromosome are therefore not always inherited together. The chance of markers being separated at meiosis is increased with increasing distance and when the regions between the two markers are particularly prone to chiasm formation, so called 'recombination hotspots'[42]. After many mitotic generations markers are shuffled many times so that only those that happen to be in regions where recombination has not occurred between the variants remain associated. If two markers are preferentially associated with one another, due to infrequent recombination between them, they can be referred to a being in linkage disequilibrium (LD). Markers which are more strongly associated are described as having higher linkage disequilibrium. One measure of LD, D', is calculated as the deviation of the observed frequency of the markers from the expected value, divided by the theoretical maximum deviation. Regions of the genome that are inherited together are known as haplotype blocks. These can be defined as sets of consecutive markers where there is little or no evidence for historical recombination. One approach to this relies on constructing a confidence interval based on the population value of D' [43]. Haplotype blocks allow the use of marker polymorphisms to 'tag' variants on the genome, as the genotype of variants near the tagging variant can be inferred from the genotype of the tag. Mapping of these patterns of inheritance is the aim of the International HapMap Project[44] (Figure 1.2).

The size of haplotype blocks is in part due to the distribution of recombination hotspots but is also a consequence of the number of generations which have elapsed, as a greater number of meioses would be expected to result in recombination at a greater number of loci. If a new population is established from a small number of individuals, then a loss of genetic variation occurs, the founder effect, and the haplotype structure will be relatively coarse[45]. Therefore, one of the major determinants of the coarseness of haplotype structure is demographic history. This is

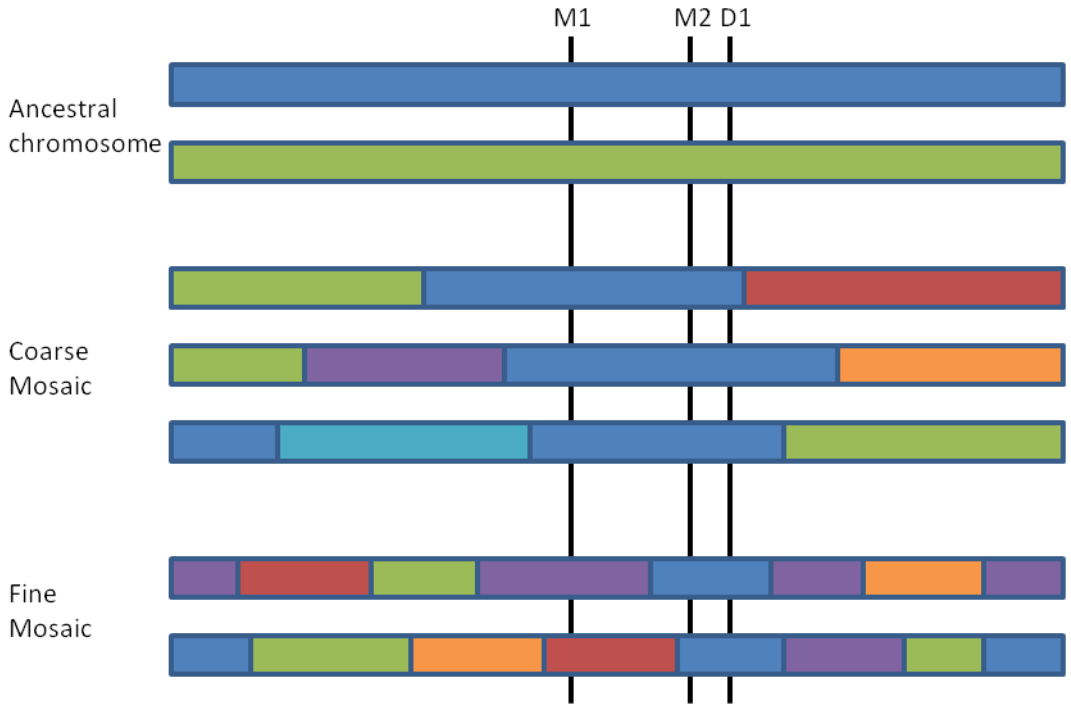
important as large haplotype blocks allow for tagging of the genome with fewer markers. A finer haplotype structure will require a greater number of tag SNPs but has the advantage that the disease associated polymorphism is in linkage disequilibrium with a smaller region of the genome, allowing finer mapping of the locus (Figure 1.3).

Figure 1.2 The haplotype map



The construction of a haplotype map occurs in three steps. **(a)** SNPs are identified in DNA samples from multiple individuals. **(b)**Adjacent SNPs that are inherited together are compiled into haplotypes. **(c)**Tag SNPs identify each haplotype. By genotyping the three tag SNPs shown in this figure, researchers can identify which of the four haplotypes shown here are present in each individual. Figure reproduced from the International HapMap project[44].

Figure 1.3 Decay of association with time



The hypothetical ancestral chromosome contains all markers in association with the disease locus. Over time recombination events ensure that each chromosome is a mosaic. After a relatively small number of generations the mosaic is quite coarse and both markers M1 and M2 are associated with the disease locus. After many generations the mosaic is fine and only marker M2, which is closer to the disease locus, remains in association. Figure adapted from Cunnington and Keavney[46].

In human populations the haplotype mosaic is relatively coarse, allowing tagging of the genome with a relatively small number of SNPs. The coarseness of the mosaic does vary to some degree between ethnic groups, however. Those populations that have been through a population bottleneck and therefore have a relatively small founder population in the not-too-distant past have had much less time for shuffling of the haplotypes and so will have a coarser mosaic. For this reason European populations need fewer SNPs (*ca.* 300,000) to tag the majority of common polymorphisms than African populations (*ca.* 1,000,000)[43], as Africa has been populated by humans for considerably longer than Europe and the European founder population will have only contained a small number of the haplotypes present in the original African gene pool. The International HapMap Project, by providing a map of human haplotypes in several reference populations, allows selection of appropriate marker polymorphisms for the

tagging of common polymorphisms. Of note, the HapMap project focused on describing the patterns of association between common single nucleotide polymorphisms, that is those with a MAF of >5%. Due to the sample size in the HapMap cohorts, the ability to define haplotypes for variants with a MAF <5% was limited by the small number of carriers of rarer alleles. As mentioned above a significant proportion of the variation between individuals is a consequence of structural variants. It appears however that common short indels and the commoner structural variants are in LD with and behave in a similar fashion to common SNPs, suggesting that they are both ancestral, *i.e.* each variant has arisen once in human history rather than being the consequence of multiple similar mutation events[47]. Structural variations in segmental duplications, however are less well captured by tagging SNPs[48].

Despite these cautions a large number of genome wide association studies (GWAS), using SNPs selected with the aid of the HapMap, have demonstrated important associations between diseases and genetic loci and identified new genes associated with these diseases[30].

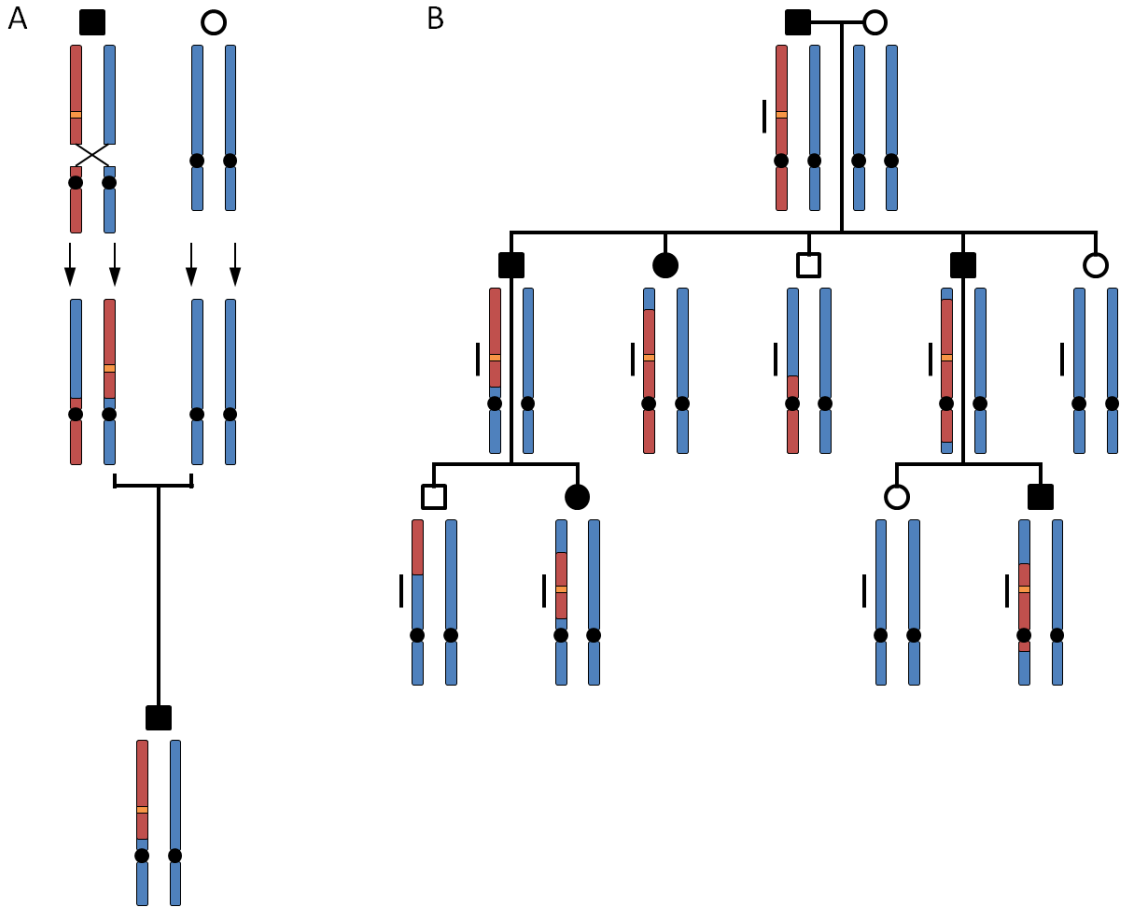
1.2 Mendelian AF

Some genetic variants can be considered to be causative of AF, *i.e.* a single gene defect results in AF in all or most carriers: these are monogenic or Mendelian defects. In most families the inheritance is autosomal dominant with penetrance; almost all family members with a single defective copy of the gene will develop AF. The genes responsible for AF in these families can be identified by traditional genetic linkage studies or, more recently, by directly sequencing a large portion of the genome, either the coding regions (exome sequencing) or a large number of candidate genes, or more recently the whole genome, in affected members.

Linkage studies (Figure 1.4) are performed by genotyping multiple genetic markers throughout the genome in both affected and unaffected family members in order to identify a haplotype, a region of a chromosome that is inherited as a complete block, which co-segregates with the disease. Once the disease-associated haplotype is identified known genes within the region can be sequenced.

The first published study of Mendelian AF used this approach to identify a region in chromosome 10 that was common to the affected members of three families with AF and left ventricular systolic dysfunction[9]. Further studies have identified regions on chromosomes 5 and 6 and a second region on chromosome 10[49-51] (Table 1). Unfortunately the regions identified are large and identification of the responsible variant in each of these cases has not, as yet, proven possible.

Figure 1.4 Linkage analysis.



Filled shapes represent affected individuals, squares are male and circles are female.

Panel A. An affected male carries a disease causing gene (orange) located on a chromosome (red). During meiosis there is recombination between the disease carrying chromosome and its normal partner (blue) so that his affected son inherits only part of the disease carrying chromosome.

Panel B. In a large family, many recombination events occur. Only those family members who have inherited that part of the chromosome in which the gene is located will be affected. The region of the chromosome which carries the gene can therefore be narrowed down to the region which is present in all of the affected family members and none of the unaffected (vertical black line). This allows sequencing of genes in the shared haplotype region.

Some families are not large enough to perform linkage analysis, as a minimum number of family members are required to produce a statistically significant result. Under these circumstances it is possible to adopt a candidate gene sequencing approach. Genes are identified which are thought to be involved in AF (candidate genes) and

these are sequenced in affected and unaffected family members. This allows identification of variants which segregate with disease, *i.e.* they are present in affected family members but not in unaffected members. Some investigators have extended this approach further, by sequencing candidate genes in cohorts of unrelated individuals with lone AF of early onset; followed up with genotyping of any variant in affected and unaffected family members, if possible. The candidate gene approach can introduce significant bias, as variants will only be found in those genes that the investigators have selected for analysis. It is possible, if not likely, that there will be disease causing variants located outside the sequenced regions which will be missed. The sequencing of a large proportion of any individual's DNA will identify rare and unreported variants. As described above there are a great number of variants in every genome. Many of these variants cause significant changes to the amino acid sequence of the encoded proteins and it is estimated that each of us typically carries around 100 genuine loss-of-function mutations with around 20 genes totally inactivated[52]. It is difficult, therefore, to be certain that any new variant identified will be causative of disease. In spite of these difficulties it has been possible to identify a number of mutations which have some evidence supporting their role in AF. These include mutations causing gain or loss of function in sodium[20,53-58], potassium[21,24,26,59-67] and calcium channels[68], gap junctions[23,69-71], endocrine peptides[22], developmentally important transcription factors[72-77] and even nucleoporins[78], which regulate transport of large molecules between the nucleus and cytoplasm (see Figure 1.5 and Table 1.1).

AF is characterised by a disturbance in the normal electrical activity in the atrium and therefore the ion currents which are responsible for the normal action potential in sino-atrial cells and atrial myocytes are of particular interest. The five phases of the action potential (numbered 0-4) are the result of complex interaction between several membrane currents, each of which is generated by one or more ion channels. The ion channels are often heteromultimers of several subunits, which are encoded by individual genes. There are also other ion currents which have tonic effects on the membrane potential and do not vary in their activity over the time course of a single action potential. In the case of sodium and calcium channels the pore-forming α-subunit consists of 4 covalently bound domains encoded by a single gene, with regulatory β-subunits. In the case of potassium channels the pore forming domains are

not covalently linked and therefore different α-subunits combine to form heteromultimers. The main genes responsible are listed by each current in Figure 1.5. Several of these genes have been identified as being associated with AF, either through gain- or loss-of-function mutations or through GWAS hits, see Tables 1.1 and 1.2.

In the atria, the non-voltage-gated $I_{K,Ca}$ Ca^{2+}-gated, I_{KACh} acetylcholine-gated and I_{KATP} ATP-gated currents serve to hyperpolarise the cell throughout the action potential and therefore shorten action potential duration and reduce automaticity. $I_{K,Ca}$ and I_{KACh} are not present in ventricular myocytes. Gap junctions, formed by connexins 40 and 43, allow the transmission of electrical impulses between cells[79].

In the atrial myocyte, the resting potential in phase 4 is maintained by the I_{K1} inward rectifying current. Depolarisation (phase 0) is usually triggered by depolarisation of an adjoining cell and is transmitted through the connexin channels that link adjoining myocytes. Phase 0 is very rapid due to the I_{Na} sodium current. There is a brief repolarisation (phase 1) as a result of the I_{to1} transient outward current, followed by a plateau (phase 2), which is a result of the balance of the $I_{Ca,L}$, I_{Ks} and I_{Kur} delayed ultra-rapid current. The I_{Kur} is unique to atrial myocytes. Repolarisation (phase 3) is mediated by the I_{Kr} current.

In the sino-atrial cell, there is a gradual depolarisation throughout phase 4 as a result of the I_f funny current and the $I_{Ca,T}$ transient calcium current, which at threshold triggers the $I_{Ca,L}$ long-lasting calcium current and depolarisation (phase 0). Repolarisation, (phase 3), is mediated by the I_{Ks} slow rectifying current and the I_{Kr} delayed rapid rectifying current. Phases 1 and 2 are absent.

Figure 1.5 The main currents contributing to the action potential in sino-atrial and atrial myocytes.

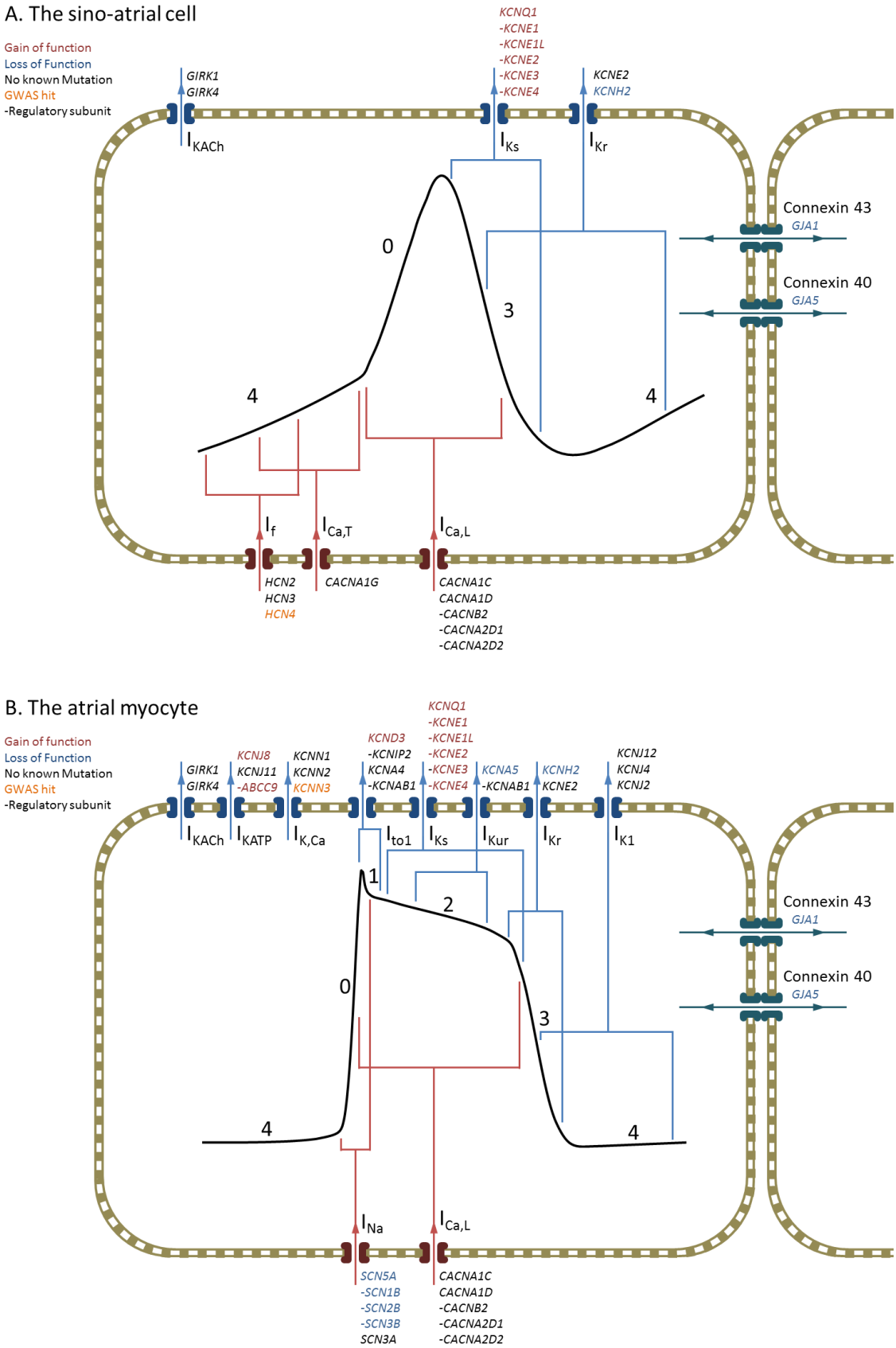


Figure legend overleaf

The sino atrial cell and the atrial myocyte are represented with a curve representing the normal action potential (the change in membrane potential over time). Depolarising currents are (red) and repolarising currents (blue) are indicated with their periods of maximum activation. The main genes responsible for each current are listed with colour coding to indicate Mendelian inheritance of gain of function (red) or loss of function (blue) mutations or nearby GWAS hits (yellow).

Table 1.1 Loci identified as causing Mendelian AF

Gene	Chromosomal region	Inheritance	Variant	Functional effect and phenotypic details	Other associated diseases
Unknown	10q22-q24	AD	Unknown	Molecular effect unknown. 19 affected subjects in 3 families, diagnosed with permanent (18) or paroxysmal (1) AF at age 2-46. Two subjects in one family had left ventricular systolic dysfunction[9].	
Unknown	6q14-q16	AD	Unknown	Molecular effect unknown. 8 of 34 family members diagnosed with paroxysmal AF at age 21-72, progressing to permanent AF in 4 older subjects. Two family members without AF had peripartum cardiomyopathy[49].	
Unknown	10p11-q21	AD	Unknown	Molecular effect unknown. Haplotype present in 9 of 34 samples. AF present in all except 3 youngest. Age at onset 40-58[50].	
Unknown	5p15	AD	Unknown	Molecular effect unknown. 8 of 27 family members affected diagnosed at age 18-42. Increase in signal-averaged P-wave duration (203±21ms vs. 116±12ms) in cases and carriers. Two affected members had mild LVH[51].	
KCNQ1	11p15.5	AD	R14C, S140G, V141M, S209P, R231H, R231C	Gain of function in KCNQ1-KCNE1 and KCNQ1-KCNE2 I_{Ks} channels. R231H and R231C found in multiple families, V141M found in a single proband with *in utero* AF[21,59,80-83].	Long QT, Jervell and Lange-Nielsen Syndrome, Short QT syndrome

Gene	Chromosomal region	Inheritance	Variant	Functional effect and phenotypic details	Other associated diseases
KCNE1	21q22.12	?AD	G25V, G60D	Gain-of-function in KCNQ1-KCNE1 I_{Ks} channels. Two families identified by candidate gene sequencing of 209 Danish probands with AF onset at <40 years old[24].	Jervell and Lange-Nielsen Syndrome, Long QT
KCNE2	21q22	AD	R27C	Gain on function in KCNQ1-KCNE2 I_{Ks} channels. Symptomatic premature atrial complexes in 5 and paroxysmal AF in 3 affected members of two families[60].	Long QT
KCNE3	11q13-q14	AD	R53H, V17M	Reduced KCNE3 inhibition of I_{Ks} channels. R53H effect unknown. R53H in one family, V17M in one individual[61,84].	Brugada Syndrome 6
KCNE4	2q36	n/a	E141A	Molecular effect unknown. Identified in a single individual in a cohort of 80 European AF probands[25].	
KCNE1L	Xq23	n/a	L65F	Failure of KCNE1L to suppress KCNQ1-KCNE1 I_{Ks} current. Single mutation found in candidate gene study of 158 patients with AF[26].	
KCND3	1p13	n/a	A2E, K214R, A545P	A545P leads to gain of function in I_{TO} channels. Identified in a single individual in a cohort of 209 Danish probands[85]. No channel dysfunction effect detected for other mutations, which were identified in a single individual in a cohort of 80 European AF probands[25].	Spinocerebellar ataxia-19

Gene	Chromosomal region	Inheritance	Variant	Functional effect and phenotypic details	Other associated diseases
KCNH2	7q36	n/a	E444K	Loss of function in I_{Kr}. Identified in a single individual in a cohort of 80 European AF probands[25].	Long QT
KCNJ2	17q24.3	AD	V93I	Gain of function of I_{K1}. Mutation found in 6 members of one family. 4 had AF at age 50-57; two had 'uncertain phenotypes'[67].	Andersen cardiodysrhythmic periodic paralysis, Short QT
KCNJ8	12p12.1	n/a	S422L	Gain of function of K_{ATP} current. Two individuals identified in a candidate gene study of a cohort of 325 probands.	?Prinzmetal angina
KCNA5	12p13.32	Some AD	Y155C, E375X, D469E, P488S, T527M, A576V, E610K, 71-81del	Loss of function of I_{Kur}. Different mutations identified in different individuals and families. 71-81del results in loss of function and interferes with tyrosine kinase modulation of I_{Kur}[62,64,65,86].	
		n/a	E48G, A305T, D322H, G586V	Gain of function in I_{Kur}. Each identified in a single individual in two cohorts[25,86].	
ABCC9	12p12.1	n/a	T1547I	Loss of ADP-induced potassium channel opening. Single mutation identified in an individual with paroxysmal AF from age 43.[66]	Dilated cardiomyopathy, Hypertrichotic osteochondrodysplasia

Gene	Chromosomal region	Inheritance	Variant	Functional effect and phenotypic details	Other associated diseases
SCN5A	3p22.2	AD	M138I, T220I, R222Q, R340Q, R367H, E428K, H445D, N470K, A572D, L618F, E655K, fs851, E446K, R814W, E1053K, T1131I, D1275N, V1279I, F1520L, F1596I, D1595H, R1826C, I1835T, N1986K, V1951M	Loss (or presumed loss) of function of I_{Na}. Only D1275N and R222Q reported in more than one family. Incomplete penetrance of DCM and AF with phenotypes often, but not always, overlapping. Reduction in I_{Na} may be dependent on presence of common H558R polymorphism of SCN5A[20,53-56,87-89].	Brugada, Long QT, DCM, Heart block
		AD	S216L, A997S, R1193Q, K1493R, R1626H, M1785T, D1819N, F2004L	Gain of function of I_{Na}. Some reported in multiple individuals with lone AF. Many reported in families with sudden infant death syndrome or long QT syndrome[57,87,89-92].	
SCN1B	19q13.11	n/a	R85H, D153N	Reduction in I_{Na}. Two individuals, one with lone paroxysmal AF, one with AF and moderate aortic stenosis; neither had LVSD[58].	Brugada, Epilepsy
			R214Q	Molecular function not known. Three individuals in a cohort of 214 Danish probands with lone AF.	
SCN2B	11q23.3	n/a	R28W, R28Q	Reduction in I_{Na}. Two individuals, one with lone paroxysmal AF one with hypertension and paroxysmal AF; neither had LVSD[58].	
SCN3B	11q24.1	n/a	R6K, L10P, A130V, M161T	Reduction in I_{Na}. Each variant identified in a single individual with lone AF[93,94].	Brugada

Gene	Chromosomal region	Inheritance	Variant	Functional effect and phenotypic details	Other associated diseases
RYR2	1q42-q43	AD	Deletion in exon 3	Presumed reduction in Ca^{2+} efflux from SR. 16 members of 2 unrelated families. Phenotype of dilated cardiomyopathy with ventricular tachycardia, atrial standstill and AF[68].	Arrhythmogenic right ventricular dysplasia, Catecholaminergic polymorphic VT
JPH2	20q13.12	n/a	E169K	Increased Ca^{2+} efflux due to loss of JPH2-mediated stabilisation of RyR2. Observed in two individuals with hypertrophic cardiomyopathy and early onset AF[95].	Hypertrophic cardiomyopathy
GJA5	1q21.1	Somatic	G38D, P88S, A96S, M163V	Reduced connexin 40-mediated cell-cell coupling and/or transport of connexin protein to the cell surface. All but A96S were somatic mutations, detected by sequencing tissue from 15 individuals with lone AF[23].	
		AD	Q49X, I75F, V85I, L221I, L229M	I75F results in loss of connexin 40 gap junction function. The molecular effects of the other mutations are unknown. Missense mutations segregate with AF in 5 unrelated Chinese families[70,71,96].	
GJA1	6q22.31	Somatic	c.932delC	Reduced connexin 43-mediated cell-cell coupling and transport of connexin protein to the cell surface. Somatic mutation in a single individual with lone AF[69].	Atrioventricular septal defect, Hallermann-Strieff syndrome, Hypoplastic left heart, Oculodentaldigital dysplasia, Syndactyly

Gene	Chromosomal region	Inheritance	Variant	Functional effect and phenotypic details	Other associated diseases
GATA4	8p23.1-p22	AD	G16C, H28D, Y38D, S70T, P103A, S160T	All shown to result in reduced transcription of reporter gene vs. wild type in cell culture. Other GATA4 mutations associated with congenital heart disease. Mutations identified in single families in three cohorts of 130, 150 and 160 Han Chinese probands with lone AF[73,74,97]	ASD, Atrioventricular septal defect, VSD
			M247T, A411V,	Molecular function not known. Each identified in a single patient, one with lone AF and one with ASD and VSD[98].	
GATA6	18q11.2	AD	Q206P, Y265X, Y235S, G469V	Y235S and G469V decreases expression of downstream reporter gene, the molecular effect of Q206P and Y265X is unknown. Mutations found in multiple members of single families identified by candidate gene sequencing of three cohorts of 110, 138 and 140 Han Chinese probands with lone AF. One family member carrying Q206P had an atrial septal defect (ASD), two family members carrying Y265X had ventricular septal defects (VSD) and one family member carrying Y235S had an ASD and one had a VSD[75-77].	ASD, Atrioventricular septal defect, Tetralogy of Fallot, Persistent truncus arteriosus, Pancreatic agenesis
NPPA	1p36-p35	AD	c.456-457delAA	Increased ANP levels with shortened APD and ERP. Seen in 11 of 15 members in a single family[22].	

Gene	Chromosomal region	Inheritance	Variant	Functional effect and phenotypic details	Other associated diseases
TBX5	12q24.21	AD	G373A	Gain of function in TBX5 presumed to increase NPPA, GJA5, KCNJ2 and TBX3 expression. 12 affected members in a single family with atypical Holt-Oram syndrome[72].	
LMNA	1q22	AD	A278T, T488P	Molecular function unknown. A278T segregated in a single family with lone AF of onset at 21-72 years old and mild muscle weakness[99]. T488P was identified in a single individual in cohort of 268 probands[27].	Dilated cardiomyopathy, Malouf Syndrome, Emery-Dreifuss muscular dystrophy, Limb girdle muscular dystrophy, Charcot-Marie-Tooth disease, Lipodystrophy, Restrictive dermopathy, Hutchinson-Gilford Progeria, Manibuloacral Dysplasia
NUP155	5p13	AR	R391H	Loss of function of nuclear pore complex. Segregated in an autosomal recessive fashion with AF in single family[78].	
NKX2-5	5q35.1	AD	F145S, N19D, F186S	Reduced transcriptional activity. Each present in a single family[100,101].	ASD, Tetralogy of Fallot, Hypoplastic left heart syndrome, VSD, Conotruncal heart malformation.

ASD – atrial septal defect, APD – action potential duration, ERP – effective refractory period, LVH – left ventricular hypertrophy, LVSD – left ventricular systolic dysfunction, VSD – ventricular septal defect. Genetic variants are given as amino acid substitutions, except where preceded by "c." to indicate cDNA changes. For some of the loci, such as for the S140G mutation in KCNQ1, there is very good evidence, with statistically significant LOD scores on linkage analysis and functional data from cellular models. For other loci there is significantly less evidence and must be treated with a degree of circumspection. Also listed are those non-AF disease associations for each gene reported in the Online Mendelian Inheritance in Man database[102]

1.3 AF inherited as a complex trait

Few of the variants listed in Table 1.1 have been identified in more than one family; partly because deleterious mutations will tend to be removed from the population by negative selection, but also because rare mutation will be lost over time by the process of genetic drift. Therefore highly penetrant variants of this type are unlikely to be important causes of the total burden of AF in the population. This does not, however, rule out the possibility that lower penetrance common variants in the same genes could be responsible for a significant proportion of AF risk. In common with many other complex diseases, it has been hypothesised that the heritability of AF in the general population results from the interaction of many individual risk variants: that is, it is inherited as a complex trait. One method of identifying the variants that contribute to this overall risk is to look for associations between common genetic variants, usually SNPs, and AF in case-control studies. The initial studies of this type focussed on variants which lay in or near candidate genes for AF. Unfortunately many of these studies were underpowered and therefore were at significant risk of type-II errors; that is they were likely to miss significant associations. Conversely, the large number of small studies and subgroup analyses performed means that the associations found are likely to reflect type-I errors, since they are underpowered to detect real associations, and the statistical significance threshold was set too low in the context of millions of common SNPs throughout the genome. Indeed, an attempt to replicate in two large European cohorts the findings of all the candidate gene association studies of AF published before 2007 failed to replicate a single association[103] and none of the associations identified in candidate gene studies have subsequently reached significance in genome wide studies.

The systematic identification of SNPs by the SNP consortium, the mapping of common variation by the HapMap project and the availability of microarray genotyping platforms have allowed extension of association studies to the whole genome. Such GWAS typically look for associations between a disease state and 300,000 to 2.5M SNPs and reliably capture 80-90% of variants in the genome that have a minor allele frequency (MAF) greater than 1%[104]. High-throughput techniques allow the inclusion of large numbers of cases and controls, and therefore stringent statistical significance

levels with replication in a similar sized cohort to the discovery cohort (typically ~2000 cases in each cohort), thus avoiding the pitfalls of the earlier candidate gene studies.

A number of GWAS have performed in order to identify regions associated with AF. The first study, performed in an Icelandic population, identified a signal at chromosome 4q25, which has been identified as the strongest signal in all subsequent GWAS to date[105]. Further studies in populations of European descent have identified associations at 16q22 and 1q21[106-108]. In 2012 a large meta-analysis of all of the previously published AF GWAS data in individuals of European ancestry was published, including 6,707 cases and 52,426 controls (Figure 3, Table 1.2)[109]. The results were further replicated in 5,381 cases and 10,030 controls of European ancestry and 843 cases and 3,350 controls of Japanese descent. This analysis confirmed the three previously described associations at 4q25, 16q22 and 1q21, and identified a further six loci which were replicated in the European cohort: 1q24, 7q31, 9q22, 10q22, 14q23 and 15q24. These genetic regions are discussed below.

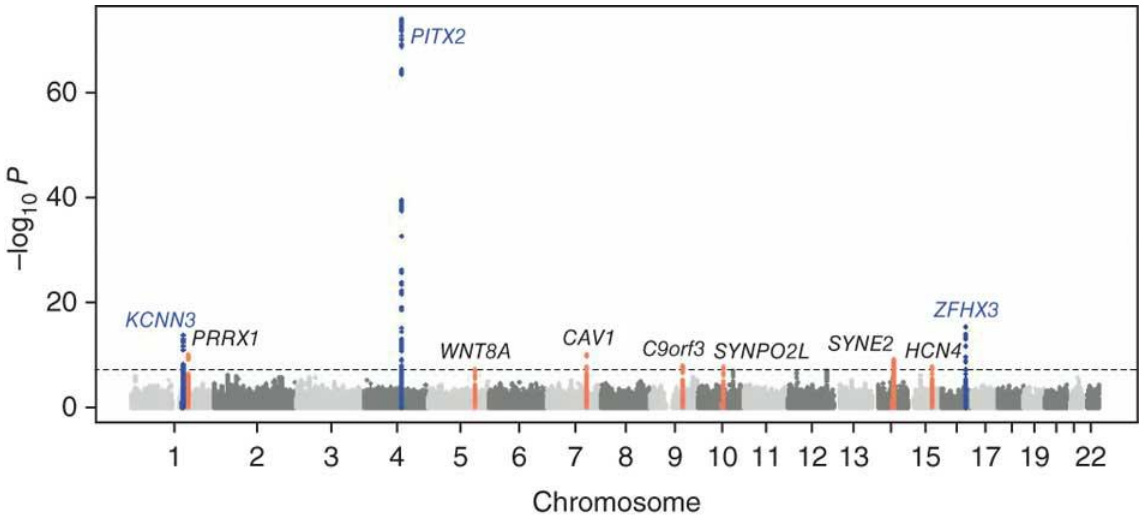
Table 1.2 AF loci identified by GWAS studies

Study	Phenotype	Cases/controls in discovery sample	Cases/controls in replication sample	Region	SNP risk allele	Nearest gene	Odds ratio [95% CI]	P value	Number of SNPs
Gudbjartsson, 2007[105]	AF and atrial flutter	550/4,476	3,336/17616	4q25	rs2200733-T	*PITX2*	1.72 [1.59-1.86]	$3x10^{-41}$	316,515
				4q25	rs10033464-T	*PITX2*	1.39 [1.26-1.53]	$7x10^{-11}$	
Gudbjartsson, 2009[106]	AF	2,385/33,752	2,427/3,379	4q25	rs2200733-T	*PITX2*	1.42 [NR]	$1x10^{-14}$	303,136
				16q22	rs7193343-T	*ZFHX3*	1.21 [1.14-1.29]	$1x10^{-10}$	
Benjamin, 2009[107]	AF	3,413/37,105	2,145/4,073	4q25	rs17042171-A	*PITX2*	1.65 [NR]	$4x10^{-63}$	2.5 million (imputed)
				16q22	rs2106261-T	*ZFHX3*	1.25 [NR]	$2x10^{-15}$	
Ellinor, 2010[108]	Lone AF	1,335/12,844	1,164/3,607	4q25	rs6843082-G	*PITX2*	2.03 [1.79-2.30]	$3x10^{-28}$	2.5 million (imputed)
				1q21	13376333-T	*KCNN3*	1.52 [1.40-1.64]	$2x10^{-21}$	

Study	Phenotype	Cases/controls in discovery sample	Cases/controls in replication sample	Region	SNP risk allele	Nearest gene	Odds ratio [95% CI]	P value	Number of SNPs
Ellinor 2012[109]	AF	6,707/52,426	5,381/10,030	4q25	rs6817105-C	*PITX2*	1.64 [1.55-1.73]	$2x10^{-74}$	2.5 million (imputed)
				15q24	rs7164883-G	*HCN4*	1.19 [1.14-1.24]	$3x10^{-17}$	
				16q22	rs2106261-T	*ZFHX3*	1.24 [1.17-1.30]	$3x10^{-16}$	
				1q21	rs6666258-C	*KCNN3*	1.18 [1.13-1.23]	$2x10^{-14}$	
				1q24	rs3903239-G	*PRRX1*	1.14 [1.10-1.17]	$8x10^{-14}$	
				14q23	rs1152591-A	*SYNE2*	1.13 [1.09-1.17]	$6x10^{-13}$	
				7q31	rs3807989-G	*CAV1*	1.11 [1.09-1.15]	$4x10^{-12}$	
				9q22	rs10821415-A	*C9orf3*	1.11 [1.08-1.15]	$4x10^{-11}$	
				10q22	rs10824026-A	*SYNPO2L*	1.15 [1.10-1.20]	$4x10^{-9}$	

NR – not reported

Figure 1.6 Manhattan plot from Ellinor *et al.* 2012[109].



Each point plotted represents an individual SNP with AF. Position on the x-axis indicates position of the SNP in the genome and SNPs are colour-coded alternately light and dark grey to indicate on which chromosome they lie. Height on the y-axis represents the strength of the evidence for the association expressed as the negative log of the p-value of the association test. The horizontal dashed line indicates the p-value cut-off threshold for GWAS of $5x10^{-8}$, which corrects for multiple testing. The three associations identified previously are marked in blue with other significant associations in orange. The association with SNPs near *WNT8A* in chromosome 5 was not replicated.

1.3.1 The 1q21 region

One of three loci independently identified in a lone AF population lies in 1q21[108], 6 megabases downstream of the locus of the 1q21.1 microdeletion syndrome, which is associated with cardiac developmental abnormalities[110]. The associated SNPs are within intron 1 of *KCNN3*, the gene which codes for the small conductance calcium-activated potassium channel SK3. The calcium-activated potassium current, $I_{K,Ca}$, mediated by SK channels, is a key effector in atrial repolarisation[111]. Activation of SK channels produces membrane hyperpolarisation suppresses membrane excitability[112]. SK channels are expressed at different levels in different cardiac tissues, being predominantly found in atrial myocytes. As described in Table 1.1 and Figure 1.5, highly penetrant mutations in other potassium channel genes (*KCNQ1*, *KCNE2*, *KCNJ2*, *KCNA5*)[21,62,65,67] have been found in familial forms of AF. Members of the SK channel family form heteromultimers; thus relatively small effects on expression of one family member such as *KCNN3* might, through disturbing interactions between channel

subunits, have marked effects on atrial electrophysiology. SK channel trafficking to the cell membrane is a requirement for the shortening of the action potential duration during burst pacing in a rabbit model[113]. Action potential shortening decreases the refractory period of the myocyte[114], which is thought to have an important role in the initiation and maintenance of AF. Endothelial derived hyperpolarising factor (EDHF) signalling in blood vessels requires SK3 activity, which initiates gap junction-conducted arteriolar dilatation[115]. This observation links SK3 to the gap junction protein Connexin-40, which has been shown to be implicated in monogenic AF[23].

1.3.2 The 1q24 region

Of the associations identified by meta-analysis but not in an individual GWAS, the strongest was seen with the genotype of the SNP rs3903239 in chromosome 1q24, near the gene *PRRX1*. *PRRX1* is a transcription factor which is widely expressed in mesenchymal tissue during embryogenesis and especially in the endocardial cushions and atrio-ventricular valves. *Prx1* null mouse models have significant great vessel abnormalities[116] and deficiencies in pulmonary vascularisation[117].

1.3.3 The 4q25 region

The locus with the strongest associations in all of the published GWAS of AF is in 4q25. The lead AF risk SNPs on 4q25 are approximately 120 kilobases 5' from the nearest gene, *PITX2*. The *PITX2* gene encodes the PITX2 protein which has three major isoforms in humans, PITX2a, PITX2b and PITX2c, and which plays a roles in the development of the heart, as well as the eye, gut, spleen, lungs and in limb myogenesis[118-121]. PITX2a and PITX2b share a promoter region and are differentially spliced. PITX2a consists of exons 1, 2, 5 and 6; PITX2b also includes exon 3[122]. PITX2c has a different promoter region and consists of exons 4, 5 and 6. Proteins often consist of several domains, which are conserved parts of the protein sequence which can function independently of the rest of the protein chain. All three isoforms of PITX2 share the same C-terminus and homeobox domain but have differing N-terminus regions. In humans a fourth isoform, PITX2d, has been identified which consists of the c-terminal and part of the homeobox domain only, and has a dominant-negative action

on the effect of the other isoforms[123]. PITX2c is the only isoform to be asymmetrically expressed in the left lateral plate mesoderm, indicating its role in the development of the left-right axis.[122] PITX2c is the only isoform to be identified in RNAseq studies of human atrium thus far[124], however a recent qPCR study indicated that other isoforms of PITX2 are expressed, predominantly PITX2a, at similar levels in left atrial appendage (LAA) and right atrial appendage (RAA)[125].

PITX2c regulates expression of downstream genes to suppress development of the left-sided sino-atrial node[126,127], control cardiac left-right asymmetry and left atrial development[128], and control development of the pulmonary vein myocardium, which is an important source of triggered activity that initiates paroxysmal AF[4]. The importance of PITX2a and PITX2b in cardiac development is less well established. PITX2 isoforms form heterodimers, potentiating their transcriptional effects[123], and providing a potential mechanism whereby genetic variation influencing transcription of non-*c* isoforms of PITX2 could indirectly affect the biological activity of the *c* isoform.

PITX2 isoforms are not expressed in easily accessible tissues other than skeletal muscle, which has made large cohort studies of eQTLs regulating PITX2 expression difficult. A recent analysis of *PITX2c* expression did not find any association with AF risk variants in a cohort of 239 LAA samples, but did identify elevated *PITX2c* expression in patients in AF at the time of tissue collection compared to those in sinus rhythm[129]. A genome-wide eQTL mapping study in left and right atrial samples from 64 European individuals found no association between *PITX2* expression and AF risk variants[130]. So far no study has examined the relationship between *PITX2a* and *PITX2b* expression and AF-associated SNPs.

1.3.4 The 7q31 region

Another SNP identified by meta-analysis, rs3807989, is in *CAV1* in chromosome band 7q31, which encodes caveolin-1, a signal transduction protein. The same SNP was found to be significantly associated with P-R interval on the ECG, itself an indicator of AF risk, and also with insulin resistance in hypertensive individuals in previous GWA studies[131,132]. *CAV1* is expressed in atrial myocytes. *CAV1* knockout mice, in which the gene has been inactivated, develop a dilated cardiomyopathy with pulmonary

hypertension and demonstrate a hypertensive, insulin resistant phenotype; suggesting multiple pathways through which AF risk could be mediated.

1.3.5 The 9q22 region

Another SNP identified by meta-analysis, rs10821415, in chromosome band 9q22, is situated in an intron of *C9orf3*, also known as *APO*, which codes for aminopeptidase O. Aminopeptidase O is a protease which cleaves angiotensin III to angiotensin IV. It is mainly expressed in the pancreas, placenta, liver, testis, and heart[133]. Angiotensin II is the main effector in the renin-angiotensin system, and regulates blood pressure and fluid and electrolyte homeostasis by its action on the AT_1 and AT_2 receptors[134]. It is cleaved to angiotensin III which retains 38% of its affinity to AT_1 and 81% of its affinity to AT_2. By contrast angiotensin IV has no affinity for AT_1 and only 1% of angiotensin II's affinity for AT_2, making the degradation of angiotensin III a key 'off-switch' in the renin-angiotensin system.

1.3.6 The 10q22 region

Another SNP identified by meta-analysis, rs10824026, is located between the genes *SYNPO2L* and *MYOZ1* in chromosome 10q22, within the region identified in the first published report of Mendelian AF[9]. Both *SYNPO2L* and *MYOZ1* encode Z-disc proteins of unknown function in cardiac and skeletal muscle[135,136]. A mouse knockout of Myoz1 did not demonstrate a significant cardiac phenotype [137], but there is some uncertainty whether Myoz1 is expressed in the mouse heart[136].

1.3.7 The 14q23 region

Another SNP identified by meta-analysis, rs1152591, is in an intron of *SYNE2* in chromosome band 14q23. *SYNE2* codes for multiple tissue specific isoforms of the protein nesprin-2, which forms complexes linking the nucleoskeleton to various structures including the actin cytoskeleton[138]. *SYNE2* mutations can cause Emery Dreifuss muscular dystrophy (EDMD) by interrupting nucleoskeleton-cytoskeleton bonds. The phenotype of EDMD in individuals with *SYNE2* mutations includes early-

onset AF[139]. Atrial tissue is subject to physical stresses and an alteration in the patterns of cytoskeletal links caused by altered *SYNE2* expression may play a role in AF susceptibility.

1.3.8 The 15q24 region

The last SNP identified by meta-analysis, rs7164883 in chromosome band 15q24, lies in the first intron of *HCN4*, which codes for an isoform of the hyperpolarisation-activated cyclic nucleotide-gated channel, also known as the pacemaker channel. This generates the I_f current in the sino-atrial node. *HCN4* mutations have been associated with conduction defects[140-142] and *Hcn4* knockout mice demonstrate bradycardia and heart block[143]. Interestingly HCN4 channels have been shown to interact with caveolin-1 and HCN4 channels with a disrupted caveolin binding domain have more positive activation potentials and slowed deactivation[144], providing a further possible explanation of mechanism of action of the *CAV1* variants.

1.3.9 The 16q22 region

An association to be identified in individual GWAS in two European populations is in chromosome band 16q22[106,107]. One of the SNPs, rs7193343, is also associated with cardioembolic ischaemic stroke[106], which is likely to be mediated by AF. The SNPs lie in intron 1 of *ZFHX3* (previously known as *ATFB1*). *ZFHX3* is the only protein-coding gene within 500kb of the GWAS hit SNPs in chromosome 16q22. *ZFHX3* codes for a transcription factor (TF) which is widely expressed and is reported in all 16 tissues covered by the Body Map 2.0 project[145]. *ZFHX3* induces the expression of platelet-derived growth factor receptor beta (*PDGFRB*)[146]. The PDGF signalling pathway has been suggested to be particularly important in atrial fibrosis, and thus potentially in arrhythmogenic atrial structural remodelling. Normal atrial fibroblasts express higher levels of PDGF pathway genes, including PDGF receptors, than ventricular fibroblasts, and atrial fibroblasts from congestive heart failure patients show enhanced expression of PDGF receptors compared with ventricular fibroblasts from such patients[147]. The mouse equivalent, *Zfhx3*, is expressed in mouse hearts and, in mice where *Pitx2*

expression in the heart is suppressed, expression levels of *Zfhx3* have been shown to be reduced.

There are two known splice variants, *ZFHX3 A* and *ZFHX3 B*. The two transcripts have different promoter regions[148]. The *ZFHX3 A* transcript differs from the *B* isoform by the addition of 420 amino acids at the N-terminus (see Figure 1). The amino acid sequence which is common to both isoforms contains an ATP-binding domain, four homeodomains and multiple zinc-finger motifs, which bind to DNA and RNA. The N-terminal portion of the sequence specific to *ZFHX3 A* contains a DEAD box-like motif and a DEAH box-like motif, these motifs are associated with RNA helicase activity and transcriptional regulation, suggesting that *ZFHX3 A* and *B* have different affinities for RNA and may interact with different downstream targets[148]. Absence of *ZFHX3* expression and mutations near the ATP binding domain are associated with an increase in malignant activity in hepatocellular, gastric, breast and prostate carcinoma [149-152] and *ZFHX3* interacts with p53 at the p21(Waf1/Cip1)promoter [153]. This suggests an important role in the regulation of the cell cycle and of cellular proliferation.

In cultured mouse HL1 cardiomyocytes, overexpression and underexpression of *Pitx2* regulates *Zfhx3* expression but overexpression of *Zfhx3* does not affect *Pitx2* levels, suggesting that *Pitx2* is an upstream regulator of *Zfhx3* expression[154].

Of the associations above, only those in 4q25, 16q22, 1q24 and 7q31 were replicated in a Japanese cohort. Alternate SNPs with stronger associations were identified in 1q21, 1q24 and 9q22. In the case of the 1q21 association the alternate Japanese SNP lies 375 kilobases from the lead European SNP[109]. No GWA study has been performed to date in African or African-American populations, where the overall incidence of AF is lower. A small study of the 4q25, 16q22 and 1q21 loci in African-Americans found that the SNPs with the strongest association differed from those found in Europeans in all three regions, suggesting that the AF association is mediated by different mechanisms in different populations[155].

1.4 Clinical implications of the genetic determinants of AF

The primary motivation behind attempts to identify genetic variation which is associated with AF and the mechanisms by which these variants operate is to improve

our ability to manage patients with AF. In this context there are two major benefits that can be derived from an increased understanding of the genetic basis of AF. One possible implication of identifying genetic determinants of AF lies in the possibility of using an individual's genotype to inform decision-making in AF management, thus improving the ability to tailor the treatment of specific individuals (personalised medicine). The other, probably more important, benefit is an increased understanding of the pathophysiological mechanisms underpinning AF. The involvement of a gene in AF development indicates that the protein product and, by extension, other proteins with which it interacts may be involved in the maintenance of normal sinus rhythm. This gives insight into how other factors, genetic and environmental, which affect these protein networks modify disease risk, and raises the possibility of deliberately altering activity in these networks, by pharmacological and other means, to aid in AF management. Insights derived from genetic studies, therefore, may result in developments which improve the diagnosis, monitoring and treatment of AF of any aetiology.

1.4.1 Personalised medicine

As has been previously stated the common variants associated with AF have modest effect sizes, with odds ratios of between 1.11 and 1.64 in the largest study[109]. An important question therefore arises; do these variants provide useful information to the clinician treating the individual patient? A large prospective study of 26 946 middle aged Swedish individuals found that genotype of the 4q25 and 16q22 SNPs predicted incidence and prevalence of AF with a risk magnitude that was similar to clinical risk factors. However, the addition of the genotyping information to the clinical risk did not result in a significant improvement in risk prediction[156]. A fine mapping study identified two further SNPs (rs17570669 and rs3853445) in 4q25 which appear to be independently associated with AF, suggesting multiple independent regulatory elements within the region. Individuals who were homozygous for the risk allele at all three SNPs were six times as likely to develop AF as those with the most common genotype. Only 1.2% of the subjects in the study had this particular combination, limiting the usefulness of the markers for disease prediction for most individuals[157].

Genotyping of common variants may be useful in specific circumstances. Genotype of the 4q25 risk SNPs predicts risk of new AF after coronary artery bypass grafting and the addition of genotype to clinical risk factors provides a small but significant improvement in risk prediction[158]. 4q25 variant rs10033464 has also been associated with an increase in both early and late AF recurrence after catheter ablation (OR 1.994, 95% CI 1.036 to 3.837, p= 0.039 and OR 4.182, 95% CI: 1.318 to 12.664 respectively)[159], and an increase in risk of AF recurrence on antiarrhythmic drug therapy (OR 3.27, 95% CI 1.77 to 6.04)[160]. These findings have been demonstrated in small studies but they have, as yet, not been replicated.

1.4.2 Biomarker development

The identification of the mechanisms by which variants affect disease risk may also allow the identification of downstream biomarkers which indicate disease risk. A biomarker is a "characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention"[161]. Biomarkers have been useful in aiding the diagnosis and monitoring of many diseases including myocardial infarction[162] and heart failure where, for example, measurement of serum brain natriuretic peptide (BNP) provides a simple blood test which can indicate left ventricular dysfunction[163].

1.4.3 Drug development

One of the primary benefits of identifying genetic variants associated with AF is that they provide significant insights into the mechanisms underlying the pathophysiology of AF. The strength of the association and the relative effect of the variant do not indicate the potential usefulness of the drug target. For example, in the case of plasma cholesterol level, common variants at the HMGCR locus in 5q13 are associated with only modest changes in LDL cholesterol but statins, which inhibit the HMGCR's product, HMG-CoA reductase, are a cornerstone of lipid lowering therapy[164]. By identifying loss- or gain-of function mutations in key ion channels, we are provided with a valuable model of the effects of pharmacological blockade or enhancement of that current. One of the most promising targets is the I_{Kur} current. It is unique to atrial

myocytes and so blocking agents are unlikely to cause ventricular arrhythmia. Vernakalant is a novel class III antiarryhthmic with a short half life with I_{Kur} blocking action. It has been approved in Europe for iv administration for cardioversion of acute AF. It is not a specific blocker of I_{Kur}, however, and concerns have been raised that there is an increased risk of ventricular arrhythmia[6].

Blockade or activation of the other currents implicated in Mendelian AF may prove to be more problematic. Loss of function in I_{Na} channels has been identified in familial AF and Brugada syndrome[165] whereas gain of function has been identified in AF and long QT syndrome type 3[166]. Clearly the maintenance of sinus rhythm requires the activity of the Na^+ channel to be finely regulated and antiarrhythmic agents would be expected to have a significant proarrhythmic potential, as is seen in the case of the commonly used class I agents[167]. Similarly, gain of function in I_{Ks} channels has been associated with AF but loss of function mutations have been identified in Long QT syndrome[168]. I_{Ks} blockade is one of the mechanisms of action of amiodarone and dronedarone, both of which prolong the Q-T interval[169,170]; more specific agents, such as HMR1556 and L-768673 have been trialled in animal models but are not yet sufficiently developed for human studies[171].

GWAS can provide insights into the potential effects of currently licensed drug treatments. *HCN4* encodes the major subunit of the ion channel responsible for the pacemaker current I_f. The association of AF with SNPs near *HCN4* may indicate the importance of I_f in AF risk. Treatment with the I_f-blocker ivabradine in the SHIFT trial was associated with a modest but significant increase in the incidence of AF, compared with controls[172]. AF is reported as a rare (<1:10 000) complication of ivabradine treatment, but AF incidence is not widely reported in the major trials of ivabradine to date[173-177]. In Chapter 3, I present a meta-analysis of the available studies which have assessed the effects of ivabradine treatment, in order to quantify any excess risk of AF.

GWAS hits also suggest possible targets for novel drug treatments. Inhibition of SK channels in isolated guinea pig and rat hearts protects against and terminates AF[178]. The $I_{K,Ca}$ current in which ZFHX3 plays an important role is more prominent in atrial myocytes, suggesting that blocking agents would be atrium-specific. The association with SNPs near the gene which codes for aminopeptidase O is also interesting. Inhibition of the renin-angiotensin-aldosterone system, either by angiotensin

converting enzyme (ACE) inhibitors or angiotensin receptor blockers, is effective as secondary prevention of AF as well as primary prevention in heart failure[179]. Further investigation of the role of APO and Angiotensin III and IV may produce further targets for drug therapy.

A less conventional target for future pharmacological agents might lie in the control of transcription in the adult atrium. PITX2, ZFHX3, GATA4, GATA6 and TBX5 are important transcription factors in regulation of cardiac development. A mouse model, in which Pitx2c expression was suppressed after embryonic day nine, had normal left-right atrial morphology but abnormal adult cardiac function[180], suggesting an ongoing role for PITX2c beyond early cardiac development. This may be true of other transcription factors; treatments which favourably affect these molecular pathways might prove to be useful in the management of AF.

1.5 Assessing the effects of risk loci on expression

1.5.1 Variation and regulation of gene expression

Phenotypic variation between individuals arises from a combination of genetic and non-genetic factors. Both genetic and non-genetic factors can operate, however, via common mechanisms. Insights derived from a greater understanding of genetic influences on disease can therefore inform our understanding of the disease in a wider sense. Alterations in the DNA sequence may result in changes to the amino acid sequence of a protein and thereby its structure ('quality') or by affecting the amount of protein produced ('quantity').

Prior to the advent of the GWAS, the majority of variants definitively associated with disease were coding variants which were either missense, and altered the structure of the protein produced, or nonsense, and resulted in haploinsufficiency. Nonsense variants result in protein only being produced from one allele, haploinsufficiency, and therefore cause a significant decrease in expression. This is the case, for example, for fibrillin-1 in Marfan syndrome[181]. This was largely due to the greater certainty with which one can ascribe pathological effects to variants which alter a protein's function, and thereby have a large phenotypic effect. Candidate gene studies, as described in Table 1.1, have identified many variants. The yield of this approach in identifying AF

risk variants has been low, however, with no pathological variant identified in the majority of subjects examined. A recent study identified that just 7.6% of alleles of candidate genes (identified by their role in cardiac physiology, rather than by GWAS or linkage analysis) in AF subjects contained a novel or very rare variant (minor allele frequency < 0.1% in the Exome Variant Server[182]), compared with 4.1% of alleles in subjects in a reference database[183]. Of course, not all individuals with apparently unprovoked AF will have a genetic substrate, but it is likely, given the low yield of the candidate gene approach and the high degree of heritability of lone AF, that a significant proportion of the genetic determinants of AF are not in the coding regions examined by this approach. The importance of non-coding regions in AF heritability has been highlighted by the results of GWAS, in which none of the most significantly associated SNPs are in protein-coding regions (Table 1.2).

Disease-causing mutations in non-coding regions have been well described for a number of disorders, including complex diseases[184-186]. The non-coding regions of the genome are coming under increasing scrutiny in order to identify functional elements. One of the largest projects to date has been the ENCODE project, which has mapped a large number of regulatory elements including transcription factor motifs, chromatin patterns, histone modifications, non-coding RNAs and DNA methylation, in order to annotate some of the functions of the 99% of the genome which is does not code for protein[187-191]. Another recent large project, the FANTOM project has created an atlas of 40 000 enhancer candidates across 800 human cell and tissue types[192]. Systematic attempts have also been made to map associations between gene expression and genotype genome-wide[193-196]. These approaches have yielded significant results in identifying the effects of variants identified by GWAS on gene expression. Not all genotype-gene expression associations are stable across all tissue types. Unsurprisingly therefore, the greatest success in identifying effects on gene expression of GWAS-identified SNPs has been achieved when the tissue in which expression is measured is relevant to the disease of interest[193]. This tissue specificity may well explain the lack of associations seen so far between AF risk SNPs and gene expression, as the most relevant tissue, namely atrium, is not easily accessible, making the collection of large cohorts of samples difficult. Additionally, the true target tissue, in which regulation of expression mediates AF risk, may not be easily identified. For example, angiotensin

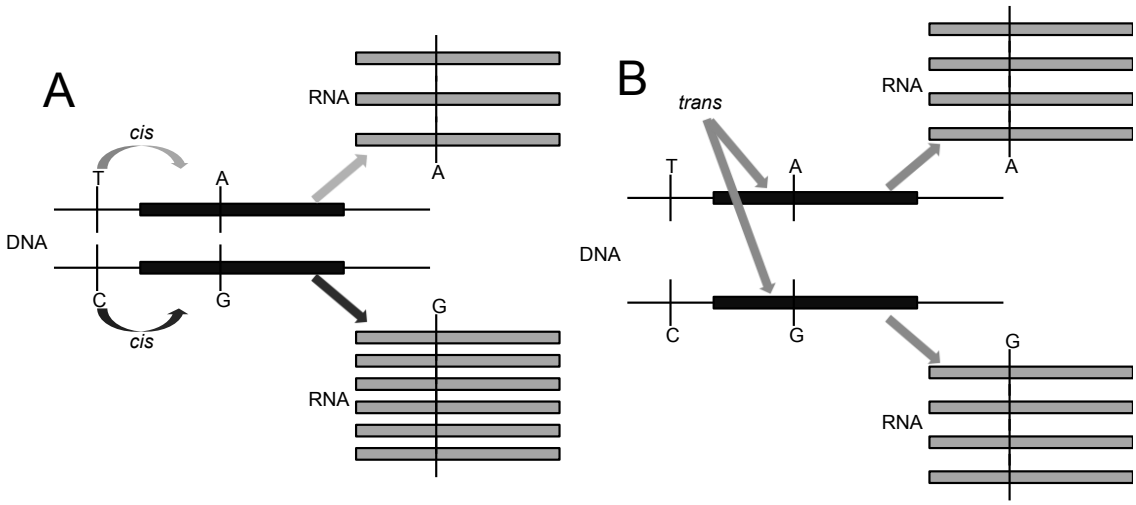
conversion by aminopeptidase O, encoded by the gene *C9orf3* could operate in any of a wide range of tissues to affect circulating angiotensin levels, and not necessarily in atrial tissue. Furthermore, the effects of any of the identified variants may be restricted to a subpopulation of cells within the atrium; cardiomyocytes, fibroblasts, endothelial cells, pacemaker cells, smooth muscle cells and epicardial cells are all present within the atria and have different gene expression profiles[197]. An alteration in the functional properties of just one cell type may result in an increase in the risk of AF.

1.5.2 Cis- *versus* trans-*acting effects on expression*

Gene expression varies considerably between individuals and variation in the expression levels of many genes has a heritable component[193,198-200]. Genetic influences on gene expression may act in *cis* or in *trans*. *Cis*-acting elements act on expression of the allele of the gene which lies on the same chromosome; *i.e.* they act in an allele-specific manner. These include regulatory elements such as promoters and enhancers and are usually close to the transcription start site of the gene which they regulate, but may lie several hundred kilobases distant[201]. *Trans*-acting factors, on the other hand, are diffusible molecules, which act on both alleles of a gene and are not necessarily transcribed near the gene which they regulate and are often located on a different chromosome altogether (Figure 1.8).

The MuTHER study investigators estimated that *cis*-acting variants are responsible for 30-36% of the heritability of gene expression, of which 60% was explained by common genetic variants (MAF>1%)[193]. This estimate may not be accurate as it is based on the assumption that all effects within 1Mb of the transcription start site operate in *cis* and all other effects operate in *trans*. This assumption does not necessarily hold. To give an extreme example; the inactivation of the X chromosome in female cells is dependent on the non-coding RNA *JPX* activating the non-coding RNA *XIST*, which lies just 92kb away, in trans[202]. *XIST*, in turn, inactivates the entire 153 million bp X chromosome in *cis*[203].

Figure 1.8 *Cis* and *trans* effects on gene expression



Depicted is the situation in which an individual is heterozygous for a *cis*-acting polymorphism with alleles T and C and also heterozygous for a transcribed polymorphism with alleles A and G. In **A**) the *cis*-acting T/C polymorphism affects expression of the allele on the same chromosome, resulting in an imbalance in the expression of RNA with A and G alleles. In **B**) *trans*-acting influences act on both chromosomes and so do not result in an imbalance in the relative amount of RNA from each allele. Adapted from Teare DM *et al.* 2011[204]

The discovery of *trans*-acting associations with gene expression is made difficult by the ability of such influences to operate independently of the location of the target. A genome-wide approach is therefore required for association studies, which necessitates large sample sizes in order to overcome the problem of correction for multiple testing. As the AF risk SNPs identified by GWAS do not lie in known transcripts, it is unlikely that they alter production of a diffusible *trans*-acting factor: they are likely, therefore, to operate in *cis* on nearby genes.

SNPs may influence gene transcript levels in *cis* via a number of mechanisms. The commonest mechanism of action is that the SNP lies within a promoter region close to the transcription start site and act to alter levels of transcription by affecting the binding of transcription factors which initiate transcription[205]. Indeed, genome-wide studies of SNP effects demonstrate that expression-associated SNPs are clustered symmetrically around the transcription start site[193,196]. Not all regulatory SNPs lie in the promoter region, however, and many exert their effect by altering distant regulatory

regions[206-208]. SNPs have also been shown to influence transcript levels by altering mRNA stability[209,210], mRNA processing efficiency[211-213], DNA methylation[214,215], and RNA splicing[216]. As described above, genetic variants in close proximity to one another may also be in LD. Associations between a variant and gene expression may, therefore, either operate directly, via one or more of the mechanisms above, or simply be associated, via LD, with another variant nearby. It is likely therefore, depending on the local LD structure, that multiple variants will be associated with gene expression, making the identification of the true functional variant within a haplotype block difficult.

1.5.3 Mapping cis-acting effects on expression

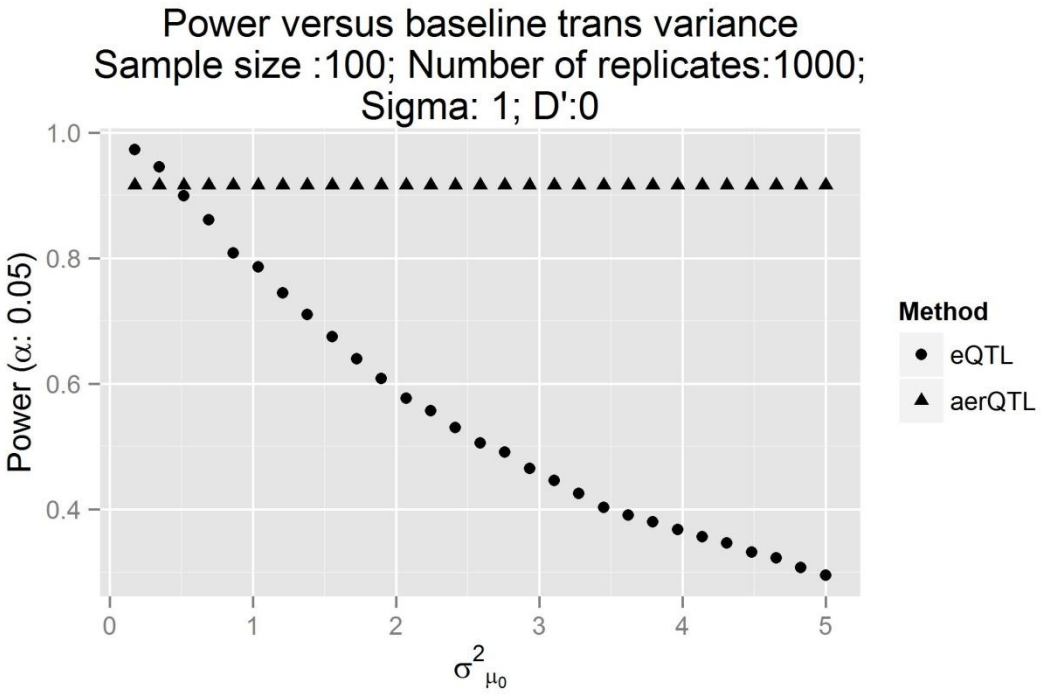
Traditionally, studies investigating the influence of polymorphisms upon gene expression have used *in vitro* methods, relying on synthetic reporter constructs transfected into cell lines. These techniques are limited by the difficulties in capturing the entire regulatory region for a gene, by the differences between the cell culture and *in vivo* environments and by the low power of *in vitro* assays to detect *in vivo* effects[217]. The most common approach to investigating expression *in-vivo* has been to compare total expression levels between individuals with different genotypes of the polymorphism under investigation. This approach treats total expression levels as a quantitative trait that is modified by the putatively active SNP (expression quantitative trait locus, or eQTL analysis)[218]. Expression can also respond to environmental, hormonal or other influences acting in *trans*. This increases inter-individual variability and may obscure the effect of cis acting polymorphisms[185].

An alternative approach to mapping *cis*-acting genetic influences on expression is to make use of the allele-specific nature of *cis*-acting effects and to measure the difference in expression between the two alleles of a gene, or more precisely, the allelic expression ratio (AER). The main benefit of using the AER to map eQTLs (aerQTL mapping) is that *trans*-acting influences would be expected to act on both copies of the gene and therefore not affect the AER.

Therefore, if there is substantial contribution to the variability in gene expression from *trans*-acting factors the traditional eQTL approach may have lower power to detect *cis*

acting effects than methods such as aerQTL analysis, which are more are specific for *cis*-acting effects. (Figure 1.9). The curves in the figure show the rapid drop-off in statistical power as the variance of the *trans* acting influences increases for the eQTL approach. This demonstrates, therefore, that when the *trans* acting influences are very similar between samples, eQTL is the more powerful approach but when there are significant differences in *trans*-acting influences between samples, aerQTL is more powerful.

Figure 1.9 The effect of *trans*-acting variation on power to detect *cis*-acting effects



The power has been simulated for 100 samples with a *cis*-acting polymorphism (minor allele frequency 0.2) that causes a 2.7 fold increase in expression and a transcribed marker with a minor allele frequency of 0.1. The mean effect of the *trans*-acting influence is 0 (i.e. there is no change in expression).

Allelic expression analysis relies on the ability to differentiate between both alleles in an individual, i.e. individuals that are not heterozygous for a transcribed polymorphism do not contribute to the analysis. Nevertheless, work on the *CDKN2A*/*CDKN2B*/*ANRIL* locus on chromosome 9p21 and genome wide approaches to eQTL and aerQTL analysis

have clearly illustrated that allelic expression analysis can achieve substantially more power than classical eQTL approach[185,219]. Since gene expression is often regulated in a tissue specific manner, the effects of polymorphisms can vary in different tissues: however, to a significant extent, the effects are consistent across a wide range of tissues, with 50-83% of effects being shared between tissue types[193,220]. It is therefore possible to gain some insight into the effects in the target tissue by studying *cis*-acting effects in readily accessible tissues, for example whole blood, provided the gene of interest is expressed. To take account of possibly significant tissue specific effects on expression, it is also important to replicate findings in the tissue of interest, namely atrial tissue, where this is available. Not all the genes of interest in the AF associated regions are expressed in blood, including the strongest candidate gene identified by GWAS, *PITX2*[105], and therefore alternative tissues are required to investigate gene expression. Atrial tissue is the ideal tissue to study, and right atrial appendage (RAA) and left atrial appendage (LAA) samples can be collected at the time of cardiac surgery with no additional risk to the tissue donor. Insights may also be gained from studying other available tissue types, where the candidate genes are expressed. One such possibility is to use arterial tissue, harvested from the internal mammary artery (IMA) at the time of cardiac bypass surgery. This also has the additional benefit that tissue samples from the IMA and RAA will often be available from the same patient.

1.5.4 Measurement of total expression and allelic expression ratio

Total expression can be measured by a number of technologies, all of which have been used for eQTL mapping. The commonest approaches involve measurement of mRNA levels using quantitative polymerase chain reaction (qPCR)[129,185], microarrays[193,221,222] or next-generation RNA sequencing (RNAseq)[195]. Microarray and RNAseq approaches are genome-wide, which means that the detection of effects is not restricted to a small number of transcripts. qPCR, however, has some key advantages. First, it is much cheaper, if only a few genes are of interest, to measure expression by qPCR than to perform either an array or RNAseq experiment. In addition, qPCR has a lower error rate than either RNAseq or microarray studies when determining differences in expression of transcripts which are expressed at low levels[223]. In order to minimise the bias in qPCR experiments, a set of guidelines has been drawn up for the design and

conduct of qPCR experiments, the *M*inimum *I*nformation for Publication of *Q*uantitative Real-Time PCR *E*xperiments (MIQE guidelines)[224], which have become the standard for the performance of qPCR. According to the recommendations set out in the MIQE guidelines, total expression is measured and normalised to the geometric mean of the total expression levels of at least three reference genes, which have been selected as having stable expression across the samples studied. What is used for comparison between groups, therefore, is the normalised total expression, which I will henceforth refer to as total expression.

Allelic expression imbalance can also be measured using a number of different approaches, including microarray analysis[219,225], qPCR[226,227], and PCR with MALDI-TOF mass spectrometry on the Sequenom platform[228,229]. The Sequenom platform allows reliable measurements of allelic expression ratio from small amounts of starting material with the additional option of performing multiple assays in a multiplex reaction

1.5.5 Statistical approaches to mapping of cis-acting variants

A standard approach to mapping eQTLs is to measure total expression and then perform a linear regression of the logarithm of the expression versus genotype of the study subjects, who will have two, one or no copies of the variant allele. This regression provides an estimate of the effect size of the allele, β, and a p value, the estimated probability of falsely rejecting the hypothesis that genotype does not affect expression.

Mapping aerQTLs requires a more complex approach. The most important difference from eQTL mapping is that the two alleles within each individual are being compared, rather than total expression levels between individuals. In the simplest case the putative *cis*-acting SNP is also the transcribed marker which is used to identify each allele of the gene. Under these circumstances it is easy to determine whether the putative *cis*-acting SNP results in allelic expression imbalance (AEI), by measuring whether the AER deviates significantly from 1:1, when measured using the transcribed SNP.

The situation is more complex when determining the effect of a SNP which is not the transcribed marker. In this case it is necessary to perform phasing of the genotypes of each individual across the studied region in order to recreate the haplotype of each of their chromosomes. It is then possible to know whether a putatively *cis*-acting variant is in phase (the minor alleles are on the same chromosome) or out of phase (the minor alleles are on opposite chromosomes) with the transcribed marker. Once phase is accounted for, the same analysis can be performed as before by measuring whether the phase-corrected AER deviates significantly from 1:1. The power of the analysis is dependent on the number of subjects who are heterozygous at both the transcribed SNP and at the *cis*-acting SNP and so the power to detect associations will always be greatest when the effect of transcribed SNP itself is assesed.

The phasing of the transcribed and *cis*-acting SNPs can be performed at the same time as the analysis or can be done in advance. However it is not always possible to determine the phase without ambiguity. Therefore rather than a single phase, the probability that each of the possible phases occurs can be computed. Often, however, only the most likely phase is used. Simulation data suggest that the power of the aerQTL mapping is diminished to a modest degree by pre-phasing and by using the only the 'best estimate' phase, but this loss of power is minimal where the two SNPs are in some degree of linkage disequilibrium (*i.e.* D'>0) and use of 'best estimate' phase significantly reduces computational requirements when many SNPs are considered[204].

1.6 Summary and overall project aims

Atrial fibrillation displays a significant degree of heritability. A large number of high-penetrance variants have been identified which can cause AF in a Mendelian fashion, but these are rare and are unlikely to contribute significantly to AF risk at a population level. GWAS have identified 9 chromosomal regions where variants are associated with AF risk. These variants do contribute to some, if not all, of the heritability of AF in the general population and knowledge of the mechanism of these associations would provide insights into the pathophysiology of AF.

It is therefore essential to investigate the AF risk variants identified by GWAS to establish their mechanisms of action. As none of the variants lie in transcribed regions, it is likely that they operate by regulating expression of nearby genes in *cis*. Genes which are regulated by AF risk SNPs provide insights into the biological processes which underlie the initiation and maintenance of AF and may identify therapeutic targets.

One particularly interesting candidate gene is *HCN4*, which encodes the major subunit responsible for the I_f current, as the I_f current blocking agent, ivabradine, is licensed for use and may influence AF risk. In Chapter 3 I present a study which aimed to identify whether inhibition of the I_f current by ivabradine was associated with an increase in AF risk and to quantify any risk identified.

The identification of effects of AF-associated SNPs on expression of candidate genes can be performed in easily accessible tissues. More than half of *cis*-acting eQTLs operate across tissue types and so effects identified in easily accessible tissue can indicate effects in disease relevant tissue. In Chapter 4 I present a study which aimed to identify associations between genotype and expression of candidate genes in peripheral blood.

Candidate genes in AF risk regions may not be expressed in peripheral blood. Additionally, the site of action of variants which are associated with AF risk is likely to be the atrium. It is therefore of interest to identify the effects of AF risk variants on gene expression in atrial and other tissues. In Chapter 5 I present a study which aimed to identify associations between expression of candidate genes in atrial and internal mammary artery tissue and AF risk SNPs.

Chapter 2

Methods

2 Methods

2.1 Materials and Samples

2.1.1 South African (SA) cohort

DNA and peripheral blood RNA samples were collected from 310 healthy adult volunteers at a blood donor clinic at the University of the Western Cape, Cape Town, South Africa. Samples were anonymised, but limited demographic data including age, gender, and self-reported ethnicity were collected. Informed consent was obtained from all participants and the study was approved by the University of Cape Town Faculty of Health Sciences Research Ethics Committee.

The self-reported ethnicity of the SA cohort was: 200 Cape mixed-ancestry; 67 African black; 19 Indian; 10 white; 4 other/unknown. 42% were male, with median age 20 years (range 17-60, lower quartile 19, upper quartile 23). These samples had been collected and largely extracted before the beginning of the project[185].

2.1.2 North East (NE) cohort

DNA and peripheral blood RNA samples were collected from 405 participants recruited on attendance for cardiac catheterisation at the Freeman Hospital, Newcastle and the James Cook University Hospital, Middlesbrough. DNA, peripheral blood and cardiac tissue RNA samples were collected from 159 participants recruited on attendance for cardiac surgery at the Freeman Hospital, Newcastle; the James Cook University Hospital, Middlesbrough; the Northern General Hospital, Sheffield; the Wythenshawe Hospital, Manchester; or Manchester Royal Infirmary. Samples were anonymised but information on previous AF status was recorded. Written informed consent was obtained from all participants and the study was approved by the NRES Committee, North East – Sunderland. Study No: 12/NE/0072.

2.1.3 Labware

All experiments were performed using standard sterile nuclease-free plasticware from the following laboratory suppliers: Thermo Fisher Scientific (USA) and Eppendorf

(Germany) and Sigma-Aldrich (USA). Barrier pipette tips (Starlab, Belgium) were used for all pre-PCR preparation steps, which were performed in designated laminar-flow hoods.

2.2 Nucleic Acid Preparation

2.2.1 DNA extraction and quantification.

For the SA cohort, peripheral blood samples for DNA analysis were collected in 5mL EDTA tubes using standard venesection from veins in the antecubital fossa. Samples were stored at -80°C until extraction and were extracted within two months. DNA was extracted using a standard phenol/chloroform methods by technicians at the Department of Molecular Biology and Human Genetics, University of Stellenbosch, Cape Town.

For the NE cohort, peripheral blood samples were collected in two 4ml EDTA tubes from the arterial sheath at the time of cardiac catheterisation or from the central venous catheter at the time of cardiac surgery. Samples were stored at -20°C until extraction and were extracted within 4 months. DNA was extracted from 150 samples using a standard phenol/chloroform method by a technician at Newcastle University (see acknowledgements). The remainder were extracted using a QIAamp DNA Blood Maxi Kit (Qiagen, Germany), according to the manufacturer's instructions by me and a technician at Newcastle University.

DNA concentrations and purity were measured by A260/A280 absorbance using a NanoDrop ND-8000 Spectrophotometer (NanoDrop Technologies, USA). Nucleic acids and proteins have absorbance maxima at 260 and 280 nm, respectively. The absorbance at 260nm can therefore be used to quantify nucleic acid concentrations and a high ratio of absorbance at 260nm *vs.* 280nm indicates low levels of protein contamination. Samples were vortexed prior to measurement and 1µl of solution was used in at least 2 replicates.

DNA was diluted to a stock concentration of 200ng/µl with TE. Further dilutions were prepared for downstream applications, as necessary.

2.2.2 DNA visualisation

After each PCR reaction carried out in Newcastle University, PCR products were visualised by gel electrophoresis to determine product size and to exclude the presence of alternative PCR products or PCR products in no-template controls. 2% w/v Agarose gel was made using 2g of SeaKem LE Agarose powder (Cambrex, USA) per 100mL of 1xTAE, containing 1ml of 1000x GelRed per ml of gel.

PCR products were mixed in a ratio of 5:1 with 6x Orange G loading dye (Sigma Aldritch, UK) and 3-5µl of mixture was loaded alongside 3-5µl 100bp DNA ladder. Electrophoresis was carried out in 1x TAE buffer for 30-60 minutes at ~100V on a 10cm gel. Gel images were then recorded using a GelDoc-It Gel imaging system (UVP, USA).

2.2.3 RNA extraction

Peripheral blood RNA samples were collected in two or four 2.5ml PAXgene blood RNA tubes in the same way as the DNA samples. After incubation at room temperature for 24 hours, they were stored at 4°C for up to 48 hours, then at -80°C until extraction. RNA was extracted from the blood using the Qiagen Paxgene blood RNA extraction kit (Qiagen, Germany) according to the manufacturer's specifications. I performed 300 extractions and the remainder were performed by a technician at Newcastle University (see acknowledgements). The DNase treatment step was omitted from the protocol.

Tissue RNA samples were stored in RNAlater (Invitrogen, USA) at 4°C from the time of collection for up to 48 hours, then at -80°C until extraction. Right atrial appendage (RAA) and left atrial appendage (LAA) samples did not require dissection prior to analysis. Internal mammary artery tissue was, for the most part, supplied as dissected for use in grafting, *i.e.* attached to a 5-10mm strip of thoracic wall, including parietal pleura and intercostal muscle. These samples required careful dissection to ensure that only arterial tissue was included in the sample used for extraction. RNA was extracted from up to 30mg of tissue using a Qiagen fibrous tissue mini kit and a TissueRupter tissue homogeniser (Qiagen, Germany) according to the manufacturer's protocol for manual extraction without the use of a QiaCUBE.

2.2.4 DNA removal

For the peripheral blood samples, contaminating DNA was removed from the eluted RNA using a DNAfree DNase kit (Ambion, USA) with a modified protocol as described below. Up to 44µl RNA was incubated with 5µl DNAfree 10x buffer, 1µl recombinant DNase I solution (2 Units/µl) and water to a total volume of 50µl for 30 minutes. A further 1µl of DNase I solution was added and incubated for 30 minutes and then a further 1µl of DNase I was added and incubated for 30 minutes. 10µl of DNase inactivation reagent were added and the mixture kept at room temperature for 2 minutes, mixing occasionally. The sample was then centrifuged at 13000g for 90 seconds before the supernatant was pipetted into a clean storage tube. For tissue samples DNase treatment was performed during extraction, as per the manufacturer's protocol. Effective DNase treatment of the RNA extracted from tissue is important to minimise the risk of bias in the aerQTL experiments. The PCR amplicons used for the determination of allelic expression were wholly contained within a single exon. This is to allow the use of genomic DNA from heterozygous individuals for normalization, but carries the risk that the results could be biased by contamination with genomic DNA. For the eQTL experiment, DNA contamination is less problematic as qPCR probe-primer sets were chosen to overlap exon boundaries and therefore not to produce PCR products from a genomic DNA template.

2.2.5 RNA quantification

RNA was quantified by A260/A280 absorbance using a NanoDrop ND-8000 Spectrophotometer (NanoDrop Technologies, USA). Samples were vortexed prior to measurement and 1µl of solution was used in at least 2 replicates. For the peripheral blood RNA samples, the elution buffer BR5 from the blood extraction kit was used as a no DNA sample for calibration. Previous testing in our laboratory has yielded reliable quantification but not purity assessment using the NanoDrop spectrophotometer, due to absorption of light of wavelength 260nm by the buffer.

2.2.6 RNA quality assessment

A selection of RNA samples were assessed by capillary electrophoresis using the Agilent 2100 bioanalyser and the RNA pico LabChip (Agilent, USA), according to the manufacturer's protocol. RNA integrity numbers (RIN) are calculated by the Agilent software on the basis of the electropherogram, RIN of greater than 7.5 were considered acceptable[230](REF, PMID 16448564). Contamination[230] with genomic DNA was assessed by performing PCR without the reverse transcription step and analysing the products on an agarose gel. Fleige and Pfaffl[231] measured the effect of RNA integrity on the results of qPCR and found that for amplicon sizes of greater than 400bp, a RIN of greater than 5 was desirable, but that performance of relative quantification assays where the PCR amplicon was small (<250bp) was independent of RNA quality. Therefore RIN greater than 5 and the use of short (<150bp) amplicons for PCR minimises the risk of bias due to RNA degradation, and so the results of the experiments performed in this study are unlikely to be biased by poor RNA quality.

2.2.7 Reverse transcription

Up to 2.5µg of RNA were reverse transcribed using a SuperScript VILO cDNA synthesis kit (Invitrogen, USA) in 20µl, according to the manufacturer's instructions. cDNA was diluted to a concentration of 25ng/µl of RNA template. cDNA was checked by PCR amplification using primers specific to the gene *GUSB*, designed to straddle an intron, and to amplify fragments of different sizes for cDNA and contaminating genomic DNA.

2.3 Assays

2.3.1 Polymerase chain reaction

PCR setup was performed in a designated pre-PCR area in a laminar flow hood using barrier pipette tips. PCR reactions were performed in 96-well plates (Thermo Fisher Scientific, UK) sealed with adhesive lids and a thermal cycler (BioRad DNA engine Tetrad 2, BioRad, USA). Except where specified, PCR was performed using Qiagen HotStar Taq DNA polymerase kit (Qiagen, Germany). The following were combined in a

PCR tube: 1µl cDNA or DNA template, 1.5µl 10x buffer, 0.9µl 25nM $MgCl_2$, 1.2µl 10nM dNTPs, 1.5µl of a 5µM solution of forward and reverse primer, 0.075µl 5U/µl HotStar Taq and water to a total volume of 15µl. Optimal annealing temperatures for each primer pair were established by performing a temperature gradient PCR with annealing temperatures of 50-70°C. Agarose gel electrophoresis of PCR products was performed for all plates to confirm successful PCR in sample wells and exclude PCR product in the no-template controls.

2.3.2 TaqMan genotyping

TaqMan genotyping was carried out in 5µl reactions according to the standard manufacturer's protocol on a 7900HT Real-Time PCR instrument (Applied Biosystems), using custom TaqMan SNP Genotyping assays (Life Technologies). The assays are listed in Appendix 1.

2.3.3 Sequenom genotyping

Sequenom experiments were carried out on my behalf by the High-Throughput Genomics Group, Wellcome Trust Centre for Human Genetics, Oxford. Multiplex SNP genotyping was performed by primer extension and MALDI-TOF mass spectrometry using iPLEX Gold SNP technology (Sequenom). SNP assays were designed using Sequenom's MassARRAY Assay Design v3.0 Software (multiplex details and primer sequences are available in Table A1.2). PCR was performed using 25ng in a 7µl reaction volume for 35 cycles using standard iPLEX methodology. Spectra were analysed using MassARRAY Typer v4.0 Software (Sequenom). Spectra and plots were manually reviewed and auto-calls were manually adjusted if required. Assay quality control was performed using 92 CEPH samples to assess genotype concordance with HapMap genotypes. Individual samples with low genotype call rates (<95%) and SNP assays with poor quality spectra/cluster plots were excluded. Correspondence to Hardy-Weinberg proportions was checked for each SNP and SNPs were excluded from analysis where the H-W p value was less than 0.001.

2.3.4 Quantification of total gene expression using real-time PCR

Quantification of total gene expression was performed by real time PCR, according to the MIQE guidelines[224]. Assays were performed using TaqMan gene expression probes and reagents (Life Technologies, USA) and run on a 7900HT Real Time PCR System (Applied Biosystems). Commercially available validated VIC- and FAM-labelled assays were used to measure expression of the primary candidate gene in each region and reference genes, (Life Technologies). Primer details are given in Table A1.1. PCR was performed in four replicates according to the manufacturer's protocol.

Relative total expression was analysed using the comparative cycle threshold (Ct) method. Ct values for the target gene were normalised to the mean Ct value of three reference genes[232]. Delta Ct values were excluded where the standard error of the technical replicates was high. Outliers were excluded using Grubb's test.

In order to be able to compare gene expression between different samples it is important that the reference genes used to normalise expression are expressed stably between samples (*i.e.* that levels of expression are similar between samples and are not influenced by outside factors); this is also true where expression is to be compared between tissues, where expression levels of reference genes may vary [233]. One measure of reference gene variability is the *M* value, defined as the average pairwise variation of a particular gene with all other control genes, which provides an indication of the degree to which expression of a reference gene differs from that of the other reference genes measured[232]. Values of *M* greater than 1.5 are considered an indication of low stability of expression. *M* values were calculated using the geNorm algorithm in the Excel software package (Microsoft, USA)[234].

2.3.5 Allelic expression analysis

Allelic expression analysis was performed following the method published by Ding and Cantor (Figure 2.1). This method can be described as follows[235]. RNA is reverse transcribed, using random hexamer primers, to cDNA. The cDNA then undergoes PCR amplification using primers designed to generate a small (~100bp) amplicon which contains the transcribed SNP. As the sequence of the two alleles of the transcribed SNP differs by only one base, amplification efficiency should be very similar for both alleles

and therefore the relative amount of each amplicon will remain the same throughout amplification. After removal of any excess dNTPs, the PCR products are then used as the basis of an extension reaction where a single mass-modified nucleotide is added to an extension primer, which binds to the PCR product immediately adjacent to the transcribed SNP. The mass modified nucleotides differ from ordinary nucleotides such that they have sufficiently different molecular weights to allow differentiation by mass spectroscopy. The relative amounts of the extended extension primers are then measured by mass spectrometry. Any method which compares the relative abundance of two alleles is vulnerable to bias due to systematic differences in amplification efficiency. In order to account for this, the results are normalised to the results from a genomic DNA template of one or more heterozygous individuals, where the starting proportion of the two alleles is known to be 1:1.

As above, Sequenom experiments were carried out on my behalf by the High-Throughput Genomics Group, Wellcome Trust Centre for Human Genetics, Oxford. Quantification of the allelic expression ratio was performed by primer extension and MALDI-TOF spectrometry using iPLEX Gold with identical parameters to the genotyping assay. Spectra were analysed using MassARRAY Typer v4.0 Software (Sequenom) and allelic expression ratios were estimated as the ratios of the area under the peak representing allele 1 to that representing allele 2. Measurements were performed in four replicates using 25ng cDNA template. Results from the amplification of genomic DNA were used as an equimolar reference to normalise the cDNA values. The allelic expression assay is prone to bias from experimental factors, such as small differences in PCR efficiency between two alleles being assayed or small differences in the efficiency of the extension reaction. Normalisation of the assay is therefore required using a mixture of both alleles in known concentrations. Genomic DNA provides a perfect template for this normalisation, as in DNA from heterozygous individuals there will be the same amount of starting material from each allele (*i.e.* one copy of each from every cell in the source material). The allelic expression assays we used therefore are not able to distinguish between cDNA and contaminating genomic DNA.

Figure 2.1 Sequenom analysis of AER

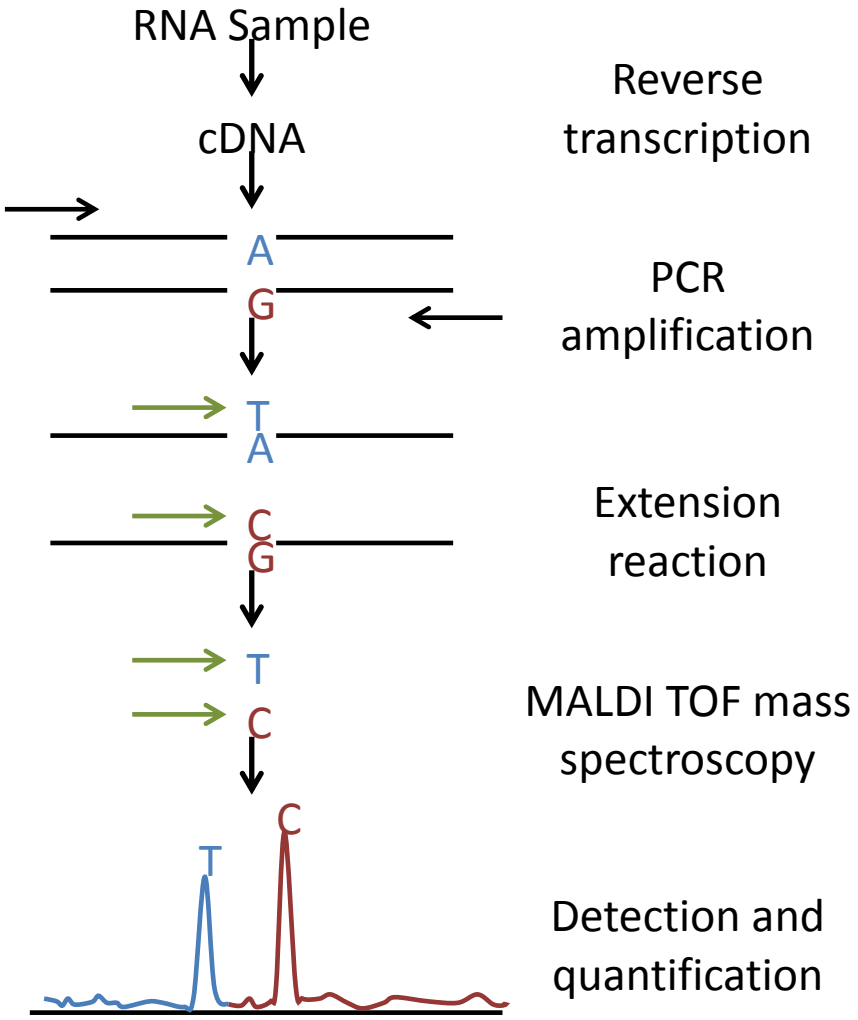
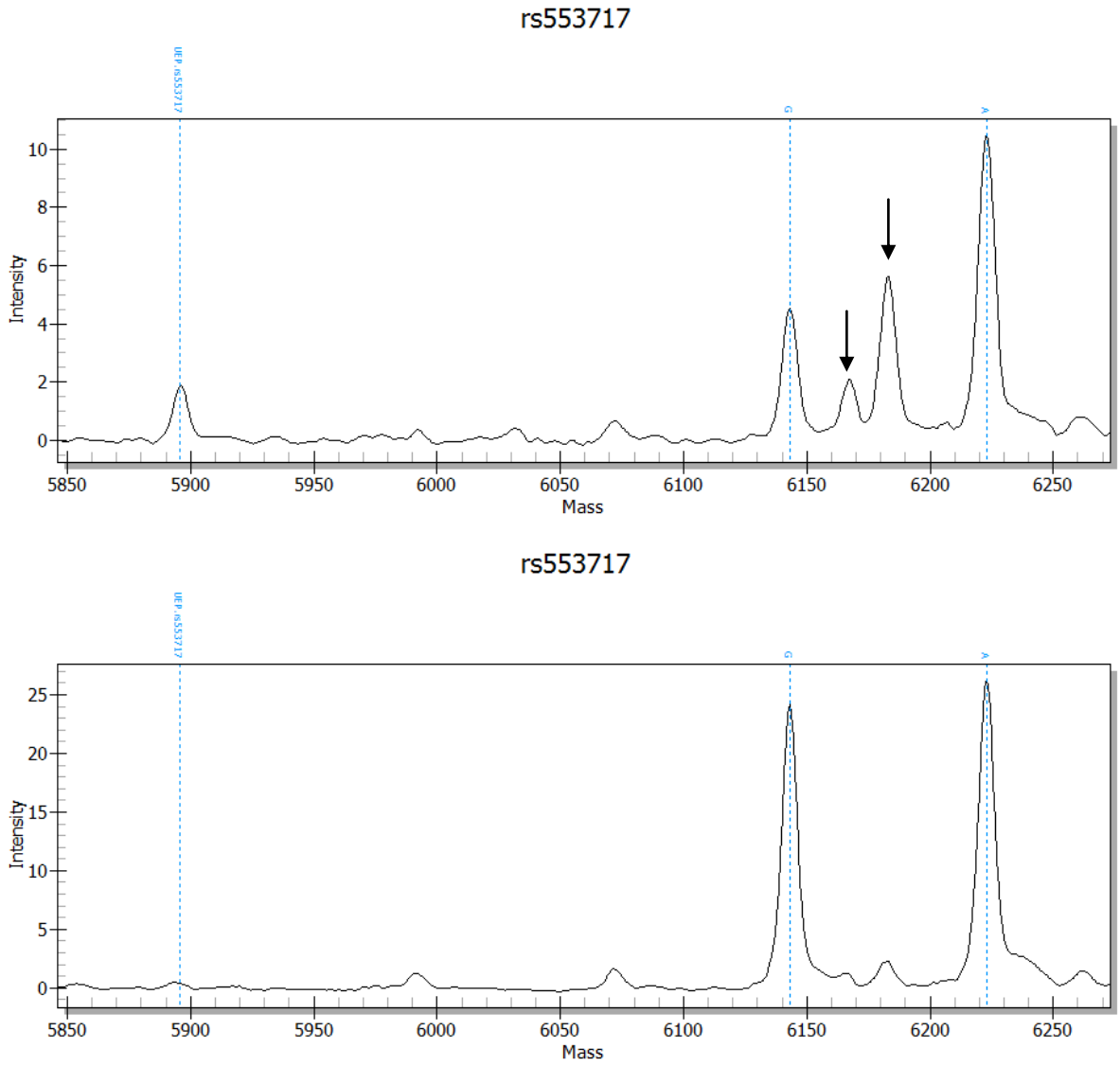


Figure adapted from Ding and Cantor[235]. A transcribed SNP with alleles A and G is amplified by PCR using primers which generate a small PCR amplicon (black arrows). The PCR products are then used as a template for the addition of a single mass modified nucleotide to an extension primer (green arrow). The extension primers are then detected by mass spectroscopy.

A variety of different methods of normalisation of allelic expression assays have been reported, including normalisation of each individual using their own genomic DNA[236], normalisation of a cohort using the mean allelic ratio from the genomic DNA assays of the entire cohort[237,238] and normalisation using a subset of the allelic ratios from genomic DNA[239]. Normalisation of each individual to their own genomic control is not superior than normalisation using the mean allelic ratio from multiple genomic controls[185], which was the method used here. Mass spectra were examined for evidence of salt peaks which could interfere with measurement of area under the peak. The extension primers undergo a Clean Resin desalting step to remove K^+ and

Na^+ ions before MALDI-TOF analysis. 6mg of Clean Resin (Sequenom), which binds salt ions, is added to each reaction well with 16µl of water and then centrifuged to bring the resins and salts to the bottom of the well. This is required because inadequate desalting can result in K^+ and Na^+ ions remaining bound to the extension primer. The subsequent mass of these ions, 39 and 23 Da, respectively, is very close to the difference in mass between extension products from different alleles and can result in errors in measurement of area under the peak and therefore of AER (Figure 2.2). AER values were excluded where the standard error of the log AER of the technical replicates was high (>0.5). Outliers were excluded using Grubb's test.

Figure 2.2 Salt peaks



Top: Mass spectrum with clear salt peaks. Additional peaks are clearly visible (arrows) with masses of 23Da and 39Da greater than the G allele peak. These represent the G allele of the

extension primer with bound Na^+ and K^+ respectively Bottom: small salt peaks are visible which do not interfere with assay measurements.

2.4 Statistical Analysis

This was conducted using established methods[185]. Briefly, total expression was treated as a quantitative trait and its association with each SNP assessed using linear regression of the log-transformed normalised expression values on the genotype assuming no dominance or interactions between the effects of different SNPs. These methods allow the assessment of the effects of single SNPs or groups of SNPs and to adjust for the effects of other sites. Multiple testing was taken into account by calculating the familywise error rate (FWER); associations with a FWER below a threshold of 0.05 were called significant. The FWER is the probability of making one or more false discovery, or type I error, among all the associations tested. Total and allelic expression data was used to estimate the proportion of total expression variance that is due to *cis* and *trans*-acting effects[185].

2.4.1 Power calculations

Using traditional eQTL analysis, a sample size of 200 would lead to over 90% power to detect a polymorphism with a minor allele frequency of 0.1 that accounts for 20% of the total variation in expression. As outlined in Teare et al (2011)[204], study power for the AER approach depends upon effect size, minor allele frequency of the studied SNPs, degree of linkage disequilibrium between the transcribed marker and putative functional SNP, and also ability to infer the phase of the SNPs in question. We have calculated that a sample size of 200 individuals and a minor allele frequency of 0.10 for the transcribed marker will have in excess of 80% power to detect the effect of a biallelic *cis* acting polymorphism with a minor allele frequency of 0.1 where there is a 20% difference in expression between the two alleles. Thus, even if there are substantial *trans*-acting effects on the variance of gene expression, limiting the power of eQTL analysis, aerQTL analysis will retain high power to detect small effects on expression in the number of subjects to be studied. In these calculations other parameters such as the variance were estimated from the results of previous studies

[185,240,241] and the power of the allelic expression study was determined using previously described methods[204]. The most common transcribed markers in *KCNN3* and *ZFHX3* that we propose to use in the AER analyses have minor allele frequencies of 0.47 and 0.41 respectively. There are two transcribed markers in PITX2c with minor allele frequencies of 0.03 and 0.1. Therefore, for all the three strongest candidate genes in the GWAS-associated regions, transcribed markers exist which will give high power to detect even small *cis* acting effects on gene expression in the number of subjects we propose to study.

2.4.2 Phasing and linkage disequilibrium assessment

Genotype phase was estimated for each cohort separately using the BEAGLE v3.3.2 Genetic Analysis Software Package [242]. LD was calculated using phased genotypes in HaploView v4.2 and haplotype blocks were identified using the confidence interval algorithm[43,243].

2.4.3 aerQTL analysis

We analysed allelic expression ratios using a previously published approach[185]. We restricted ourselves to biallelic markers and code one arbitrarily chosen allele as 0 and the other as 1. We designate with g the phase-known and with T the phase-unknown genotype of an individual. The latter can be ascertained through genotyping. We assume that the amount of mRNA originating from a single allele follows a lognormal distribution where the variance does not vary between the two alleles. The log of the ratio of the expression levels of both alleles, I can therefore be assumed to be normally distributed.

For an individual that is heterozygous for m transcribed polymorphisms, m ratios can be determined. We designate the vector of the logarithms of these ratios as

$$I = (I^1, \dots, I^m).$$

Under the assumptions above, the components of I are normally distributed with $I^k \sim N(\mu_k(g), \sigma_k)$ where the means $\mu_k(g)$ depend on the genotype g but the variance σ_k is genotype independent but may depend on the site used to measure the allelic expression ratio. We model the expected value as a linear combination of the influences of the typed polymorphisms:

$$\mu_k(g) = \sum_{i=1}^{n} \beta_i \, h_{ik}(g)$$

where β_i represents the effect of the i^{th} *cis*-acting marker; and h_{ik} characterises the phase between transcribed and putative *cis*-acting markers:

$$h_{ik}(g) = \begin{cases} 1 & \text{if the genotype at markers } i \text{ and } k \text{ is } 11/00 \\ -1 & \text{if it is } 10/01 \\ 0 & \text{otherwise} \end{cases}$$

In order to assess the association between a specific SNP and allelic expression, let us consider a set of L individuals. For an individual $l (l = 1, \dots, L)$ we can measure the unphased genotype T_l and a vector representing the log of the allelic expression ratios I_l. Up to a multiplicative constant the likelihood of observing certain pattern of imbalance in this set of individuals given their genotyping results is:

$$L = \prod_{l} f(I_l | T_l)$$

with

$$f(I_l | T_l) = \sum_{g} f(I_l | g)$$

where g describes the most likely phased genotype given the genotyping results T_l and $f(I_l|g)$ describes the density of the distribution of I_l given the genotype g. g was estimated using the BEAGLE Genetic Analysis Software Package

We assume the allelic expression ratios measured at different sites are conditionally independent given the genotype. Therefore:

$$f(I_l|g) = \prod_k f\left(I_l^k|g\right)$$

where $f\left(I_l^k|g\right) = f\left(I_l^k; \mu_k(g), \sigma_k\right)$ and $f\left(I_l^k; \mu_k(g), \sigma_k\right)$ denotes the density of a normal distribution with the individual expression ratio I_l^k as variate, a genotype dependent mean $\mu_k(g)$ and a variance σ_k^2. Therefore L depends on $\beta_i(i = 1, \dots, n)$ and $\sigma_k(k = 1, \dots, m)$, and maximisation of this likelihood allows assessment of the effects of single SNPs or groups of SNPs and to adjust for the effects of other markers by comparing nested models using likelihood ratio tests.

The effect sizes are presented as β in all figures (i.e. the log effect size). The fold-change can therefore be determined by e^{β}

2.4.4 eQTL analysis

The association between total expression and SNP genotype was assessed using linear regression of the log transformed normalised expression values on the genotype, assuming no dominance or interactions between the effects of different SNPs. PCR plate was included as a categorical variable.

In order to allow comparison with the results of aerQTL analysis, the effect sizes are presented as β in all figures (i.e. the log effect size). The fold-change can therefore be determined by e^{β}

2.4.5 Correction for multiple testing

For both total and allelic expression multiple testing was taken into account by calculating a significance threshold as 0.05 divided by the number of tests performed[244].

2.4.6 Estimation of proportion of variation due to cis acting effects

From our allelic and total expression data we also estimated the proportion of the total expression variance that is due to *cis*-acting effects. This assumes that *cis*- and *trans*-acting factors act in an additive manner, do not interact, are independent, and

that there is random mating, no segregation distortion and the locus is not subject to imprinting. Given these assumptions, we estimate the variance due to *cis*-acting effects $V(c)$, as

$$\hat{V}(c) = \frac{1}{2(n-1)} \sum_{i=1}^{n} e_{Ti}^2 \left(\frac{1-r_i}{1+r_i} \right)^2$$

where r_i is the allelic expression ratio for individual i, and the proportion of the total variation due to *cis*-acting effects can be estimated as $\frac{2\hat{V}(c)}{\hat{V}(e_T)}$, where $V(e_T)$ is the estimated total variance, *i.e.*

$$\hat{V}(e_T) = \frac{1}{n-1} \sum_{i=1}^{n} (e_{Ti} - \overline{e}_T)^2$$

with

$$\overline{e}_T = \frac{1}{n} \sum_{i=1}^{n} e_{Ti}$$

and e_{Ti} represents the total expression level for individual $i185$.

Statistical analyses were performed in the R software package version 2.15.1 “roasted marshmallows”[245] using custom scripts.

Chapter 3

Atrial fibrillation associated with ivabradine treatment: meta-analysis of randomised controlled trials

3 Atrial fibrillation associated with ivabradine treatment: meta-analysis of randomised controlled trials

3.1 Abstract

A recent meta-analysis of genome-wide association studies (GWAS) of atrial fibrillation (AF) identified associations near the gene *HCN4*. In the atria the protein HCN4 is the predominant subunit of the ion channel responsible for the I_f pacemaker current. The I_f channel-blocker ivabradine is licensed for use to limit the heart rate in patients with angina pectoris and chronic heart failure. The sick sinus syndrome, which is characterised by sinus bradycardia, is commonly associated with AF, and AF in these circumstances may represent an escape rhythm from profound bradycardia or sinus arrest. AF is reported as a very rare (less than 1 in 10,000) side effect of ivabradine treatment but the magnitude of the risk of AF is unknown. This study quantified the risk of AF associated with ivabradine treatment for any indication by meta-analysis of clinical trial data.

Medline, Embase, Web of Knowledge and the Cochrane central register of controlled trials were searched for double-blinded randomised controlled trials of ivabradine with a follow-up period of at least four weeks. For studies where AF data were unpublished AF incidence was sought from the European Medicines Agency (EMeA) website and by personal communication. The main study outcome was defined as incident AF during the trial follow-up period. Studies were appraised for risk of bias using components recommended by the Cochrane Collaboration. Meta-analyses were performed of relative risk of AF and absolute risk difference of AF per year of treatment.

AF incidence was reported in the published manuscript for one of ten studies identified, AF incidence was reported in the safety data on the EMeA website for one study and AF incidence was provided by personal communication for four studies. We also included in our analysis safety data from the 'Overall Oral Safety Set' (OOSS), a combination of clinical trials with aggregated safety data reported by the EMeA.

Ivabradine treatment was associated with a relative risk increase of AF of 15% compared with controls (95% CI 7%-24%, p=0.0027). From this we estimated that 208 patient years of treatment (95% CI 122-667) would be required to cause one new case of AF.

By demonstrating that I_f blockade increases the risk of incident AF, this study confirms the role of HCN4 in the pathogenesis of AF identified by GWAS. In particular, this result indicates that the risk variants are likely to act by reducing the activity of the I_f current. In addition, this study has shown that AF is a substantially more common side effect of ivabradine treatment than 1 in 10,000, the risk reported in the product literature. The incidence of AF has not routinely been reported in clinical trials of ivabradine. The risk of AF needs to be taken into consideration when weighing the balance of risk and benefits of ivabradine treatment.

3.2 Introduction

Genome-wide association studies have identified associations between genetic variants in the region of the gene *HCN4* and both heart rate and atrial fibrillation (AF).[28,109] HCN4 is the predominant ion channel subunit in the hyperpolarisation-activated, cation non-selective (HCN) family in human atria[246]. HCN ion channels are responsible for the I_f current, which leads to the spontaneous depolarisation of cardiac nodal tissue and therefore regulates the heart rate[246]. The effect of these variants on the I_f current is unknown. Bradycardia and AF are both components of sinus node dysfunction and permanent pacing prevents AF in patients with bradycardia due to sinus node dysfunction. In addition, the pulmonary venous myocardium, which is an important source of AF initiation and maintenance[4], demonstrates an I_f current[247].

AF risk variants may operate in one of two ways. The AF risk and heart rate associated variants lie on either side of exon one of *HCN4* and are in perfect LD in the 1000 genomes CEU population[248]. It is possible, therefore that they are in LD with a coding variant which alters I_f activity. There are six missense, five synonymous and one frameshift variant reported in *HCN4* exon one in build 138 of dbSNP[39], of which one missense variant, rs143090627, and one synonymous variant, rs201193660, have a minor allele frequency (MAF)>1%. The functional effects of these polymorphisms are

not known. MutationTaster[249] identifies rs201193360 as a likely disease-causing variant, due to the presence of regulatory features, including a DNase hypersensitive site and methylation sites at this position. The alternative hypothesis is that the AF-associated variants regulate expression of *HCN4*. HCN4 forms heteromultimers with HCN2, which is also expressed in human atria[246]. Co-expression of HCN4 and HCN2 results in an I_f current with faster gating kinetics and activation at more depolarised potentials than when either isoform is expressed alone. In addition, the kinetics are faster when HCN4 and HCN2 are expressed in the ratio, 10:1, which is close to the ratio of expression in human atria, than at other ratios[250]. This raises the possibility that either increased or decreased expression of HCN4 will have the same effect on the I_f current, namely slower gating kinetics resulting in reduced automaticity and bradycardia. Sick sinus syndrome (SSS) is characterised by sinus bradycardia but is also associated with atrial fibrillation in the tachycardia-bradycardia syndrome. The longer time period between sinus beats in sinus bradycardia facilitates the conditions for developing AF by causing an increase in atrial ectopy and an increase in the dispersion in atrial refractoriness (i.e. the degree of heterogeneity in the length of the refractory period of the atrial myocytes), which are the main initiators of AF[251].

Ivabradine (Procoralan®, Servier) is a heart-rate-lowering drug which acts by specifically inhibiting the pacemaker I_f current[252]. Ivabradine was approved for use by the European Medicines Agency (EMeA) in 2005 for use in the treatment of stable angina pectoris in patients with normal sinus rhythm who are not able to tolerate beta-blocker therapy. In 2010 the indication was extended to include treatment in patients with uncontrolled angina symptoms and a heart rate in excess of 60 beats per minute (bpm) despite beta-blocker therapy, following the results of the results of the BEAUTIFUL trial.[174] A new indication was approved in 2012, following the results of the SHIFT trial, for the treatment of chronic heart failure (New York Heart Association class II to IV) with systolic dysfunction, in patients in sinus rhythm and whose heart rate is greater than 75 bpm. To be used in combination with standard therapy, including beta-blocker therapy, or when beta-blockers are contraindicated or not tolerated.[172] Use of ivabradine is not yet approved by the US Food and Drug Administration.

Atrial fibrillation is common in patients with coronary artery disease and cardiac failure[6] and so incident AF in the target population for ivabradine treatment would not

necessarily be attributed to the drug. Atrial fibrillation is reported as a very rare (less than 1 in 10 000) side-effect of ivabradine treatment in the British National Formulary and the product literature.[173] The incidence of AF in patients randomised to ivabradine treatment *vs*. placebo was significantly increased in the SHIFT trial however [risk difference (RD) 1.76%, 95% confidence interval (CI) 0.41%-3.31%, p=0.03], suggesting that the risk may be higher than previously thought.

3.3 Aims

1. To determine whether treatment with ivabradine, which inhibits the HCN4-mediated I_f current, was associated with an increase in the risk of AF.

2. To quantify that risk, both in terms of relative risk increase and number needed to harm per year.

3.4 Methods

The study was performed in accordance with the recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Table 3.1).[253]

3.4.1 Eligibility criteria, information sources and search protocol

The following studies were included in the analysis: randomised controlled double-blind trials of ivabradine, for any indication and of any size, which followed patients up for at least 4 weeks. Studies were not necessarily excluded if the AF incidence was not reported in the published manuscript, but instead attempts were made to obtain the AF data by contacting the study authors and by searching the European Medicines Agency (EMeA) website. Both placebo-controlled studies and non-inferiority studies where an alternative drug treatment was used as a comparator were included. The following were excluded: open or single-blinded studies, studies where the follow-up period was less than 4 weeks and studies where there was significant risk of bias, as assessed by the methods recommended by the Cochrane collaboration. A systematic search was performed, without language restriction, for randomized clinical trials of ivabradine treatment for any indication, using the search term 'Ivabradine' and the

document type ‘clinical trial’ or ‘randomized controlled trial’. Medline, Embase, the Web of Knowledge from 1980 to October 2013 and the Cochrane Central Register of Controlled Trials were searched. The reference lists of published trials, review articles and meta-analyses were also examined to identify other eligible trials. The scientific discussions of the EMeA were read to identify unpublished studies and to identify the incidence of AF in published studies where this information was not included in the published report. Contact was made directly with the authors and, in the case of industry-sponsored trials, the sponsors, of studies which did not report the AF incidence in the original manuscript.

Table 3.1 PRISMA checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	64
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	65
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	66
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	66
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	68
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	68
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	69
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	69
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if	73

		applicable, included in the meta-analysis).	
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	73
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	73
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	73
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	74
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	74
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	74
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	74
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	74
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	74
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	77

Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	80
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	79-80
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	77
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	80
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	82
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	84
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	83
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	ii

3.4.2 Selection and quality assessment

Studies were assessed on the basis of their title or abstract and those studies which appeared to meet the eligibility criteria were selected for full text review. Trials were assessed for eligibility and risk of bias using the components recommended by the Cochrane Collaboration. These include: sequence generation, *i.e.* the method by which subjects were assigned to study groups; allocation concealment, whether patients were aware of the order in which participants would be assigned to each arm of the study; blinding of participants, personnel and outcome assessors after allocation; incomplete outcome data; selective outcome reporting, such as the reporting of only AF that resulted in hospitalisation, rather than all new cases; and other sources of bias such as the exclusion of patients at high risk of AF [254]. Any trial with a high or unclear risk of bias, by any of the Cochrane criteria was considered to be at high risk of bias; all other studies were considered to be at low risk of bias. Non-inclusion of AF incidence in the study report could be considered to be an example of selective outcome reporting. However where complete AF incidence data were available from another source, namely by personal communication or in the EMeA documentation, non-reporting of AF incidence alone was not considered to be evidence of a high risk of bias.

3.4.3 Outcome measures and data collected

The main outcome measure was incident AF reported during the trial follow-up period. Where data on AF incidence were not reported, the authors were contacted by email on at least two occasions. AF incidence reported in the EMeA scientific discussions where trials were submitted as evidence for licensing was identified and, in the case of trials sponsored by the manufacturer, contact was made with the manufacturer directly. Study data was collected on data collection forms which recorded reference data, ethical approval, randomisation, blinding, control agent (placebo or other drug), follow-up duration, inclusion and exclusion criteria and sponsorship and funding information as well as numbers of patients in the ivabradine and control arms with incident AF.

3.4.4 Statistical analyses

Meta-analysis of intention to treat outcomes was performed using the metafor library in R statistical software package version 3.0.2.[245,255] A two-step approach[256] was used by estimating heterogeneity (τ^2) using the empirical Bayes estimator[257] and then fitting a generalised linear mixed-effects model. The confidence interval, and therefore p-value, for τ^2 was calculated using methods described by Veichtbauer.[258] We modelled the effect of ivabradine as a random effect as this best accounts for both within-study and between study variance.[259] The Knapp and Hartung adjustment was used to account for the uncertainty of τ^2. [260] When estimating absolute risk difference we included length of follow-up as a factor in the model, assuming a risk difference of zero at zero days of follow-up, as the absolute risk of developing AF can reasonably be expected to be related to length of exposure. This allows an estimation of the absolute risk difference (or number needed to harm) per year of treatment. Heterogeneity was assessed using the I^2 statistic, considering I^2>50% to be significant[261] as well as reporting the p-value for unexplained heterogeneity. Publication bias was estimated using the rank test,[262] the regression test[263] and by visual analysis of funnel plots. In order to determine the effect of treatment indication a meta-analysis of relative risk, including treatment indication as a categorical moderator variable was also performed.

3.5 Results

3.5.1 Study selection and characteristics of included trials

The initial search strategy identified 84 published articles. Titles and abstracts were reviewed and papers that were not clinical trials, were not of sufficient duration of follow-up, or which reported supplementary data from trials published elsewhere were excluded. The full texts of the remaining 34 articles were assessed for eligibility. Eight papers were translated into English with the assistance of Dr Pogoryelova, see acknowledgements. After studies were excluded on the basis of study design (1 study), follow-up duration (2), randomisation (9), study blinding (10), and duplication (2) our search strategy yielded 10 double-blind randomised controlled trials that enrolled 20,022 patients for a mean of 1.5 (range 0.076-1.88) years follow-up, for a total of

30,090 patient years of follow-up (Figure 3.1, Table 3.2).[172,174-177,264-269] Risk of bias of the included studies is summarised in Table 3.3.

Figure 3.1 Flow chart of study selection.

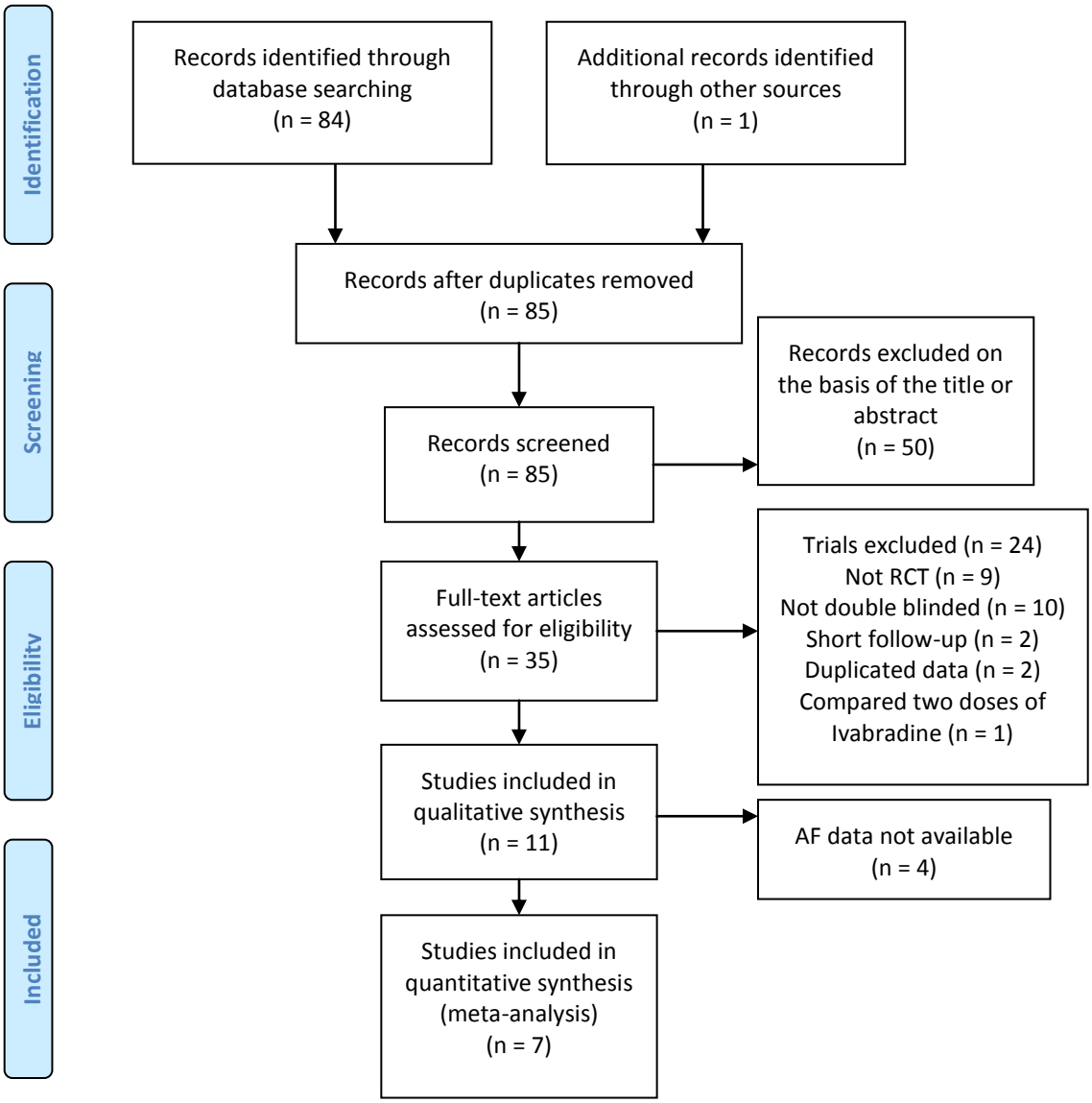


Table 3.2. Eligible trials for which AF data was sought.

Author	Year	Acronym	Number of participants	Indication	Control	Ivabradine dose (bd)	AF data source
Tardif et al.	2005	INITIATIVE	939	Angina	Atenolol	7.5mg/10mg	Included in OOSS
Ruzyllo et al.	2007		1195	Angina	Amlodipine	7.5mg/10mg	Included in OOSS
Fox et al.	2008	BEAUTIFUL	10 907	Heart Failure	Placebo	5mg/7.5mg	EMeA
Tardif et al.	2009	ASSOCIATE	875	Angina	Placebo	5mg/7.5mg	No data
Fasullo et al.	2009		153	Anterior STEMI	Metoprolol	5mg/7.5mg	No data
Swedberg et al.	2010	SHIFT	6492	Heart Failure	Placebo	2.5mg-7.5mg	Original paper
Nerla et al.	2012		61	Type II diabetes	Atenolol/ placebo	5mg	Personal communication
Dominguez-Rodriguez et al.	2012		27	NSTE-ACS	Placebo	5mg	Personal communication
Cappato et al.	2012		21	IST	Placebo	5mg	Personal communication
Villano et al.	2013		46	Microvascular angina	Ranolazine/ placebo	5mg	Personal communication
EMeA Overall Oral Safety Set	2005	OOSS	3936		Atenolol/ amlodipine /placebo	5mg-10mg	EMeA

The AF data for the INITIATIVE study and from Ruzyllo et al. were not available separately, but were included in the overall oral safety set from the EMeA with three other studies. The number of participants includes only those with complete follow-up data. EMeA –European Medicines Agency, STEMI – ST elevation myocardial infarction, NSTE-ACS – non-ST elevation acute coronary syndrome, IST – inappropriate sinus tachycardia, bd – twice daily.

Table 3.3 Risk of bias

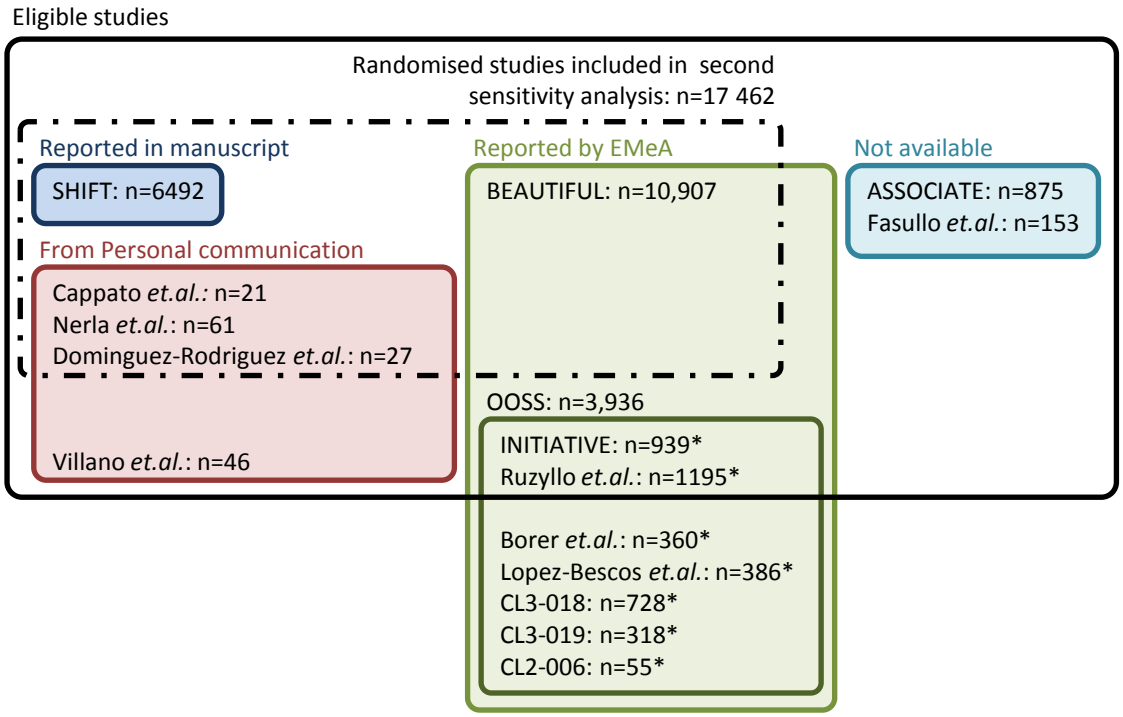
Study	Sequence generation	Allocation concealment and blinding	Outcome data included in final analysis	Selective outcome reporting	Other sources of bias	Overall risk of bias
Tardif et al. 2005 (INITIATIVE)	Randomised	Double-blinded placebo controlled	94%	No AF data reported	None	Low
Ruzyllo et al. 2007	Randomised	Double-blinded, allocation concealed	97%	No AF data reported	None	Low
Fox et al. 2008 (BEAUTIFUL)	Randomised	Double-blinded placebo controlled	100%	No AF data reported	None	Low
Tardif et al. 2009 (ASSOCIATE)	Randomised	Double-blinded placebo controlled	100%	No AF data reported	None	Low
Fasullo et al. 2009	Randomised	Double-blinded, allocation concealed	99%	No AF data reported	None	Low
Swedberg et al. 2010 (SHIFT)	Randomised	Double-blinded placebo controlled	99%	AF incidence reported	None	Low
Nerla et al. 2012	Randomised	Double-blinded allocation concealed	92%	No AF data reported	None	Low
Dominguez-Rodriguez et al. 2012	Randomised	Double-blinded placebo controlled	100%	No AF data reported	None	Low
Cappato et al. 2013	Randomised	Double-blinded placebo controlled	90%	No AF data reported	None	Low
Villano et al.	Randomised	Double-blinded, allocation concealed	100%	No AF data reported	None	Low
EMeA Overall Oral Safety Set (OOSS)	All randomised studies	Double blinded, allocation concealed	unknown	AF data reported	Multiple studies considered together	Low

The inclusion of multiple studies in the OOSS would be more likely to increase imprecision rather than introduce bias. To confirm this, a sensitivity analysis was performed with exclusion of the OOSS.

Of the 10 studies identified, AF incidence was only reported in one, the SHIFT trial. The authors of the other studies were contacted by email to request the unpublished AF data. Of the five studies not sponsored by the manufacturer, four authors replied with AF data. None of the authors of studies sponsored by the manufacturer provided AF data. We contacted the manufacturer directly, who declined to provide the unpublished data, citing ongoing in-house analyses. The original scientific discussion[270] for ivabradine, published by the EMeA prior to marketing authorisation in 2005, did not report safety data, including AF incidence, for each of the submitted trials separately. It did report AF incidence in a combined data set which included the INITIATIVE study and the study by Ruzyllo *et.al.* in addition to two phase-three clinical trials (for which we were unable to identify a published report), a 1 year dose-comparison study[271] and two very short (four doses and two weeks, respectively) dose-ranging phase-two studies. All of the studies included in this overall oral safety set (OOSS) were randomised, controlled and double-blinded. None of the trials for which we obtained AF incidence data separately were included in the OOSS. A subsequent update to the scientific discussion, published in 2010, considered the evidence from the BEAUTIFUL and ASSOCIATE trials. AF incidence was reported for the BEAUTIFUL study, but not the ASSOCIATE study. In order to provide the most accurate estimate of the effect of AF the OOSS was included as a single study in the meta-analysis and two sensitivity analyses were performed. In the first, the OOSS was excluded; and in the second, trials where treatment was not compared with placebo, but instead with an alternative agent, such as amlodipine or a beta blocker, were excluded.

In the full data set for analysis including the OOSS, therefore, there were 21 490 patients, followed up for a mean of 1.43 years to provide a total of 30 755 patient years of follow up. When the OOSS was excluded the data set included 17 508 patients, followed up for a mean of 1.67 years to give a total of 29 385 patient years of follow up. In the placebo-controlled trials only data set there were 17 462 patients followed up for a mean of 1.68 years to give a total of 29 380 patient years of follow-up (Figure 3.2)

Figure 3.2 Sources of AF incidence

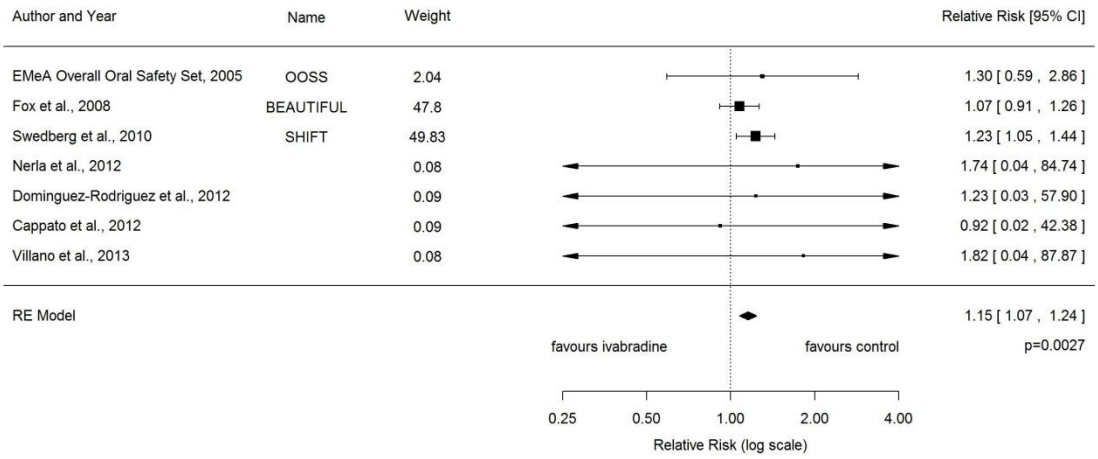


Numbers of individuals with complete follow-up data from each study are given. For the individual studies within the OOSS, the numbers of patients which were included in the data set from each study were not reported. Total numbers of patients enrolled are given (*). There is no overlap between the OOSS and studies where AF incidence is identified from other sources

3.5.2 Incidence of atrial fibrillation

Compared with controls, treatment with ivabradine was associated with a relative risk (RR) of AF of 1.15 times that of alternative treatments [95%CI=1.07-1.24, p=0.0027] (Figure 3.3). The results were largely similar if the OOSS was excluded, RR=1.15 [95%CI=1.06-1.25, p=0.0065] and if only placebo controlled trials were examined, RR=1.15 [95%CI=1.05-1.26, p=0.015]. When treatment indication was included in the model, the association between treatment indication and AF risk was not significant (p=0.99), although this analysis was limited as the majority of patient years of follow-up (17 462 /17 508) were in the two trials (SHIFT and BEAUTIFUL) for which the main treatment indication was heart failure.

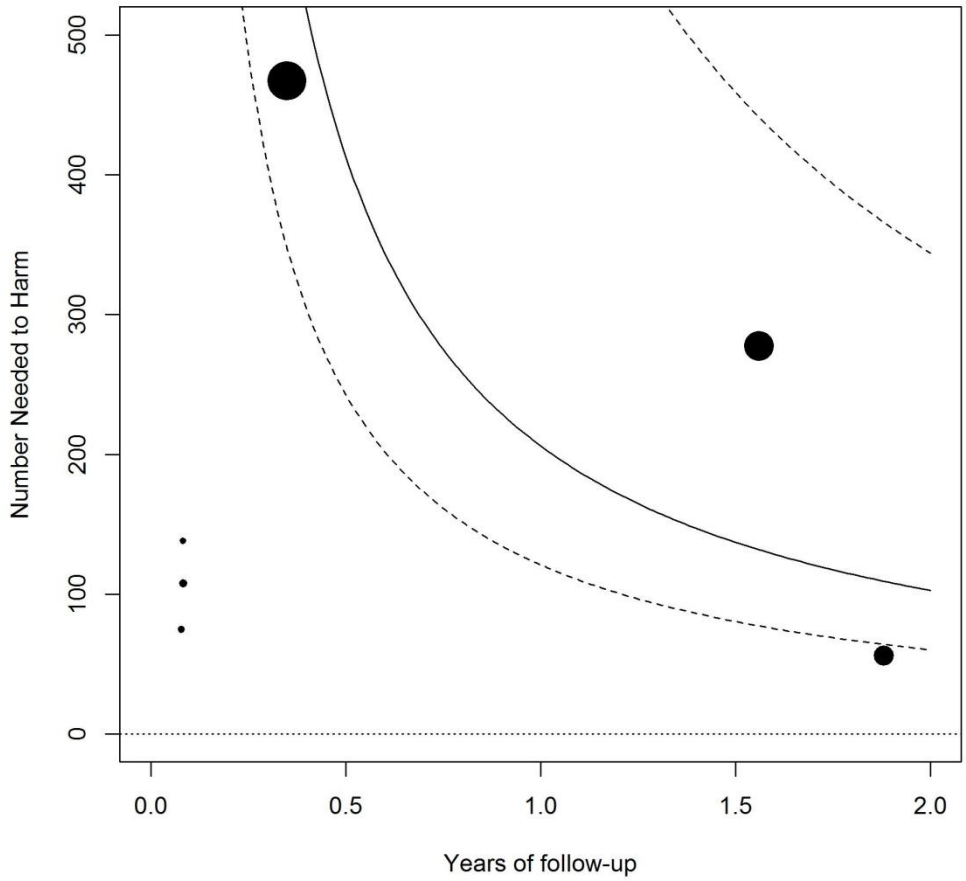
Figure 3.3 Forest plot of relative risk of AF in all trials with available data.



RE Model – random effects model.

When the absolute risk of AF was compared between ivabradine and control groups, ivabradine treatment was associated with a number needed to harm (NNH) of 208.3 per year of treatment [95%CI=122.0-666.7, p=0.013]. This effect was very similar when the OOSS was excluded, NNH per year was 208.3 [95%CI=113.6-1250, p=0.028], and when only placebo-controlled trials were included, NNH per year was 208.3 [95%CI=105.3-10 000, p=0.048] (Figure 3.4). As ivabradine is indicated for chronic treatment of angina or cardiac failure and treatment is likely to continue for more than one year, this NNH represents a 50-fold higher risk than that reported in the product literature, which classifies AF as a “very rare” (i.e. less than 1:10 000) complication; if one year of treatment is considered.

Figure 3.4 Number needed to harm per year of follow-up.

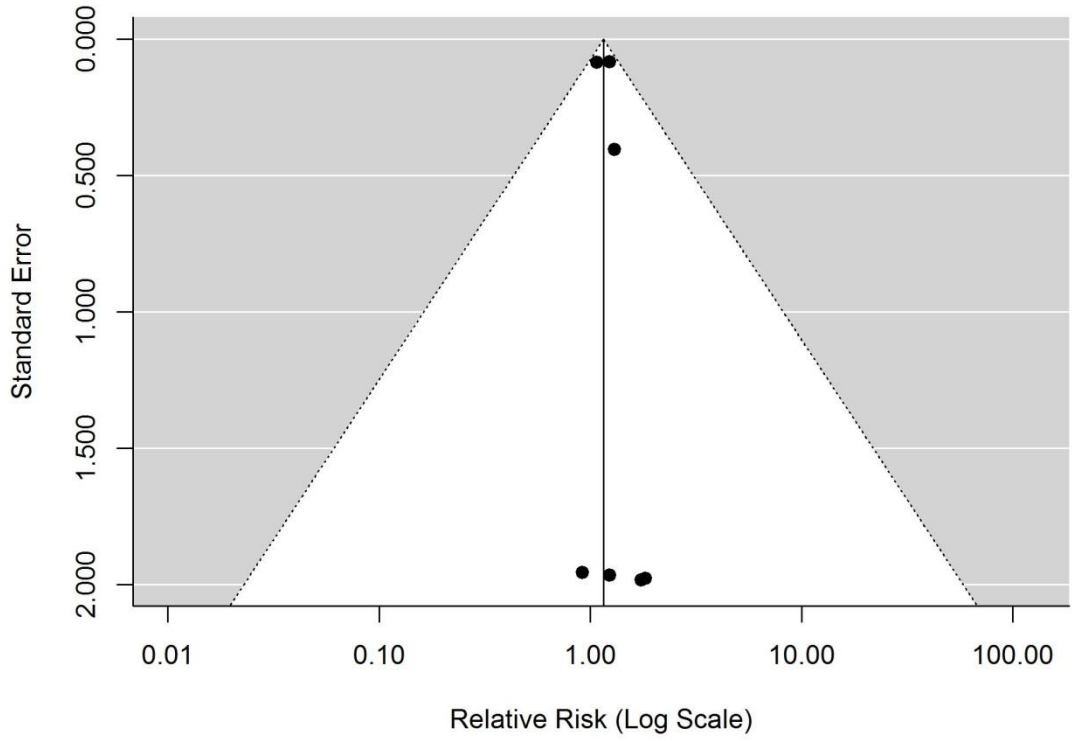


Individual studies are plotted as circles with size proportional to the weighting in the meta-analysis model. The best fit line (continuous line) with 95% CI (dashed lines) is shown.

The heterogeneity of relative risk in the trials included was low, I^2=0.00%, with the observed heterogeneity explained by within-study variance (p=0.96). The same was true in the analysis of absolute risk difference, I^2=0.00%, where the residual heterogeneity after accounting for length of follow-up was explained by within-study variance (p=0.86). This explains the similarity of the meta-analysis results between sub-groups of trials and suggests that the effect of ivabradine on AF risk is independent of treatment indication.

There was no evidence of publication bias, rank test p=1.00, regression test p=0.64, and examination of the funnel plot did not suggest any missing data (Figure 3.5).

Figure 3.5 Funnel plot of relative risk vs. standard error.



The plot is largely symmetrical, suggesting that there is no publication bias.

3.6 Discussion

The results of this study strongly indicate the importance of the I_f current in the pathophysiology of AF. These data also, using as much of the data as was available, quantify the risk of AF associated with ivabradine treatment. Use of ivabradine causes an approximately 15% increase in the risk of developing AF. This risk is greater than previously reported in the product safety information.

The associations between genetic markers near *HCN4*, which is responsible for the I_f current, and both AF and heart rate mirror the effects of the I_f channel blocker ivabradine. Although they have different aims, design and methods of analysis, the results of GWAS, in this case, indicate similar findings to the results of a Mendelian randomisation experiment. Genetic variation is inherited randomly and is present before any disease develops and therefore must be causative of any association with

disease. Genetic variants can therefore be used to identify whether a phenomenon, in this case reduced I_f current activity, is causative of, rather than simply associated with, a disease state[272]. If the genetic variants identified by GWAS can be demonstrated to result in reduced I_f activity, then this indicates that reduced I_f activity causes AF, rather than the other way around. The effect of AF risk variants on expression of *HCN4* and any subsequent effects on the I_f current are not known. Further investigation of the effects of the risk SNPs on expression of *HCN4* and whether any transcribed polymorphisms in strong LD with the AF risk SNPs alter the kinetics of the I_f current would therefore be of interest.

AF is a common complication of coronary artery disease and cardiac failure. Patients with these conditions have a higher rate of complications from AF than patients with AF alone[273]. These risks however need to be balanced against the benefits of ivabradine. In the SHIFT study, ivabradine treatment resulted in a reduction in the composite endpoint of death or hospitalisation with worsening heart failure; the number needed to treat (NNT) was 26, which was largely driven by a reduction in hospitalisations.[172] A combined analysis of the individual data from the SHIFT trial and those patients in the BEAUTIFUL trial with a baseline heart rate ≥70 bpm also identified a reduction in the same endpoint, the NNT was 40.7.[274] Interestingly, in the same analysis, which looked at patients with a higher baseline heart rate, the increase in AF incidence was greater than we found in our meta-analysis, the NNH was 58 (albeit over a longer follow-up of 19-22 months). This raises the possibility that it is the patient group with the most to gain from ivabradine treatment that have the greatest risk of developing AF and that the absolute increase in the risk of AF with ivabradine treatment in this group is of a similar size to the absolute decrease in risk of hospitalisation.

The results of this analysis also highlight the importance of large safety data sets and ongoing post-marketing surveillance of licensed and approved medications to identify side-effects which might otherwise be dismissed as complications of the underlying disease in a patient group with significant co-morbidity.

3.6.1 Limitations

It was not possible to obtain the incidence of AF for all of the trials which we identified, and the funnel plot shows a relative lack of mid-sized trials, which reflects this. The use of the OOSS from the EMeA scientific discussion may have resulted in an underestimate of the heterogeneity of the data available. Five studies, possibly quite different in various ways from each other, were combined to form a single data set that was treated in the analysis as a single trial. This is likely to result in an overestimate of within-study variance and an underestimate of between-study variance. The sensitivity analyses in which the OOSS is excluded, however, did not yield significantly different results to the primary analysis.

As the AF data for the largest trials which studied patients with angina, rather than heart failure, was not available separately for each trial, it was not possible to definitively assess the effect of treatment indication on AF risk. It would be possible, given the data presented here alone, that AF is less common as a side effect of ivabradine treatment in patients with angina pectoris rather than heart failure. However, the SIGNIFY trial, which was conducted to examine the effects of ivabradine in patients with stable coronary artery disease without clinical heart failure, also identified a significant increase in AF incidence with ivabradine treatment *vs.* placebo[275].

We are unable to provide data comparing the AF risk profile of patients who developed AF to those that did not in either the ivabradine treatment or control groups. It is possible that the increase in AF risk is restricted to a small high-risk subgroup within the overall study population for each trial.

3.7 Conclusions

Ivabradine treatment is associated with a 15% increase in the relative risk of atrial fibrillation. This confirms the role of the I_f current in the pathophysiology of AF. This provides further evidence that AF risk variants identified in chromosome 15 may act by altering activity of *HCN4*. We estimate that 208 patient-years of treatment with

ivabradine would be required to cause one new case of AF. The risk of AF needs to be taken into consideration when weighing the balance of risk and benefits of ivabradine treatment.

Chapter 4

Influence of polymorphisms at AF-associated loci on gene expression in peripheral blood

4 Influence of polymorphisms at AF-associated loci on gene expression in peripheral blood

4.1 Abstract

Genome-wide association studies (GWAS) have identified genetic variants in nine chromosomal regions that are associated with atrial fibrillation (AF). The mechanisms underlying these associations are unknown. None of the genetic variants identified lie in coding regions and so the most likely hypothesis is that they act by influencing expression of nearby genes. This study investigates the association between genetic variants at GWAS-identified loci and expression of nearby candidate genes in peripheral blood.

Tagging SNPs were genotyped to capture common genetic variation at each locus. Allelic expression ratios of candidate genes were measured using the Sequenom platform and associations with phased genotypes were assessed using a likelihood ratio test. For candidate genes where a significant association between genotype and allelic expression was seen, total expression was measured using qPCR and associations between genotype and total expression were determined.

Expression of candidate genes at five of nine loci; *KCNN3* in 1q21, *CAV1* and *CAV2* in 7q31, *C9orf3* and *FANCC* 9q22, *SYNE2* and *ESR2* in 14q23 and *ZFHX3* in 16q22; was identified in peripheral blood. The risk [T] allele of the AF-associated SNP rs13376333 in chromosome 1q21 was associated with a 0.85-fold decrease in expression of *KCNN3* ($p=2x10^{-4}$). The risk [G] allele of the AF-associated SNP rs3807989 in chromosome 7q31 was associated with a 0.91-fold decrease in expression of *CAV1* ($p=4x10^{-5}$). The risk [T] allele of rs1152591 in chromosome 14q23 was associated with a 1.12-fold increase in expression of SYNE2 ($p=4x10^{-27}$) and a 0.82-fold reduction in *CAV1* expression ($p=5x10^{-3}$). A coding SNP, rs1051614, in *KCNN3* was associated with a 0.83-fold reduction in *CAV1* expression ($p=2x10^{-4}$).

This study has identified associations between AF-associated variants at three loci and expression of *KCNN3*, *CAV1*, and *SYNE2*, identifying them as important candidate genes

in the development and maintenance of AF. The data also identify *cis*-acting variants which affect *ZFHX3* and *C9orf3* expression.

4.2 Background

Genome-wide association studies (GWAS) have identified a number of genetic variants which are associated with an increased risk of AF. The latest meta-analysis of GWAS studies identified and replicated associations in 9 chromosomal regions, as previously summarised in Table 1.2 in Chapter 1[105-109]. Three of these regions have been identified at GWAS significance in multiple independent studies, namely in 1q21, 4q25, and 16q22.

None of the lead GWAS SNPs for AF are situated within a known spliced transcript, and so these polymorphisms are not likely to result in alteration of a diffusible factor which might affect expression in *trans*. It is likely, therefore that the mechanism of action of these variants is by affecting gene expression of nearby genes in *cis*.

The candidate genes in the regions identified perform a wide range of functions: there are cell structural proteins, (*SYNE2*, *MYOZ1* and *SYNPO2L*), ion channels (*KCNN3* and *HCN4*), transcription factors (*PITX2*, *ZFHX3* and *PRRX1*), cell-signal transduction proteins (*CAV1* and *CAV2*) and genes which regulate endocrine function (*ESR2*, *C9orf3*), all of which could have a plausible role in the pathophysiology of AF. It is important to demonstrate an association between the variants of interest and expression of any candidate gene, as genetic influences can operate over large distances, even between chromosomes, and so the genes responsible for AF risk may not simply be those closest to the variants identified. Perhaps more importantly, identifying an association between a risk allele and gene expression also provides information on the direction of the effect, i.e. whether the variant up-regulates or down-regulates expression. This provides a guide as to the likely effects of any pharmacological intervention which increases or decreases activity of the protein product.

Polymorphisms may also modify expression of particular transcript variants. Cardiovascular risk-associated variants at chromosome 9p21 are associated with expression of some, but not all, of the transcript variants of the long non-coding RNA

ANRIL[276]. The transcripts which are associated with cardiovascular risk variants demonstrate stronger *trans*-acting influences on pro-atherosclerotic gene networks[277]. It is therefore important to consider the relationship between risk variants and expression of not only the gene, but also individual transcripts, which may have non-overlapping or even opposite effects.

As described in the introduction, associations between genetic variation and gene expression can be assessed by measuring total expression (eQTL mapping) or by measuring allelic expression ratios (aerQTL mapping). Previous studies have demonstrated the power of the aerQTL approach in identifying genetic variants that are associated with gene expression[229,278].

Genome-wide eQTL studies have been performed in a number of different tissues and in lymphoblastoid cell lines (LCLs)[193,196,279]. Associations have been demonstrated between the AF risk SNP rs1152591 in Chr 14q23 and *ESR2* expression in LCLs, and adipose tissue, between the AF risk SNP rs3807989 and *CAV1* in LCLs, adipose tissue and skin, and between rs3807989 and *CAV2* in LCLs[193,194,196]. However no associations have been demonstrated at the other seven loci which could explain the mechanisms of action of these variants. Moreover the 14q23 association is of doubtful significance since *ESR2* is expressed at very low levels in atrial tissue.

As discussed in Chapter 1, there are differences in genetic architecture between different human populations. This provides the opportunity for fine mapping of genetic associations. If, in a region of interest, the pattern of linkage disequilibrium is finer in a particular population, then that population in which the haplotype blocks are smaller can be used to map the association to a smaller region, and therefore reduce the number of variants which may explain the association.

4.3 Aims

1. To identify which AF risk loci candidate genes are expressed in whole blood.

2. To perform aerQTL mapping of candidate genes at those loci.

3. To perform eQTL mapping for selected primary candidate genes in each region.

4. To determine whether multiple loci are independently associated with gene expression in any genetic region

5. To determine whether AF-associated variants correlate with gene expression.

6. In conjunction with the results presented in Chapter 6, to determine the concordance of eQTL and aerQTL mapping in blood with cardiac tissues, and thus to investigate the utility of peripheral blood as a surrogate marker for cardiac expression.

7. To determine whether there are differences in genetic architecture and genetic influences on expression between North of England and South African populations

8. To determine the power of trans-ethnic fine-mapping to refine the localisation of aerQTLs in the genetic regions identified in multiple independent GWAS of AF.

4.4 Materials and methods

4.4.1 Participants and materials

Study participants and ethical approvals were as set out in Chapter 2.

Nucleic acid extraction and cDNA synthesis was performed as described in Chapter 2.

RNA integrity was determined on a random subset of samples using the Agilent 2100 bioanalyser, as set out in Chapter 2.

4.4.2 Determination of candidate gene expression in peripheral blood

The main candidate genes at each AF risk locus were selected on the basis of proximity to the risk variants and an assessment of the biological plausibility of the gene's role in AF risk. The following genes were identified at each locus: *KCNN3* at 1q21, *PRRX1* at 1q24, *PITX2* at 4q25, *CAV1* at 7q31, *C9orf3* at 9q22, *SYNPO2L* at 10q22, *SYNE2* at 14q23, *HCN4* at 14q23 and *ZFHX3* at 16q22. Primers to determine the expression patterns of candidate genes were designed using Primer3 software v0.4.0[280,281]. PCR was carried out with cDNA from 25ng RNA, using HotStar Taq reagents (Qiagen, Germany), according to the manufacturer's protocols and PCR products were visualised by gel electrophoresis.

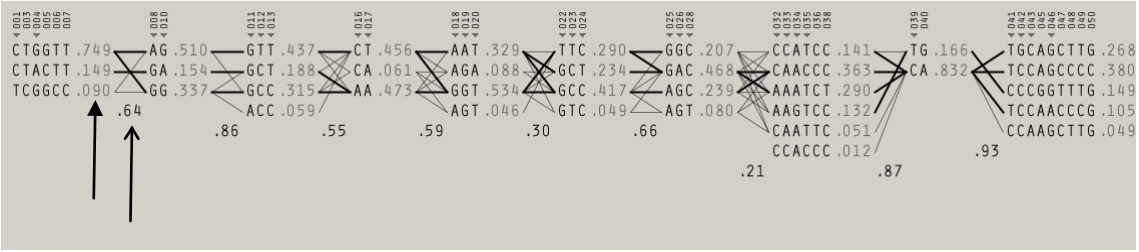
4.4.3 Selection of transcribed SNPs for allelic expression analysis

Using the UCSC genome browser [282,283] transcribed SNPs with minor allele frequency (MAF) >0.05 were identified as suitable candidates for use in the allelic expression assays. The following transcribed SNPs were selected: rs1051614, in *KCNN3,* rs1049334, rs1049337 and rs9920 in *CAV1,* rs8940 in *CAV2,* rs4657 in *C9orf3*, rs4647554 in *FANCC*, rs1152582 and rs7161192 in *SYNE2*, rs1256049, 4986938 and rs928554 in *ESR2*, and rs740178 and rs10852515 in *ZFHX3*.

4.4.4 Selection of mapping SNPs

Tag SNPs required to capture common variation in the core regions of interest (15kb either side of those candidate genes expressed in blood) were selected using HaploView 4.0 tagger software using the following parameters: HapMap population: CEU, minimum minor allele frequency: 0.05, pairwise tagging, r^2 threshold: >0.8. SNPs previously reported to be associated with disease phenotypes were forcibly included in the set of selected SNPs[108,109,131,284-287]. Details of included SNPs are given in the results section, below.

Figure 4.1 Selection of mapping SNPs



Example output from the HaploView software. Haplotype blocks have been identified for SNPs (labelled 001 to 050) in this region by the software on the basis of allele frequencies in the CEU population. The frequency of each haplotype in the population is shown (closed arrow), as is the degree of LD between each haplotype (open arrow). The thick and thin lines indicate haplotypes which are continuous in at least 10% and 1% of the population respectively. The small triangles beside the SNP number indicate the minimum SNPs which need to be genotyped in order to reconstruct the haplotypes in this population.

4.4.5 Genotyping

Multiplex SNP genotyping was performed on the NE cohort, using the Sequenom platform as set out in Chapter 2. 202 SNPs were typed in 9 separate multiplex reactions (Table A1.2). Genotyping was also performed for the chromosome 1q21 and 16q22 SNPs in the SA cohort. Primer sequences are shown in Appendix 1. For one SNP, rs3807989, which could not be typed on the Sequenom platform, genotyping was carried out using a custom TaqMan SNP Genotyping assay (C_2972990_10, Life Technologies), as described in Chapter 2.

4.4.6 Measurement of allelic expression ratios

Allelic expression ratios (AER) in NE cohort RNA samples were determined for all genes within 100kb of the GWAS-associated SNP at each locus. Due to the limited amount of RNA available, AER were determined in the SA cohort only for *KCNN3* and *ZFHX3*, the candidate genes which were expressed in peripheral blood and for which associations have been identified in individual GWAS studies as well as by meta-analysis. Assays for the measurement of allelic expression were identical to those used for genotyping the transcribed SNPs, as the amplicons for each were wholly contained within a single exon. Primer sequences are shown in Appendix 1.

Quantification of the allelic expression ratio was performed using the Sequenom platform as set out in Chapter 2. The gDNA ratios were relatively homogenous with little inter-individual variability compared to cDNA ratios (Figure 4.5). AER values were excluded where the standard error of the technical replicates was high (S.E. >0.5). Outliers were excluded using Grubb's test.

4.4.7 Quantification of normalised total gene expression using real-time PCR

Quantification of total gene expression was performed as described in Chapter 2. Real-time PCR assays were performed in two separate assay arrangements. *KCNN3* and *ZFHX3* transcript levels were normalised to *ACTB*, *GAPDH* and *RN18S1*. *ZFHX3a*, *CAV1*, *C9orf3* and *SYNE2* levels were normalised to *ACTB*, *GAPDH* and *HPRT1*. *RN18S1* was replaced by *HPRT1* as a reference gene in the latter experiments. This was because

absolute expression of *RN18S1* was found to be substantially higher than the other reference genes (~8900-fold), raising the possibility that excessive amplification of RN18S1 might interfere with amplification of the other transcript in a duplex PCR reaction. In the *KCNN3* and *ZFHX3* experiments, neither transcript was measured in duplex with *RN18S1*, and therefore measurement of *KCNN3* and *ZFHX3*, would not have been affected by high expression of *RN18S1*. Assay details are given in Table A1.1. Assays were performed using cDNA synthesised from 25ng of RNA in a 15µl reaction volume.

4.4.8 Statistical analyses

Phasing was performed using the BEAGLE software package and aerQTL analysis using a likelihood ratio test in the R software package for transcripts with a single transcribed SNP or in the Matlab software package for transcripts with more than one transcribed SNP, as described in Chapter 2. Associations between total expression and SNP genotype were assessed as described in Chapter 2.

In order to correct for the testing of multiple associations, the threshold for statistical significance at each locus was determined using the Bonferroni correction[244]. For *trans*-acting associations the threshold was determined using a Bonferroni correction for the total number of SNPs typed at all loci. For example, the significance threshold for *cis* acting effects of the 16 SNPs in chromosome 7q31 was set at 0.05/16: p=0.0031, whereas *trans*-acting effects were assessed for all 202 SNPs and so the threshold for significance was set at 0.05/202: p=0.00024. Full results are presented in Appendix 2.

For those genes where total expression and allelic expression ratios were measured, the relative contribution from *cis*-acting influences to the total variance was calculated as described in Chapter 2.

4.4.9 Analysis of linkage disequilibrium

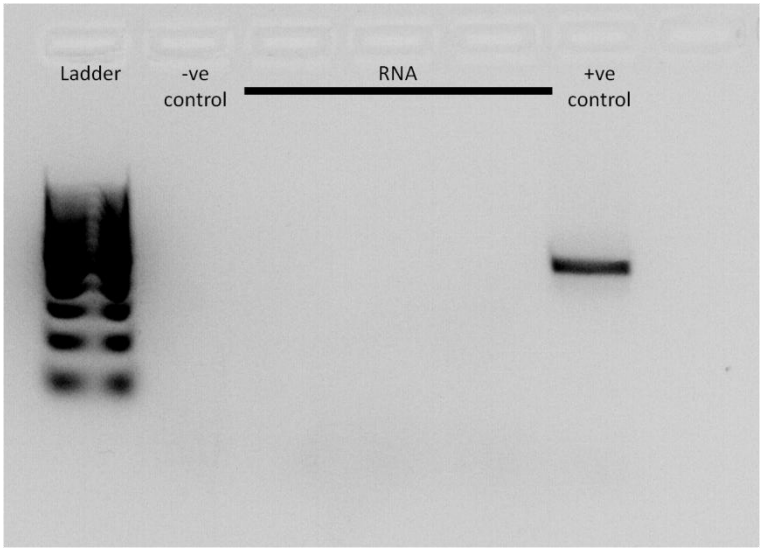
LD was calculated using phased genotypes in HaploView v4.2 and haplotype blocks were identified using the confidence interval algorithm[43,243].

4.5 Results

4.5.1 Quality control

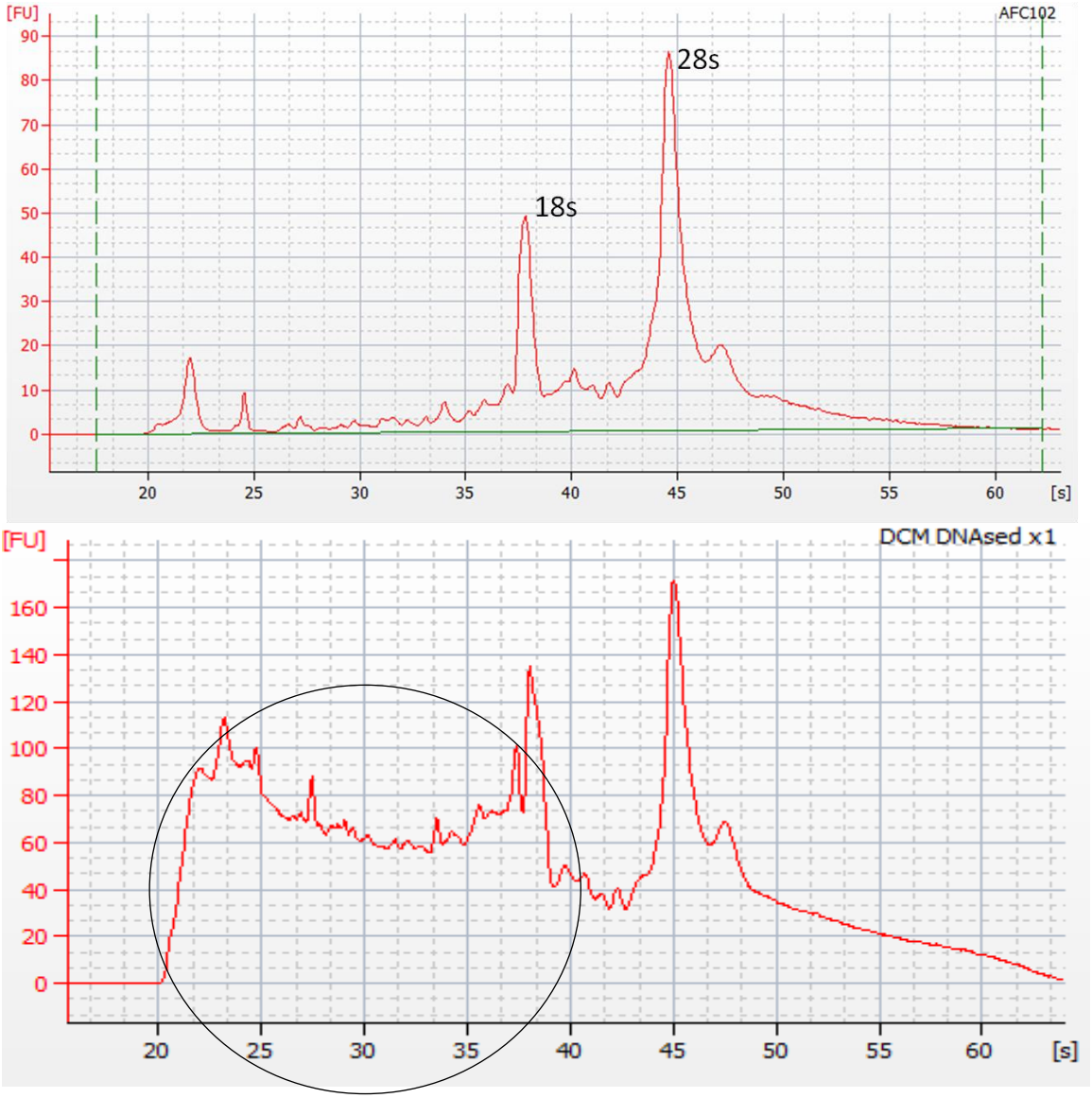
PCR without reverse transcription demonstrated absence of genomic DNA contamination of the extracted RNA (Figure 4.2). The median RNA integrity number for the samples analysed by bioanalyser was 7.85 (Figure 4.3). Visual inspection of mass spectroscopy traces for the allelic expression assay of CAV1 using rs9920 revealed failure of the extension reaction in 468 of 1744 assays (Figure 4.4). This may be the result of an unrecognised SNP interfering with the binding of the unextended primer and therefore could introduce bias into the assessment of AER. This assay was excluded from further analysis.

Figure 4.2 Demonstration of adequate DNase treatment



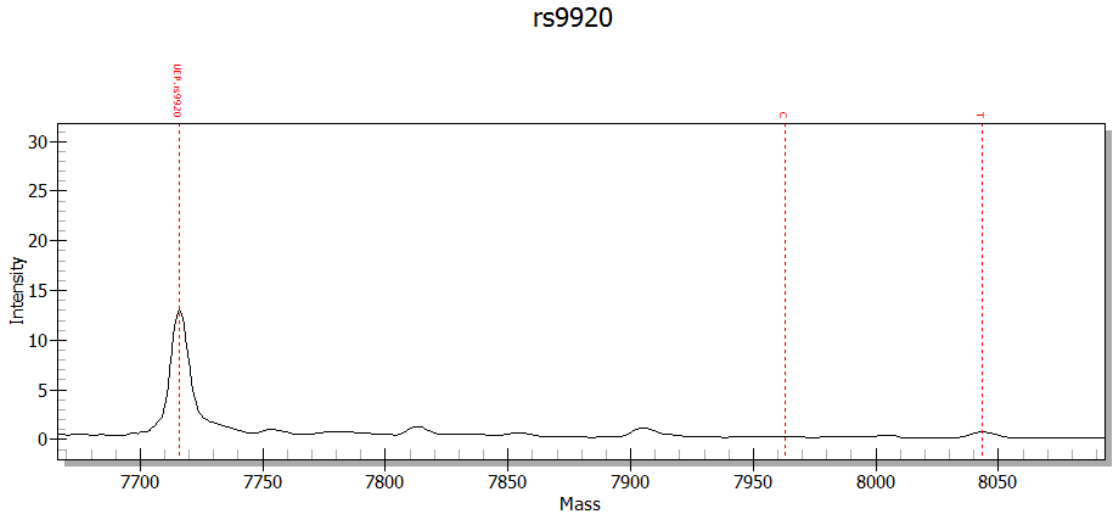
Gel electrophoresis of PCR products with the following templates: water (-ve control), RNA without a reverse transcription reaction (three central wells) and genomic DNA (+ve control).

Figure 4.3 Agilent Bioanalyser electropherogram



Time in seconds (indicating fragment size) is plotted on the x-axis and fluorescence (indicating amount of RNA) on the y-axis. An example electropherogram from an RNA sample with a RIN of 8.6 is shown in the upper panel. There are clear 18s and 28s peaks in a ratio of 1:2, with no evidence of degraded RNA, indicating good RNA quality. By contrast the sample in the lower panel is significantly degraded with a RIN of 4.4; 18s and 28s peaks remain visible but there is a significant amount of RNA in shorter fragments, circled.

Figure 4.4 Example of failed extension reaction

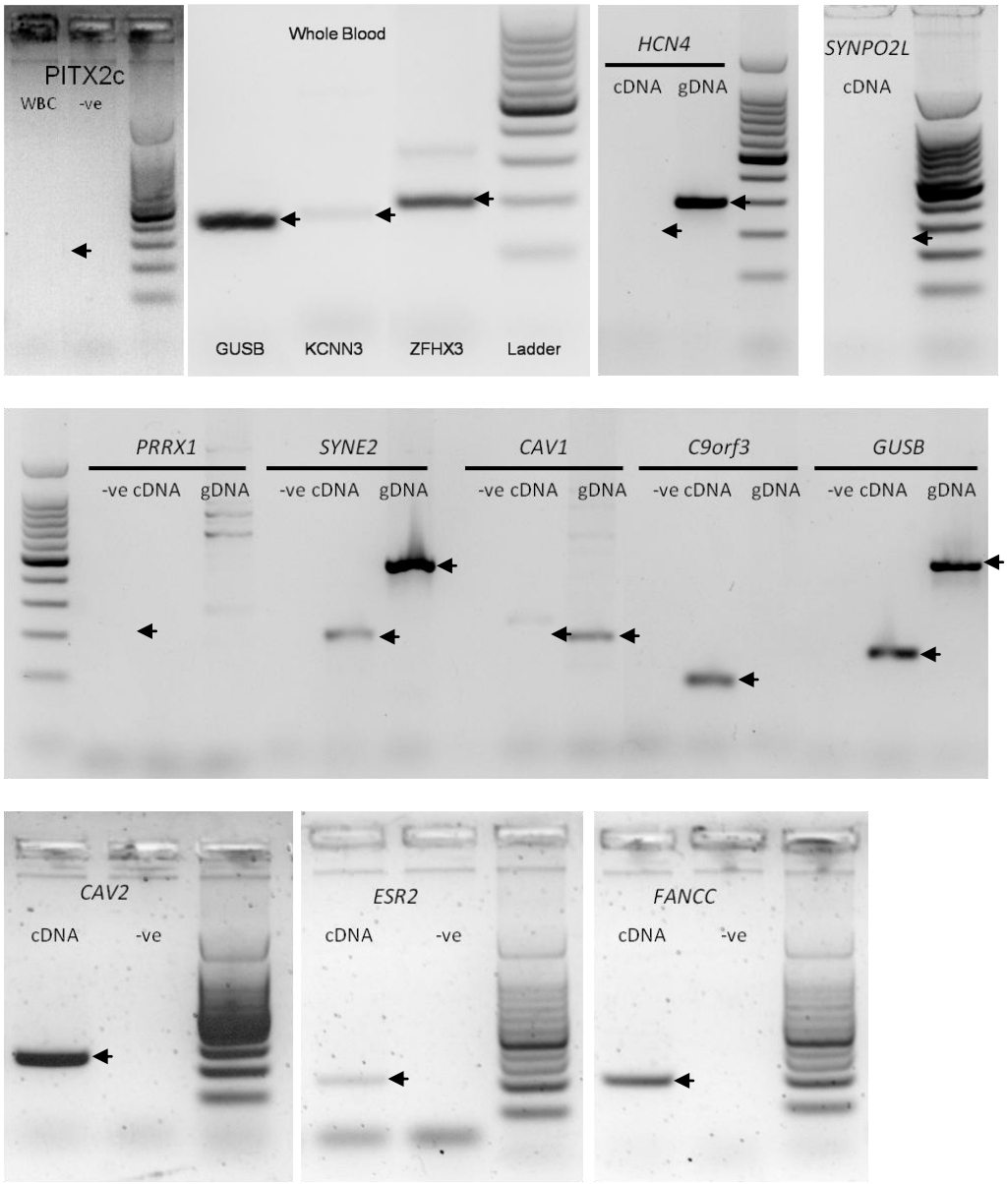


A clear spectral peak can be seen for the unextended primer on the left of the image, marked 'UEP.rs9920'. By contrast there are no peaks at either of the expected positions for the C or T alleles, indicating that the extension reaction has failed.

4.5.2 Expression profiling

In five regions we found that the main candidate genes were expressed in blood: *KCNN3* in chromosome 1q21, *CAV1* in chromosome 7q31, *C9orf3* in chromosome 9q22, *SYNE2* in chromosome 14q23, and *ZFHX3* in chromosome 16q22. Expression of the following nearby genes in each of these three regions was also confirmed; *CAV2* in chromosome 7q31, *FANCC* in chromosome 9q22, and *ESR2* in chromosome 14q23 (Figure 4.5).

Figure 4.5 Expression profiles of candidate genes



PCR products after electrophoresis on agarose gel alongside a 100bp DNA ladder. Arrows indicate the expected position of DNA bands.

4.5.3 Inter individual variation in expression

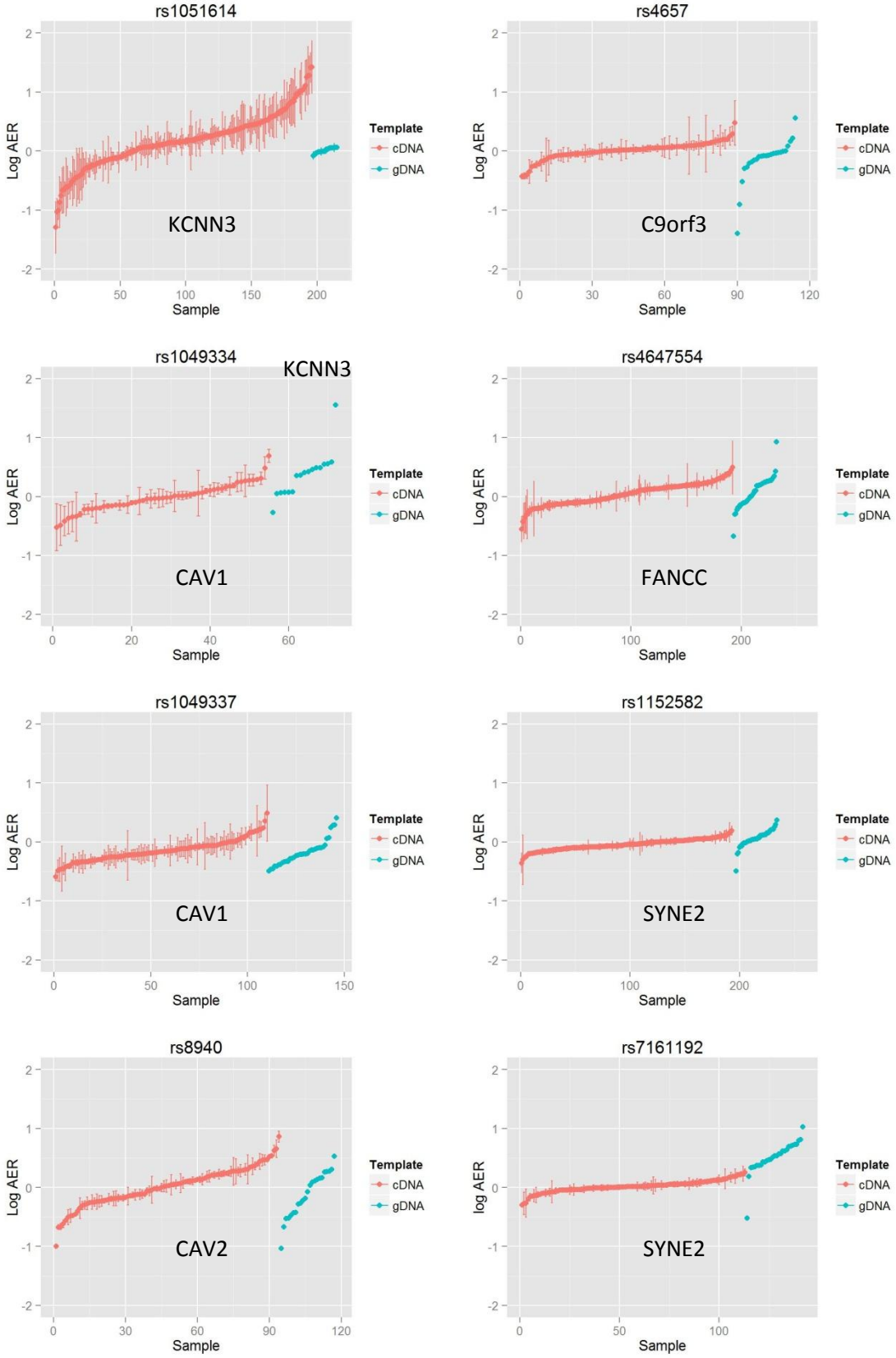
All of the genes for which total expression was measured demonstrated variability in expression levels in our sample. Total expression levels demonstrated greater variability for *KCNN3* and *CAV1* levels than *C9orf3*, *SYNE2*, or *ZFHX3* levels. Allelic expression ratio (AER) ranges differed for each gene, depending on which transcribed SNP was used (Table 4.1, Figure 4.6). This is not wholly surprising as the individuals who are heterozygous at one transcribed SNP may not be heterozygous at another, and so the allelic expression ratios for each transcribed SNP represents the values for a subset of the total cohort. The number of heterozygotes for the transcribed SNP, rs1256049 was lower than at the other loci, due to the low MAF (MAF=0.034 in the NE cohort *cf.* MAF=0.031 in the HapMap CEU cohort). For *KCNN3*, a relatively large proportion (17.4%) of the total variance in expression was due to *cis*-acting effects.

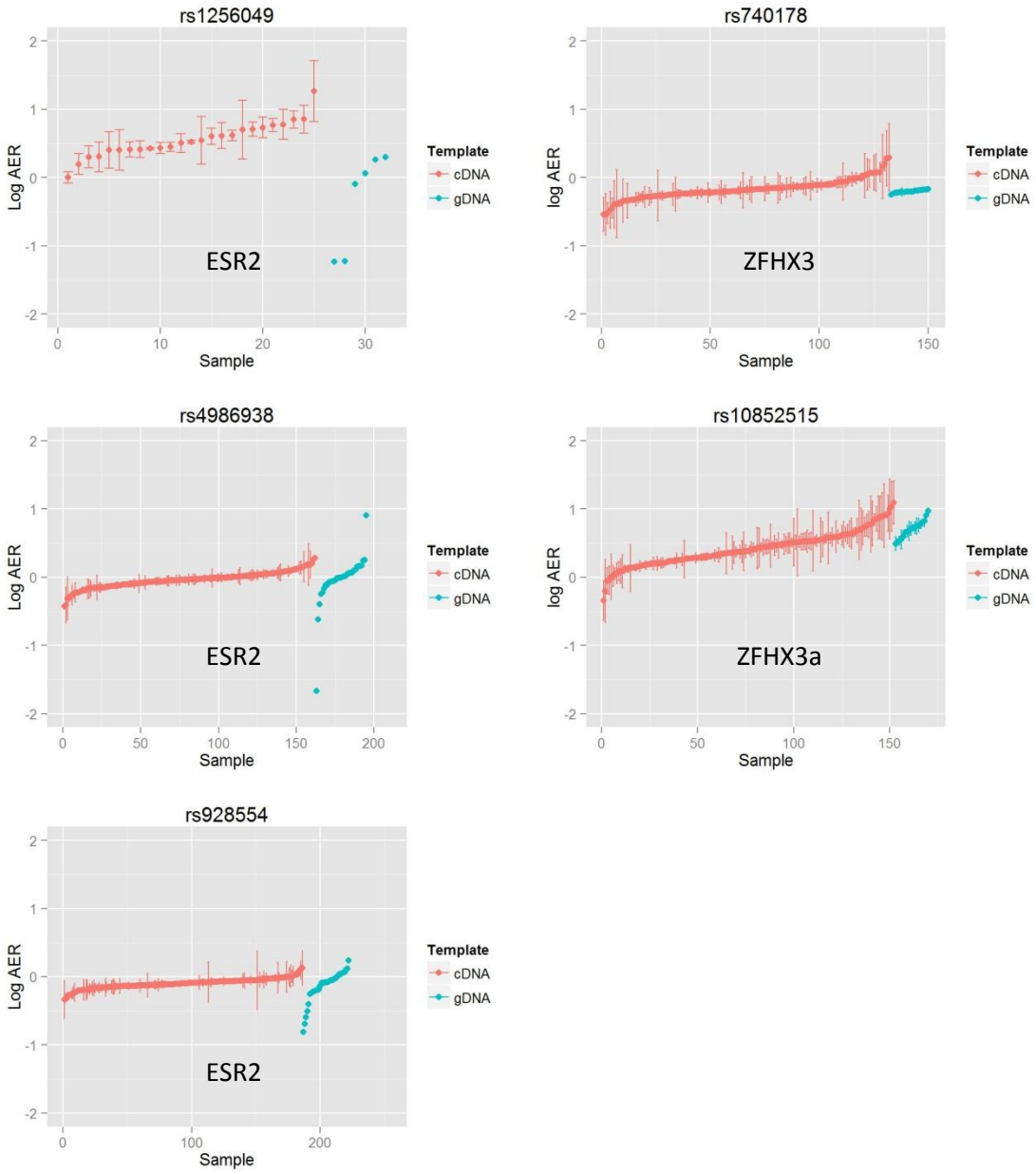
Table 4.1 Variance of normalised total and allelic expression

Gene	Total expression IQR	Transcribed SNP	AER 1st and 3rd Quartiles	Proportion of variance due to *cis*-acting effects
KCNN3	1.91-fold	rs1051614	0.84-1.42	17.4%
CAV1	2.18-fold	rs1049334	0.62-0.84	8.7%
		rs1049337	0.94-1.20	
CAV2		rs8940	0.92-1.42	
C9orf3	1.55-fold	rs4657	0.95-1.09	6.3%
FANCC		rs4647554	0.83-1.13	
SYNE2	1.50-fold	rs1152582	0.83-0.95	6.0%
		rs7161192	0.51-0.57	
ESR2		rs1256049	0.87-1.24	
		rs4986938	0.89-1.02	
		rs928554	0.97-1.07	
ZFHX3	1.61-fold	rs740178	0.96-1.10	3.7%
		rs10852515	0.63-0.86	

Inter-quartile ranges are given for total expression and the 1st and 3rd quartiles for the allelic expression ratio measured using each transcribed SNP.

Figure 4.6 Variance in allelic expression ratios



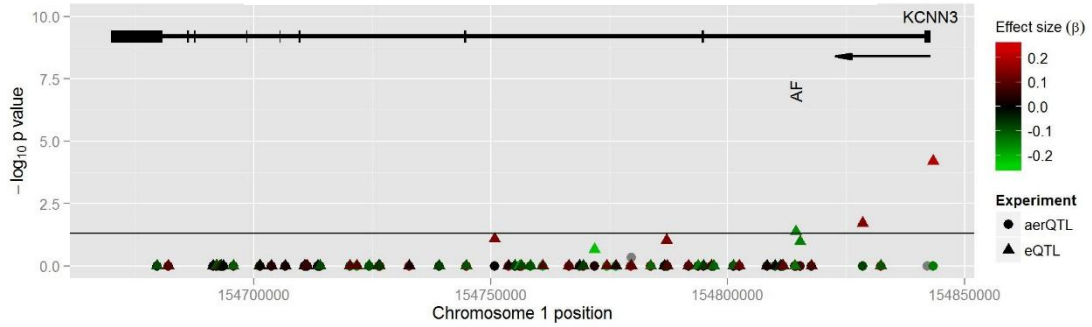


Log allelic expression ratios are shown for each of the transcribed SNPs listed in Table 4.1. For each SNP, the log AER for each heterozygous individual is plotted in red with error bars representing the standard error of the technical replicates. The AER measured in genomic DNA samples used for normalisation are plotted in blue. AER values were normalised to the mean allelic ratio measured in genomic DNA, as described in Chapter 2.

4.5.4 AF-associated SNPs on chromosome 1q21 are associated with increased KCNN3 expression

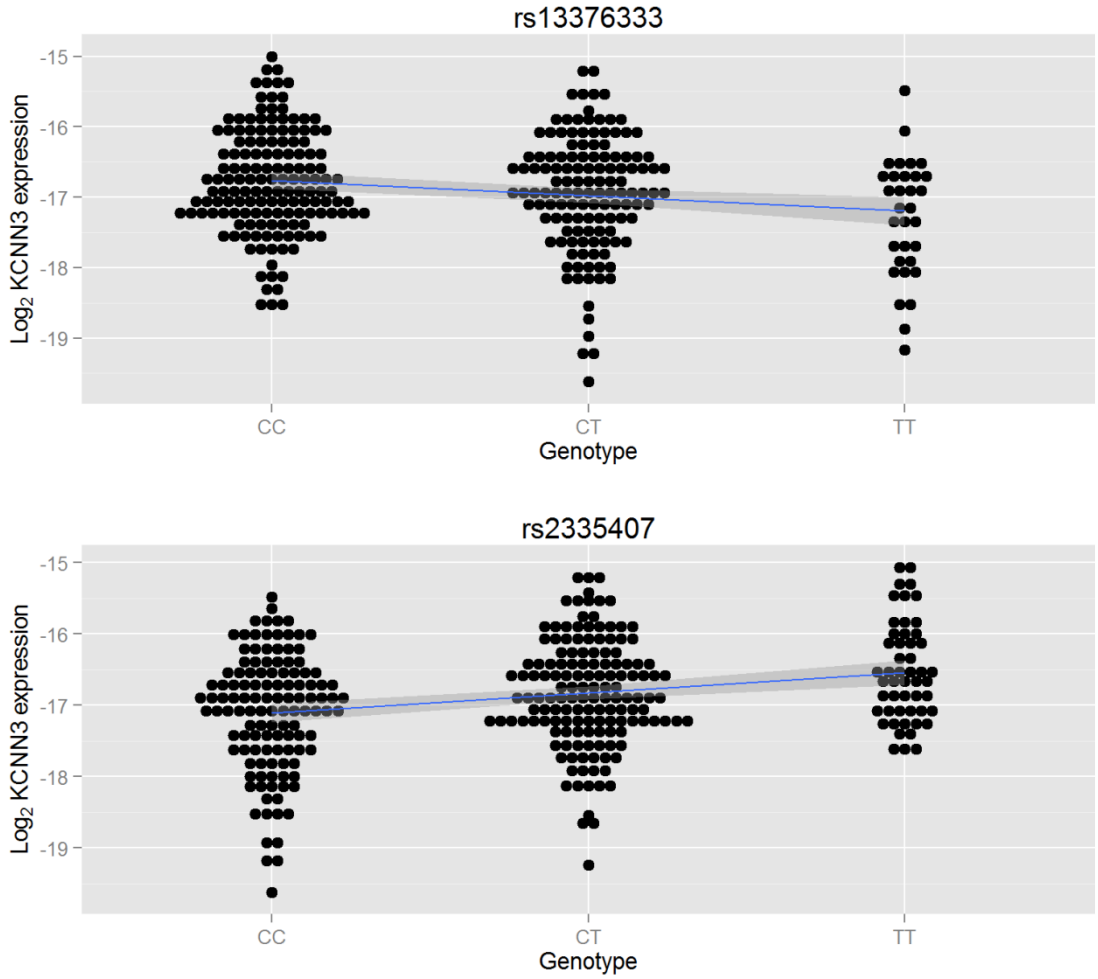
Allelic expression ratios were measured for *KCNN3* on 1q21. No associations were demonstrated between the allelic expression ratios of *KCNN3* and any SNP in the NE cohort. However, eQTL mapping of *KCNN3* expression identified 6 SNPs that were significantly associated with *KCNN3* expression. The strongest association was with the minor allele of rs2335407, which was associated with a 1.24-fold increase in expression of *KCNN3* ($p=3.1x10^{-7}$). The risk [T] allele of one the lead AF SNPs, rs13376333, was associated with a 0.85-fold decrease in *KCNN3* expression ($p=2.1x10^{-4}$) (Figures 4.7 and 4.8, Table A2.1). After correction for the effect of the strongest eQTL SNP, rs2335407, the minor allele of a second SNP, rs7532286, remained significantly associated with a 1.13-fold increase in *KCNN3* expression.

Figure 4.7 Associations at chromosome 1q21 in the NE cohort



Individual variants are plotted by chromosomal position on the x-axis and significance of the association (as -log p value) on the y-axis. Effect size associated with the minor allele (log fold-change, β) is represented by colour, with increased expression in red and reduced expression in green. The results of aerQTL experiments are represented as circles and of eQTL experiments as triangles. The chromosomal positions of *KCNN3* and the GWAS hit SNP for AF are marked.

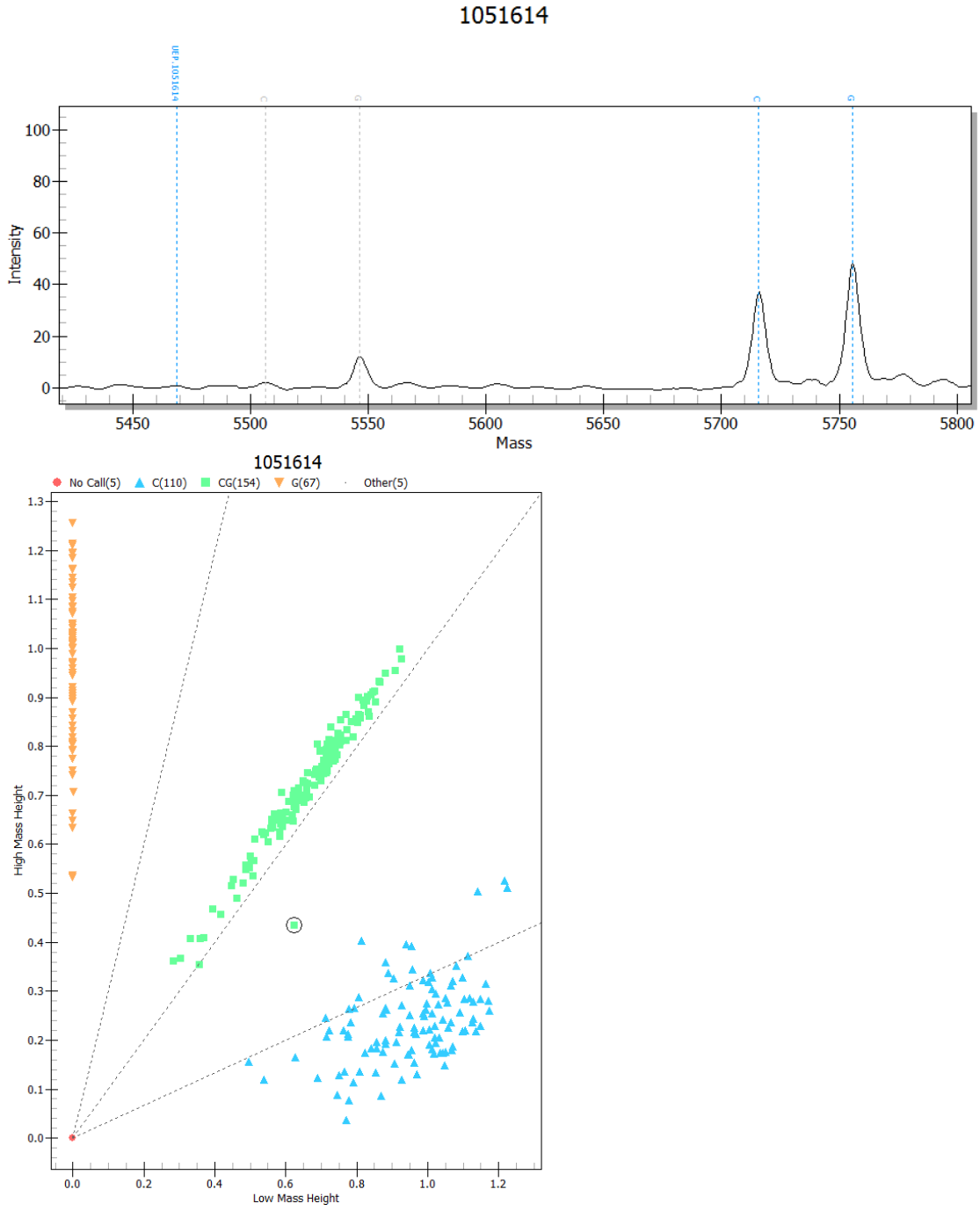
Figure 4.8 Normalised total *KCNN3* expression by genotype.



Top: total expression of *KCNN3* is shown by genotype at the AF-associated SNP, rs13376333. **Bottom**: expression of *KCNN3* by genotype of rs2335407, the most strongly associated eQTL SNP for *KCNN3*.

The discrepancy between the results of the eQTL and the aerQTL experiment raises the possibility that there was a technical problem with the assay using the transcribed SNP rs1051614. However, on inspection of the mass spectra, there were no obvious salt or interference peaks and the assay peaks are large relative to the background signal. Furthermore in the cluster plot, the genotyped samples fall into three reasonably well-defined clusters, with little variation in the ratio of peak heights in heterozygotes, suggesting good assay performance. One sample, circled in Figure 4.9, did not fall within any cluster and was problematic in several assays; it was therefore excluded from analysis in all assays.

Figure 4.9 rs1041615 assay performance

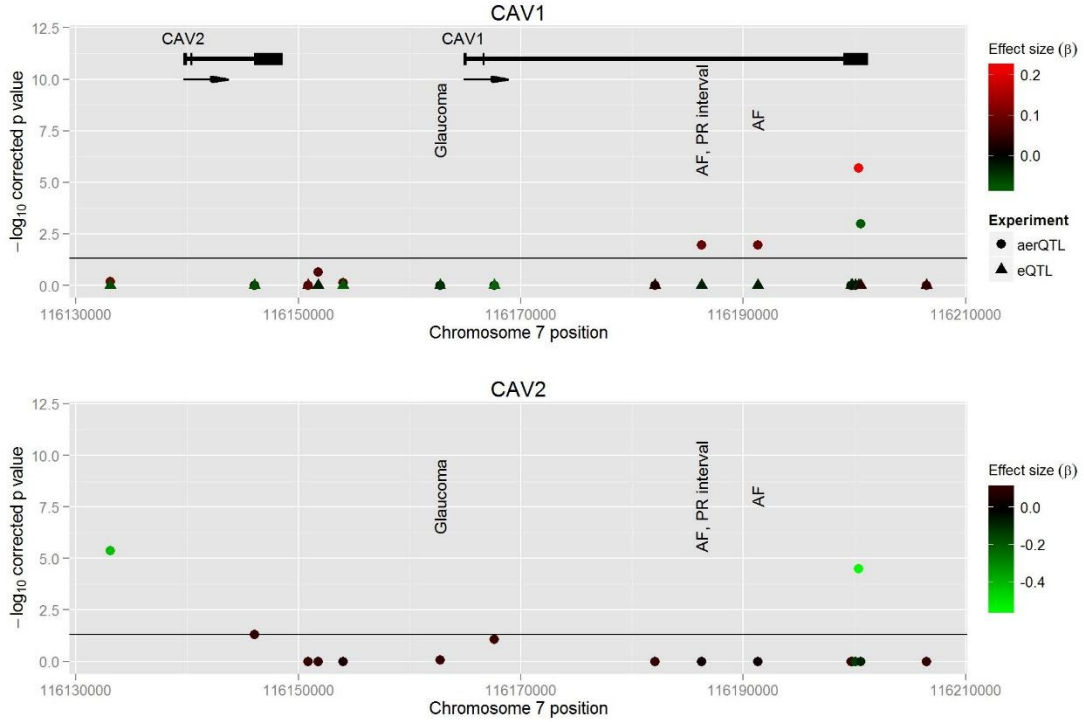


Top: mass spectrometry trace of rs1051614 in cDNA. **Bottom:** genotyping cluster plot from genomic DNA. Symbols indicate genotype automated calls as per the key at the top of the plot. The circled individual did not cluster in many of the genotyping assays (possibly indicating contamination with DNA) and was excluded from all analyses.

4.5.5 AF-associated SNPs in chromosome 7q31 are associated with decreased CAV1 expression

Allelic expression ratios were measured for two genes at chromosome 7q31, *CAV1* and *CAV2*. Associations were seen between 7 of 16 SNPs and *CAV1* expression. The strongest association was observed for the minor allele of the transcribed SNP rs1049334 ($p=7.5x10^{-09}$), which was associated with a 1.26-fold increase in expression. The risk [G] allele of the GWAS associated SNP for AF, rs3807989, was associated with a 0.91-fold decrease in expression ($p=4.2x10^{-5}$). Five of 16 SNPs were associated with *CAV2* expression. The most significant association was seen for rs17138749, which caused a 0.65-fold decrease in expression ($p=1.6x10^{-8}$) and lies 6.5kb upstream of the *CAV2* transcription start site. The AF associated GWAS SNPs were not associated with *CAV2* expression. The transcribed SNP rs1049334 had significant effects on both *CAV1* and *CAV2* expression, but in opposite directions, the minor allele was associated with a 1.26-fold increase in *CAV1* and a 0.57-fold decrease in *CAV2*. eQTL mapping of *CAV1* in the 7q31 locus did not identify any SNPs significantly associated with total expression (Figure 4.10, Table A2.2).

Figure 4.10 Associations at chromosome 7q31

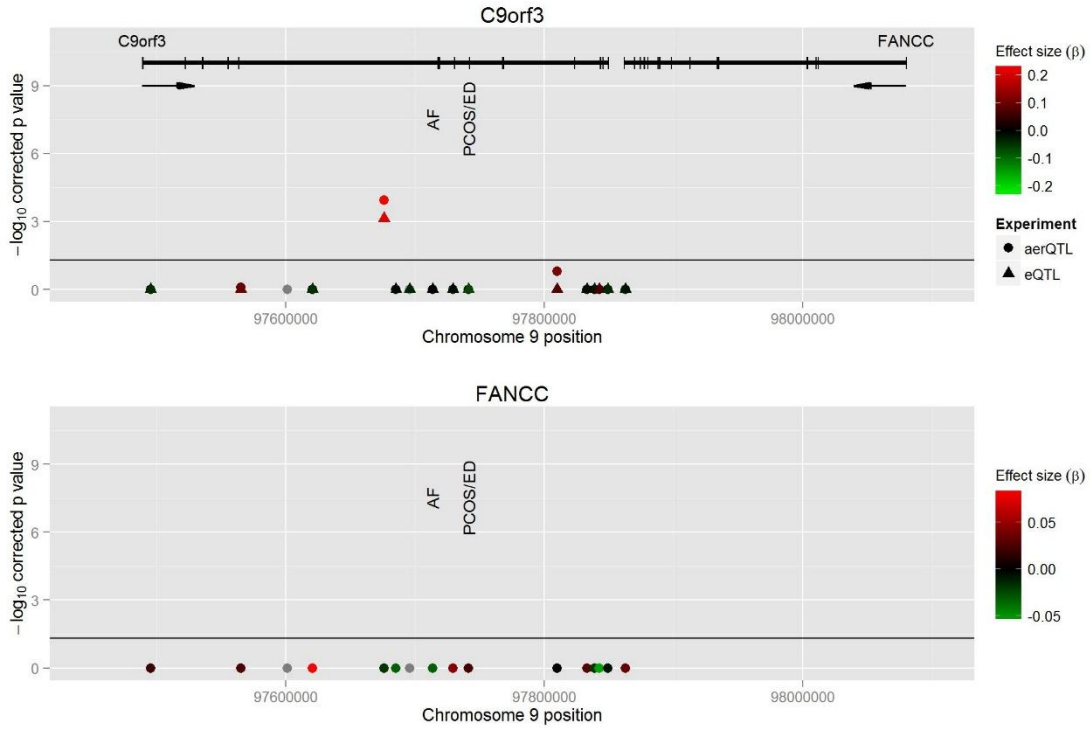


Top: Associations between genotype and expression of *CAV1*. **Bottom:** Associations between genotype and expression of *CAV2*.

4.5.6 Variants in chromosome 9q22 are associated with C9orf3 expression

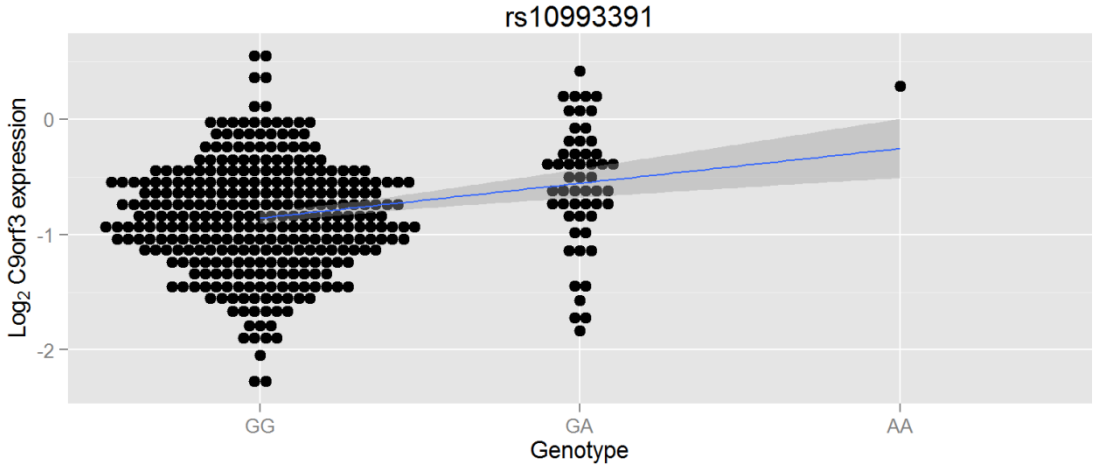
Allelic expression ratios were measured for two transcripts in the chromosome 9q22 region, *C9orf3* and *FANCC*. Three of 17 SNPs were significantly associated with *C9orf3* expression. The strongest association was seen for rs10993391, the minor [G] allele which was associated with a 1.27-fold increase in expression ($p=4.1x10^{-7}$). No SNPs were associated with *FANCC* expression. The AF associated SNP, rs10821415, was not associated with *C9orf3* or *FANCC* expression. eQTL mapping of *C9orf3* expression confirmed a 1.23-fold increase in expression with rs10993391 ($p=3.6x10^{-6}$), but not rs356122 or rs7048941 (Figures 4.11 and 4.12, Table A2.3).

Figure 4.11 Associations at chromosome 9q22



Top: Associations between genotype and expression of *C9orf3*. **Bottom:** Associations between genotype and expression of *FANCC*.

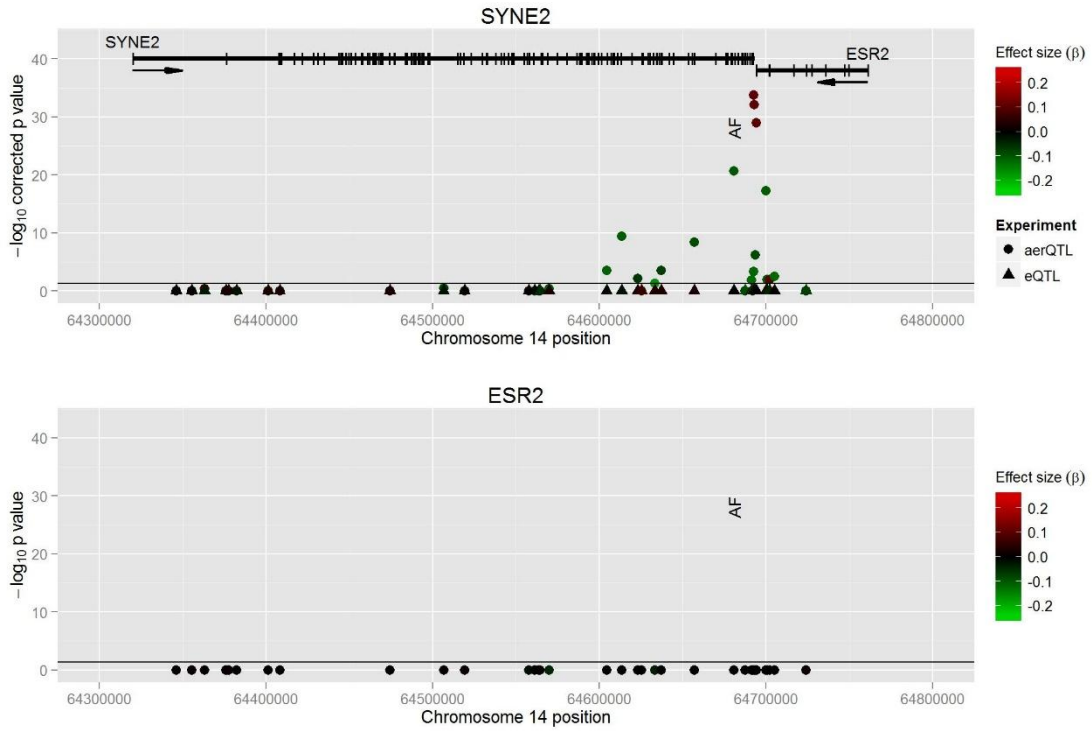
Figure 4.12 *C9orf3* total expression by genotype at the 9q22 locus



4.5.7 AF associated variants in chromosome 14q23 are associated with increased SYNE2 expression

Allelic expression ratios were measured for two genes at the chr 14q23 loci *SYNE2* and *ESR2*. Allelic expression of *SYNE2* was significantly associated with 24 of 37 SNPs. The strongest association was seen with the minor allele of the transcribed SNP rs1152582, which was associated with a 1.13-fold increase in expression ($p=6.2x10^{-37}$). The risk [T] allele GWAS hit SNP for AF, rs1152591, was associated with a 1.12-fold increase in expression of *SYNE2* ($p=7.5x10^{-24}$). After correction for the effects of the most significantly associated aerQTL SNP, rs11852582, a single SNP, rs11158533, was significantly associated with a 1.14-fold increase in SYNE2 expression. No SNPs were significantly associated with *ESR2* expression. eQTL mapping of *SYNE2* expression did not identify any significant associations (Figure 4.13, Table A2.4).

Figure 4.13 Associations at chromosome 14q23



Top: Associations between genotype and expression of *SYNE2*. **Bottom:** Association between genotype and *ESR2* expression.

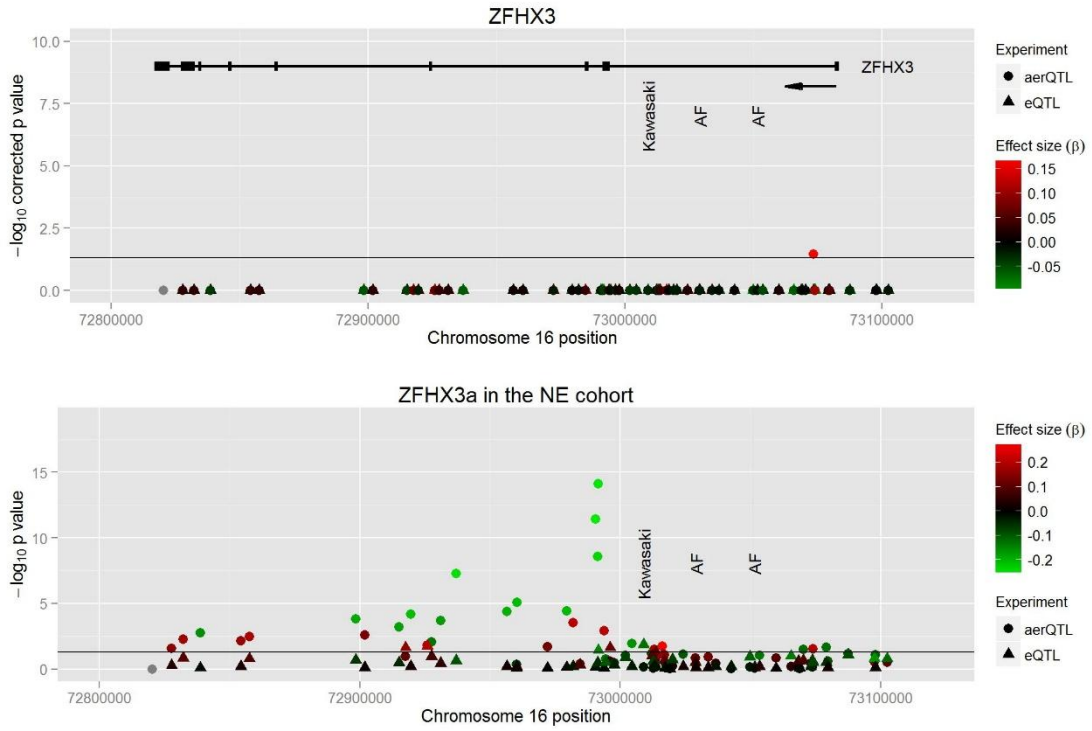
4.5.8 Variants on chromosome 16q22 are associated with ZFHX3 expression in a transcript-specific manner

aerQTL mapping of both transcripts of *ZFHX3* together identified a single significantly associated SNP, rs8060701, the minor allele of which was associated with a 1.18-fold increase in expression ($p=1.3x10^{-4}$). After correcting for the effect of rs8060701, no further associations were seen in either cohort. eQTL mapping did not identify any significant associations between SNP genotype and expression of *ZFHX3*.

aerQTL mapping of *ZFHX3* isoform *a* in the NE cohort revealed several strongly associated SNPs. The most strongly associated SNP was the transcribed SNP, rs10852515, which was associated with a 0.76-fold reduction in expression ($p=7.5x10^{-15}$). After correction for the effects of rs10852515, no other variants were independently associated with expression (Figure 4.14, Table A2.5). The aerQTL SNP associated with expression of both transcripts, rs8060701, was not independently

associated with *ZFHX3a* expression, rs10852515 and rs8060701 were not in significant LD in the SA cohort (r^2=0) or the NE cohort (r^2=0.005).

Figure 4.14 Associations with expression at chromosome 16q22

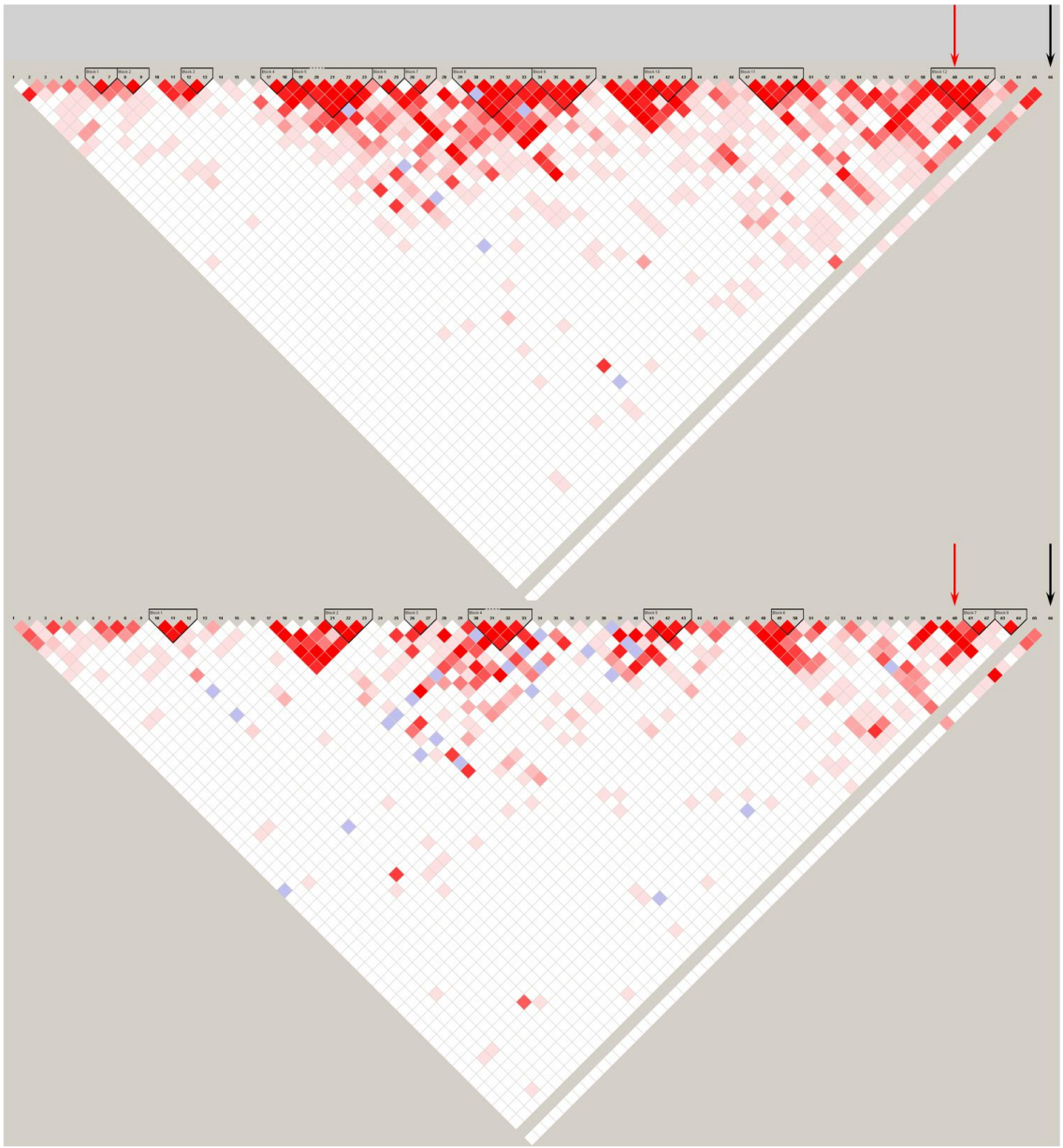


Top: Associations between genotype and expression of ZFHX3. **Bottom:** Associations between genotype and *ZFHX3a* expression.

4.5.9 Patterns of linkage disequilibrium differ between the NE and SA cohorts

Patterns of LD at 1q21 and 16q22 in the two populations are shown in Figures 4.15 and 4.16. The pattern of LD and the haplotype blocks, calculated using the confidence intervals method, differed between the two populations. In both regions the SA population had a finer haplotype structure with fewer and smaller haplotype blocks identified. At 1q21, 12 haplotype blocks were identified in the NE cohort, but in the SA cohort the pattern of LD was much finer with only 8 small LD blocks identified. There was a significant degree of LD between the lead eQTL SNP and the AF risk SNP in both cohorts (r^2=0.67 and r^2=0.41 for the NE and SA cohorts, respectively).

Figure 4.15 Linkage disequilibrium at chromosome 1q21

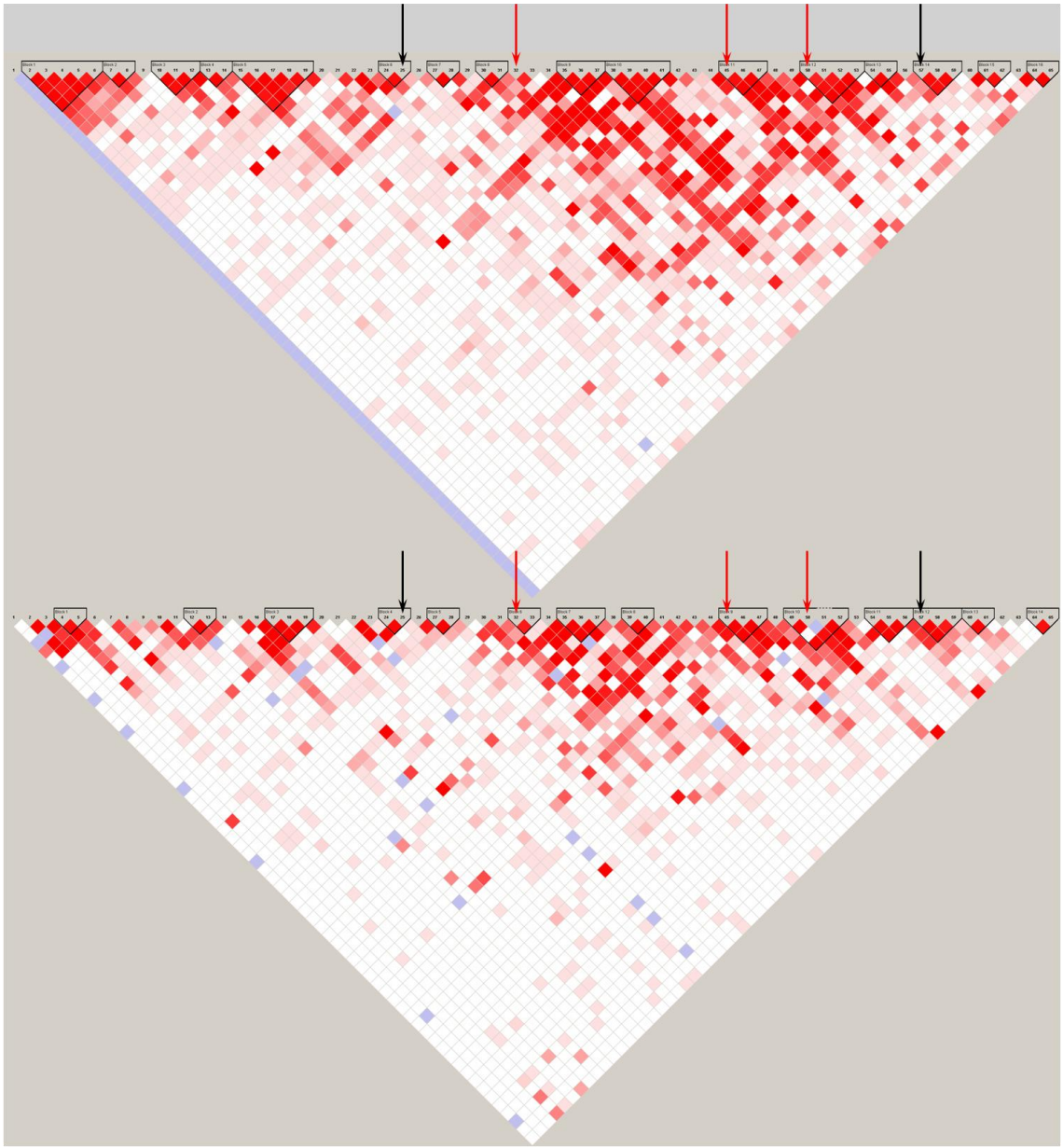


LD between individual SNPs, at the top, is represented by colour with red indicating D'=1 and white indicating D'=0. The AF associated SNP, rs13376333 is indicated by a red arrow and the lead eQTL SNP, rs2335407 by a black arrow. **Top:** LD in the NE cohort. **Bottom:** LD in the SA cohort.

At the 16q22 locus, 16 haplotype blocks were identified in the NE cohort: whereas in the SA cohort, the pattern of LD was finer than in the NE cohort; 14 smaller haplotype

blocks were identified. The GWAS hit SNPs for AF were in LD in both cohorts (r^2=0.85 and r^2=0.69 in the NE and SA cohorts, respectively). The aerQTL SNP for *ZFHX3,* rs8060701, did not lie in a large haplotype block in either cohort. The aerQTL lead SNP for *ZFHX3a* was in LD with a number of SNPs upstream of its position (to the left in Figure 4.14) in the NE cohort, but less so in the SA cohort.

Figure 4.16 Linkage disequilibrium at chromosome 16q22

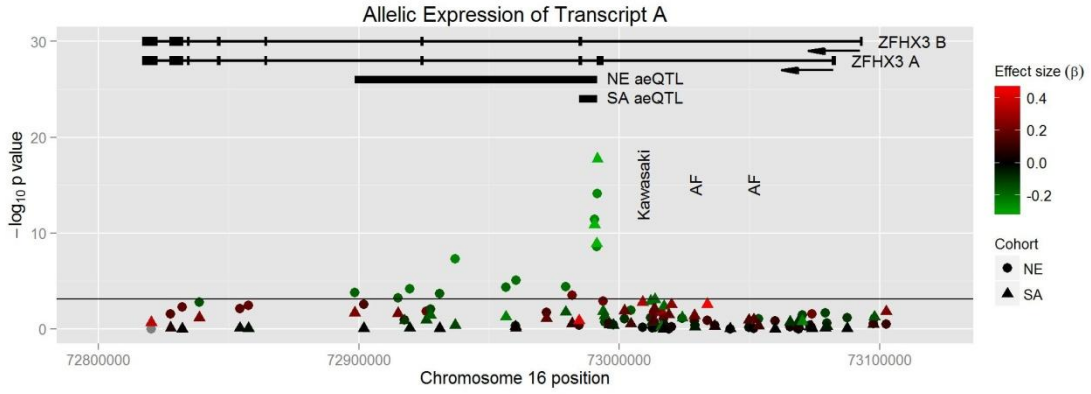


Black arrows indicate the aerQTL lead SNP, rs10852515, centre left and the eQTL SNP, rs8060701, right. Red arrows indicate GWAS hit SNPs for Kawasaki disease, AF and AF from left to right, respectively. **Top:** LD in the NE cohort. 16 haplotype blocks were identified. The GWAS hit SNPs for AF were in significant LD. The aerQTL lead SNP is in LD with a number of SNPs upstream of its position (to the left). The eQTL lead SNP lies in a small haplotype block, but is not otherwise in strong LD with nearby SNPs. **Bottom:** LD in the SA cohort. The pattern of LD is finer than in the NE cohort, 14 smaller haplotype blocks were identified. The aerQTL lead SNP is in less strong LD with upstream SNPs.

4.5.10 Trans-ethnic fine mapping allows finer localisation of aerQTL signals

In order to investigate the differences in aerQTLs between different populations, aerQTL mapping was repeated in the SA cohort. Again in the chromosome 1q21 region, none of the genotyped variants were significantly associated with expression of KCNN3. In the 16q22 region, the association between *ZFHX3* expression and rs8060701 was replicated (p=0.003) No other SNPs were identified as being associated with *ZFHX3* expression. The association between *ZFHX3a* expression and rs10852515 was also replicated ($p=1.7x10^{-18}$). Again, no SNPs were significantly associated with *ZFHX3a* expression after correcting for the effect of the lead SNP. Fewer SNPs in a smaller chromosomal region were significantly associated with *ZFHX3a* expression in the SA cohort. The aerQTL in the NE cohort included 12 significantly associated SNPs in a 93kb region. By contrast, in the SA cohort the aerQTL was confined to three SNPs in a 7kb region (Figure 4.17, Table A2.6). In both cohorts, the significantly associated variants were in LD with the lead aerQTL SNP, rs10852515.

Figure 4.17 Comparison of aerQTLs between the NE and SA cohorts for *ZFHX3a*

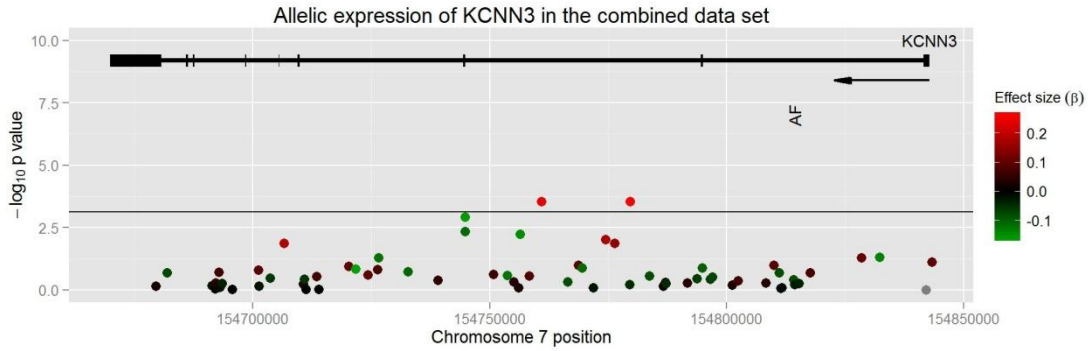


The significantly associated SNPs from analysis of the NE cohort (circles) are distributed over a 93kb region (black bar, top), whereas the significantly associated SNPs from the analysis of the SA cohort lie within a 7kb region.

4.5.11 Analysis of a combined data set provides greater power to identify aerQTLs

In order to maximise the power to detect aerQTLs at the 1q21 and 16q22 regions, the data from the NE and SA cohorts were combined to form a single data set. The phased haplotypes were used from the previous analysis. In the combined data set, two SNPs, rs11584635 and rs2798601, were significantly associated with a 1.27-fold and 1.30-fold increase in allelic expression of *KCNN3*, respectively ($p=2.9x10^{-4}$, $p=3.0x10^{-4}$) (Figure 4.18, Table A2.7). After correction for the effects of rs11584635, rs2798601 was not significantly associated with allelic expression.

Figure 4.18 Associations with expression at chromosome 1q21 in the combined data set

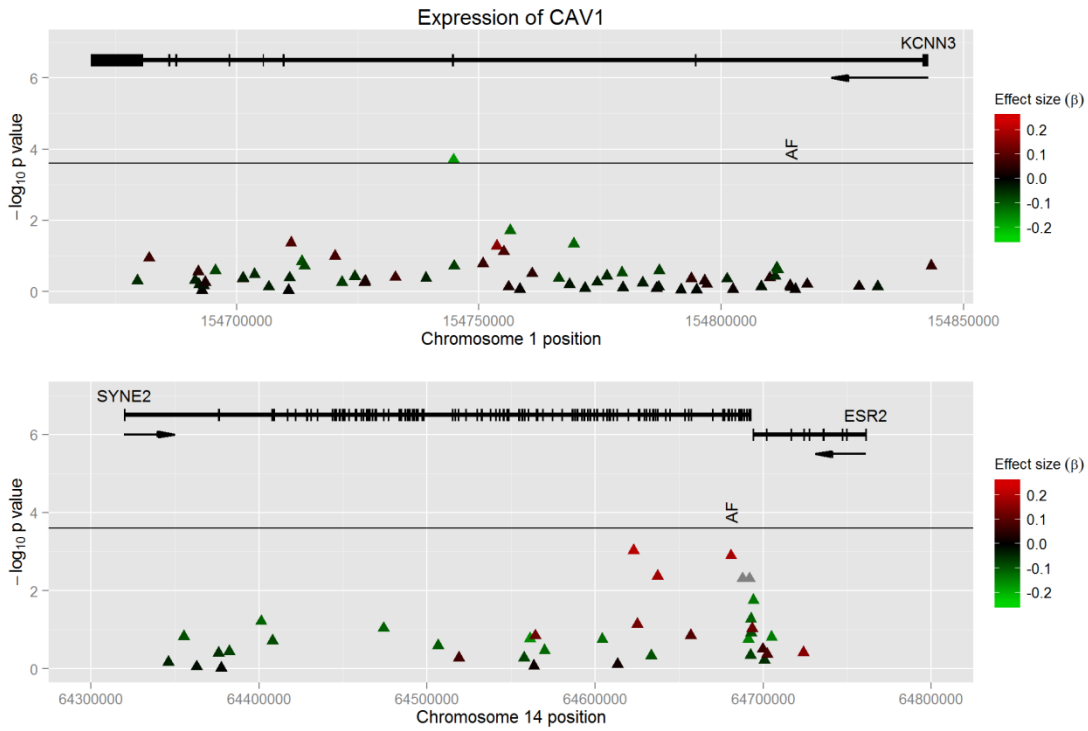


.

4.5.12 Trans-*acting eQTLs*

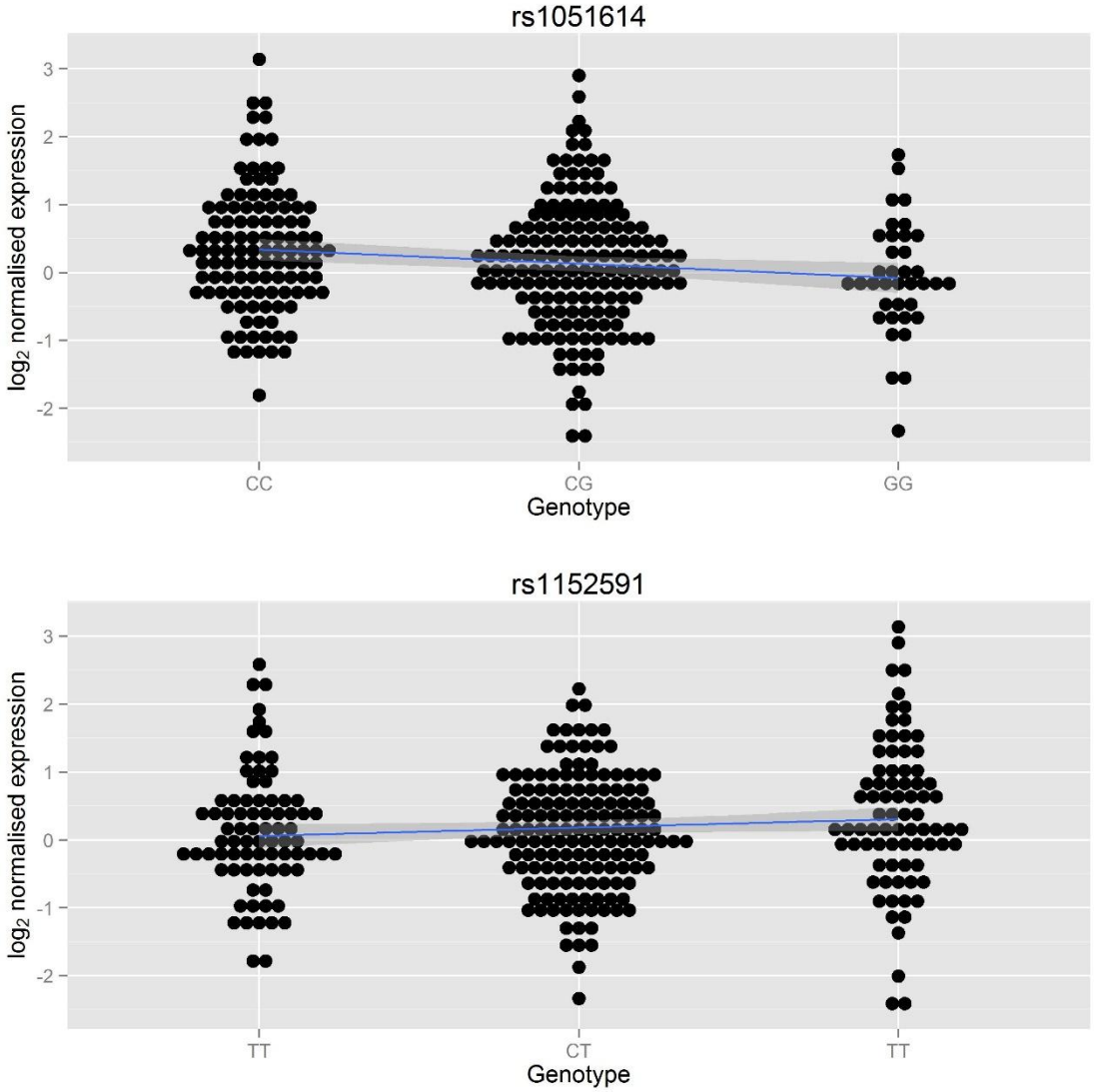
Trans eQTL mapping was also performed to identify genotype-gene expression associations between AF risk regions. There were two findings of interest. First, the minor allele of the transcribed SNP, rs1051614, in chromosome band 1q21 was associated with a 0.83-fold decrease in expression of *CAV1* ($p=2.0x10^{-4}$). Second, the risk allele of rs1152582 at 14q23 was associated with a 0.82-fold reduction in *CAV1* expression ($p=1.3x10^{-3}$),which was the same dissection of effect of the AF risk allele of rs3807989 on chromosome 7, but this did not reach the significance threshold for trans-acting associations (Figures 4.19 and 4.20, Table A2.8).

Figure 4.19 Associations with CAV1 expression in *trans*



Top: associations between expression and markers in chromosome 1. **Bottom:** associations between expression and markers in chromosome 14.

Figure 4.20 Normalised total gene expression of CAV1 by genotype



Top: Expression of CAV1 in chromosome 7q31 by genotype at the KCNN3 transcribed SNP rs1051614. **Bottom:** Expression of CAV1 in chromosome 7q31 by genotype at the KCNN3 transcribed SNP rs11152591. This association did not reach significance when allowing for multiple testing.

4.6 Discussion

This chapter describes a comprehensive investigation of the associations with expression phenotypes that may underlie the associations between common genetic variants and AF, using peripheral blood. The results demonstrate associations between GWAS hit SNPs for AF and gene expression at several loci. Risk variants for AF are associated with increased expression of *KCNN3* and *SYNE2*, and reduced expression of

CAV1. This strongly suggests that altered expression of *KCNN3*, *CAV1* and *SYNE2* may be important molecular genetic contributors to AF susceptibility. In addition, these results demonstrate an association between a GWAS hit SNP for primary open-angle glaucoma and increased expression of *CAV2*, suggesting that up-regulation of *CAV2* expression plays an important role in the pathogenesis of primary open angle glaucoma. The associations between disease associated SNPs and *KCNN3*, *CAV2* and *SYNE2* expression have not previously been reported in primary human tissues (Table 4.2).

Table 4.2 Previously published associations between AF GWAS SNPs and gene expression

SNP	Chromosome Band	Gene	LCL eQTL	Human tissue eQTL	eQTL or aerQTL in this study
rs6817105	4q25	*PITX2*			
rs7164883	15q24	*HCN4*			
rs2106261	16q22	*ZFHX3*			
rs6666258	1q21	*KCNN3*			Yes
rs3903239	1q24	*PRRX1*			
rs1152591	14q23	*SYNE2*	Yes[193]		Yes
		ESR2	Yes[193,194]	Yes[193,194]	
rs3807989	7q31	*CAV1*	Yes[193,194,196]	Yes[193,194]	Yes
		CAV2	Yes[193]		
rs10821415	9q22	*C9orf3*			
rs10824026	10q22	*SYNPO2L*			

4.6.1 eQTL and aerQTL mapping at AF-associated loci

On chromosome 1q21, the associations that were seen in the eQTL experiment differed significantly from those seen in the aerQTL experiment. The eQTL experiment

identified a strong association between a SNP, rs2335407, in the promoter region of *KCNN3* and *KCNN3* expression. This site has a score of 4 (minimal evidence for binding) in RegulomeDB, due to the presence of a DNase hypersensitivity binding cluster and transcription factor binding[288]. A variant at this position would be expected to act in *cis* by altering the binding of transcription factors to the promoter region. The absence of a signal at this SNP in the aerQTL experiment, even after the addition of the SA cohort to improve power, is surprising, therefore. There are three possible explanations for this discrepancy. The first explanation is that the allelic expression assay using rs1051614 has failed to measure expression of *KCNN3* adequately. This seems unlikely as the mass spectrometry traces show no evidence of salt or other interference peaks (Figure 4.18), the genotyping cluster plots using the same primers show clean clusters of genotypes, and an aerQTL was detected elsewhere in the gene, at rs11584635. Another possible explanation of this discrepancy is that the effect of this SNP is in *trans*. It has already been pointed out that aerQTL mapping is not able to detect *trans*-acting effects, as they would be expected to operate on both alleles. If rs2235407 affects expression in trans, therefore, the effect would not be detected by the aerQTL assay. This implies that there is a diffusible factor, such as a non-coding RNA, produced at or near rs2235407. There are no transcripts which fit this description annotated in the UCSC or Ensembl databases[145,283] or identified by RNAseq[289], however to explain this.

The most likely explanation, therefore is that the two experiments rely on the measurement of different transcripts. The TaqMan assay for *KCNN3* is situated on the exon 9-10 boundary whereas rs1051614 is situated in exon 3. Both assays measure all three RefSeq transcripts, but the UCSC genome browser reports an unconfirmed truncated splice variant which would not be detected by the qPCR assay. It is possible therefore that, if a sufficiently large amount of this splice variant were present, that the effect of rs2335407 on expression of the RefSeq transcripts would be masked in the allelic expression experiment. In order to determine whether this is the case, it would be necessary to repeat the AER experiment using one or more transcribed SNPs which are located in the distal exons of *KCNN3*, *i.e.* they would not be affected by the truncated splice variant. Ideally, one would use a transcribed SNP from Exon 9 or 10, where the TaqMan assay binds. This experiment was not performed as part of the

original study as the association with the AF GWAS SNP had not been identified and so this experiment lay beyond the scope of the original hypothesis.

4.6.2 Transcript-specific mapping of aerQTLs

The current study identified very different effects on gene expression when considering both transcripts of *ZFHX3* compared to analysis of *ZFHX3a* expression alone. SNPs at exon boundaries have been shown to result in exon skipping and production of functional proteins without affecting total gene expression levels[216]. Alterations in the relative expression of two transcript variants of a single gene may therefore not be detectable using an assay which measures both transcripts. Previous genome wide microarray eQTL analyses have identified the same association as this study between rs8060701 and *ZFHX3*[193,196], but the microarray probe is not transcript-specific, and would not be able to detect the effect of rs10852515 on *ZFHX3a.* Splice variants can have distinct functions. For example *Pitx2c* but not *Pitx2a* or *Pitx2b*, is expressed asymmetrically in the left lateral plate mesoderm in vertebrates[122] and is required for normal anatomical development of the atria[290]. This considerably complicates the analysis of mechanisms of action of disease-associated variants; candidate genes may have a great number of annotated, and possibly even more un-annotated, transcript variants; for example a recent analysis of *SYNE2* identified 14 isoforms, in contrast to the 4 RefSeq transcripts identified[138]. It is possible therefore, that the discrepancy between the results of the eQTL and aerQTL analyses of *SYNE2* is a result of the chosen assays binding to different isoforms of the gene. This study demonstrates that aerQTL mapping allows detection of transcript-specific effects and provides a useful technique for the investigation of the effects of genetic variants on genes where multiple transcript variants are present.

4.6.3 Trans-ethnic fine mapping

Fine localisation of variants associated with phenotypic traits can be limited by linkage disequilibrium in the region of interest. Greater haplotype diversity and fine LD structure therefore allows improved localisation of eQTLs. African populations, and in

particular the Cape Mixed Ancestry population of South Africa, as a result of a mixed ancestral background, have a greater haplotype diversity than European populations. Therefore a trans-ethnic fine-mapping approach comparing different populations would be expected to allow better identification of functional variants than mapping in a single population. [291,292]. This approach has been demonstrated previously[293]. The SA cohort used for this experiment has previously been used for fine-mapping of aerQTLs of the long non-coding RNA, ANRIL, at chromosome 9p21, where the refinement of the aerQTL was less dramatic than the effect shown here at 16q22. This demonstrates that the power of trans-ethnic fine-mapping to refine the aerQTL signal is dependent on the differences in LD structure between the two populations in the region(s) of interest, and will be variable throughout the genome

Combined analysis of data from different cohorts has also been previously shown to improve the power to detect eQTLs[221]. The current study also demonstrates the utility of this approach. Combined analysis of the NE and SA cohorts, by increasing the number of individuals studied, allowed identification of a *cis*-acting aerQTL for KCNN3 which was not possible in either cohort when analysed separately.

4.6.4 Trans acting eQTLs between chromosomal regions

The current study identified a single *trans*-acting variant in *KCNN3* which appeared to be associated with *CAV1* expression. rs1051614 is a synonymous variant, but it is situated within 1000bp of four non-synonymous coding variants. It is possible, therefore, that the change in *CAV1* expression could be effected, directly or indirectly, by a conformational change in the SK3 Ca^{2+}-gated K^{+} channel. The degree of LD between rs1051614 and nearby coding variants has not been established in the HapMap or 1000Genomes studies, however. The alternative explanation is that the *trans* acting eQTL operates via altered expression of *KCNN3*, which in turn affects *CAV1* expression. This seems less likely as no change in *KCNN3* expression was demonstrated to be associated with genotype at rs1051614.

No other SNPs in the chromosome 1q21 region, including the AF associated SNP rs13376333, were associated with or had a tendency to be associated with *CAV1* expression. In light of this, the association seen between rs1051614 and *CAV1*

expression must be treated with some degree of circumspection and replication and/or functional analysis would be of interest.

The association between the AF-associated SNP in chromosome 14q23, rs1152591, and *CAV1* expression did not reach the threshold for significance allowing for multiple testing. The direction of effect was in keeping with the results for the chromosome 7q31 SNPs, which supports the finding, but replication and functional evidence would be important to rule out association by chance alone.

4.6.5 The utility of peripheral blood for investigating AF risk variants

It is important to note that in most cases, the increased risk of AF will not be mediated by changes in gene expression in blood, but instead in another, less-accessible tissue, probably in the atrium itself. A number of factors might preclude the detection of associations between risk SNPs that are in fact causally related to AF risk and gene expression in blood. First, candidate genes may not be expressed in blood at sufficient levels to accurately measure expression differences. Second, eQTLs identified in blood may not be present in other tissues and *vice versa*: the MuTHER study estimated that 56–83% of eQTLs were shared across tissues, with the remainder displaying tissue specificity[193]. Therefore eQTL mapping in peripheral blood is an imperfect surrogate for eQTL mapping in the relevant tissues. Nevertheless, peripheral blood remains the most easily accessible human tissue and even though the absence of an eQTL in blood does not exclude an eQTL in another tissue, the presence of an eQTL provides evidence that a variant regulates expression and that this regulatory effect may operate in other tissues.

Primary candidate genes in only five of nine regions associated with AF were expressed in peripheral blood. Therefore, in almost half of the regions of interest, eQTL and aerQTL mapping in peripheral blood was not possible, providing no insight to the mechanisms of action of these risk variants. This was the case for variants in chromosome 4q25, the most strongly associated AF risk locus in all of the GWAS published to date. In addition no association between gene expression and AF risk variants was demonstrated at two of the five regions studied. Fine mapping of expression-genotype associations in blood has, therefore, allowed identification of

potential risk mechanisms in three of the nine regions identified. This is a modest improvement on the results from microarray studies in LCLs, which identified eQTLs for AF risk SNPs at two of the risk regions, *CAV1* and *CAV2* in chromosome band 7q31 and *ESR2* in chromosome band 14q22[193,196]. The difficulty in identifying risk mechanisms is not, perhaps, wholly surprising. In the MuTHER study, GWAS hits for inflammatory and auto-immune diseases were more likely to be identified as eQTLs in LCLs than in adipose tissue or skin[193]. eQTL mapping in atrial tissue, therefore, is more likely to identify AF risk mechanisms, as is demonstrated in Chapter 6.

4.6.6 Mechanisms of action of AF-associated SNPs

The AF risk allele was associated with a decrease in expression of *KCNN3*. *KCNN3* codes for SK3, one of three members of the Ca^{2+}-gated K^+ channel family. All three subunits are expressed in human atrium and from heteromultimers via the interaction of coiled-coil domains in the C termini[111,294]. The SK channels significantly contribute to the repolarisation phase of the cardiac action potential[111]. In a mouse model, over-expression of SK3 resulted in shortened action potential duration, increased inducibility of atrial arrhythmias and increased frequency of sudden death[295,296]. Any perturbation of SK3 levels may result in altered stoichiometry of SK2/SK3 heteromultimers, which, in turn, may affect repolarisation in ways which are not easily predictable. Either over- or under-expression, therefore, may result in loss of function of the SK3/SK2 ion channel, and therefore provide the same arrhythmic substrate.

This study demonstrates an association between decreased *CAV1* but not *CAV2* expression and the risk alleles of AF-associated SNPs in chromosomes 7 and 14. *CAV1* and *CAV2* are widely expressed in human tissues and encode caveolins 1 and 2, which are involved in the formation of caveolae. Caveolae are specialised invaginations of the cell membrane with an important role in the transcytosis of macromolecules, signal transduction and cholesterol transport in various cell types[297-299]. Atrial fibrosis is an important component in the pathophysiology of AF[300]. A recent study of *CAV1* and *CAV2* expression in human right atrial appendage demonstrated that *CAV1* but not *CAV2* expression was reduced in patients with AF, which is also demonstrated in Chapter 5, and that TGF-β1 stimulation of atrial fibroblasts *in vitro* suppressed *CAV1*

expression. A *CAV1* gain-of-function assay showed attenuation of the TGF-β1 pro-fibrotic signalling pathway, suggesting that *CAV1* is an important anti-fibrotic mediator which is protective against AF[301]. The chromosome 7 AF-associated SNP was also associated with *CAV1* but not *CAV2* expression in the adipose tissue and skin arms of the MuTHER study[193].

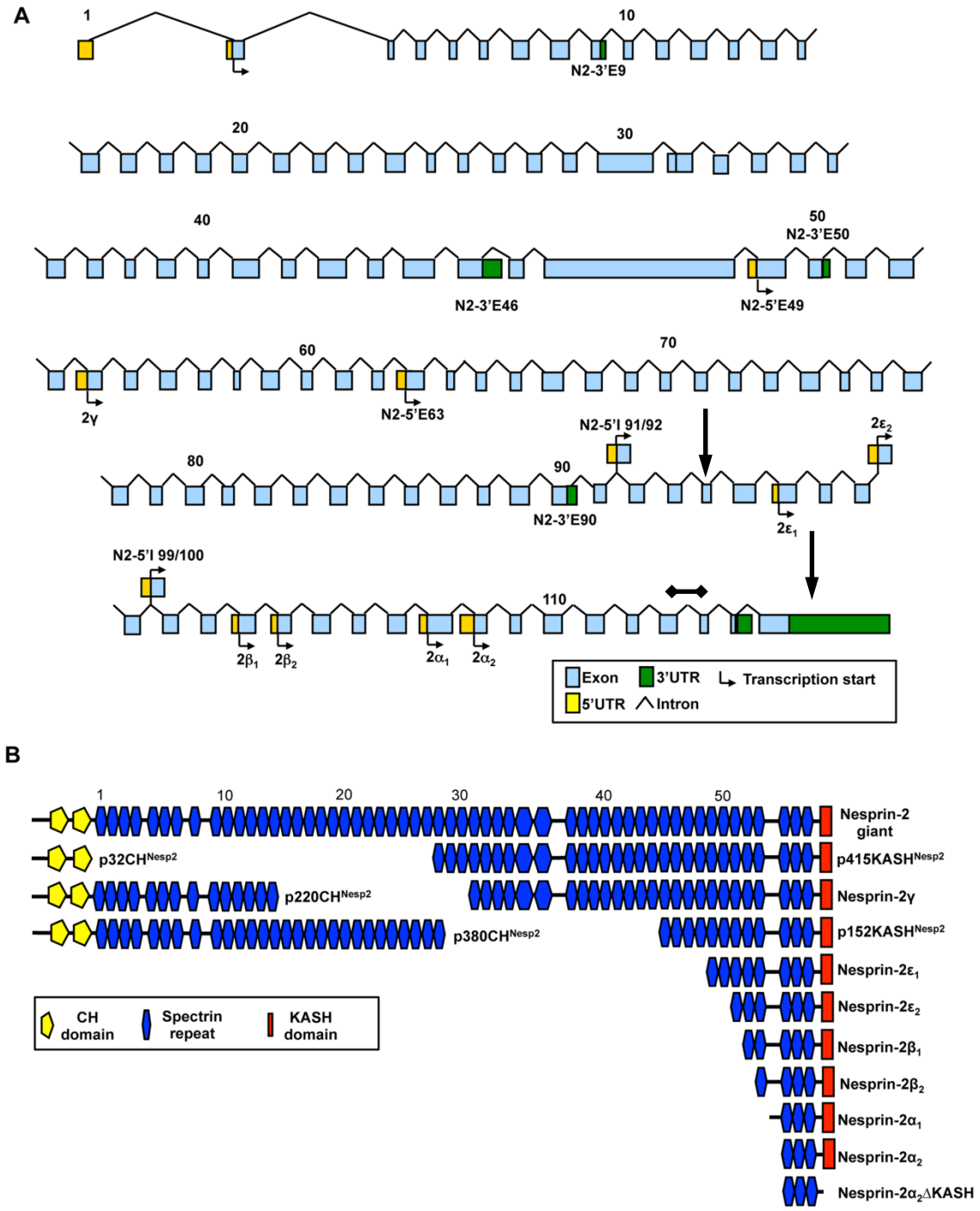
The importance of relative expression of *CAV1* and *CAV2* is indicated by identification of opposite effects on expression of a variant in the region, suggesting that some regulatory sites act to increase expression of one gene and suppress expression of another.

We have demonstrated variants at chromosome 9q22 are associated with *C9orf3* expression in blood, but that the GWAS-associated SNP for AF is not. Given that blood is not the primary site of expression of *C9orf3*, it is possible that the AF-associated SNP regulates expression in another tissue type. It has been estimated that only 56-83% of eQTLs are present across multiple tissue types and genome-wide attempts to identify eQTLs at GWAS hit sites have found greater numbers of associations in tissues which are relevant to the disease phenotype[193].

The results from aerQTL mapping and eQTL mapping of CAV1, the regulatory variants identified in the aerQTL experiment were very significantly associated with allelic expression ($p=6x10^{-37}$) but not at all associated with total expression. This is a little difficult to explain by differences in power between the two approaches alone. There are two likely explanations for this. First, the power of the eQTL mapping may be adversely affected by very large variance in total expression, from a combination of *trans*-acting and experimental factors. However, the variance in *SYNE2* expression and the estimated proportion of the total variance due to *cis*-acting factors was similar to that seen for other transcripts, making this unlikely. Second, it is possible that the two assays did not measure the same transcripts. The qPCR probe and rs1152582 both measure the same four RefSeq transcripts of *SYNE2*, but are not situated on the same exon. It has recently been shown that *SYNE2* encodes multiple transcripts which have diverse structural roles within the cell and that these are expressed in tissue-specific patterns (Figure 4.19)[138]. These new isoforms were identified in brain, skeletal muscle and HeLa cDNA libraries; more isoforms may be found in other tissue types. It is

possible that the AF risk SNP is associated with an altered relative abundance of these different transcripts, possibly in a tissue specific manner. The allelic expression assays used here may, therefore, capture an effect on expression that does not act on the exons which are bound by the qPCR assay.

Figure 4.19 Nesprin 2 isoforms



Reproduced from Rajgor *et.al.*[138]. Many new transcript variants of *SYNE2* are identified, which encode multiple different Nesprin-2 isoforms. CH domain: calponin homology domain, which binds directly to F-actin. Spectrin repeat: a structural domain, common in cytoskeletal

proteins. KASH domain: Klarsicht/ANC-1/Syne homology transmembrane domain, which acts as a nuclear envelope targeting motif. **A:** genomic map of *SYNE2* with position of transcription start sites and UTSs identified to date. The transcribed SNPs rs7161192 and rs1152582 are located in exon 94 and the exon 116 3'UTR, respectively (arrows). The qPCR probe binds to the exon 113-114 boundary (bar). **B:** Nesprin-2 isoforms. There are a large number of protein isoforms which have differing structural components and patterns of expression.

It was not possible to demonstrate an association between expression of *ZFHX3* or *ZFHX3a* and AF risk SNPs. This study did demonstrate a transcript-specific aerQTL for *ZFHX3* expression. Changes in the relative expression of the two isoforms of *ZFHX3* have been shown to determine the rate of differentiation of C2C12 myoblasts, the *B* transcript accelerated myogenic differentiation and the *A* isoform resulted in persistence of an undifferentiated phenotype [302]. It is likely that the portion of the gene common to both transcripts, i.e. that part which makes up the *B* isoform, is responsible for tumour-suppressor functions, as it contains the site of interaction with the oestrogen receptor in breast cancer and is the most common site for somatic mutations in gastric carcinoma [149,303]. This suggests that transcript-specific regulation of *ZFHX3* expression plays an important role in control of differentiation and cellular proliferation.

4.7 Conclusions

Results of experiments in this chapter have identified multiple associations between gene expression and genetic variants associated with AF. AF risk variants are associated with increased expression of *KCNN3* and *SYNE2* and decreased expression of *CAV1*. This provides further evidence for the roles of *CAV1*, *SYNE2*, and *KCNN3* in the pathophysiology of AF, and identifies them as targets for further study.

Chapter 5

Influence of polymorphisms at AF-associated loci on gene expression in cardiac and vascular tissues

5 Influence of polymorphisms at AF-associated loci on gene expression in cardiac and vascular tissues

5.1 Abstract

The previous chapter investigated the associations between genotype of AF risk variants and total and allelic expression of candidate genes in peripheral blood. Not all of the candidate genes identified in the AF risk regions are expressed in peripheral blood. In addition, it is possible that the effects of AF risk variants are tissue specific. This chapter examines the effect of AF risk variants on total expression and AER in left atrial appendage, right atrial appendage and internal mammary artery tissue, using the methods described in Chapter 2.

Total expression of 16 transcripts and allelic expression ratio of 15 transcripts were measured in RAA, LAA and IMA. Expression of *PITX2* was significantly increased in left versus right atrial appendage ($p=4x10^{-4}$), whereas expression was greater in the right atrial appendage for *HCN4* ($p=4x10^{-5}$), *C9orf3* ($p=3x10^{-4}$) and *FANCC* ($p=4x10^{-4}$). A history of AF was associated with reduced expression of *SYNE2* (p=0.003), *HCN4* (p=0.01) and *CAV1* (p=0.03). AF risk variants were associated with increased expression of non-c isoforms of *PITX2* ($p=6x10^{-4}$), and decreased expression of *MYOZ1* ($p=6x10^{-15}$), *CAV1* ($p=6x10^{-8}$), *C9orf3* ($2x10^{-5}$), and *FANCC* ($p=9x10^{-8}$).

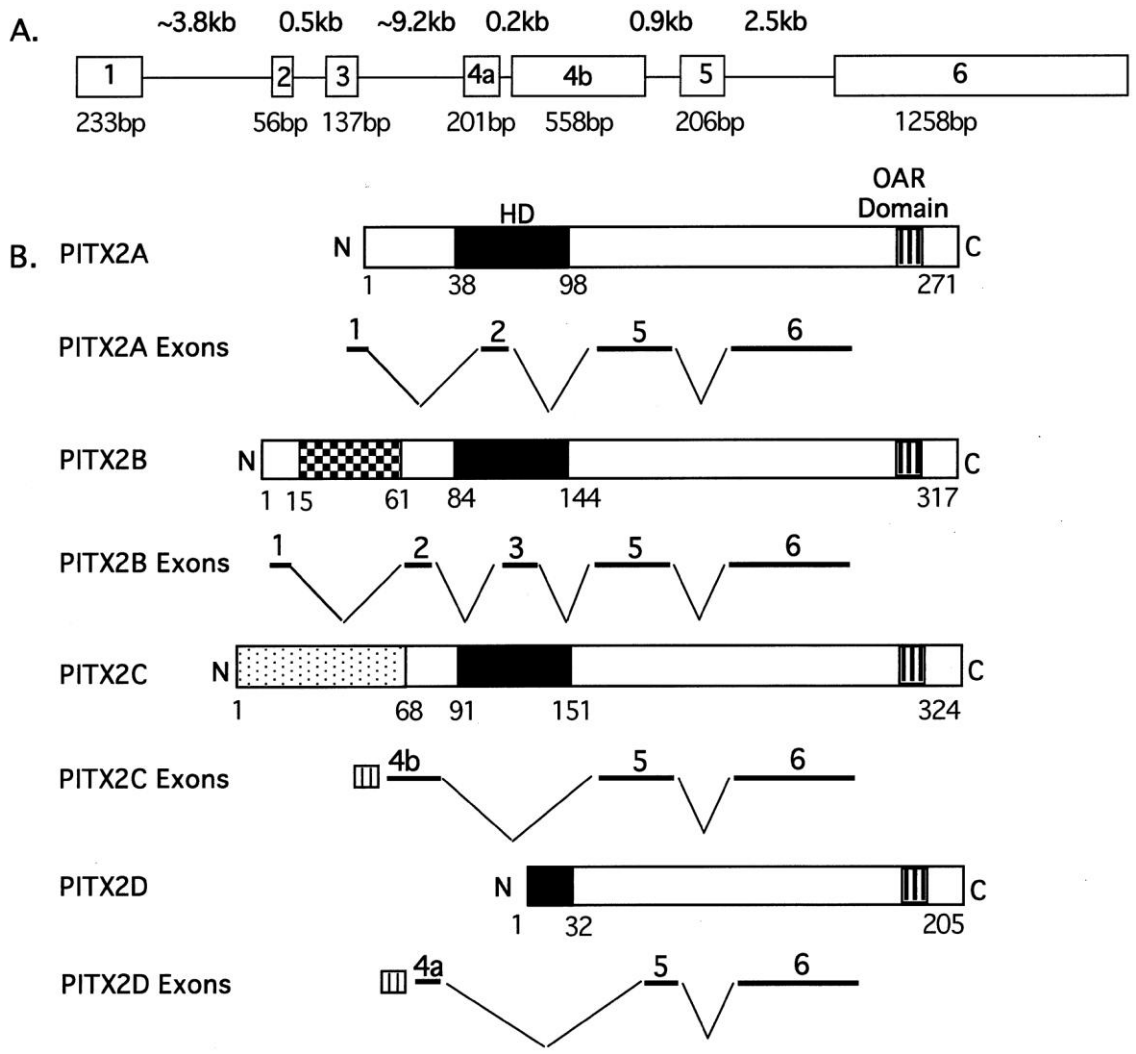
These results confirm the association between decreased levels of *CAV1* and AF risk variants, identified in Chapter 4. Additionally they demonstrate an association between AF risk variants and increased expression of the non-c isoforms of *PITX2* as well as decreased expression of *MYOZ1*, *CAV1* and *FANCC*.

5.2 Background

The previous chapter investigated the associations between genetic variation in regions which have been associated with AF by GWAS studies and gene expression in peripheral blood. As discussed in that chapter, peripheral blood, whilst having the virtue of being easily accessible, is an imperfect surrogate for investigating the effects of genetic variation on gene expression in more disease relevant tissues. In four of nine regions identified by GWAS[109], the primary candidate genes were not adequately expressed in peripheral blood to allow eQTL or aerQTL mapping. In addition, in two of the regions studied in the previous chapter, no effect on expression of the AF-associated SNPs could be identified. One of the more likely explanations for this is that the effects of the AF-associated variants are specific to a tissue more relevant for disease, namely the atrium, and therefore have no measurable effect in peripheral blood.

The importance of transcript-specific effects on expression was discussed in Chapter 4. Of the candidate genes identified in each of the nine regions two, *HCN4* and *MYOZ1*, have a single RefSeq transcript variant[283]. Most of the candidate genes have multiple transcripts, which encode different proteins. For *ZFHX3* and *SYNE2*, there is evidence that the different isoforms have different functions or patterns of expression[138,302]. PITX2 is a special case however; firstly because it is the candidate gene for the strongest AF risk locus, and secondly because there is specific evidence that the different isoforms have differing roles in cardiac tissue. The *PITX2* gene encodes four major isoforms of the PITX2 protein in humans, PITX2a, PITX2b, PITX2c and PITX2d, which play roles in development of the heart, as well as the eye, gut, spleen, lungs and in limb myogenesis[118]. PITX2a and PITX2b share a promoter region and are differentially spliced. PITX2a consists of exons 1, 2, 5 and 6; PITX2b also includes exon 3[122]. PITX2c has a different promoter region and consists of exons 4, 5 and 6. All three isoforms share the same C-terminus and homeobox domain (Figure 5.1). PITX2d also has a different transcription start site and consists of part of exon 4 and exons 5 and 6; it does not contain the nucleic acid binding homeobox domain, and has an inhibitory effect on the other isoforms[123].

Figure 5.1 PITX2 structure



A, genomic organization of the *PITX2* gene; intron sizes are shown on the *top*, and exon sizes are at the *bottom*; exons are numbered. *B*, the protein structure is shown with the location of the homeodomain (*HD*) and 14-amino acid conserved OAR domain. Chequered and stippled *boxes* denote the differences in the N-terminal region of the isoforms. The exons that code for the respective proteins are shown *below* each isoform. PITX2C and PITX2D RNA is transcribed using an internal promoter shown as a *striped box* flanking exon 4. Figure from Cox *et al.*[123]

PITX2c is the only isoform to be asymmetrically expressed in the left lateral plate mesoderm, consistent with its role in the development of the left-right axis.[122] PITX2c is the only isoform to be identified in RNAseq studies of human atrium thus far[124], however a qPCR study demonstrated that the other isoforms of PITX2 are also expressed, predominantly PITX2a, without significant differences between LAA and

right atrial appendage (RAA)[125]. As a result, PITX2c has been considered the most important cardiac isoform, and the roles of PITX2a and PITX2b in the developing and adult atrium have been less well studied.

The collection of disease-relevant tissue for eQTL mapping of GWAS hits for cardiac disease is necessarily problematic. For many other organ-specific diseases, tissue biopsy is a safe and routine part of clinical care and surplus biopsy material can be used for RNA analysis. For cardiac disease, with the possible exception of allograft rejection monitoring, there is no definite indication for cardiac biopsy, due to the relatively high risk and low diagnostic yield[304]. Alternative sources of tissue that have been used for eQTL experiments include post-mortem samples and explanted organs from cardiac transplant recipients[130,305,306]. A more easily available source of cardiac tissue for analysis is provided at the time of cardiac surgery. All patients requiring cardiopulmonary bypass at the time of cardiac surgery have at least one cannula placed in the right atrium to drain blood from the heart. The most common access point is through the right atrial appendage (RAA), and tissue is routinely discarded from the cannulation site to facilitate closure. In addition, many patients undergoing surgery involving the mitral valve have the left atrial appendage (LAA) removed to prevent formation of left atrial appendage thrombus and subsequent stroke in the event that they have or develop AF[307]. A third source of tissue at the time of cardiac surgery is the internal mammary artery (IMA). The left IMA (LIMA) is preferentially used to bypass atherosclerotic lesions in the left anterior descending artery[308]. When the LIMA is harvested, the terminal portion is removed and discarded in order to provide a clean end to the artery with which to perform an anastomosis. None of these approaches allows the collection of cardiac tissue from healthy individuals. Collection of tissue at the time of cardiac surgery has two advantages. First, large numbers of cardiac operations are performed in many centres, meaning that collection of a sufficiently large cohort for eQTL mapping is practicable. Second, there is no additional risk to the tissue donors as the tissue collected would be routinely discarded in the course of their usual care.

5.3 Aims

1. To determine which of the candidate genes at each locus are expressed in RAA, LAA and IMA.

2. To compare expression levels of candidate genes between RAA, LAA and IMA.

3. To perform eQTL and aerQTL mapping of candidate genes.

4. To determine the concordance of the eQTL and aerQTL mapping results with the results from peripheral blood.

5.4 Methods

5.4.1 Participants and ethics statement

Study participants and ethical approvals were as set out in Chapter 2. Participants were recruited from volunteers undergoing cardiac surgery at five centres in the north of England. 52 of the participants were also included in the NE cohort used for peripheral blood analysis in the previous chapter.

5.4.2 DNA and RNA extraction

Nucleic acid extraction and cDNA synthesis was performed as described in Chapter 2.

RNA integrity was determined on a random subset of samples using the Agilent 2100 bioanalyser, as set out in Chapter 2.

5.4.3 Selection of transcribed SNPs for allelic expression analysis

For regions investigated in peripheral blood, the same transcribed SNPs were used for allelic expression ratio (AER) analysis. For genes expressed in atria but not peripheral blood, the following SNPs were identified for use in AEI analysis: rs58473244 and rs3820416 in *PRRX1*, rs913257 and rs7552922 in *GORAB*, rs2739200 in *PITX2a/b*, rs6533526 and rs75911264 in *PITX2* (all isoforms), rs60632610, rs4746139, rs3812629 and rs34163229 in SYNPO2L, rs41280400 in MYOZ1, rs529004, rs35177144 and rs3743496 in HCN4. All of the transcribed polymorphisms in PITX2 were present either

in exons common to PITX2a and PITX2b or in exons common to all three isoforms. Therefore it was not possible for AER to be assessed for any transcript individually.

5.4.4 Selection of mapping SNPs

For regions investigated in peripheral blood, the same mapping SNPs were used. Tag SNPs were selected for the other regions using the same approach described in Chapter 2.

5.4.5 Genotyping

The assays described in Chapter 4 were used to genotype SNPs in the regions studied previously. Genotyping of 91 SNPS in the other regions was performed in 4 multiplex reactions on the Sequenom platform as described in Chapter 2. For one SNP, rs10824026, where genotyping could not be performed on the Sequenom platform, genotyping was attempted using a custom TaqMan SNP genotyping assay (C__31474386_10), as described in Chapter 2. As this approach was also unsuccessful genotypes for this variant were imputed using BEAGLE software[242] and phased genotype data in BEAGLE format from phase 1 of the 1000 Genomes Project[309].

5.4.6 Measurement of allelic expression ratio and total expression

The assays described in Chapter 4 section 4.4.6 were used to measure AER in the regions studied previously. Measurement of AER of the other transcripts was performed in 3 multiplexes as set out in Chapter 2, Section 2.3.5. Assays were identical to those used for genotyping as the amplicons for AER measurement were contained wholly within a single exon. Primer sequences are shown in Appendix 1. The TaqMan assays described in Chapter 4, Section 4.4.7 were used to measure total expression in the regions previously studied. In addition, normalised gene expression was measured for the non-primary candidate genes in each region, and all candidate genes in the regions not previously studied, *i.e. PRRX1*, *GORAB*, *PITX2*, *CAV2*, *FANCC*, *SYNPO2L*, *MYOZ1*, *HCN4*, *ESR2*, and *ZFHX3a*. As *PITX2a* and *PITX2b* are reported as having differing expression patterns to *PITX2c*[122] and given that *PITX2c* has been identified as

the most important isoform for cardiac development, a single assay was selected to measure *PITX2a* and *PITX2b* expression together (*PITX2a/b*) and an alternative assay to measure *PITX2c* expression. In all, total expression of 16 transcripts was measured in two separate experimental plate arrangements, with expression levels normalised to expression of the reference genes *ACTB*, *GAPDH* and *HPRT1*. Assay details are given in Table A1.1.

5.4.7 Statistical analyses

M values, indicating stability of reference genes, were calculated using the geNorm algorithm in the Excel software package (Microsoft, USA)[234]. All statistical analyses were performed as described in chapters 2 and 4. Phasing was performed using the BEAGLE software package, aerQTL analysis was performed using the R and Matlab software packages and eQTL mapping was performed using the R software package as described in Chapter 2. The significance threshold for regulatory effects on expression was determined using the Bonferroni correction as described previously. The relative contribution from *cis*-acting influences on total expression was determined as described in Chapter 2.

5.5 Results

5.5.1 Comparison of reference gene expression between RAA, LAA and IMA

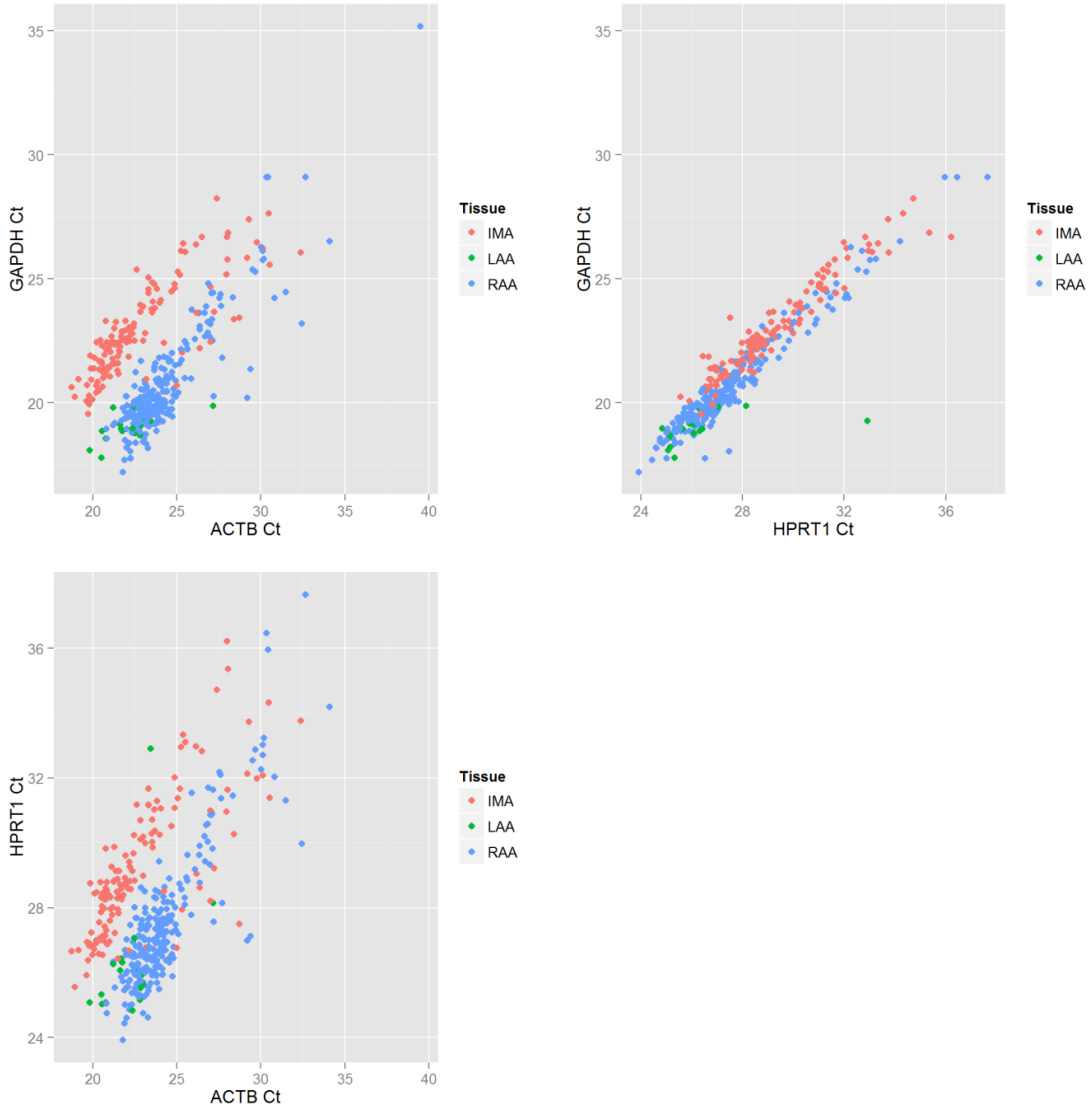
Reference gene Ct values were well correlated within each tissue type (Figure 5.2). ACTB showed a greater degree of variability (as indicated by higher *M* values) than HPRT1 and GAPDH, particularly between tissues. Expression levels of *ACTB* were not stable (*M*>1.5) in LAA and across tissue types, therefore expression levels were normalised to GAPDH and HPRT1 levels alone when making comparisons between tissue types.

Table 5.1 Comparison of reference gene stability within and between tissues

	M value			
Gene	RAA	LAA	IMA	All Tissues
ACTB	1.191	1.958	1.386	2.222
GAPDH	0.799	1.357	0.963	1.490
HPRT1	0.796	1.222	1.014	1.405

ACTB expression was less stable in all tissues and when samples were compared across tissue types.

Figure 5.2 Stability of reference gene expression

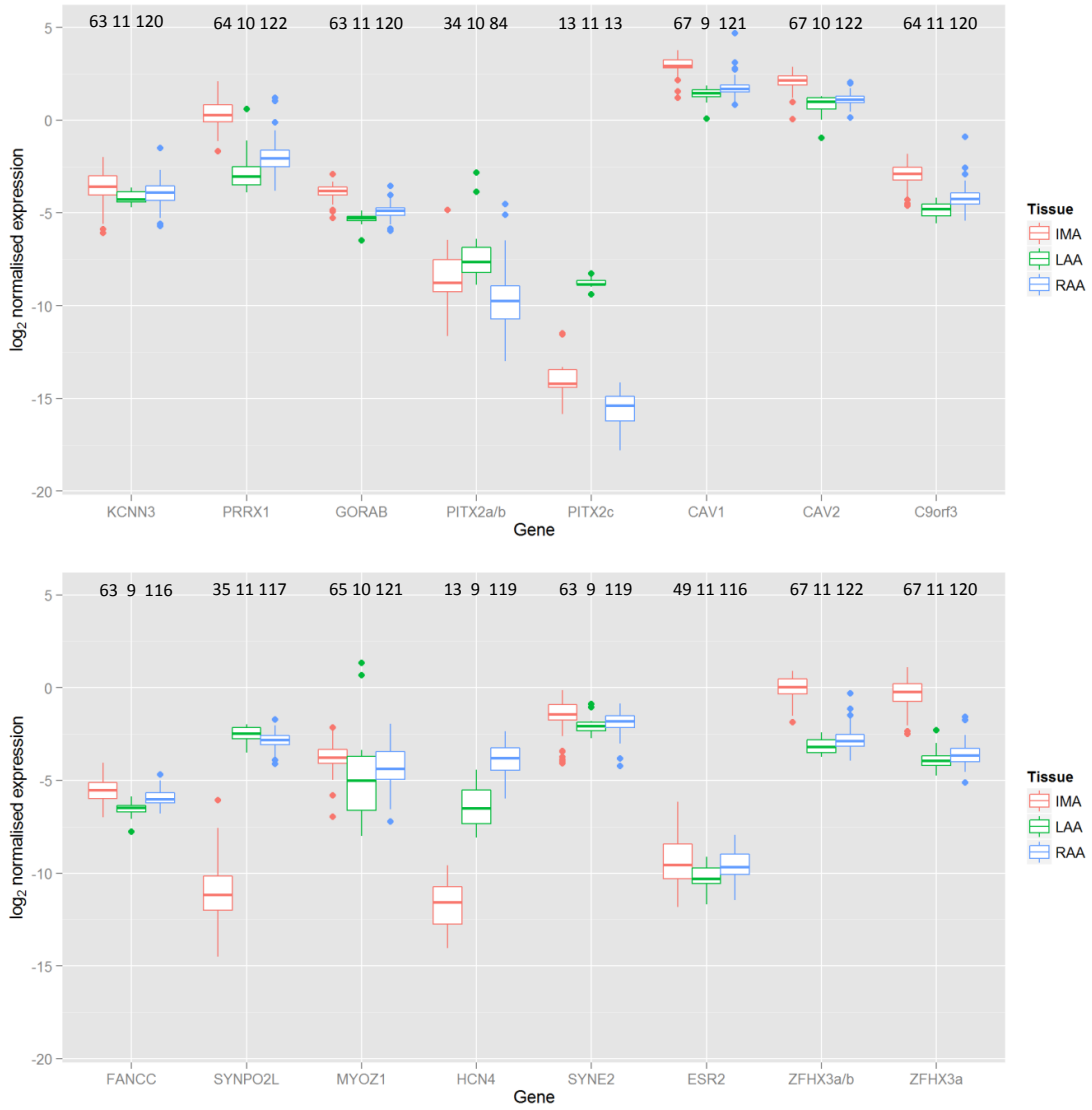


Comparison of Ct values of the three reference genes by tissue type. Each point represents a single sample of RNA from IMA (red), LAA (green), or RAA (blue).

5.5.2 Comparison of total expression of candidate genes between RAA, LAA and IMA

With the exception of PITX2 and ESR2, expression levels in left and right atrium were more alike than in IMA. The significance levels for the differences in expression are presented in Table 5.2. PI*TX2a/b* and *PITX2c* were expressed at increased levels in LAA compared with RAA. *C9orf3*, *FANCC* and *HCN4* were expressed at lower levels in LAA compared to RAA. Expression levels of *PITX2c* in RAA and IMA, *HCN4* in IMA and *SYNPO2L* in IMA were too low to allow reliable measurement of expression levels and therefore they were not included in further analyses (Figure 5.3, Table 5.2).

Figure 5.3 Comparison of candidate gene expression by tissue



Boxplots are shown indicating expression of candidate genes in each of the three tissue types. The number of observations on which each boxplot is based are indicated above.

Table 5.2 Comparison of gene expression between tissues

Gene	fold-change: LAA vs RAA	p value	fold-change: IMA vs RAA	p-value
KCNN3	0.86	0.08	1.22	0.03
PRRX1	0.63	0.11	**5.05**	**$3.6x10^{-44}$**
GORAB	0.72	0.002	**2.03**	**$4.7x10^{-34}$**
PITX2a/b	**6.30**	**$4.6x10^{-4}$**	**2.21**	**$2.8x10^{-4}$**
PITX2c	**120**	**$5.1x10^{-11}$**	**3.80**	**$4.0x10^{-4}$**
CAV1	0.79	0.03	**2.37**	**$2.0x10^{-37}$**
CAV2	0.78	0.1	**1.96**	**$2.9x10^{-29}$**
C9orf3	**0.64**	**$2.6x10^{-4}$**	**2.40**	**$3.3x10^{-28}$**
FANCC	**0.62**	**$4.1x10^{-4}$**	**1.31**	**$5.1x10^{-5}$**
SYNPO2L	1.22	0.05	**0.004**	**$2.6x10^{-26}$**
MYOZ1	0.80	0.70	**1.41**	**$2.0x10^{-4}$**
HCN4	**0.17**	**$3.5x10^{-5}$**	**0.004**	**$2.7x10^{-11}$**
SYNE2	0.93	0.55	1.30	0.003
ESR2	0.62	0.01	1.27	0.12
ZFHX3a/b	0.79	0.02	**6.92**	**$1.3x10^{-64}$**
ZFHX3a	0.84	0.21	**9.63**	**$1.3x10^{-53}$**

Expression levels between tissue types were compared using an unpaired t-test. Mean fold-change and p values are shown for comparisons of gene expression between LAA and RAA and IMA and RAA. Values which reached statistical significance threshold of $1.0x10^{-3}$, calculated using the Bonferroni correction for 48 comparisons (16 genes in three tissues), are highlighted **in bold.**

5.5.3 Associations between AF status and gene expression.

The current study did not aim to detect changes in gene expression associated with AF, therefore only a small number of study participants had a history of AF (17/123 RAA samples, 6/12 LAA samples and 7/67 IMA samples). None of the associations between AF status and expression level remained significant after a Bonferroni correction. However, in participants with a history of AF compared to those without, expression of *CAV1*, *HCN4* and *SYNE2* was reduced in RAA and expression of *CAV1* was reduced in LAA at a nominal significance level of 0.05 (Table 5.4).

Table 5.4 Change in gene expression associated with a history of AF

Gene	Fold change in RAA	p value	Fold change in LAA	p value	Fold change in IMA	p value
KCNN3	0.9 (0.73-1.12)	0.34	1.07 (0.72-1.59)	0.70	0.76 (0.52-1.1)	0.13
PRRX1	0.83 (0.67-1.03)	0.08	1.29 (0.92-1.8)	0.11	0.96 (0.77-1.21)	0.74
GORAB	0.95 (0.79-1.14)	0.57	1.15 (0.86-1.52)	0.29	1.05 (0.82-1.35)	0.66
PITX2a/b	1.38 (0.81-2.34)	0.22	1.25 (0.47-3.33)	0.54	0.74 (0.2-2.78)	0.59
PITX2c			0.77 (0.48-1.22)	0.22		
CAV1	**0.85 (0.72-0.99)**	**0.03**	**0.84 (0.72-0.98)**	**0.03**	1.12 (0.85-1.46)	0.38
CAV2	0.97 (0.84-1.12)	0.68	0.94 (0.73-1.22)	0.59	1.21 (0.94-1.56)	0.13
C9orf3	0.96 (0.79-1.16)	0.65	1.5 (0.96-2.34)	0.07	1.01 (0.82-1.26)	0.90
FANCC	0.87 (0.72-1.05)	0.14	1.04 (0.77-1.41)	0.73	1.11 (0.78-1.6)	0.52
SYNPO2L	1.05 (0.89-1.24)	0.52	1.06 (0.65-1.74)	0.77		
MYOZ1	1.16 (0.74-1.82)	0.50	0.84 (0.11-6.48)	0.85	1.17 (0.84-1.63)	0.31
HCN4	**0.64 (0.46-0.88)**	**0.01**	1.19 (0.33-4.29)	0.76		
SYNE2	**0.76 (0.64-0.91)**	**0.003**	1.28 (0.58-2.8)	0.46	1 (0.66-1.52)	0.98
ESR2	0.78 (0.57-1.06)	0.11	1.58 (0.64-3.9)	0.27	0.58 (0.27-1.28)	0.16
ZFHX3a/b	0.95 (0.8-1.12)	0.50	1.28 (0.98-1.67)	0.06	1.04 (0.86-1.25)	0.68
ZFHX3a	0.97 (0.81-1.16)	0.71	1.29 (0.91-1.84)	0.12	1.07 (0.7-1.63)	0.72

Expression levels between AF cases and controls were compared using an unpaired t-test. Mean fold changes are presented as expression in individuals with a history of AF compared to those without. 95% confidence intervals for the fold change are given in brackets. Significant associations ($p<0.05$) are highlighted **in bold.**

5.5.4 Variants in chromosome 1q21 are associated with KCNN3 *expression*

Total expression and allelic expression ratios of *KCNN3* were measured in RAA, LAA and IMA. No associations were detected by eQTL analysis. However, aerQTL mapping of this region identified six SNPs which were significantly associated with expression of *KCNN3* in RAA, of which two were significantly associated with expression in IMA and one in LAA. The most strongly associated SNP in all three tissues was the transcribed

SNP rs1051614. The minor allele was associated with a 0.73-fold decrease in *KCNN3* expression in RAA, a 0.70-fold decrease in LAA and a 0.74-fold decrease in IMA ($p=5.4x10^{-21}$, $p=6.3x10^{-8}$ and $p=4.9x10^{-5}$, respectively). The AF-associated variant, rs13376333, was not associated with expression in any tissue (Figures 5.4 & 5.5, Tables A2.9 & A2.10).

Figure 5.4 Associations with total expression of *KCNN3* in chromosome band 1q21

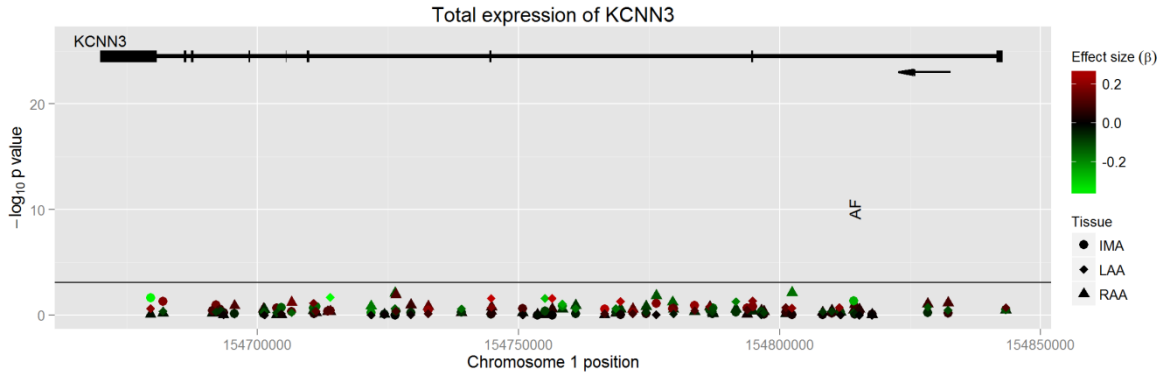
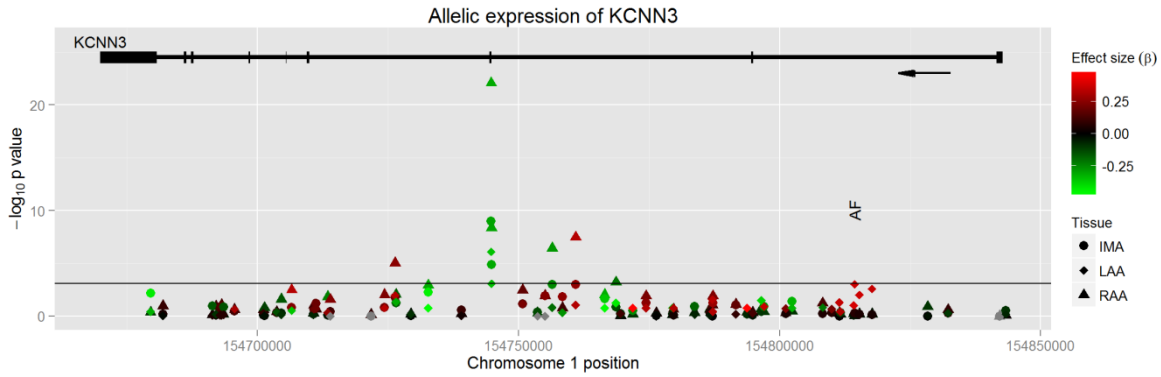


Figure 5.5 Associations with allelic expression ratio of *KCNN3* in chromosome band 1q21



5.5.5 An AF-associated variant in chromosome band 1q24 is associated with allelic expression of **GORAB** *in LAA*

Total expression and allelic expression ratios were measured for *PRRX1* and *GORAB* in all three tissues. There were an insufficient number of heterozygotes for the transcribed SNP rs3820416, which distinguished *PRRX1* allelic expression, in the LAA cohort to allow further analysis of AER for this gene in left atrial tissue.

No associations with *PRRX1* expression in any tissue were detected by eQTL analysis. However aerQTL analysis identified ten variants that were significantly associated with expression of *PRRX1* in RAA, of which 6 were significantly associated with expression in IMA. The most significantly associated SNP in both tissues was the transcribed SNP, rs3820416, of which the minor allele was associated with a 0.51-fold decrease in expression in RAA and a 0.50-fold decrease in expression in IMA ($p=5.7x10^{-35}$, and $p=7.5x10^{-24}$, respectively). The AF-associated SNP, rs3903239, was not associated with expression of *PRRX1* in either tissue (Figures 5.6 & 5.7, Tables A2.11 & A2.12).

The minor allele of a single SNP, rs12038255, was significantly associated with a 1.27-fold increase in expression of *GORAB* in RAA ($p=3.1x10^{-5}$), but not in LAA or IMA. Mapping using aerQTL analysis identified eight variants that were associated with expression of *GORAB* in RAA, of which three were also associated with expression in IMA. For a further three variants, not associated with expression of *GORAB* in RAA or IMA, an association was found with expression in LAA. The minor allele of the most strongly associated SNP in RAA and IMA, rs913257, was associated with a 1.25-fold increase in expression in RAA and a 1.36-fold increase in expression in IMA ($p=1.8x10^{-12}$ and $p=3.9x10^{-7}$, respectively). The most strongly associated SNP in LAA was rs502612, of which the minor allele was associated with a 0.87-fold decrease in *GORAB* expression (p=0.010). The risk [G] allele of the AF-associated SNP, rs3903239, was associated with a 1.18-fold increase in expression of *GORAB* (p=0.049) (Figures 5.6 & 5.7, Tables A2.13 & A2.14).

Figure 5.6 Associations with total expression in chromosome band 1q24

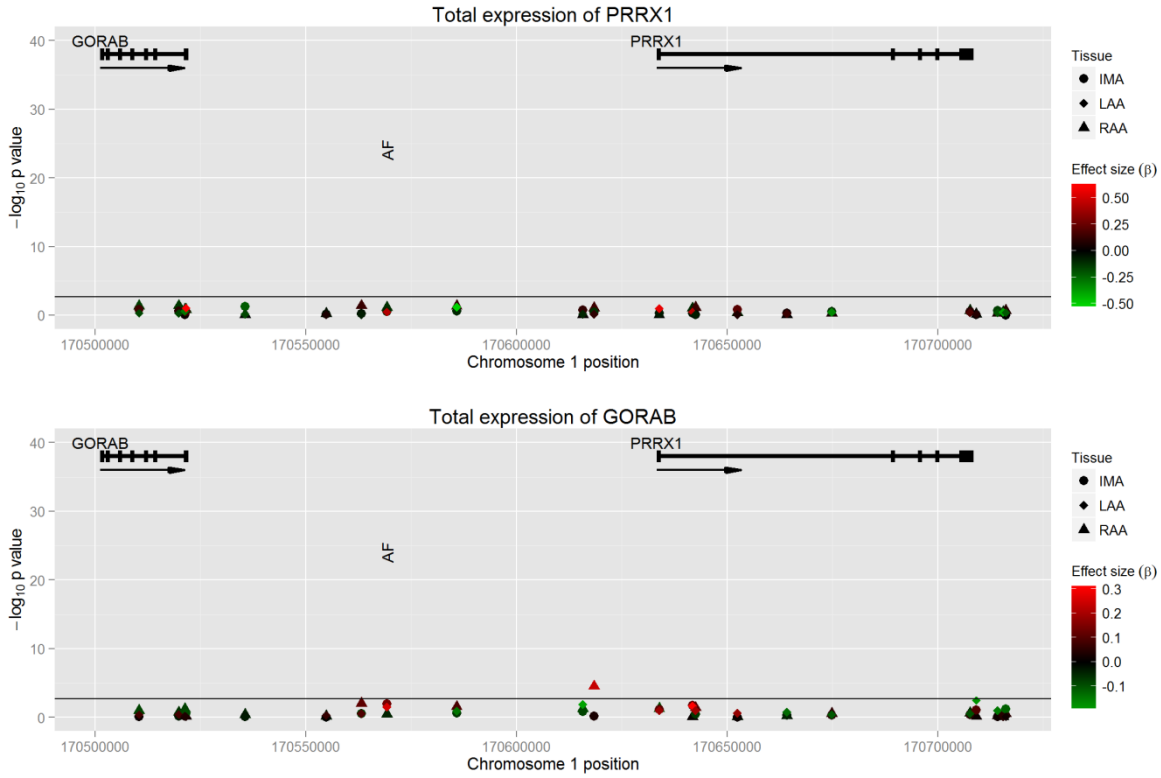
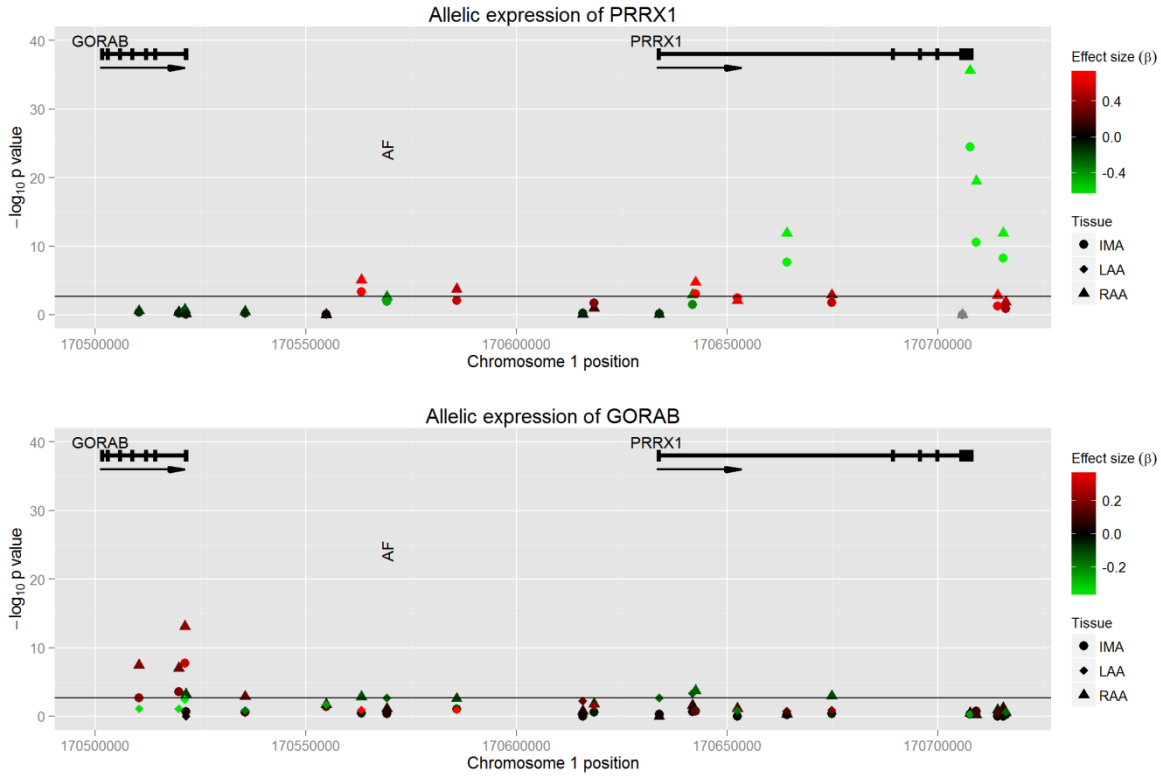


Figure 5.7 Associations with allelic expression ratio in chromosome band 1q24



5.5.6 AF risk variants in chromosome band 4q25 are associated with increased expression of the non-c isoforms of **PITX2** *in RAA*

Total expression of *PITX2* was measured using two different assays, as described in the methods section. One assay was specific for the c isoform, *PITX2c*, and the other was specific for the non-c isoforms, *PITX2a* and *PITX2b* (this assay is hereafter referred to as *PITX2a/b*). There were no individuals in the study population who were heterozygous for the transcribed markers that were specific to *PITX2c* or *PITX2a/b*. It was only possible, therefore, to measure the allelic expression ratio of all three *PITX2* isoforms together. Four SNPs were significantly associated with expression of *PITX2a/b* in the RAA samples. The strongest association was seen for the risk allele [A] of the AF-associated SNP, rs17042171, which was associated with a 2.01-fold increase in expression (p=0.0006). The risk allele [T] of the AF-associated SNP rs2200733 was associated with a 1.98-fold increase in expression (p=0.0010) and the risk allele [G] of the AF-associated SNP rs6843082, was also associated with a 1.79-fold increase in expression (p=0.0011). The other significantly associated SNP was rs16997168, the minor allele of which was associated with a 1.94-fold increase in expression (p=0.0007) and was in significant LD with rs17042171 (r^2=0.34). No association was seen for the AF risk SNP rs10033464. After correction for the effects of rs17042171, no significant association was seen for any other genotyped SNP. There were no significant associations between genotype and expression of *PITX2a/b* or *PITX2c* in LAA or IMA.

There were no significant associations between allelic expression ratio of *PITX2* and any of the typed SNPs (Figures 5.8 & 5.9 Tables A2.15-A2.17).

Figure 5.8 Associations with total expression in chromosome band 4q25

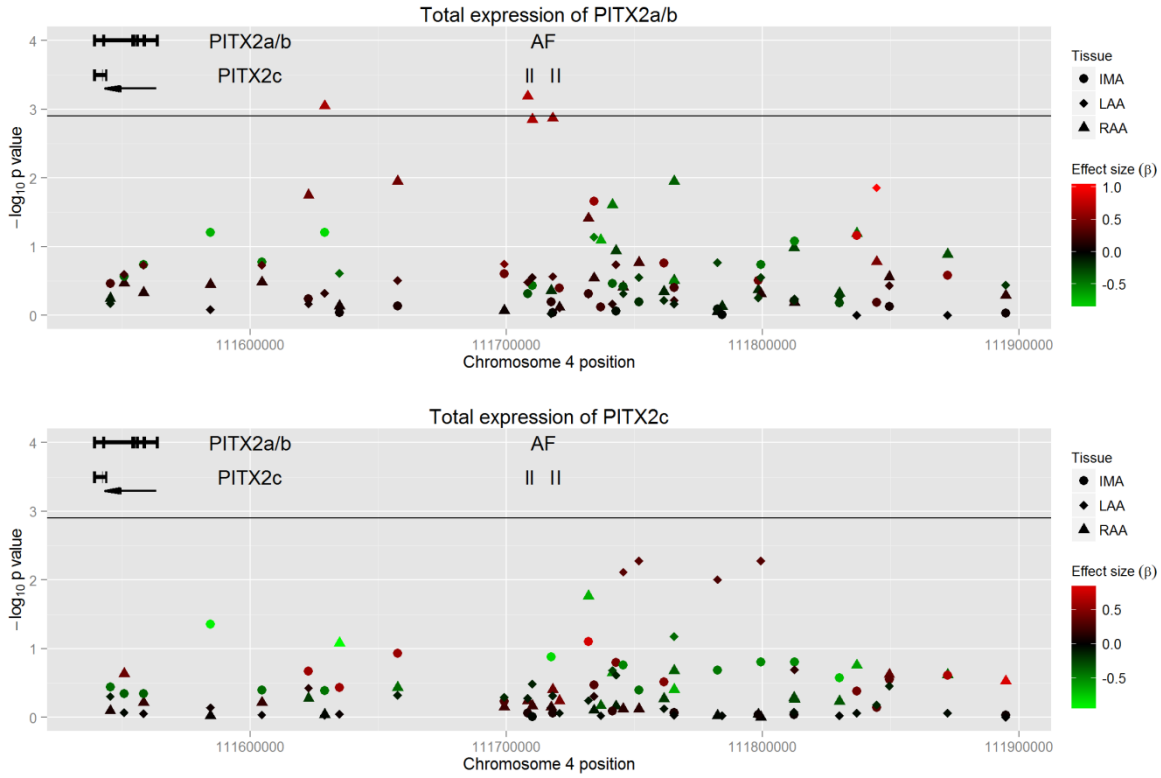
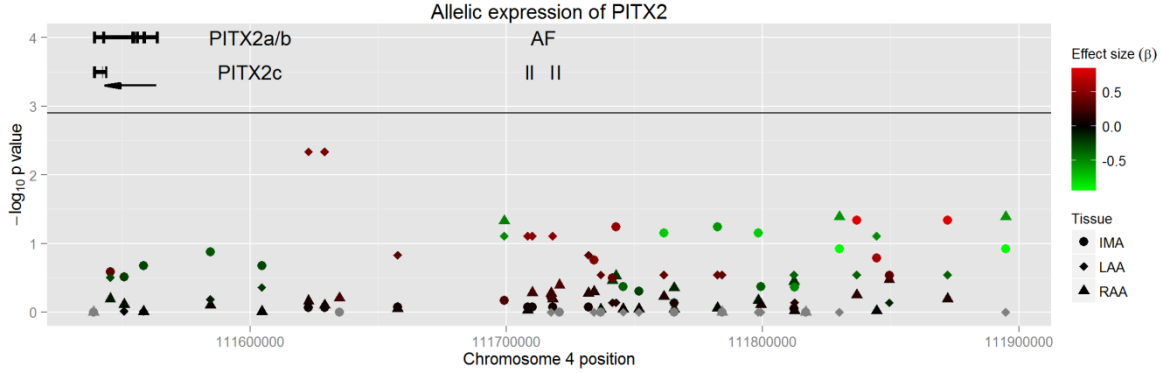


Figure 5.9 Associations with allelic expression ratio in chromosome band 4q25



5.5.7 AF associated SNPs in chromosome band 7q31 are associated with decreased expression of CAV1 *in right atrial appendage*

Total expression and allelic expression ratios were measured for *CAV1* and *CAV2* in RAA and IMA. There were an insufficient number of individuals heterozygous at the transcribed marker, rs8904, to measure allelic expression ratio of *CAV2* in the LAA cohort.

eQTL analysis identified a single SNP, rs9920, which was associated with expression of *CAV1* in RAA, the minor [C] allele being associated with a 1.22-fold increase in expression (p=0.0040). Two of 11 SNPs were associated with expression of *CAV2* in RAA. The most strongly associated was rs9920, the minor allele of which was associated with a 1.21-fold increase in expression (p=0.0007). The other significantly associated SNP was the PR interval-associated SNP, rs1773845, the risk [A] allele of which was associated with a 0.90-fold reduction in expression (p=0.0037). No SNPs were associated with total expression of *CAV1* or *CAV2* in either LAA or IMA.

aerQTL analysis demonstrated that five of 11 SNPs were associated with expression of *CAV1* in RAA. The most strongly associated SNP was rs1049337 with the minor [T] allele being associated with a 0.75-fold reduction in expression ($p=5.1x10^{-19}$). The risk [G] allele of the AF-associated SNP, rs3807989, was associated with a 0.89-fold reduction in expression ($p=5.9x10^{-8}$). Two SNPs were significantly associated with *CAV1* expression in IMA. Again, the strongest association was with rs1049337, the minor allele being associated with a 0.71-fold reduction in expression ($p=1.5x10^{-7}$).

Using aerQTL analysis, three of 11 SNPs were significantly associated with expression of *CAV2* in RAA. The strongest association was with the transcribed marker, rs8940, the minor allele of which was associated with a 0.89-fold reduction in expression ($p=2.5x10^{-8}$). The AF and PR interval-associated SNPs were not associated with expression. No SNPs were significantly associated with expression in IMA (Figures 5.10 & 5.11, Tables A2.18- A2.21).

Figure 5.10 Associations with total expression in chromosome band 7q31

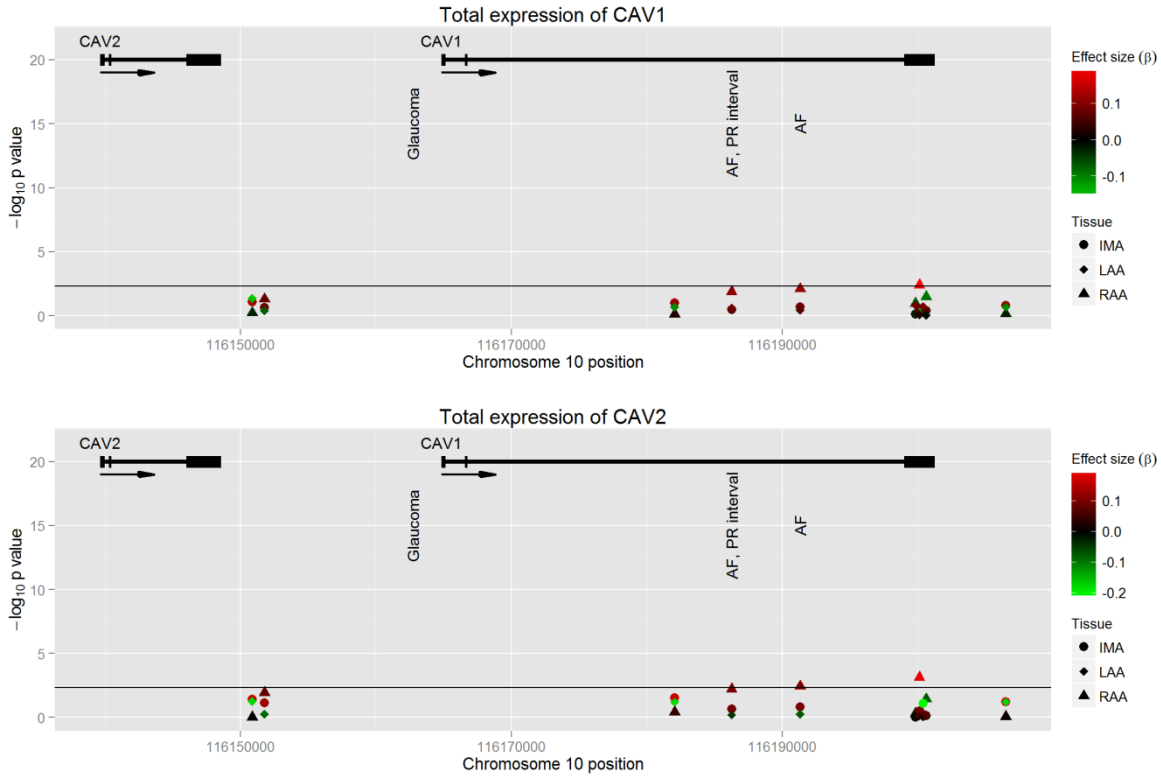
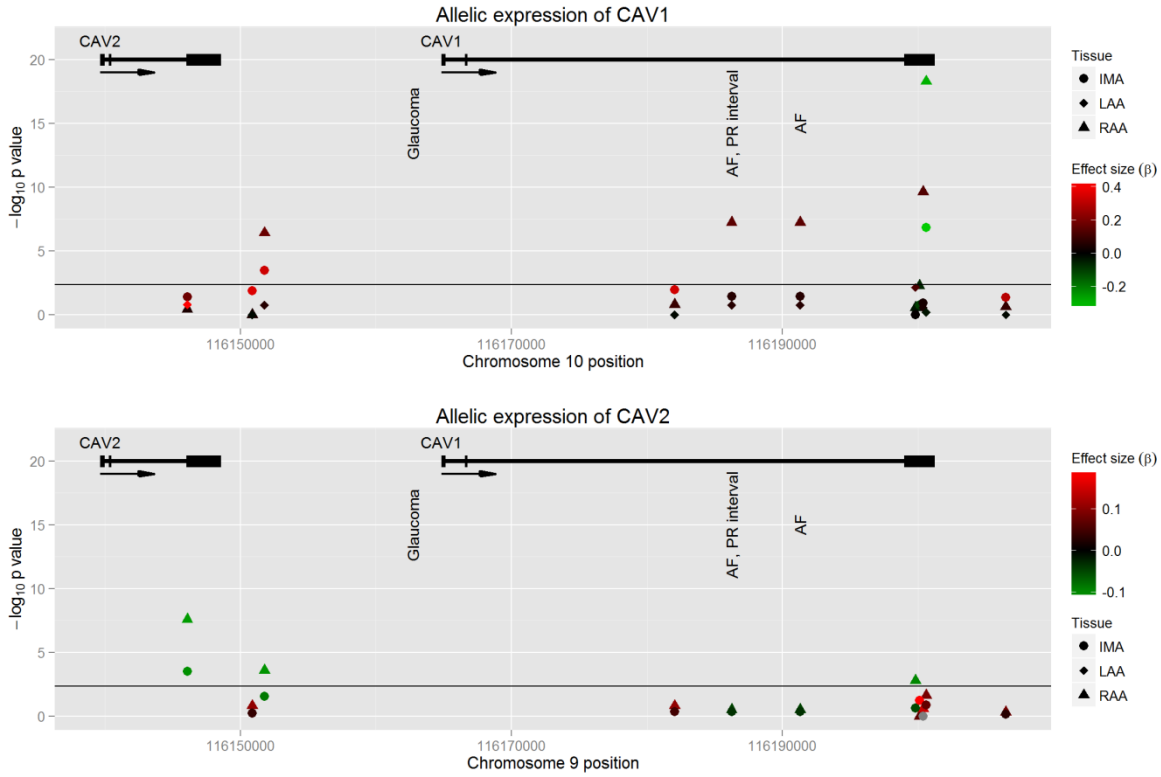


Figure 5.11 Associations with allelic expression ratio in chromosome band 7q31



5.5.8 An AF risk SNP in chromosome band 9q22 is associated with decreased expression of C9orf3 *and* FANCC

Total expression and allelic expression ratios of *C9orf3* and *FANCC* were measured in all three tissues. No significant associations were identified between genotype at any of the SNPs and expression of either *C9orf3* or *FANCC* in any tissue by eQTL analysis.

Using aerQTL analysis, significant associations with expression of *C9orf3* were identified for five of 15 SNPs. The most significant association was with the transcribed SNP, rs4657, the minor [C] allele of which was associated with a 0.91-fold decrease in expression ($p=7.7x10^{-12}$). The risk [A] allele of the AF-associated SNP, rs10821415, was associated with a 0.91-fold reduction in expression ($p=1.5x10^{-5}$).The minor allele of rs4657 was also associated with a 0.91-fold reduction in expression in LAA ($p=1.6x10^{-5}$). No SNPs were significantly associated with allelic expression in IMA.

Four SNPs were significantly associated with expression of *FANCC* in RAA. The AF-risk SNP, rs10821415, was the most strongly associated, the risk [A] allele being associated with a 0.94-fold reduction in expression ($p=8.9x10^{-8}$). The same risk allele was associated with a 0.96-fold reduction in expression in IMA (p=0.0019). Only the transcribed SNP, rs4647554, was associated with expression in LAA (p=0.0030) (Figures 5.12 & 5.13, Tables A2.22-A2.25).

Figure 5.12 Associations with total expression in chromosome band 9q22

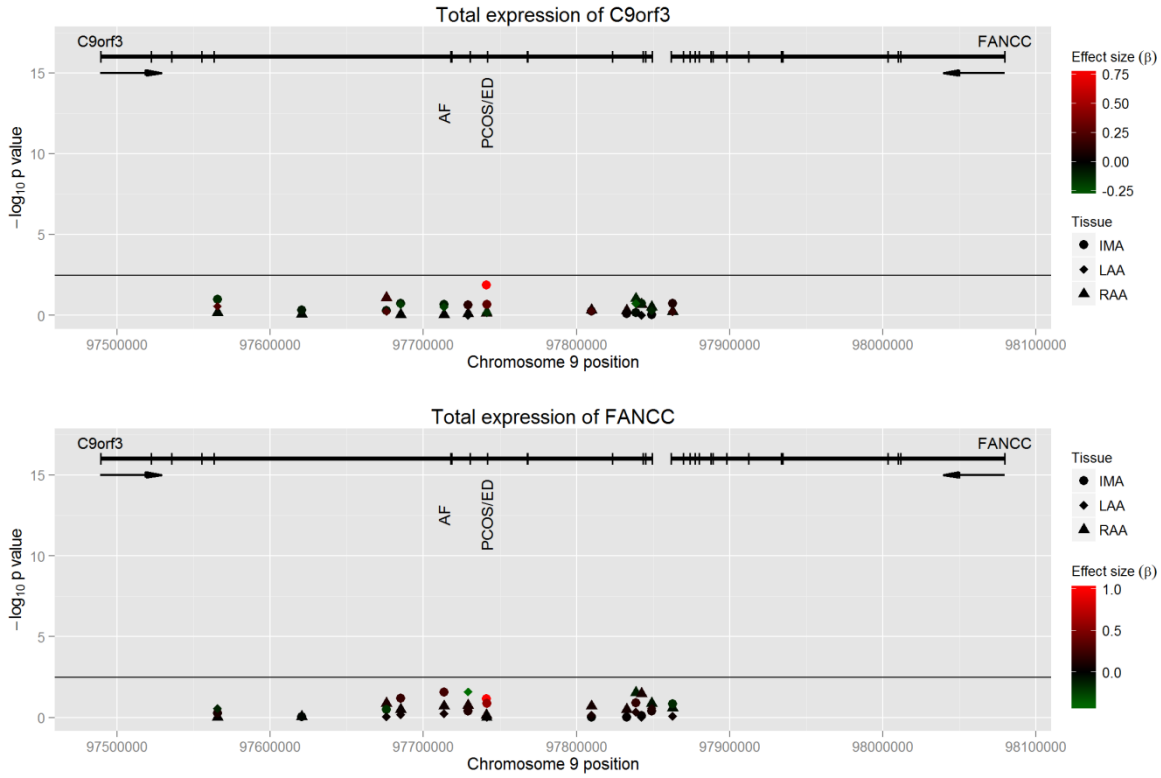
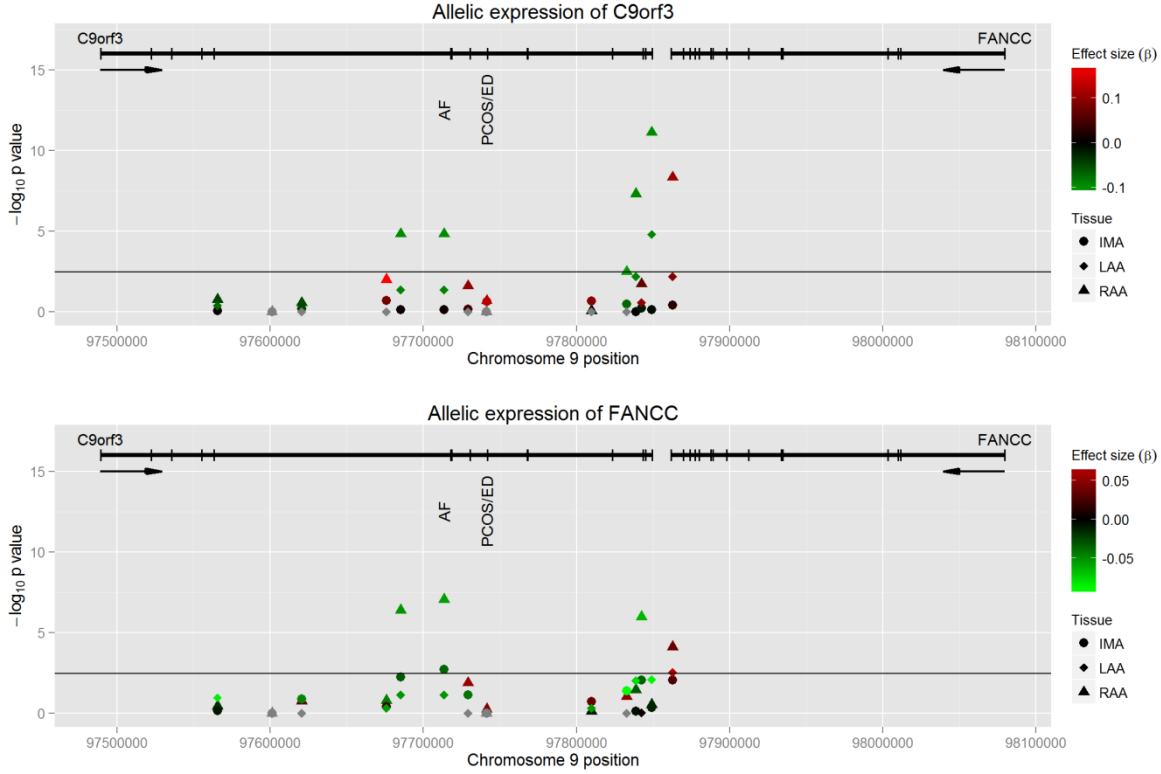


Figure 5.13 Associations with allelic expression ratio in chromosome band 9q22



5.5.9 An AF risk SNP in chromosome band 10q22 is associated with increased expression of **MYOZ1**

Total expression and allelic expression ratios were measured for *MYOZ1* and *SYNPO2L* in RAA. There were an insufficient number of individuals heterozygous for the transcribed SNP rs41280400, used to measure allelic expression ratios of *MYOZ1*, in the LAA and IMA cohorts to permit further analysis. It was not possible to type the AF-associated SNP in this region, rs10824026, using either Sequenom or TaqMan assays. Genotypes for this variant were therefore imputed as described in the methods section of this chapter.

By eQTL analysis, eight of nine variants were associated with expression of *MYOZ1* in RAA. The strongest association was for rs3182629, the minor [A] allele of which was associated with a 2.58-fold increase in expression ($p=4.7x10^{-15}$). The risk allele [A] of the AF risk SNP, rs10824026, was associated with a 0.39-fold decrease in expression ($p=5.5x10^{-15}$). The AF risk allele was associated with an even greater 0.14-fold reduction in expression in LAA (p=0.0008), but there were no SNPs associated with expression in IMA. No variants were significantly associated with expression of *SYNPO2L* in any tissue.

By aerQTL analysis, seven of nine variants were associated with expression of *MYOZ1* in RAA. The AF risk SNP, rs10824026, was among the most strongly associated with the risk allele being associated with a 0.26-fold reduction in expression ($p=5.7x10^{-5}$). As a result of the small numbers of heterozygous individuals in the sample and strong linkage disequilibrium, several SNPs had identical p-values and effect sizes.

Five of nine variants were associated with expression of *SYNPO2L* in RAA. The strongest association was seen for the transcribed SNP, rs4746139, the minor allele of which was associated with a 0.87-fold reduction in expression ($p=6.5x10^{-8}$). The AF risk SNP was not associated with expression. The effect size of the *SYNPO2L* transcribed SNPs is small, this modest effect size and the fact that the transcribed SNPs are not in perfect LD with the AF risk SNP explains the lack of association of the AF risk SNPs with *SYNPO2L* AER, despite the strong LD in the region. There were no associations between any variant and allelic expression in LAA or IMA (Figures 5.14 & 5.15, Figures A2.26-A2.29).

Figure 5.14 Associations with total expression in chromosome band 10q22

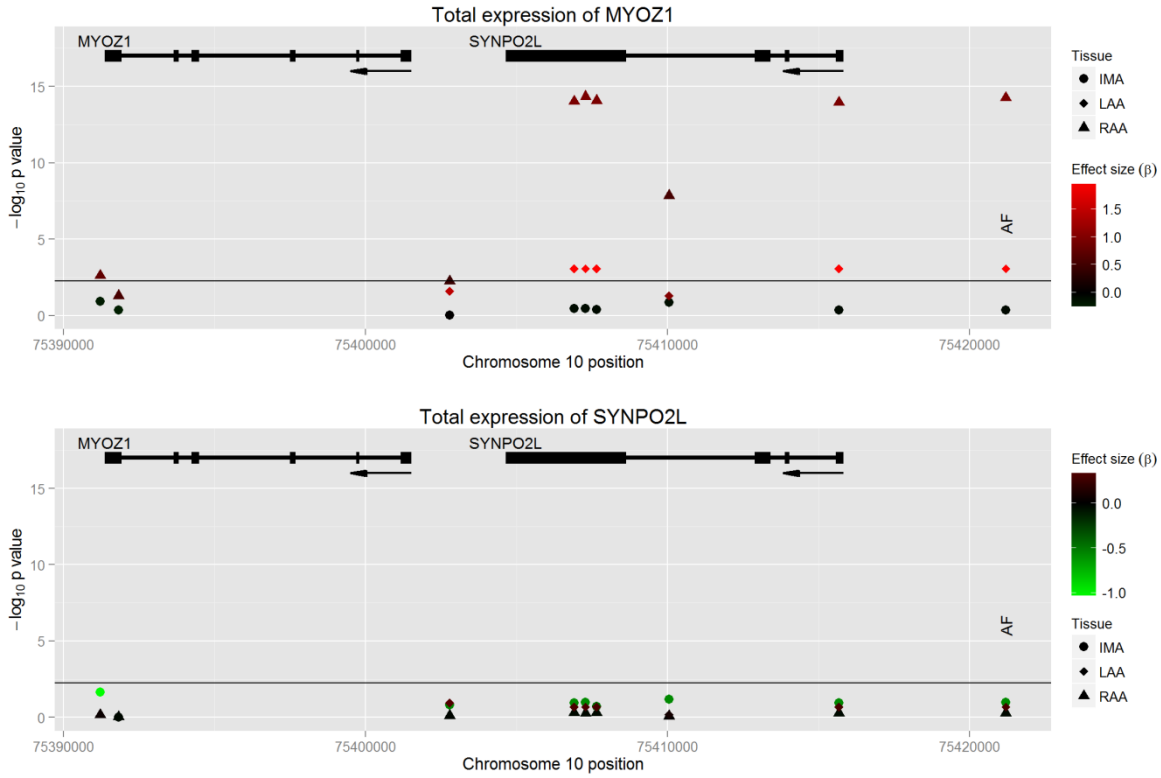
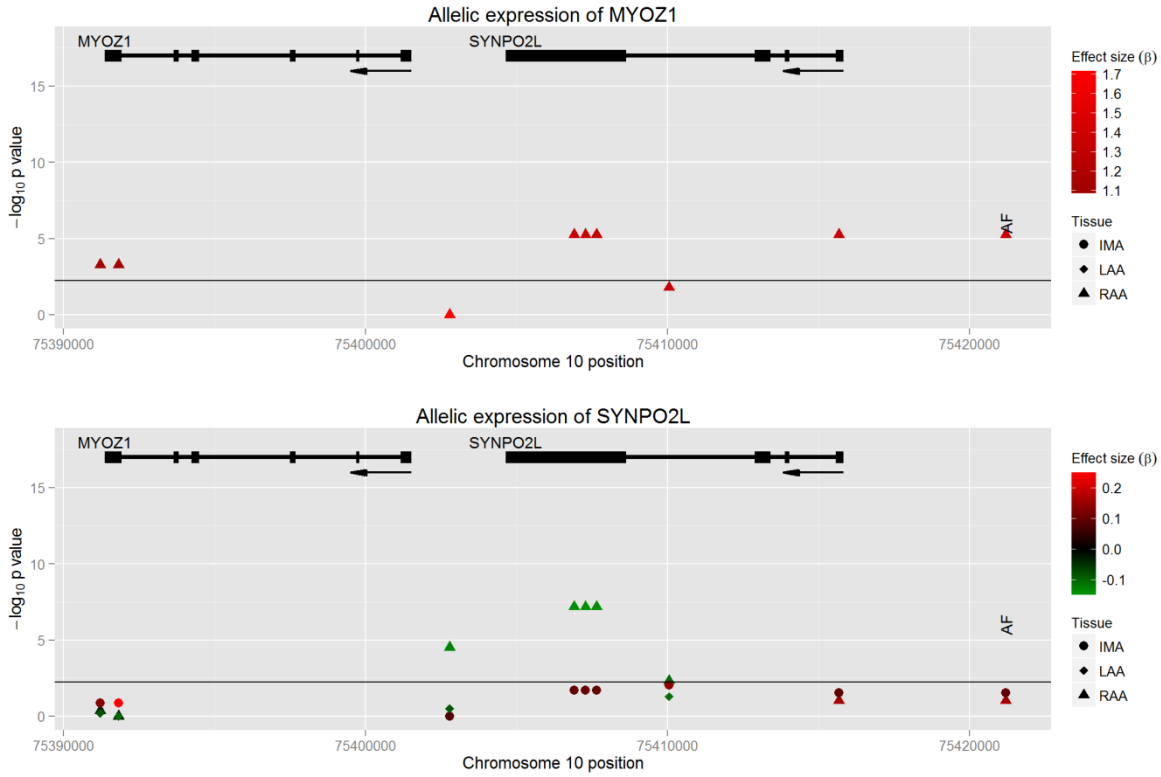


Figure 5.15 Associations with allelic expression ratio in chromosome band 10q22



5.5.10 Variants in chromosome band 14q23 are associated with allelic expression of **SYNE2**

Total and allelic expression of *SYNE2* and *ESR2* were measured in all three tissues. There were no associations between any SNP and *SYNE2* expression in RAA or IMA. eQTL analysis demonstrated that the minor allele of a single SNP, rs4986938, was associated with a 1.59-fold increase in expression of *SYNE2* in LAA (p=0.0014). No SNPs were associated with expression of *ESR2* in any tissue.

Using aerQTL analysis, associations were shown with expression of *SYNE2* in RAA in five of 33 SNPs. The strongest association was seen for rs1256064, the minor allele of which was associated with a 0.89-fold decrease in expression ($p=8.6x10^{-6}$). Three SNPs were associated with expression in IMA, the minor allele of the transcribed SNP, rs1152582, being associated with a 1.12-fold increase in expression ($p=1.1x10^{-6}$). No variants were associated with expression in LAA. The AF risk SNP was not associated with expression in any tissue.

The minor allele of a single variant, rs17101651, was associated with a 0.88-fold decrease in expression of *ESR2* in RAA. No variants were associated with expression in LAA or IMA. The AF risk variant was not associated with expression in any tissue (Figures 5.16 & 5.17, Tables A2.30-A2.33).

Figure 5.16 Associations with total expression in chromosome band 14q23

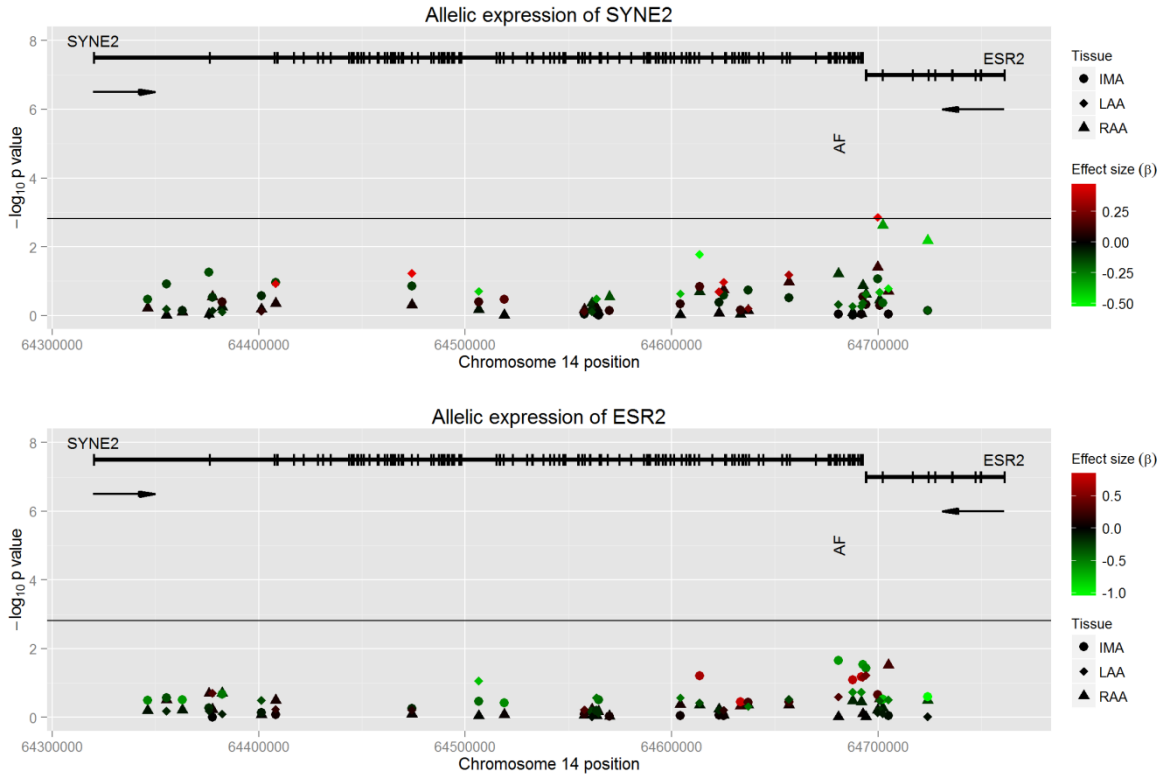
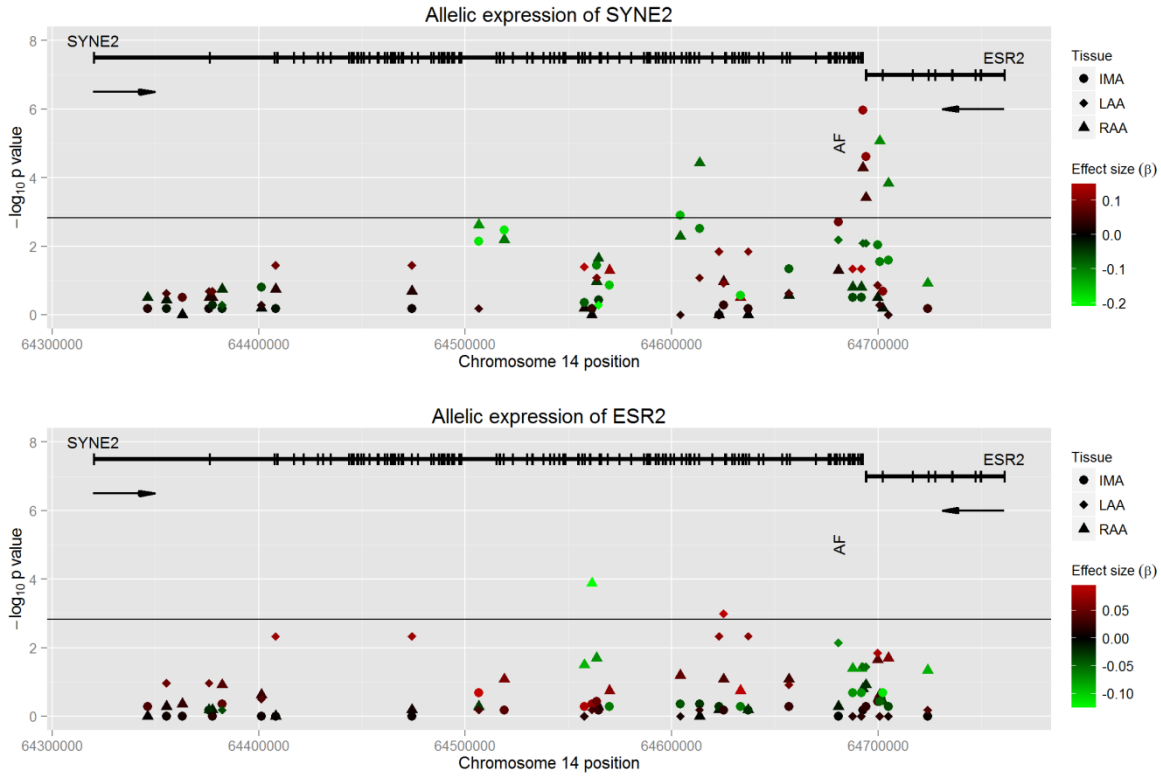


Figure 5.17 Associations with allelic expression ratio in chromosome band 14q23



5.5.11 Variants in chromosome band 15q24 are associated with expression of **HCN4**

Total expression and allelic expression ratios were measured for *HCN4* in all three tissues. Expression levels of *HCN4* in IMA were too low to permit further analysis. Using eQTL analysis, the minor allele of a single SNP, rs3826046, was associated with a 1.60-fold increase in expression of *HCN4* in RAA (p=0.007). No variants were associated with expression of *HCN4* in LAA.

Using aerQTL analysis, ten SNPs were significantly associated with expression of *HCN4* in RAA and 12 SNPs were significantly associated with expression in LAA. Of these, there were nine SNPs that were significantly associated with expression of *HCN4* in both tissues. The minor allele of rs488156, the most strongly associated SNP in RAA, was associated with a 1.11-fold increase in expression of *HCN4* ($p=3.8x10^{-6}$). The minor allele of rs3743496, the most strongly associated SNP in LAA, was associated with a 0.97-fold decrease in expression (p=0.003). The AF risk SNP, rs7164883, was not associated with allelic expression in any tissue (Figures 5.18 & 5.19, Tables A2.34 & 2.35).

Figure 5.18 Associations with total expression in chromosome band 15q24

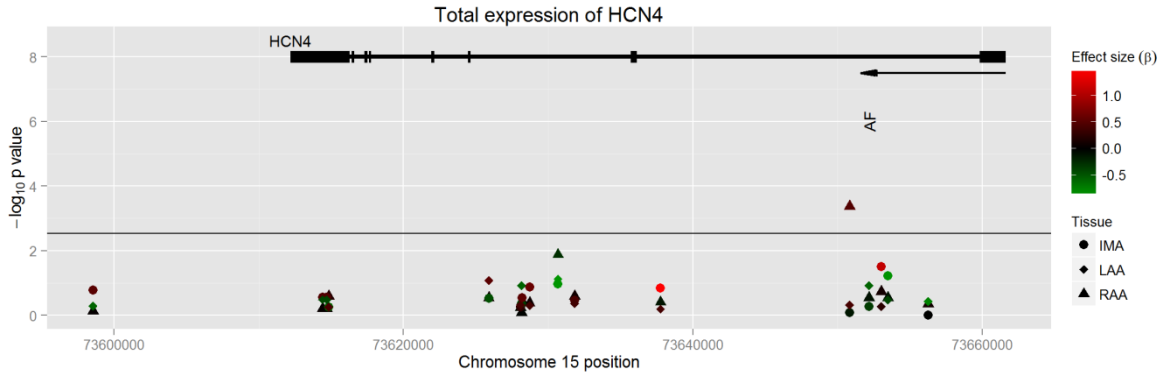
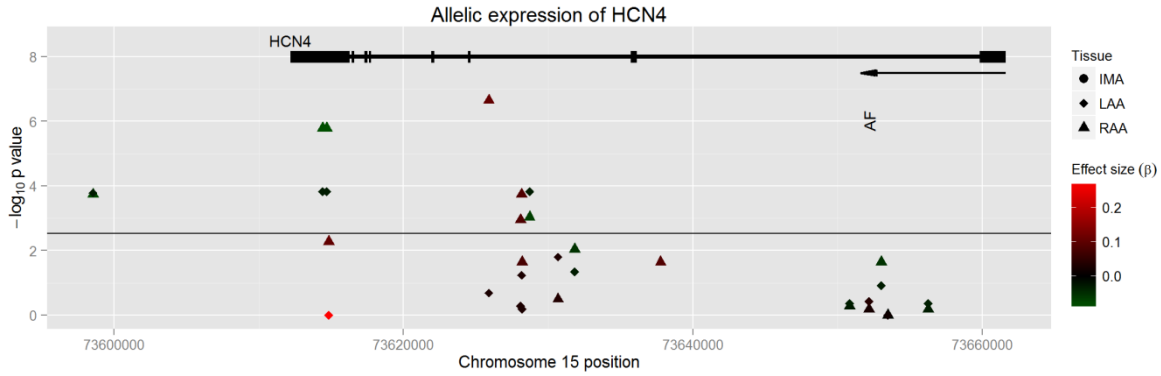


Figure 5.19 Associations with allelic expression ratio in chromosome band 15q24



5.5.12 No variants in chromosome band 16q22 are associated with expression of ZFHX3 or ZFHX3a

Total expression was measured for *ZFHX3* transcript A, (*ZFHX3a*) and both A and B transcripts of *ZFHX3* together in all three tissues. There were insufficient heterozygous individuals at either of the transcribed SNPs, rs740178 and rs10852515, to permit measurement of the allelic expression ratios of *ZFHX3* or *ZFHX3a* in the LAA cohort. No association between genotype and expression of either *ZFHX3* or *ZFHX3a* was seen in any tissue by using either eQTL or aerQTL analysis (Figures 5.20 & 5.21, Tables A2.36-A2.39).

Figure 5.20 Associations with total expression in chromosome band 16q22

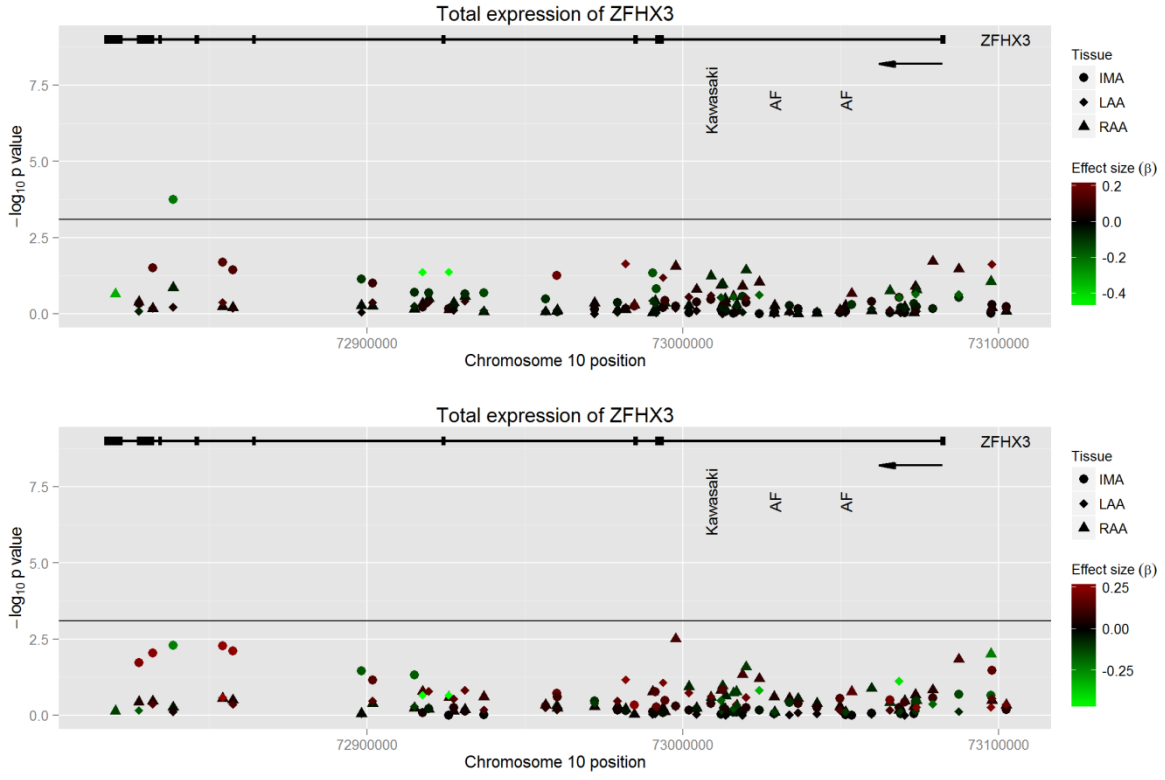
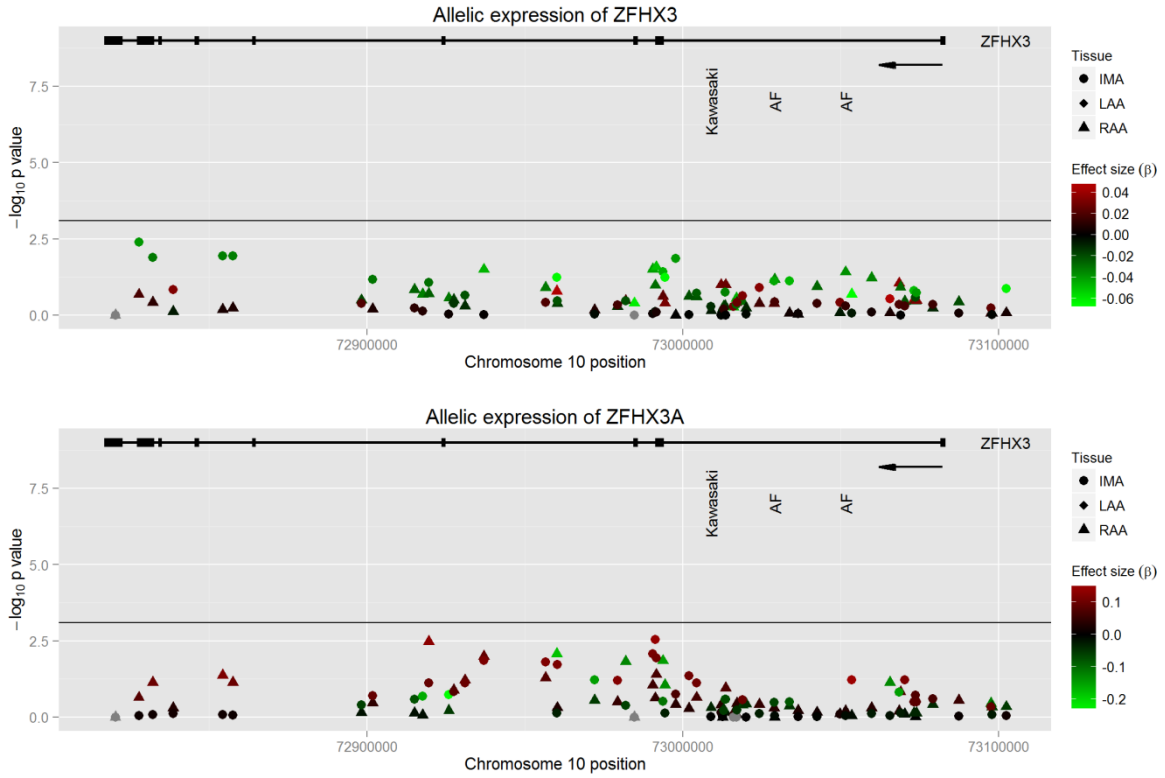


Figure 5.21 Associations with allelic expression ratio in chromosome band 16q22



5.6 Discussion

5.6.1 Patterns of expression of candidate genes in LAA, RAA and IMA

It is unsurprising that the levels of expression of the candidate genes studied here were significantly different between atrial tissue, which consists largely of cardiac muscle cells, and internal mammary artery tissue, which is mostly made up of smooth muscle cells. The largest differences were seen for *HCN4*, which encodes the major subunit of the I_f pacemaker channel and is not known to play an important role in smooth muscle electrophysiology; and *SYNPO2L* which encodes a structural component of the Z-disc in striated muscle[135].

Increased expression in IMA was noted for the transcription factors *ZFHX3*, *PRRX1* and *PITX2*. The two greatest differences identified in LAA vs RAA expression were also reported in the top 20 differences reported by a recent RNAseq study of LAA and RAA by Hsu *et.al*. *HCN4* was over-expressed 5.88-fold in the RAA compared to LAA in the present study, compared to 7.53-fold over-expression in the RAA samples from Hsu

et.al. PITX2c levels were over-expressed 120-fold in the LAA in the present study compared to 116-fold in the study by Hsu *et.al.* Expression levels of the oestrogen receptor ESR2 were very low in all tissues, suggesting that it is unlikely to play an important role in AF pathogenesis.

Previous studies of PITX2 expression in atrial tissue have not demonstrated levels of non-c isoform expression commensurate with my findings. A previous RNAseq study identified only *PITX2c* in adult human LAA and RAA[124]. This study used 3' library selection to isolate RNA, a technique that would be expected to favour detection of the 3' end of the transcript, which all three isoforms share. Bias against the longer, non-c transcripts may, as acknowledged by the authors of that paper, have thereby been introduced[124]. We prepared cDNA using random hexamer primers, which would not be likely to result in the same bias. In addition RNAseq has a higher error rate than qPCR for transcripts such as *PITX2* expressed at low levels[223]. A qPCR study reported that the c isoform represented >99% of *PITX2* expression in LAA and approximately 90% of *PITX2* expression in the RAA; with *PITX2c* expression being 100-fold greater in LAA compared to RAA[125]. However a later study by the same group presented a western blot of PITX2, using a non-isoform specific antibody, showing a difference in protein levels between LAA and RAA of only 2-fold[310]. This discrepancy may be due to expression of non-c isoforms of PITX2 detected at the protein but not the RNA level.

The original description of MYOZ1 noted that its protein product, Calsarcin-2, is only transiently expressed in fetal development in mouse heart and that expression levels were not detected by Northern blotting, in adult human heart and mouse atria. In the current experiment we demonstrated expression of MYOZ1 at levels approximately 32-fold lower than the mean expression of the selected reference genes. The discordance between these and previous results may simply derive from the greater sensitivity of the qPCR approach. Alternatively MYOZ1 expression may vary throughout the heart, with greater expression in the atrial appendage than in the material used in the experiment by Frey *et al*. A microarray-based genome-wide eQTL mapping experiment performed using LAA and RAA detected sufficient levels of MYOZ1 to allow identification of an eQTL in both tissues. The BLAST tool was used to ensure that the primer sequences used were specific to *MYOZ1* and would not wrongly detect *MYOZ2* expression[311].

5.6.2 Associations between gene expression and AF status.

Previous studies have examined gene expression changes in human atria with AF. An eQTL mapping study of *PITX2c* expression in LAA identified an increase in *PITX2c* expression in patients who were in AF at the time of surgery compared to those who had a history of AF but were in sinus rhythm. However, there were no significant differences between either group and the smaller number of patients who had no history of AF[129]. Conversely, two small studies found that *PITX2c* levels were lower in atrial tissue from patients in AF than in atrial tissue from patients in sinus rhythm[125,267]. *PITX2b* potentiates the action of *PITX2c* on gene expression[123], and so upregulation of non-*c* isoforms is likely to result in greater *PITX2c*-mediated transcriptional activity, which may result in increased AF risk.

This study was designed to detect eQTL associations, and so was not powered to detect associations between AF status and gene expression. Nevertheless, this study has replicated a previously identified[301] reduction in *CAV1* expression in association with AF at the $p<0.05$ significance level (Table 5.4). The same study found no difference in *CAV2* expression between AF cases and controls, which is consistent with the results presented here. I found no significant differences between AF cases and controls in expression of either *PITX2c* or *PITX2a/b* in LAA or *PITX2a/b* in RAA. Due to the limited phenotypic data available, it was not possible to divide patients with a history of AF into groups by rhythm status at the time of surgery, and the small number of patients with a history of AF would have rendered results from such a subgroup analysis unreliable.

5.6.3 eQTL and aerQTL mapping of genetic effects on expression in tissue

This study has systematically mapped genetic effects on gene expression at loci associated with AF risk by measuring both total expression and allelic expression ratios. The results from Chapter 4 have demonstrated that, in general, aerQTL mapping performs better than eQTL mapping in identifying *cis*-acting genetic variants. As previously discussed the performance of the aerQTL approach is dependent on the

presence of one or more transcribed SNPs with a minor allele frequency high enough that a sufficient number of individuals in the population studied are heterozygotes. This is the case with the results from this chapter, where no significant associations were found between total expression and genotype at most of the loci studied, despite there being significant associations with allelic expression at the same loci. The size of the cohorts studied in this chapter were smaller than those in the experiments using peripheral blood, showing that the aerQTL approach is able to demonstrate effects in quite small cohorts whereas eQTL mapping is less powerful.

The importance of identifying transcribed SNPs is demonstrated by the findings at 4q25, however. By using transcript specific primers, it has been possible to show that AF-associated SNPs affect expression of the non-c isoforms of *PITX2*, but not *PITX2c*. Transcribed SNPs specific for the c and non-c isoforms of *PITX2* were genotyped, but the minor alleles of these variants were not present in the study population. It was therefore not possible to measure transcript-specific effects on allelic expression ratio, and no effects of the AF-associated SNPs on allelic expression ratio could be identified when all of the *PITX2* isoforms were measured together. The minor allele frequency of the transcript-specific transcribed SNPs is reported as very low in the 1000 genomes study, and so our findings are not entirely surprising[309].

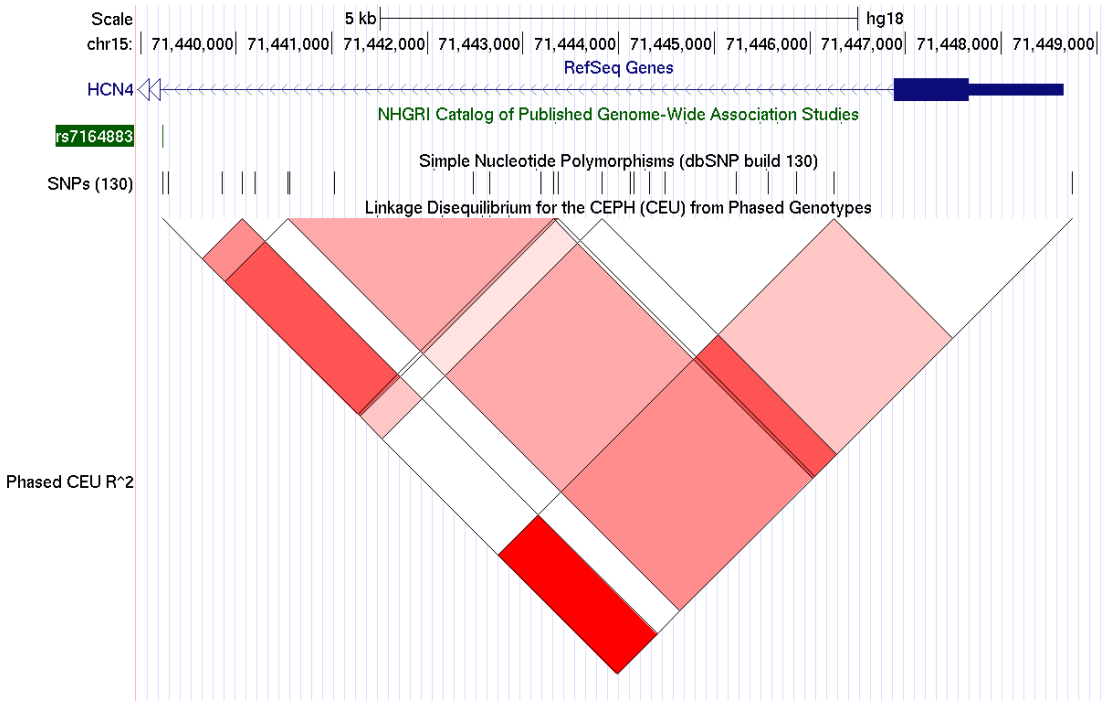
5.6.5 Mechanisms underlying genetic associations with AF

No association was demonstrated between expression of *HCN4* and the AF risk variants. As has been shown in Chapter 3, inhibition of the I_f current increases the incidence of AF. *HCN4* therefore remains the most likely candidate gene in this region. It has been demonstrated here that expression of *HCN4* differs between right and left atrial appendage. This is unsurprising as *HCN4* is primarily expressed in nodal tissue. The sino-atrial (SA) node is a relatively compact structure in the high right atrium, and *HCN4* expression has been shown to vary within the right atrium with higher expression seen at the site of the anatomical SA node[312]. Differential SNP effects on expression in apparently similar tissues have been demonstrated previously. For example; rs143383, which is associated with risk of osteoarthritis, has been shown to affect expression of *GDF5* in the knee, but not the hip[313]. It is likely, therefore that

regulation of *HCN4* expression differs between the SA node and RAA, and any SNP effect may be specific to nodal tissue.

Alternatively, the AF risk SNP rs7164883 is in LD with a coding region of the nearest gene, HCN4 (Figure 5.22). It is possible, therefore, that the effect of rs7164883 is not mediated by reduced expression of *HCN4*, but instead is a consequence of linkage disequilibrium with a coding variant which alters the physiological properties of the I_f channel. There are 17 coding variants listed in build 138 of dbSNP in exon 1 of *HCN4*. None are sufficiently common that LD values are available in the HapMap or 1000 genomes data sets. Fine mapping of the AF association signal at this locus with typing of low frequency variants may allow identification of coding risk variants which would merit further study.

Figure 5.22 Linkage disequilibrium between AF risk variant and the first exon of HCN4



LD is indicated by the inverted mountain plot with darker shades of red indicating higher r^2 values. The AF risk variant, rs7164883, is indicated in green at the far left of the plot. LD values are taken from the CEU population in phase II of the HapMap project.

The role of the SK3 channels encoded by *KCNN3* has already been discussed in Chapter 4. Neither of the approaches used here were able to identify an association between the AF associated SNPs and expression of *KCNN3* in any of the tissues studied here. SK3 channels are expressed in a range of non-cardiac tissues, however, including vascular smooth muscle. SK3 channels regulate blood pressure in mice by affecting arteriolar vasodilatation[314]. There are no GWAS hits for blood pressure near KCNN3, despite large studies having been performed, making effects on systemic blood pressure an unlikely mechanism of action[315-317]. SK3 channels are expressed in human pulmonary endothelium and pharmacological activation of these channels results in dilatation of bronchioles and pulmonary arteries[318]. Pulmonary hypertension increases the risk of developing AF[319], and one could speculate that subclinical changes in pulmonary artery pressure might have a similar effect. To determine whether this mechanism is important in AF risk, it would be necessary to demonstrate that the AF risk SNPs were associated with pulmonary artery pressure, measured by right heart catheterisation or echocardiography, and subsequently that variations in pulmonary pressure are associated with AF risk. This highlights the importance of tissue specificity when attempting to identify the mechanisms of action of disease-associated genetic variants. We were not able to demonstrate an association between expression and genotype in tissue taken from the atrial appendage or from a medium-sized artery, but this does not exclude an effect in small arteries, arterioles or the pulmonary vasculature.

The AF risk variants in chromosome band 1q24 were not consistently associated with expression of either *PRRX1* or *GORAB* in RAA or IMA. An association was identified between the risk allele for AF and an increase in expression of *GORAB* in LAA, but this only just reached significance at the 5% threshold and would require replication in a second cohort to be considered robust.

No association was determined between the AF risk variants in chromosome band 16q22 in any of the tissues studied. Nor was any association demonstrated between expression and genotype of a genetic variant associated with Kawasaki disease: either in atrial or arterial tissue. The inability to demonstrate an association here may, as has been suggested previously, be due to the effects of AF risk variants being limited to specific tissues or cell types, not studied here or, particularly in the case of

transcription regulators such as *ZFHX3*, being limited to specific time points in development, which are not captured in this sample from an adult population.

The associations between genetic variants and allelic expression of *SYNE2* in right atrial appendage followed a similar pattern to those identified in peripheral blood, but did not reach the same level of statistical significance in the smaller tissue cohorts. A significant association was not identified between the AF risk SNPs and allelic expression ratio. This may be for one of two major reasons. The current experiment may simply be underpowered, with too few subjects to detect a modest effect size. The other possible explanation is that the risk SNPs do not exert their influence by affecting expression of *SYNE2* in atrial tissue. The presence of AF as part of the phenotype associated with deleterious *SYNE2* mutations has already been discussed in Chapter 4 and *SYNE2* remains a likely candidate gene. It is possible that the effects of *SYNE2* on AF risk are mediated indirectly. Heart failure, with impaired or preserved ejection fraction, is associated with an increased risk of AF[320-322], and the Emery Dreifuss phenotype includes cardiomyopathy[139]. The increased incidence of AF, therefore, may be mediated by higher ventricular filling pressures rather than by a direct effect of *SYNE2* expression in the atria themselves. In order to investigate whether an intermediate phenotype of the left ventricle underlies this association, it would be necessary to establish whether the AF risk SNPs are associated with changes in left ventricular function or filling pressures. As in peripheral blood, there were no associations between the AF risk SNPs and total expression or the allelic expression ratio of *ESR2*.

The risk variants of multiple AF-associated SNPs in chromosome 4q25 upregulate expression of the a and b isoforms of *PITX2* in human RAA. The risk allele of the AF-associated SNP rs17042171 was associated with a 2.01-fold increase in expression of the *PITX2a* and *b* isoforms (p=0.0005). This is the first demonstration of a gene expression phenotype associated with AF risk SNPs in this region. These SNPs have been identified as the common variants with the strongest association with AF in multiple GWAS to date. Our observations implicate the non-c isoforms of PITX2 in lone AF. Furthermore, we demonstrate that the non-c isoforms of *PITX2* are expressed in both left and right atrium, with a higher expression in LAA vs. RAA.

The pattern of association between expression of *PITX2a* and *PITX2b* and SNP genotype is very similar to the pattern of association between AF risk and SNP genotype. With the exception of rs10033464, which was identified in a GWAS of an Icelandic population, and has not been replicated at GWAS significance in other populations[105], AF risk SNPs in the region were associated with expression and there were no SNPs associated with expression that are not in LD with the AF risk SNPs. The lack of association at rs10033464 may be due to differences in LD structure between the GWAS population and the population in the present study, in which LD between rs10033464 and the lead eQTL SNP is weak (r^2=0.013).

Two previous studies have examined the relationship between GWAS hit SNPs at 4q25 and *PITX2* expression in atrial tissue. A recent eQTL mapping study did not detect associations between AF risk SNPs and *PITX2c* expression[129]; however expression of the other isoforms was not investigated in that study. A recent genome wide eQTL mapping study using the Affymetrix U133 array, on which the *PITX2* probes are not isoform-specific, found no association between *PITX2* expression and the AF risk variants on 4q25. Neither of these studies specifically examined the non-*c* isoforms of *PITX2*.

To date, animal models involving gene targeting of *Pitx2* have not addressed the contribution of the non-c isoforms to arrhythmic susceptibility. A homozygous mouse knockout model of *Pitx2c* alone demonstrated left-right atrial isomerism and early postnatal death[323]. The heterozygote exhibited increased susceptibility to pacing induced arrhythmias and increased expression of genes associated with the sino-atrial node in the left atrium[125,127]. A *Pitx2a/b* knockout mouse model retained normal left-right atrial architecture even when combined with a hypomorphic *PITX2c* allele, but has not, to our knowledge, undergone electrophysiological testing[290]. It is possible that the *Pitx2a/b* knockout mouse would also exhibit electrophysiological abnormalities, due to interaction between Pitx2 isoforms; further investigation in this model would be of interest.

The associations identified in chromosome band 7q31 in right atrial appendage confirm the findings in peripheral blood. There is a clear association between allelic expression ratio of *CAV1* but not *CAV2.* This finding, and the observation that *CAV1*,

but not *CAV2* levels, are associated with AF status in this cohort and in previously published studies[301], provides strong evidence for the role of *CAV1* in AF risk. The observation that expression of *CAV2* in internal mammary artery is associated with one of the AF-associated SNPs in the region is interesting. The strength of the association is not very great, and the lack of an association in the allelic expression experiment, which would be expected to have greater power, raises the possibility that this arises from the play of chance. Replication in an independent cohort would be of value.

The AF risk SNPs in chromosome band 9q22 were associated with a modest but very significant effect on allelic expression of *FANCC* in both RAA and IMA and a larger effect on allelic expression of *C9orf3* in RAA only. It is not possible to determine which of the candidates plays a role in AF risk, or whether both are important. Fanconi anaemia (FA) is a disease with a broad phenotype characterised by increased susceptibility to a variety of cancers, bone marrow failure and developmental abnormalities. FA is caused by a failure of the FA complex of proteins, which plays an important role in DNA repair. *FANCC* encodes member C of the FA complex and biallelic mutation in any of the 15 FA complex proteins can cause FA[324]. FA proteins may have a wider array of functions, however; for example FANCC is required by Hsp70 to prevent apoptosis of haemopoetic stem cells[325]. Loss of function of *FANCC* in a cell line resulted in dysregulation of a large number of transcripts but most prominently caused downregulation of *CAV1* expression[326], providing a possible mechanism whereby reduced *FANCC* expression might result in increased AF risk. AF is not reported as part of the phenotype associated with *FANCC* mutations but any association may be masked by the poor prognosis associated with FA.

As discussed in Chapter 4, underexpression of C9orf3 might result in increased activity of the renin-angiotensin-aldosterone system (RAAS) by reduced breakdown of angiotensin III by the protein encoded by *C9orf3*, aminopeptidase O. The role of RAAS in the electrical and structural remodelling which is associated with AF is well established[327-329], and inhibition of RAAS prevents AF[330]. The results of the present study do not indicate whether the AF risk SNPs in this region operate primarily by upregulating RAAS or by FANCC-mediated downregulation of CAV1. The two mechanisms are not mutually exclusive, however, and it is possible that both mechanisms contribute to the overall effect.

Both of the candidate genes identified in chromosome band 10q22, *MYOZ1* and *SYNPO2L*, encode components of the Z-disc in striated muscle. *MYOZ1* encodes Calsarcin-2, which tethers calcineurin to α-actinin[136]. *MYOZ1* has been considered as a candidate gene in idiopathic dilated cardiomyopathy but no *MYOZ1* mutations were found in a cohort of 185 unrelated DCM probands[331]. A genome-wide eQTL mapping study in atrial tissue identified an association between total expression of *MYOZ1* and SNPs in the region, but did not report on the lead AF GWAS SNP itself[130]. It was not possible to genotype the AF-risk SNP, rs10824026, using two different technologies in the present study, which indicates that this is a difficult to genotype variant. The lead variant identified in the genome-wide study, rs3740293, is in strong LD with the AF risk SNP in the 1000 Genomes cohort (r^2=0.75, D'=1). Genotype at rs3740293 was determined by imputation in the present study. The findings of this study and the strong LD between rs10824026 and rs3740239 strongly indicate that decreased expression of *MYOZ1* is the mechanism by which AF risk at this locus is mediated. By contrast, neither in the present study, nor in the experiment by Lin *et.al.*, was expression of *SYNPO2L* affected by genotype of the AF-associated SNPs.

5.7 Conclusions

Results of experiments in this chapter have demonstrated novel associations between AF risk variants and gene expression in cardiac tissue. AF risk variants were associated with increased expression of *PITX2a/b* and decreased expression of *MYOZ1*, *CAV1*, *C9orf3* and *FANCC* in right atrial tissue; these associations have not been demonstrated previously. Furthermore, although this experiment was not primarily designed to detect difference between AF cases and controls, AF has been shown to be associated with reduced expression of *SYNE2*, *HCN4* and *CAV1*. These results provide clear evidence for the roles of these genes in the pathogenesis of atrial fibrillation.

Chapter 6

General discussion and future directions

6 General discussion

6.1 Preamble

In Chapter 3 the emergence of the parallel discoveries of ivabradine as an inhibitor of the I_f current, mediated by *HCN4* and of the *HCN4* region as a susceptibility locus for AF led me to investigate the effects of ivabradine on AF risk. To my knowledge, this is the first example to date where a drug has been demonstrated to have a side effect hypothesised as a result of findings from a genetics study. The identification of genetic effects which mirror drug side effects is not new, however. The best example perhaps, being that of rs1800693, which predisposes individuals to multiple sclerosis by upregulating the expression of a soluble version of the TNF receptor with decoy activity[216]. The SNP mimics the effect of TNF-blocking drugs which increase multiple sclerosis risk, and exacerbate the condition if it is present already, in patients treated with anti-TNF therapy for inflammatory conditions. As well as identifying potential side effects of drug treatment, this finding demonstrates the potential utility of genetic association studies in identifying drug targets. Current antiarrhythmic drug therapy has a relatively poor success rate in maintaining sinus rhythm. Approximately 50% of patients treated with an antiarrhythmic drug will have no improvement in their symptoms[332]. There is a clear need therefore to identify suitable targets for further study.

As has been pointed out previously, GWAS only indicate loci which are associated with a phenotype. Unless the variants identified are within a coding region, which is rarely the case, the responsible gene must be identified. An important example is provided by the link between genetic variants within introns of the gene *FTO* and obesity. *FTO* was identified as the strongest GWAS locus for body mass, and knockout of the gene in a mouse model produced the anticipated effects on body mass. However, no association in human tissue was ever demonstrated between risk variants and *FTO* expression. A recent study has identified that the obesity risk SNPs within *FTO* regulate expression of a distant gene, *IRX3*, and that modulation of *IRX3* expression levels, rather than *FTO* expression levels, affects body mass and composition[333]. This demonstrates the importance of identifying eQTLs as the first phase of the

investigation of risk variants identified by GWAS, in order that research efforts can be focussed on the most appropriate target.

In this chapter I will discuss the utility of the methods used in Chapters 4 and 5 in identifying SNP effects on expression. I will then discuss the limitations of the eQTL and aerQTL analyses described in Chapters 4 and 5, focussing on the issues of power, phenotypic characterisation, the identification of candidate genes, the site and timing of expression and the statistical methods used. I will then consider future experiments, discussing how new approaches might address some the limitations of the present study and describing experiments that might build on the results described here.

6.2 The relative utility of eQTL and aerQTL analysis in identifying SNP effects on expression

In most of the regions analysed in chapters 4 and 5 the aerQTL approach was more powerful than the eQTL approach in detecting *cis*-acting effects. For *CAV1*, *SYNE2* and *ZFHX3* aerQTL mapping detected effects which were not seen in eQTL mapping. In the chromosome 9q22 region, both methodologies detected effects on C9orf3 expression. It has previously been shown that eQTL mapping and aerQTL mapping identify the same cis-acting loci[185,293]. In this experiment we have shown that the increase in power from the aerQTL approach can prove decisive in identifying regulatory elements, where eQTL mapping has not been able to do so. This has implications for future investigations of the regulatory effects of variants at other loci, including the thousands identified by GWAS[30].

It has already been pointed out that the aerQTL method would be expected to have greater power than the eQTL approach due to the large number of *trans*-acting influences on expression which are difficult to identify and control for. We did not include potential *trans*-acting influences on expression in our analysis. Published eQTL studies in similarly sized cohorts have shown that the inclusion of age, gender and medical history in eQTL mapping models has had little effect on the results of the analysis[129,185]. This is likely because the *trans*-acting effects which have been identified contribute relatively little to the total influences which act on expression in *trans*. The proportion of the variation in expression that can be explained by known factors,

therefore, is small and their inclusion in a statistical model does not greatly impact on the power to detect regulatory effects. It is important to note that aerQTL mapping relies on linkage disequilibrium between the transcribed marker and the putative regulatory SNP. This can significantly reduce the power of this approach.

Another important reason why the aerQTL approach may have greater power to detect *cis*-acting effects lies in the regulation of overall gene expression. Gene expression is often regulated by *trans*-acting negative feedback mechanisms[334,335], which serve to maintain relatively constant transcript levels. Any change in overall expression mediated by a *cis*-acting variant may, therefore, not result in a detectable change in total transcript levels. If the *cis*-acting variant is in the heterozygous state, however, the altered ratio of the two alleles will remain. This may be the case in the regions studied here, especially in the case of *SYNE2*, where a very strong association is seen in the aerQTL experiment, but not in the eQTL experiment. By contrast, in the case of *C9orf3*, the aerQTL and eQTL experiments demonstrated very similar power to detect the effects of rs10993391 on expression, suggesting that overall transcript levels are under looser control. This raises the question: if overall transcript levels are not significantly different, as a result of *cis-trans* regulation, how does the SNP affect AF risk? The simplest explanation is that the overall levels are affected by the variable to a small degree, which is sufficient to affect AF risk on a population-wide basis, but is not detectable in a cohort of the size studied here. A second explanation is that the effects acting in *cis* and *trans* differ between tissues, such that regulation of total expression may be looser, allowing *cis*-acting effects to have a greater influence on overall transcript levels. The third explanation is that there is an effect only under certain circumstances (e.g. for components of the immune system when challenged by a pathogen), this would also be expected to be the result of *trans*-acting effects.

This study used up to three transcribed markers to perform aerQTL mapping at several different loci. The use of two transcribed markers improves the power of aerQTL mapping to identify *cis*-acting effects[185], and this methodology was used here to incorporate information from a third marker in the case of *ESR2*. There are two main constraints to the use of multiple transcribed markers in this situation. The first is resource limitation. An increase in the number of assays performed requires the use of greater amounts of RNA, which when taken from an easily accessible tissue, such as

blood, is rarely a limiting factor, but may be problematic in the case of less accessible tissues, where less total RNA may be available. The other major difficulty lies in the identification of suitable transcribed variants for analysis. SNPs are less frequent in coding regions than in non-coding regions and for some transcripts there are very few transcribed variants[336]. The usefulness of a transcribed marker for aerQTL mapping is also dependent upon its minor allele frequency (MAF), as allelic imbalance can only be measured in individuals who are heterozygous at the transcribed marker. The power of the test therefore is dependent on the number of heterozygotes which, for biallelic markers, falls from 50% of the sample when MAF=0.5 to 18% when MAF=0.1 and less than 2% when MAF=0.01[204].

6.3 Limitations

The present study did not identify effects on expression at every locus examined, leaving some of the mechanisms which underlie AF risk variants unexplained. There are a number of limitations to the study which may explain the lack of association identified. In the case of *HCN4*, it is possible that the mechanism of action of the risk SNP is not via gene expression, as the AF risk SNP is in LD with the first exon of *HCN4*. In the case of the other variants, however, the GWAS hit SNP is not in LD with a coding region.

6.3.1 Power

As described in Chapter 2, power calculations performed on the basis of simulations indicated that around 200 individuals would be required in each cohort to adequately detect a biologically meaningful effect. Only the blood cohort met this level and the RAA and IMA cohorts were somewhat smaller. This may mean that the studies performed in tissue were underpowered to detect SNP effects. Notwithstanding, the tissue cohort was the largest that I could assemble given the resource and time limitations of my MD studies, and it has yielded several novel and potentially important associations with expression.

6.2.2 Phenotypic characterisation

In order to facilitate recruitment within a constrained time period the study protocol called for a minimal amount of phenotypic data to be collected on recruitment, namely AF status. A greater amount of phenotypic data, such as BMI, blood pressure, medication history would have allowed further analysis to be performed. However, as outlined in the Introduction Chapter, these phenotypic traits and the *trans* acting influences on expression with which they might be associated would not be expected to influence allelic expression ratio. The adoption of a more complex recruitment protocol, with the attendant time, training and data protection costs, would almost certainly have resulted in recruitment of a smaller cohort, exacerbating the problems of inadequate power.

6.3.3 Selection of candidate genes

Due to the constraints of the current project, only a limited number of transcripts were analysed at each AF risk locus. Care was taken to ensure that the assays used were designed to capture as many of the transcripts of each candidate gene as possible, and in the case of *PITX2* and *ZFHX3* to measure more than one transcript of the gene. As illustrated by the case of the *IRX1* gene described above, however, careful selection of one or two candidate genes at a locus may miss the transcript which mediates the disease risk. Furthermore, for almost all of the candidate genes, several different transcripts have been identified. The results from analysis of *ZFHX3* and *PITX2* expression indicate that there are transcript-specific regulators of expression. Some of the genes studied, particularly SYNE2, have a great number of transcripts, which are differentially expressed in different tissues. It is possible, therefore, that the current study has not examined the correct gene or transcript, and that the AF risk SNPs affect expression of other transcripts, either splice variants or alternative genes, not examined here.

6.3.4 Site of expression

Tissue specificity of gene expression is well recognised, and it has been shown that 50-80% of eQTLs identified in one tissue can be replicated in another[193]. This study examined expression of transcripts in three tissues; peripheral blood, RAA and IMA. It is not surprising that by analysing atrial tissue, it has been possible to identify associations between AF risk SNPs and gene expression that were not identified in peripheral blood. It is interesting however, that in two cases, those of *SYNE2* and *KCNN3*, associations which were identified in blood could not be reproduced in RAA. It is possible that this is simply due to the reduced power to detect effects in the smaller RAA cohort. This observation is important as, due to the relative ease with which peripheral blood samples can be obtained for gene expression analysis when compared with other tissues, it will always be easier to study large cohorts of peripheral blood samples, and so conduct adequately powered investigations. The alternative explanation for the discrepancy in results between the two experiments is that there are effects on expression in peripheral blood that do not operate in atrial appendage. This might be assumed to mean that the effects are not, therefore, important in the pathophysiology of AF; given that the atria are the tissue in which AF exists. However, AF often develops as a consequence of other pathophysiological processes which in turn place strain, usually in the form of elevated atrial pressures, on the atria. It is entirely plausible therefore that a genetic variant which has no effect at all on gene expression in the atria themselves may, by affecting expression of a gene in a distant tissue, increase the risk of AF. For example, genetic variants near *PITX2* and *ZFHX3* have been identified in GWAS of stroke[337] . These are not likely to operate by affecting gene expression in the brain, but rather by affecting the risk of AF. In this case, the identification of genetic effects from a cohort of blood samples is not inferior to the use of a cohort of atrial tissue samples.

6.3.5 Timing of expression

PITX2 is a transcription factor with a well documented role in cardiac development. The function of PITX2 expression in the adult heart is less well understood. It is possible that the AF risk SNPs operate not in the adult at all, but instead affect

expression at a key stage in development which has lasting effects on AF risk in the adult. Patterns of gene expression vary throughout development. It is likely therefore that different regulatory components are active at different time points in the development of the organism. There is good evidence from a study in *Caenorhabditis elegans* nematodes that eQTLs operate at specific time points of the development cycle[338]. It is likely that the same is true of other organisms, including humans. This raises the possibility that risk variants operate at a specific time point in fetal development and that studies performed in adults would not be able to identify the target transcript.

6.4 Future considerations

One of the first steps that I would like to take in order to follow up the results presented here would be to confirm our findings using an alternative methodology. One established technique for measuring gene expression is cap analysis of gene expression (CAGE), which captures the 5' ends of RNA transcripts for quantification. This approach has the benefits of allowing differentiation between transcripts with different transcription start sites which will be useful in separating out the transcripts of those genes which exist in multiple forms. An alternative technique to confirm the results presented here is RNA sequencing (RNAseq). RNA sequencing is an increasingly affordable tool in which the entire transcriptome is sequenced. This allows not only qualitative assessment of which transcript variants are expressed in a particular tissue, but also quantification of the amount of those transcripts present. The use of RNAseq data therefore offers the opportunity to overcome one of the important limitations outlined above, namely the need to select candidate transcripts; all expressed transcripts will be captured, even those which have not previously been described, negating any bias that might arise from selecting specific transcripts for study. RNAseq also, therefore, offers the possibility of identifying *trans* acting effects from distant sites. RNAseq has a greater dynamic range than the other technologies available for genome-wide eQTL analysis. RNAseq will also capture every transcribed SNP in each individual studied. This will allow use of a far greater number of transcribed SNPs in each transcript, allowing more reliable estimation of the AER in a greater number of

individuals, thereby increasing the power of the aerQTL approach. The use of RNAseq data therefore may allow a more powerful analysis of the regions already studied. This may allow identification of eQTLs for genes which were not able to be examined here. RNAseq will be particularly useful in defining the specific transcripts of *SYNE2* and *PITX2* which are regulated by AF risk SNPs. Additionally, for several of the genes studied here, particularly *MYOZ1* and *PITX2*, the levels of expression that we have identified differ from those described in the previously published literature. The possible reasons for these differences have been discussed above, but an RNAseq study of a large number of atrial samples would be justified to provide a reliable comprehensive atlas of gene expression in human atria. Unfortunately, due to the cost and complexity of both CAGE and RNAseq, these confirmatory experiments are beyond the scope of the current MD project.

The other line of enquiry which will be of interest will be to further investigate the associations identified here. Increased expression of *CAV1* is associated with protection from AF. Overexpression of CAV1 has been shown to reduce activity in the TGF1β pathway. It seems possible therefore that agents which promote expression of *CAV1* might have a protective effect in AF. *CAV1* is upregulated by the transcription factor, PPARγ and activation of PPARγ by drugs such as pioglitazone also upregulates *CAV1* expression[339]. Small studies have suggested that pioglitazone may prevent recurrence of AF following catheter ablation and prevent progression of AF in diabetic patients[340,341]. Furthermore, telmisartan an angiotensin II receptor which is, uniquely in its class, a partial PPARγ agonist, appears to be better than other antihypertensives at preventing AF[342,343].

MYOZ1 presents a broader challenge for future investigation as very little is known about its role in cardiac tissue. In the mouse *Myoz2* but not *Myoz1* is expressed in cardiac tissue. This suggests that the mouse model is poorly suited to elucidate the role of *MYOZ1* in AF risk. Anatomical studies will allow definition of the pattern of expression of *MYOZ1* in the human heart. Tissue culture models may provide insights to the effects of increased and decreased expression and identify physiological consequences and downstream targets of altered *MYOZ1* expression.

Similarly the roles of *C9orf3* and *FANCC* in the heart are poorly understood. aerQTLs were identified for an AF risk SNP and both genes. It is possible that only one of these genes mediates the increase in AF risk and that the other is a bystander. Alternatively they may interact in some way. In order to elucidate these mechanisms further work must be done to identify the physiological role, site of action and importance of *C9orf3*. Tissue culture experiments may allow identification of downstream consequences of upregulation or down-regulation of expression.

6.5 Conclusions

By performing meta-analysis of randomised controlled trials of ivabradine, an I_f inhibitor, I demonstrated an increase in risk of incident AF in patients treated with ivabradine *vs.* controls, supporting the role of *HCN4* in AF susceptibility.

Analysis of total expression and allelic expression ratios in whole blood of candidate genes identified associations between AF risk variants and increased expression of *KCNN3* and *SYNE2* and decreased expression of *CAV1*.

Analysis of total expression and allelic expression ratios in right atrial appendage tissue identified further associations between AF risk variants and increased expression of *PITX2a/b* and decreased expression of *MYOZ1*, *CAV1*, *C9orf3* and *FANCC* in right atrial tissue. Furthermore, although the experiments were not designed to detect differences between AF cases and controls, I have shown that AF is associated with reduced expression of *SYNE2*, *HCN4* and *CAV1*.

References

References

1. Stewart S, Hart CL, Hole DJ, McMurray JJ. Population prevalence, incidence, and predictors of atrial fibrillation in the Renfrew/Paisley study. Heart 2001;86:516-21.

2. Libby P, Braunwald E. Braunwald's heart disease : a textbook of cardiovascular medicine. 8th ed. Philadelphia: Saunders/Elsevier; 2008.

3. Iwasaki YK, Nishida K, Kato T, Nattel S. Atrial fibrillation pathophysiology: implications for management. Circulation 2011;124:2264-74.

4. Haissaguerre M, Jais P, Shah DC, et al. Spontaneous initiation of atrial fibrillation by ectopic beats originating in the pulmonary veins. N Engl J Med 1998;339:659-66.

5. Jalife J, Berenfeld O, Mansour M. Mother rotors and fibrillatory conduction: a mechanism of atrial fibrillation. Cardiovasc Res 2002;54:204-16.

6. Camm AJ, Kirchhof P, Lip GY, et al. Guidelines for the management of atrial fibrillation: the Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology (ESC). Europace 2010;12:1360-420.

7. Jahangir A, Lee V, Friedman PA, et al. Long-term progression and outcomes with aging in patients with lone atrial fibrillation: a 30-year follow-up study. Circulation 2007;115:3050-6.

8. Wolff L. Familial Auricular Fibrillation. N Engl J Med 1943;229:396-8.

9. Brugada R, Tapscott T, Czernuszewicz GZ, et al. Identification of a genetic locus for familial atrial fibrillation. N Engl J Med 1997;336:905-11.

10. Rang HP. Pharmacology. 4th ed. New York: Churchill Livingstone; 2001.

11. Calkins H, Kuck KH, Cappato R, et al. 2012 HRS/EHRA/ECAS Expert Consensus Statement on Catheter and Surgical Ablation of Atrial Fibrillation: recommendations for patient selection, procedural techniques, patient management and follow-up, definitions, endpoints, and research trial design. Europace 2012;14:528-606.

12. Packer DL, Kowal RC, Wheelan KR, et al. Cryoballoon ablation of pulmonary veins for paroxysmal atrial fibrillation: first results of the North American Arctic Front (STOP AF) pivotal trial. J Am Coll Cardiol 2013;61:1713-23.

13. Christophersen IE, Ravn LS, Budtz-Joergensen E, et al. Familial aggregation of atrial fibrillation: a study in Danish twins. Circ Arrhythm Electrophysiol 2009;2:378-83.

14. Peden JF, Farrall M. Thirty-five common variants for coronary artery disease: the fruits of much collaborative labour. Hum Mol Genet 2011;20:R198-205.

15. Locatelli I, Lichtenstein P, Yashin AI. The heritability of breast cancer: a Bayesian correlated frailty model applied to Swedish twins data. Twin research : the official journal of the International Society for Twin Studies 2004;7:182-91.

16. Fox CS, Parise H, D'Agostino RB, Sr., et al. Parental atrial fibrillation as a risk factor for atrial fibrillation in offspring. JAMA 2004;291:2851-5.

17. Poulsen P, Kyvik KO, Vaag A, Beck-Nielsen H. Heritability of type II (non-insulin-dependent) diabetes mellitus and abnormal glucose tolerance--a population-based twin study. Diabetologia 1999;42:139-45.

18. Kupper N, Willemsen G, Riese H, Posthuma D, Boomsma DI, de Geus EJ. Heritability of daytime ambulatory blood pressure in an extended twin design. Hypertension 2005;45:80-5.

19. Allessie M, Ausma J, Schotten U. Electrical, contractile and structural remodeling during atrial fibrillation. Cardiovasc Res 2002;54:230-46.

20. Olson TM, Michels VV, Ballew JD, et al. Sodium channel mutations and susceptibility to heart failure and atrial fibrillation. JAMA 2005;293:447-54.

21. Chen YH, Xu SJ, Bendahhou S, et al. KCNQ1 gain-of-function mutation in familial atrial fibrillation. Science 2003;299:251-4.

22. Hodgson-Zingman DM, Karst ML, Zingman LV, et al. Atrial natriuretic peptide frameshift mutation in familial atrial fibrillation. N Engl J Med 2008;359:158-65.

23. Gollob MH, Jones DL, Krahn AD, et al. Somatic mutations in the connexin 40 gene (GJA5) in atrial fibrillation. N Engl J Med 2006;354:2677-88.

24. Olesen MS, Bentzen BH, Nielsen JB, et al. Mutations in the potassium channel subunit KCNE1 are associated with early-onset familial atrial fibrillation. BMC Med Genet 2012;13:24.

25. Mann SA, Otway R, Guo G, et al. Epistatic effects of potassium channel variation on cardiac repolarization and atrial fibrillation risk. J Am Coll Cardiol 2012;59:1017-25.

26. Ravn LS, Aizawa Y, Pollevick GD, et al. Gain of function in IKs secondary to a mutation in KCNE5 associated with atrial fibrillation. Heart Rhythm 2008;5:427-35.

27. Brauch KM, Chen LY, Olson TM. Comprehensive mutation scanning of LMNA in 268 patients with lone atrial fibrillation. Am J Cardiol 2009;103:1426-8.

28. den Hoed M, Eijgelsheim M, Esko T, et al. Identification of heart rate-associated loci and their effects on cardiac conduction and rhythm disorders. Nat Genet 2013;45:621-31.

29. Eijgelsheim M, Newton-Cheh C, Sotoodehnia N, et al. Genome-wide association analysis identifies multiple loci related to resting heart rate. Hum Mol Genet 2010;19:3885-94.

30. A Catalog of Published Genome-Wide Association Studies. 2011. at www.genome.gov/gwastudies.)

31. Hindorff LA, Sethupathy P, Junkins HA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A 2009;106:9362-7.

32. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. Nature 2001;409:860-921.

33. Frazer KA, Murray SS, Schork NJ, Topol EJ. Human genetic variation and its contribution to complex traits. Nat Rev Genet 2009;10:241-51.

34. Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. Nat Rev Genet 2006;7:85-97.

35. Cooper DN, Nature Publishing Group. Nature encyclopedia of the human genome. London ; New York: Nature Pub. Group; 2003.

36. Levy S, Sutton G, Ng PC, et al. The diploid genome sequence of an individual human. PLoS Biol 2007;5:e254.

37. Stefansson H, Helgason A, Thorleifsson G, et al. A common inversion under selection in Europeans. Nat Genet 2005;37:129-37.

38. Pelak K, Shianna KV, Ge D, et al. The characterization of twenty sequenced human genomes. PLoS Genet 2010;6.

39. Sherry ST, Ward MH, Kholodov M, et al. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 2001;29:308-11.

40. Kruglyak L, Nickerson DA. Variation is the spice of life. Nat Genet 2001;27:234-6.

41. Collins A. Allelic association: linkage disequilibrium structure and gene mapping. Mol Biotechnol 2009;41:83-9.

42. Chakravarti A, Buetow KH, Antonarakis SE, Waber PG, Boehm CD, Kazazian HH. Nonuniform recombination within the human beta-globin gene cluster. Am J Hum Genet 1984;36:1239-58.

43. Gabriel SB, Schaffner SF, Nguyen H, et al. The structure of haplotype blocks in the human genome. Science 2002;296:2225-9.

44. A haplotype map of the human genome. Nature 2005;437:1299-320.

45. Service S, DeYoung J, Karayiorgou M, et al. Magnitude and distribution of linkage disequilibrium in population isolates and implications for genome-wide association studies. Nat Genet 2006;38:556-60.

46. Cunnington MS, Keavney BD. Genetics of coronary heart disease. In: Yusuf S, Cairns J, Camm J, Fallen EL, Gersh BJ, eds. Evidence based cardiology. 3rd ed: Wiley-Blackwell; 2010:268-87.

47. Hinds DA, Kloek AP, Jen M, Chen X, Frazer KA. Common deletions and SNPs are in linkage disequilibrium in the human genome. Nat Genet 2006;38:82-5.

48. Hallast P, Nagirnaja L, Margus T, Laan M. Segmental duplications and gene conversion: Human luteinizing hormone/chorionic gonadotropin beta gene cluster. Genome Res 2005;15:1535-46.

49. Ellinor PT, Shin JT, Moore RK, Yoerger DM, MacRae CA. Locus for atrial fibrillation maps to chromosome 6q14-16. Circulation 2003;107:2880-3.

50. Volders PG, Zhu Q, Timmermans C, et al. Mapping a novel locus for familial atrial fibrillation on chromosome 10p11-q21. Heart Rhythm 2007;4:469-75.

51. Darbar D, Hardy A, Haines JL, Roden DM. Prolonged signal-averaged P-wave duration as an intermediate phenotype for familial atrial fibrillation. J Am Coll Cardiol 2008;51:1083-9.

52. MacArthur DG, Balasubramanian S, Frankish A, et al. A systematic survey of loss-of-function variants in human protein-coding genes. Science 2012;335:823-8.

53. Zimmer T, Surber R. SCN5A channelopathies--an update on mutations and mechanisms. Progress in biophysics and molecular biology 2008;98:120-36.

54. Hershberger RE, Parks SB, Kushner JD, et al. Coding sequence mutations identified in MYH7, TNNT2, SCN5A, CSRP3, LBD3, and TCAP from 313 patients with familial or idiopathic dilated cardiomyopathy. Clin Transl Sci 2008;1:21-6.

55. Cheng J, Morales A, Siegfried JD, et al. SCN5A rare variants in familial dilated cardiomyopathy decrease peak sodium current depending on the common polymorphism H558R and common splice variant Q1077del. Clin Transl Sci 2010;3:287-94.

56. McNair WP, Sinagra G, Taylor MR, et al. SCN5A mutations associate with arrhythmic dilated cardiomyopathy and commonly localize to the voltage-sensing mechanism. J Am Coll Cardiol 2011;57:2160-8.

57. Makiyama T, Akao M, Shizuta S, et al. A novel SCN5A gain-of-function mutation M1875T associated with familial atrial fibrillation. J Am Coll Cardiol 2008;52:1326-34.

58. Watanabe H, Darbar D, Kaiser DW, et al. Mutations in sodium channel beta1- and beta2-subunits associated with atrial fibrillation. Circ Arrhythm Electrophysiol 2009;2:268-75.

59. Hong K, Piper DR, Diaz-Valdecantos A, et al. De novo KCNQ1 mutation responsible for atrial fibrillation and short QT syndrome in utero. Cardiovasc Res 2005;68:433-40.

60. Yang Y, Xia M, Jin Q, et al. Identification of a KCNE2 gain-of-function mutation in patients with familial atrial fibrillation. Am J Hum Genet 2004;75:899-905.

61. Zhang DF, Liang B, Lin J, Liu B, Zhou QS, Yang YQ. [KCNE3 R53H substitution in familial atrial fibrillation]. Chin Med J (Engl) 2005;118:1735-8.

62. Olson TM, Alekseev AE, Liu XK, et al. Kv1.5 channelopathy due to KCNA5 loss-of-function mutation causes human atrial fibrillation. Hum Mol Genet 2006;15:2185-91.

63. Yang C, Kwan YW, Chan SW, Lee SM, Leung GP. Potentiation of EDHF-mediated relaxation by chloride channel blockers. Acta Pharmacol Sin 2010;31:1303-11.

64. Yang T, Yang P, Roden DM, Darbar D. Novel KCNA5 mutation implicates tyrosine kinase signaling in human atrial fibrillation. Heart Rhythm 2010;7:1246-52.

65. Yang Y, Li J, Lin X, et al. Novel KCNA5 loss-of-function mutations responsible for atrial fibrillation. J Hum Genet 2009;54:277-83.

66. Olson TM, Alekseev AE, Moreau C, et al. KATP channel mutation confers risk for vein of Marshall adrenergic atrial fibrillation. Nat Clin Pract Cardiovasc Med 2007;4:110-6.

67. Xia M, Jin Q, Bendahhou S, et al. A Kir2.1 gain-of-function mutation underlies familial atrial fibrillation. Biochem Biophys Res Commun 2005;332:1012-9.

68. Bhuiyan ZA, van den Berg MP, van Tintelen JP, et al. Expanding spectrum of human RYR2-related disease: new electrocardiographic, structural, and genetic features. Circulation 2007;116:1569-76.

69. Thibodeau IL, Xu J, Li Q, et al. Paradigm of genetic mosaicism and lone atrial fibrillation: physiological characterization of a connexin 43-deletion mutant identified from atrial tissue. Circulation 2010;122:236-44.

70. Yang YQ, Zhang XL, Wang XH, et al. Connexin40 nonsense mutation in familial atrial fibrillation. Int J Mol Med 2010;26:605-10.

71. Yang YQ, Liu X, Zhang XL, et al. Novel connexin40 missense mutations in patients with familial atrial fibrillation. Europace 2010;12:1421-7.

72. Postma AV, van de Meerakker JB, Mathijssen IB, et al. A gain-of-function TBX5 mutation is associated with atypical Holt-Oram syndrome and paroxysmal atrial fibrillation. Circ Res 2008;102:1433-42.

73. Wang J, Sun YM, Yang YQ. Mutation spectrum of the GATA4 gene in patients with idiopathic atrial fibrillation. Mol Biol Rep 2012;39:8127-35.

74. Yang YQ, Wang MY, Zhang XL, et al. GATA4 loss-of-function mutations in familial atrial fibrillation. Clin Chim Acta 2011;412:1825-30.

75. Li J, Liu WD, Yang ZL, Yang YQ. Novel GATA6 loss-of-function mutation responsible for familial atrial fibrillation. Int J Mol Med 2012;30:783-90.

76. Yang YQ, Li L, Wang J, et al. GATA6 loss-of-function mutation in atrial fibrillation. Eur J Med Genet 2012;55:520-6.

77. Yang YQ, Wang XH, Tan HW, Jiang WF, Fang WY, Liu X. Prevalence and spectrum of GATA6 mutations associated with familial atrial fibrillation. Int J Cardiol 2012;155:494-6.

78. Zhang X, Chen S, Yoo S, et al. Mutation in nuclear pore component NUP155 leads to atrial fibrillation and early sudden cardiac death. Cell 2008;135:1017-27.

79. van der Velden HM, Jongsma HJ. Cardiac gap junctions and connexins: their role in atrial fibrillation and potential as therapeutic targets. Cardiovasc Res 2002;54:270-9.

80. Otway R, Vandenberg JI, Guo G, et al. Stretch-sensitive KCNQ1 mutation A link between genetic and environmental factors in the pathogenesis of atrial fibrillation? J Am Coll Cardiol 2007;49:578-86.

81. Das S, Makino S, Melman YF, et al. Mutation in the S3 segment of KCNQ1 results in familial lone atrial fibrillation. Heart Rhythm 2009;6:1146-53.

82. Bartos DC, Anderson JB, Bastiaenen R, et al. A KCNQ1 mutation causes a high penetrance for familial atrial fibrillation. J Cardiovasc Electrophysiol 2013;24:562-9.

83. Bartos DC, Duchatelet S, Burgess DE, et al. R231C mutation in KCNQ1 causes long QT syndrome type 1 and familial atrial fibrillation. Heart Rhythm 2011;8:48-55.

84. Lundby A, Ravn LS, Svendsen JH, Hauns S, Olesen SP, Schmitt N. KCNE3 mutation V17M identified in a patient with lone atrial fibrillation. Cell Physiol Biochem 2008;21:47-54.

85. Olesen MS, Refsgaard L, Holst AG, et al. A novel KCND3 gain-of-function mutation associated with early-onset of persistent lone atrial fibrillation. Cardiovasc Res 2013;98:488-95.

86. Christophersen IE, Olesen MS, Liang B, et al. Genetic variation in KCNA5: impact on the atrial-specific potassium current IKur in patients with lone atrial fibrillation. Eur Heart J 2013;34:1517-25.

87. Darbar D, Kannankeril PJ, Donahue BS, et al. Cardiac sodium channel (SCN5A) variants associated with atrial fibrillation. Circulation 2008;117:1927-35.

88. Ellinor PT, Nam EG, Shea MA, Milan DJ, Ruskin JN, MacRae CA. Cardiac sodium channel mutation in atrial fibrillation. Heart Rhythm 2008;5:99-105.

89. Olesen MS, Yuan L, Liang B, et al. High prevalence of long QT syndrome-associated SCN5A variants in patients with early-onset lone atrial fibrillation. Circ Cardiovasc Genet 2012;5:450-9.

90. Ackerman MJ, Siu BL, Sturner WQ, et al. Postmortem molecular analysis of SCN5A defects in sudden infant death syndrome. JAMA 2001;286:2264-9.

91. Arnestad M, Crotti L, Rognum TO, et al. Prevalence of long-QT syndrome gene variants in sudden infant death syndrome. Circulation 2007;115:361-7.

92. Li Q, Huang H, Liu G, et al. Gain-of-function mutation of Nav1.5 in atrial fibrillation enhances cellular excitability and lowers the threshold for action potential firing. Biochem Biophys Res Commun 2009;380:132-7.

93. Olesen MS, Jespersen T, Nielsen JB, et al. Mutations in sodium channel beta-subunit SCN3B are associated with early-onset lone atrial fibrillation. Cardiovasc Res 2011;89:786-93.

94. Wang P, Yang Q, Wu X, et al. Functional dominant-negative mutation of sodium channel subunit gene SCN3B associated with atrial fibrillation in a Chinese GeneID population. Biochem Biophys Res Commun 2010;398:98-104.

95. Beavers DL, Wang W, Ather S, et al. Mutation E169K in junctophilin-2 causes atrial fibrillation due to impaired RyR2 stabilization. J Am Coll Cardiol 2013;62:2010-9.

96. Sun Y, Yang YQ, Gong XQ, et al. Novel germline GJA5/connexin40 mutations associated with lone atrial fibrillation impair gap junctional intercellular communication. Hum Mutat 2013;34:603-9.

97. Jiang JQ, Shen FF, Fang WY, Liu X, Yang YQ. Novel GATA4 mutations in lone atrial fibrillation. Int J Mol Med 2011;28:1025-32.

98. Posch MG, Boldt LH, Polotzki M, et al. Mutations in the cardiac transcription factor GATA4 in patients with lone atrial fibrillation. Eur J Med Genet 2010;53:201-3.

99. Beckmann BM, Holinski-Feder E, Walter MC, et al. Laminopathy presenting as familial atrial fibrillation. Int J Cardiol 2010;145:394-6.

100. Ritchie MD, Rowan S, Kucera G, et al. Chromosome 4q25 variants are genetic modifiers of rare ion channel mutations associated with familial atrial fibrillation. J Am Coll Cardiol 2012;60:1173-81.

101. Xie WH, Chang C, Xu YJ, et al. Prevalence and spectrum of Nkx2.5 mutations associated with idiopathic atrial fibrillation. Clinics 2013;68:777-84.

102. Online Mendelian Inheritance in Man, OMIM®. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD). (Accessed 2013, at http://omim.org/.)

103. Sinner MF, Lubitz SA, Pfeufer A, et al. Lack of replication in polymorphisms reported to be associated with atrial fibrillation. Heart Rhythm 2011;8:403-9.

104. Marian AJ, Belmont J. Strategic approaches to unraveling genetic causes of cardiovascular diseases. Circ Res 2011;108:1252-69.

105. Gudbjartsson DF, Arnar DO, Helgadottir A, et al. Variants conferring risk of atrial fibrillation on chromosome 4q25. Nature 2007;448:353-7.

106. Gudbjartsson DF, Holm H, Gretarsdottir S, et al. A sequence variant in ZFHX3 on 16q22 associates with atrial fibrillation and ischemic stroke. Nat Genet 2009;41:876-8.

107. Benjamin EJ, Rice KM, Arking DE, et al. Variants in ZFHX3 are associated with atrial fibrillation in individuals of European ancestry. Nat Genet 2009;41:879-81.

108. Ellinor PT, Lunetta KL, Glazer NL, et al. Common variants in KCNN3 are associated with lone atrial fibrillation. Nat Genet 2010;42:240-4.

109. Ellinor PT, Lunetta KL, Albert CM, et al. Meta-analysis identifies six new susceptibility loci for atrial fibrillation. Nat Genet 2012;44:670-5.

110. Christiansen J, Dyck JD, Elyas BG, et al. Chromosome 1q21.1 contiguous gene deletion is associated with congenital heart disease. Circ Res 2004;94:1429-35.

111. Tuteja D, Rafizadeh S, Timofeyev V, et al. Cardiac small conductance Ca2+-activated K+ channel subunits form heteromultimers via the coiled-coil domains in the C termini of the channels. Circ Res 2010;107:851-9.

112. Weatherall KL, Seutin V, Liegeois JF, Marrion NV. Crucial role of a shared extracellular loop in apamin sensitivity and maintenance of pore shape of small-conductance calcium-activated potassium (SK) channels. Proc Natl Acad Sci U S A 2011;108:18494-9.

113. Ozgen N, Dun W, Sosunov EA, et al. Early electrical remodeling in rabbit pulmonary vein results from trafficking of intracellular SK2 channels to membrane sites. Cardiovasc Res 2007;75:758-69.

114. Van Wagoner DR. Electrophysiological remodeling in human atrial fibrillation. Pacing and clinical electrophysiology : PACE 2003;26:1572-5.

115. Milkau M, Kohler R, de Wit C. Crucial importance of the endothelial K+ channel SK3 and connexin40 in arteriolar dilations during skeletal muscle contraction. FASEB J 2010;24:3572-9.

116. Bergwerff M, Gittenberger-de Groot AC, Wisse LJ, et al. Loss of function of the Prx1 and Prx2 homeobox genes alters architecture of the great elastic arteries and ductus arteriosus. Virchows Arch 2000;436:12-9.

117. Ihida-Stansbury K, McKean DM, Gebb SA, et al. Paired-related homeobox gene Prx1 is required for pulmonary vascular development. Circ Res 2004;94:1507-14.

118. Essner JJ, Branford WW, Zhang J, Yost HJ. Mesendoderm and left-right brain, heart and gut development are differentially regulated by pitx2 isoforms. Development 2000;127:1081-93.

119. L'Honore A, Ouimette JF, Lavertu-Jolin M, Drouin J. Pitx2 defines alternate pathways acting through MyoD during limb and somitic myogenesis. Development 2010;137:3847-56.

120. Zacharias AL, Lewandoski M, Rudnicki MA, Gage PJ. Pitx2 is an upstream activator of extraocular myogenesis and survival. Dev Biol 2011;349:395-405.

121. Shiratori H, Yashiro K, Shen MM, Hamada H. Conserved regulation and role of Pitx2 in situs-specific morphogenesis of visceral organs. Development 2006;133:3015-25.

122. Schweickert A, Campione M, Steinbeisser H, Blum M. Pitx2 isoforms: involvement of Pitx2c but not Pitx2a or Pitx2b in vertebrate left-right asymmetry. Mech Develop 2000;90:41-51.

123. Cox CJ, Espinoza HM, McWilliams B, et al. Differential regulation of gene expression by PITX2 isoforms. J Biol Chem 2002;277:25001-10.

124. Hsu J, Hanna P, Van Wagoner DR, et al. Whole genome expression differences in human left and right atria ascertained by RNA sequencing. Circ Cardiovasc Genet 2012;5:327-35.

125. Kirchhof P, Kahr PC, Kaese S, et al. PITX2c is expressed in the adult left atrium, and reducing Pitx2c expression promotes atrial fibrillation inducibility and complex changes in gene expression. Circ Cardiovasc Genet 2011;4:123-33.

126. Mommersteeg MT, Hoogaars WM, Prall OW, et al. Molecular pathway for the localized formation of the sinoatrial node. Circ Res 2007;100:354-62.

127. Wang J, Klysik E, Sood S, Johnson RL, Wehrens XH, Martin JF. Pitx2 prevents susceptibility to atrial arrhythmias by inhibiting left-sided pacemaker specification. Proc Natl Acad Sci U S A 2010;107:9753-8.

128. Logan M, Pagan-Westphal SM, Smith DM, Paganessi L, Tabin CJ. The transcription factor Pitx2 mediates situs-specific morphogenesis in response to left-right asymmetric signals. Cell 1998;94:307-17.

129. Gore-Panter SR, Hsu J, Hanna P, et al. Atrial Fibrillation Associated Chromosome 4q25 Variants Are Not Associated with PITX2c Expression in Human Adult Left Atrial Appendages. PloS one 2014;9:e86245.

130. Lin H, Dolmatova EV, Morley MP, et al. Gene expression and genetic variation in human atria. Heart Rhythm 2014;11:266-71.

131. Pfeufer A, van Noord C, Marciante KD, et al. Genome-wide association study of PR interval. Nat Genet 2010;42:153-9.

132. Pojoga LH, Underwood PC, Goodarzi MO, et al. Variants of the caveolin-1 gene: a translational investigation linking insulin resistance and hypertension. J Clin Endocrinol Metab 2011;96:E1288-92.

133. Diaz-Perales A, Quesada V, Sanchez LM, et al. Identification of human aminopeptidase O, a novel metalloprotease with structural similarity to aminopeptidase B and leukotriene A4 hydrolase. J Biol Chem 2005;280:14310-7.

134. Kaschina E, Unger T. Angiotensin AT1/AT2 receptors: regulation, signalling and function. Blood Press 2003;12:70-88.

135. Beqqali A, Monshouwer-Kloots J, Monteiro R, et al. CHAP is a newly identified Z-disc protein essential for heart and skeletal muscle function. J Cell Sci 2010;123:1141-50.

136. Frey N, Richardson JA, Olson EN. Calsarcins, a novel family of sarcomeric calcineurin-binding proteins. Proc Natl Acad Sci U S A 2000;97:14632-7.

137. Frey N, Frank D, Lippl S, et al. Calsarcin-2 deficiency increases exercise capacity in mice through calcineurin/NFAT activation. J Clin Invest 2008;118:3598-608.

138. Rajgor D, Mellad JA, Autore F, Zhang Q, Shanahan CM. Multiple novel nesprin-1 and nesprin-2 variants act as versatile tissue-specific intracellular scaffolds. PloS one 2012;7:e40098.

139. Zhang Q, Bethmann C, Worth NF, et al. Nesprin-1 and -2 are involved in the pathogenesis of Emery Dreifuss muscular dystrophy and are critical for nuclear envelope integrity. Hum Mol Genet 2007;16:2816-33.

140. Milanesi R, Baruscotti M, Gnecchi-Ruscone T, DiFrancesco D. Familial sinus bradycardia associated with a mutation in the cardiac pacemaker channel. N Engl J Med 2006;354:151-7.

141. Ueda K, Nakamura K, Hayashi T, et al. Functional characterization of a trafficking-defective HCN4 mutation, D553N, associated with cardiac arrhythmia. J Biol Chem 2004;279:27194-8.

142. Schulze-Bahr E, Neu A, Friederich P, et al. Pacemaker channel dysfunction in a patient with sinus node disease. J Clin Invest 2003;111:1537-45.

143. Baruscotti M, Bucchi A, Viscomi C, et al. Deep bradycardia and heart block caused by inducible cardiac-specific knockout of the pacemaker channel gene Hcn4. Proc Natl Acad Sci U S A 2011;108:1705-10.

144. Barbuti A, Scavone A, Mazzocchi N, Terragni B, Baruscotti M, Difrancesco D. A caveolin-binding domain in the HCN4 channels mediates functional interaction with caveolin proteins. J Mol Cell Cardiol 2012;53:187-95.

145. Flicek P, Ahmed I, Amode MR, et al. Ensembl 2013. Nucleic Acids Res 2013;41:D48-55.

146. Kim TS, Kawaguchi M, Suzuki M, et al. The ZFHX3 (ATBF1) transcription factor induces PDGFRB, which activates ATM in the cytoplasm to protect cerebellar neurons from oxidative stress. Dis Model Mech 2010;3:752-62.

147. Burstein B, Libby E, Calderone A, Nattel S. Differential behaviors of atrial versus ventricular fibroblasts: a potential role for platelet-derived growth factor in atrial-ventricular remodeling differences. Circulation 2008;117:1630-41.

148. Miura Y, Tam T, Ido A, et al. Cloning and characterization of an ATBF1 isoform that expresses in a neuronal differentiation-dependent manner. J Biol Chem 1995;270:26840-8.

149. Cho YG, Song JH, Kim CJ, et al. Genetic alterations of the ATBF1 gene in gastric cancer. Clin Cancer Res 2007;13:4355-9.

150. Kim CJ, Song JH, Cho YG, et al. Down-regulation of ATBF1 is a major inactivating mechanism in hepatocellular carcinoma. Histopathology 2008;52:552-9.

151. Zhang Z, Yamashita H, Toyama T, et al. ATBF1-a messenger RNA expression is correlated with better prognosis in breast cancer. Clin Cancer Res 2005;11:193-8.

152. Sun X, Frierson HF, Chen C, et al. Frequent somatic mutations of the transcription factor ATBF1 in human prostate cancer. Nat Genet 2005;37:407-12.

153. Miura Y, Kataoka H, Joh T, et al. Susceptibility to killer T cells of gastric cancer cells enhanced by Mitomycin-C involves induction of ATBF1 and activation of p21 (Waf1/Cip1) promoter. Microbiol Immunol 2004;48:137-45.

154. Franco D, Daimi H, Dominguez JN, et al. PITX2 is upstream of ZFHX3 and KCNN3: epistatic and functional relationship of GWAS atrial fibrillation-associated candidate genes. Cardiovascular Research 2012;93:S90-S.

155. Delaney JT, Jeff JM, Brown NJ, et al. Characterization of genome-wide association-identified variants for atrial fibrillation in African Americans. PloS one 2012;7:e32338.

156. Smith JG, Newton-Cheh C, Almgren P, Melander O, Platonov PG. Genetic polymorphisms for estimating risk of atrial fibrillation in the general population: a prospective study. Arch Intern Med 2012;172:742-4.

157. Lubitz SA, Sinner MF, Lunetta KL, et al. Independent susceptibility markers for atrial fibrillation on chromosome 4q25. Circulation 2010;122:976-84.

158. Body SC, Collard CD, Shernan SK, et al. Variation in the 4q25 chromosomal locus predicts atrial fibrillation after coronary artery bypass graft surgery. Circ Cardiovasc Genet 2009;2:499-506.

159. Husser D, Adams V, Piorkowski C, Hindricks G, Bollmann A. Chromosome 4q25 variants and atrial fibrillation recurrence after catheter ablation. J Am Coll Cardiol 2010;55:747-53.

160. Parvez B, Vaglio J, Rowan S, et al. Symptomatic response to antiarrhythmic drug therapy is modulated by a common single nucleotide polymorphism in atrial fibrillation. J Am Coll Cardiol 2012;60:539-45.

161. Naylor S. Biomarkers: current perspectives and future prospects. Expert review of molecular diagnostics 2003;3:525-9.

162. Hoeller R, Rubini Gimenez M, Reichlin T, et al. Normal presenting levels of high-sensitivity troponin and myocardial infarction. Heart 2013;99:1567-72.

163. Maisel AS, Krishnaswamy P, Nowak RM, et al. Rapid measurement of B-type natriuretic peptide in the emergency diagnosis of heart failure. N Engl J Med 2002;347:161-7.

164. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature 2010;466:707-13.

165. Chen Q, Kirsch GE, Zhang D, et al. Genetic basis and molecular mechanism for idiopathic ventricular fibrillation. Nature 1998;392:293-6.

166. Clancy CE, Rudy Y. Linking a genetic defect to its cellular phenotype in a cardiac arrhythmia. Nature 1999;400:566-9.

167. Echt DS, Liebson PR, Mitchell LB, et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. N Engl J Med 1991;324:781-8.

168. Splawski I, Shen J, Timothy KW, et al. Spectrum of mutations in long-QT syndrome genes. KVLQT1, HERG, SCN5A, KCNE1, and KCNE2. Circulation 2000;102:1178-85.

169. Touboul P, Brugada J, Capucci A, Crijns HJ, Edvardsson N, Hohnloser SH. Dronedarone for prevention of atrial fibrillation: a dose-ranging study. Eur Heart J 2003;24:1481-7.

170. Vaughan Williams EM. QT and action potential duration. Br Heart J 1982;47:513-4.

171. Wulff H, Castle NA, Pardo LA. Voltage-gated potassium channels as therapeutic targets. Nat Rev Drug Discov 2009;8:982-1001.

172. Swedberg K, Komajda M, Bohm M, et al. Ivabradine and outcomes in chronic heart failure (SHIFT): a randomised placebo-controlled study. Lancet 2010;376:875-85.

173. British National Formulary online. (Accessed 12/11/2013, 2013, at http://www.medicinescomplete.com/mc/bnf/current/.)

174. Fox K, Ford I, Steg PG, Tendera M, Ferrari R. Ivabradine for patients with stable coronary artery disease and left-ventricular systolic dysfunction (BEAUTIFUL): a randomised, double-blind, placebo-controlled trial. Lancet 2008;372:807-16.

175. Tardif JC, Ford I, Tendera M, Bourassa MG, Fox K. Efficacy of ivabradine, a new selective I(f) inhibitor, compared with atenolol in patients with chronic stable angina. Eur Heart J 2005;26:2529-36.

176. Tardif JC, Ponikowski P, Kahan T. Efficacy of the I(f) current inhibitor ivabradine in patients with chronic stable angina receiving beta-blocker therapy: a 4-month, randomized, placebo-controlled trial. Eur Heart J 2009;30:540-8.

177. Ruzyllo W, Tendera M, Ford I, Fox KM. Antianginal efficacy and safety of ivabradine compared with amlodipine in patients with stable effort angina pectoris: a 3-month randomised, double-blind, multicentre, noninferiority trial. Drugs 2007;67:393-405.

178. Diness JG, Sorensen US, Nissen JD, et al. Inhibition of small-conductance Ca2+-activated K+ channels terminates and protects against atrial fibrillation. Circ Arrhythm Electrophysiol 2010;3:380-90.

179. Schneider MP, Hua TA, Bohm M, Wachtell K, Kjeldsen SE, Schmieder RE. Prevention of atrial fibrillation by Renin-Angiotensin system inhibition a meta-analysis. J Am Coll Cardiol 2010;55:2299-307.

180. Tessari A, Pietrobon M, Notte A, et al. Myocardial Pitx2 differentially regulates the left atrial identity and ventricular asymmetric remodeling programs. Circ Res 2008;102:813-22.

181. Judge DP, Biery NJ, Keene DR, et al. Evidence for a critical contribution of haploinsufficiency in the complex pathogenesis of Marfan syndrome. J Clin Invest 2004;114:172-81.

182. Exome Variant Server. at http://evs.gs.washington.edu/EVS/.)

183. Olesen MS, Andreasen L, Jabbari J, et al. Very early-onset lone atrial fibrillation patients have a high prevalence of rare variants in genes previously associated with atrial fibrillation. Heart Rhythm 2014;11:246-51.

184. Antonarakis SE, Irkin SH, Cheng TC, et al. beta-Thalassemia in American Blacks: novel mutations in the "TATA" box and an acceptor splice site. Proc Natl Acad Sci U S A 1984;81:1154-8.

185. Cunnington MS, Santibanez Koref M, Mayosi BM, Burn J, Keavney B. Chromosome 9p21 SNPs Associated with Multiple Disease Phenotypes Correlate with ANRIL Expression. PLoS Genet 2010;6:e1000899.

186. Ueda H, Howson JM, Esposito L, et al. Association of the T-cell regulatory gene CTLA4 with susceptibility to autoimmune disease. Nature 2003;423:506-11.

187. Consortium EP, Bernstein BE, Birney E, et al. An integrated encyclopedia of DNA elements in the human genome. Nature 2012;489:57-74.

188. Thurman RE, Rynes E, Humbert R, et al. The accessible chromatin landscape of the human genome. Nature 2012;489:75-82.

189. Neph S, Vierstra J, Stergachis AB, et al. An expansive human regulatory lexicon encoded in transcription factor footprints. Nature 2012;489:83-90.

190. Sanyal A, Lajoie BR, Jain G, Dekker J. The long-range interaction landscape of gene promoters. Nature 2012;489:109-13.

191. Djebali S, Davis CA, Merkel A, et al. Landscape of transcription in human cells. Nature 2012;489:101-8.

192. Andersson R, Gebhard C, Miguel-Escalada I, et al. An atlas of active enhancers across human cell types and tissues. Nature 2014;507:455-61.

193. Grundberg E, Small KS, Hedman AK, et al. Mapping cis- and trans-regulatory effects across multiple tissues in twins. Nat Genet 2012;44:1084-9.

194. Nica AC, Parts L, Glass D, et al. The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. PLoS Genet 2011;7:e1002003.

195. Lappalainen T, Sammeth M, Friedlander MR, et al. Transcriptome and genome sequencing uncovers functional variation in humans. Nature 2013;501:506-11.

196. Stranger BE, Montgomery SB, Dimas AS, et al. Patterns of cis regulatory variation in diverse human populations. PLoS Genet 2012;8:e1002639.

197. Xin M, Olson EN, Bassel-Duby R. Mending broken hearts: cardiac development as a basis for adult heart regeneration and repair. Nature reviews Molecular cell biology 2013;14:529-41.

198. Cheung VG, Conlin LK, Weber TM, et al. Natural variation in human gene expression assessed in lymphoblastoid cells. Nat Genet 2003;33:422-5.

199. Cheung VG, Spielman RS, Ewens KG, Weber TM, Morley M, Burdick JT. Mapping determinants of human gene expression by regional and genome-wide association. Nature 2005;437:1365-9.

200. Morley M, Molony CM, Weber TM, et al. Genetic analysis of genome-wide variation in human gene expression. Nature 2004;430:743-7.

201. Sagai T, Hosoya M, Mizushina Y, Tamura M, Shiroishi T. Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. Development 2005;132:797-803.

202. Tian D, Sun S, Lee JT. The Long Noncoding RNA, Jpx, Is a Molecular Switch for X Chromosome Inactivation. Cell 2010;143:390-403.

203. Brockdorff N. X-chromosome inactivation: closing in on proteins that bind Xist RNA. Trends Genet 2002;18:352-8.

204. Teare MD, Pinyakorn S, Heighway J, Santibanez Koref MF. Comparing methods for mapping cis acting polymorphisms using allelic expression ratios. PloS one 2011;6:e28636.

205. Pastinen T, Hudson TJ. Cis-acting regulatory variation in the human genome. Science 2004;306:647-50.

206. Visel A, Rubin EM, Pennacchio LA. Genomic views of distant-acting enhancers. Nature 2009;461:199-205.

207. Lettice LA, Heaney SJ, Purdie LA, et al. A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. Hum Mol Genet 2003;12:1725-35.

208. Furniss D, Lettice LA, Taylor IB, et al. A variant in the sonic hedgehog regulatory sequence (ZRS) is associated with triphalangeal thumb and deregulates expression in the developing limb. Hum Mol Genet 2008;17:2417-23.

209. Capon F, Allen MH, Ameen M, et al. A synonymous SNP of the corneodesmosin gene leads to increased mRNA stability and demonstrates association with psoriasis across diverse ethnic groups. Hum Mol Genet 2004;13:2361-8.

210. Carter AM, Sachchithananthan M, Stasinopoulos S, Maurer F, Medcalf RL. Prothrombin G20210A is a bifunctional gene polymorphism. Thromb Haemost 2002;87:846-53.

211. Gehring NH, Frede U, Neu-Yilik G, et al. Increased efficiency of mRNA 3' end formation: a new genetic mechanism contributing to hereditary thrombophilia. Nat Genet 2001;28:389-92.

212. Michalova E, Vojtesek B, Hrstka R. Impaired pre-mRNA processing and altered architecture of 3' untranslated regions contribute to the development of human disorders. Int J Mol Sci 2013;14:15681-94.

213. Proudfoot NJ. Ending the message: poly(A) signals then and now. Genes Dev 2011;25:1770-82.

214. Ober C, Aldrich CL, Chervoneva I, et al. Variation in the HLA-G promoter region influences miscarriage rates. Am J Hum Genet 2003;72:1425-35.

215. Wagner JR, Busche S, Ge B, Kwan T, Pastinen T, Blanchette M. The relationship between DNA methylation, genetic and expression inter-individual variation in untransformed human fibroblasts. Genome biology 2014;15:R37.

216. Gregory AP, Dendrou CA, Attfield KE, et al. TNF receptor 1 genetic risk mirrors outcome of anti-TNF therapy in multiple sclerosis. Nature 2012;488:508-11.

217. Cirulli ET, Goldstein DB. In vitro assays fail to predict in vivo effects of regulatory polymorphisms. Hum Mol Genet 2007;16:1931-9.

218. Schadt EE, Monks SA, Drake TA, et al. Genetics of gene expression surveyed in maize, mouse and man. Nature 2003;422:297-302.

219. Almlof JC, Lundmark P, Lundmark A, et al. Powerful identification of cis-regulatory SNPs in human primary monocytes using allele-specific gene expression. PloS one 2012;7:e52260.

220. Emilsson V, Thorleifsson G, Zhang B, et al. Genetics of gene expression and its effect on disease. Nature 2008;452:423-8.

221. Stranger BE, Nica AC, Forrest MS, et al. Population genomics of human gene expression. Nat Genet 2007;39:1217-24.

222. Fairfax BP, Makino S, Radhakrishnan J, et al. Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. Nat Genet 2012;44:502-10.

223. Liu S, Lin L, Jiang P, Wang D, Xing Y. A comparison of RNA-Seq and high-density exon array for detecting differential gene expression between closely related species. Nucleic Acids Res 2011;39:578-88.

224. Bustin SA, Benes V, Garson JA, et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 2009;55:611-22.

225. Ge B, Pokholok DK, Kwan T, et al. Global patterns of cis variation in human cells revealed by high-density allelic expression analysis. Nat Genet 2009;41:1216-22.

226. Chen X, Weaver J, Bove BA, et al. Allelic imbalance in BRCA1 and BRCA2 gene expression is associated with an increased breast cancer risk. Hum Mol Genet 2008;17:1336-48.

227. Arai T, Kakuta Y, Kinouchi Y, et al. Increased expression of NKX2.3 mRNA transcribed from the risk haplotype for ulcerative colitis in the involved colonic mucosa. Hum Immunol 2011;72:587-91.

228. Cunnington MS, Plummer CJ, McComb JM. A survey of investigations performed prior to permanent pacemaker implantation. Age Ageing 2010;39:135-9.

229. Kulzer JR, Stitzel ML, Morken MA, et al. A Common Functional Regulatory Variant at a Type 2 Diabetes Locus Upregulates ARAP1 Expression in the Pancreatic Beta Cell. Am J Hum Genet 2014;94:186-97.

230. Schroeder A, Mueller O, Stocker S, et al. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol 2006;7:3.

231. Fleige S, Pfaffl MW. RNA integrity and the effect on the real-time qRT-PCR performance. Mol Aspects Med 2006;27:126-39.

232. Vandesompele J, De Preter K, Pattyn F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome biology 2002;3:RESEARCH0034.

233. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 2001;29:e45.

234. Ohl F, Jung M, Xu C, et al. Gene expression studies in prostate cancer tissue: which reference gene should be selected for normalization? Journal of molecular medicine 2005;83:1014-24.

235. Ding C, Cantor CR. A high-throughput gene expression analysis technique using competitive PCR and matrix-assisted laser desorption ionization time-of-flight MS. Proc Natl Acad Sci U S A 2003;100:3059-64.

236. Forton JT, Udalova IA, Campino S, Rockett KA, Hull J, Kwiatkowski DP. Localization of a long-range cis-regulatory element of IL13 by allelic transcript ratio mapping. Genome Res 2007;17:82-7.

237. De Luca V, Strauss J, Semeralul M, et al. Analysis of BDNF Val66Met allele-specific mRNA levels in bipolar disorder. Neurosci Lett 2008;441:229-32.

238. Sun C, Southard C, Witonsky DB, Olopade OI, Di Rienzo A. Allelic imbalance (AI) identifies novel tissue-specific cis-regulatory variation for human UGT2B15. Hum Mutat 2010;31:99-107.

239. Heighway J, Bowers NL, Smith S, Betticher DC, Koref MF. The use of allelic expression differences to ascertain functional polymorphisms acting in cis: analysis of MMP1 transcripts in normal lung tissue. Ann Hum Genet 2005;69:127-33.

240. Cunnington MS, Kay C, Avery PJ, Mayosi BM, Koref MS, Keavney B. STK39 polymorphisms and blood pressure: an association study in British Caucasians and assessment of cis-acting influences on gene expression. BMC Med Genet 2009;10:135.

241. Santibanez Koref M, Wilson V, Cartwright N, et al. MLH1 Differential allelic expression in mutation carriers and controls. Ann Hum Genet 2010;74:479-88.

242. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am J Hum Genet 2007;81:1084-97.

243. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21:263-5.

244. Dunn OJ. Multiple Comparisons among Means. J Am Stat Assoc 1961;56:52-64.

245. R: A Language and Environment for Statistical Computing. at http://www.R-project.org.)

246. Robinson RB, Siegelbaum SA. Hyperpolarization-activated cation currents: from molecules to physiological function. Annu Rev Physiol 2003;65:453-80.

247. Suenari K, Cheng CC, Chen YC, et al. Effects of ivabradine on the pulmonary vein electrical activity and modulation of pacemaker currents and calcium homeostasis. J Cardiovasc Electrophysiol 2012;23:200-6.

248. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics 2008;24:2938-9.

249. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. Nature methods 2014;11:361-2.

250. Zhang Q, Huang A, Lin YC, Yu HG. Associated changes in HCN2 and HCN4 transcripts and I(f) pacemaker current in myocytes. Biochim Biophys Acta 2009;1788:1138-47.

251. Misier AR, Opthof T, van Hemel NM, et al. Increased dispersion of "refractoriness" in patients with idiopathic paroxysmal atrial fibrillation. J Am Coll Cardiol 1992;19:1531-5.

252. DiFrancesco D. Funny channels in the control of cardiac rhythm and mode of action of selective blockers. Pharmacol Res 2006;53:399-406.

253. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339:b2535.

254. Higgins JPT, Green S, Cochrane Collaboration. Cochrane handbook for systematic reviews of interventions. Chichester, England ; Hoboken, NJ: Wiley-Blackwell; 2008.

255. Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. J Stat Softw 2010;36:1-48.

256. Raudenbush S. Analysing effect sizes: random effects models. In: Cooper HM, Hedges LV, Valentine JC, eds. The handbook of research synthesis and meta-analysis. 2nd ed. New York: Russell Sage Foundation; 2009:295-315.

257. Morris CN. Parametric Empirical Bayes Inference - Theory and Applications. J Am Stat Assoc 1983;78:47-55.

258. Viechtbauer W. Confidence intervals for the amount of heterogeneity in meta-analysis. Stat Med 2007;26:37-52.

259. Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. Psychol Methods 1998;3:486-504.

260. Knapp G, Hartung J. Improved tests for a random effects meta-regression with a single covariate. Stat Med 2003;22:2693-710.

261. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539-58.

262. Begg CB, Mazumdar M. Operating Characteristics of a Bank Correlation Test for Publication Bias. Biometrics 1994;50:1088-101.

263. Moreno SG, Sutton AJ, Ades AE, et al. Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study. Bmc Med Res Methodol 2009;9.

264. Fasullo S, Cannizzaro S, Maringhini G, et al. Comparison of ivabradine versus metoprolol in early phases of reperfused anterior myocardial infarction with impaired left ventricular function: preliminary findings. J Card Fail 2009;15:856-63.

265. Nerla R, Di Franco A, Milo M, et al. Differential effects of heart rate reduction by atenolol or ivabradine on peripheral endothelial function in type 2 diabetic patients. Heart 2012;98:1812-6.

266. Dominguez-Rodriguez A, Consuegra-Sanchez L, Blanco-Palacios G, et al. Anti-inflammatory effects of ivabradine in patients with acute coronary syndrome: a pilot study. Int J Cardiol 2012;158:160-2.

267. Chinchilla A, Daimi H, Lozano-Velasco E, et al. PITX2 insufficiency leads to atrial electrical and structural remodeling linked to arrhythmogenesis. Circ Cardiovasc Genet 2011;4:269-79.

268. Cappato R, Castelvecchio S, Ricci C, et al. Clinical efficacy of ivabradine in patients with inappropriate sinus tachycardia: a prospective, randomized, placebo-controlled, double-blind, crossover evaluation. J Am Coll Cardiol 2012;60:1323-9.

269. Villano A, Di Franco A, Nerla R, et al. Effects of ivabradine and ranolazine in patients with microvascular angina pectoris. Am J Cardiol 2013;112:8-13.

270. Procoralan : EPAR - Scientific Discussion. 2005. at http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Scientific_Discussion/human/000598/WC500035338.pdf.)

271. Lopez-Bescos L, Filipova S, Martos R. Long-term safety and efficacy of ivabradine in patients with chronic stable angina. Cardiology 2007;108:387-96.

272. Timpson NJ, Lawlor DA, Harbord RM, et al. C-reactive protein and its role in metabolic syndrome: mendelian randomisation study. Lancet 2005;366:1954-9.

273. Lip GY, Nieuwlaat R, Pisters R, Lane DA, Crijns HJ. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. Chest 2010;137:263-72.

274. Fox K, Komajda M, Ford I, et al. Effect of ivabradine in patients with left-ventricular systolic dysfunction: a pooled analysis of individual patient data from the BEAUTIFUL and SHIFT trials. Eur Heart J 2013;34:2263-70.

275. Fox K, Ford I, Steg PG, et al. Ivabradine in stable coronary artery disease without clinical heart failure. N Engl J Med 2014;371:1091-9.

276. Holdt LM, Beutner F, Scholz M, et al. ANRIL Expression Is Associated With Atherosclerosis Risk at Chromosome 9p21. Arterioscl Throm Vas 2010;30:620-U518.

277. Holdt LM, Hoffmann S, Sass K, et al. Alu Elements in ANRIL Non-Coding RNA at Chromosome 9p21 Modulate Atherogenic Cell Functions through Trans-Regulation of Gene Networks. Plos Genetics 2013;9.

278. Cunnington MS, Keavney B. Genetic mechanisms mediating atherosclerosis susceptibility at the chromosome 9p21 locus. Curr Atheroscler Rep 2011;13:193-201.

279. Dimas AS, Deutsch S, Stranger BE, et al. Common regulatory variation impacts gene expression in a cell type-dependent manner. Science 2009;325:1246-50.

280. Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3. Bioinformatics 2007;23:1289-91.

281. Untergasser A, Cutcutache I, Koressaar T, et al. Primer3--new capabilities and interfaces. Nucleic Acids Res 2012;40:e115.

282. Fujita PA, Rhead B, Zweig AS, et al. The UCSC Genome Browser database: update 2011. Nucleic Acids Res 2011;39:D876-82.

283. The UCSC Genome Browser.

284. Comuzzie AG, Cole SA, Laston SL, et al. Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. PloS one 2012;7:e51954.

285. Thorleifsson G, Walters GB, Hewitt AW, et al. Common variants near CAV1 and CAV2 are associated with primary open-angle glaucoma. Nat Genet 2010;42:906-9.

286. Holm H, Gudbjartsson DF, Arnar DO, et al. Several common variants modulate heart rate, PR interval and QRS duration. Nat Genet 2010;42:117-22.

287. Kerns SL, Ostrer H, Stock R, et al. Genome-wide association study to identify single nucleotide polymorphisms (SNPs) associated with the development of erectile

dysfunction in African-American men after radiotherapy for prostate cancer. Int J Radiat Oncol Biol Phys 2010;78:1292-300.

288. Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res 2012;22:1790-7.

289. Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 2013;14:178-92.

290. Liu C, Liu W, Lu MF, Brown NA, Martin JF. Regulation of left-right asymmetry by thresholds of Pitx2c activity. Development 2001;128:2039-48.

291. McKenzie CA, Abecasis GR, Keavney B, et al. Trans-ethnic fine mapping of a quantitative trait locus for circulating angiotensin I-converting enzyme (ACE). Hum Mol Genet 2001;10:1077-84.

292. Tishkoff SA, Reed FA, Friedlaender FR, et al. The genetic structure and history of Africans and African Americans. Science 2009;324:1035-44.

293. Campino S, Forton J, Raj S, et al. Validating discovered Cis-acting regulatory genetic variants: application of an allele specific expression approach to HapMap populations. PloS one 2008;3:e4105.

294. Tuteja D, Xu D, Timofeyev V, et al. Differential expression of small-conductance Ca2+-activated K+ channels SK1, SK2, and SK3 in mouse atrial and ventricular myocytes. Am J Physiol Heart Circ Physiol 2005;289:H2714-23.

295. Zhang XD, Timofeyev V, Li N, et al. Critical roles of a small conductance Ca2+-activated K+ channel (SK3) in the repolarization process of atrial myocytes. Cardiovasc Res 2014;101:317-25.

296. Mahida S, Mills RW, Tucker NR, et al. Overexpression of KCNN3 results in sudden cardiac death. Cardiovasc Res 2014;101:326-34.

297. Fu C, He J, Li C, Shyy JY, Zhu Y. Cholesterol increases adhesion of monocytes to endothelium by moving adhesion molecules out of caveolae. Biochim Biophys Acta 2010;1801:702-10.

298. Xu Y, Krause A, Hamai H, Harvey BG, Worgall TS, Worgall S. Proinflammatory phenotype and increased caveolin-1 in alveolar macrophages with silenced CFTR mRNA. PloS one 2010;5:e11004.

299. Harris J, Werling D, Hope JC, Taylor G, Howard CJ. Caveolae and caveolin in immune cells: distribution and functions. Trends Immunol 2002;23:158-64.

300. Kostin S, Klein G, Szalay Z, Hein S, Bauer EP, Schaper J. Structural correlate of atrial fibrillation in human patients. Cardiovasc Res 2002;54:361-79.

301. Yi SL, Liu XJ, Zhong JQ, Zhang Y. Role of caveolin-1 in atrial fibrillation as an anti-fibrotic signaling molecule in human atrial fibroblasts. PloS one 2014;9:e85144.

302. Berry FB, Miura Y, Mihara K, et al. Positive and negative regulation of myogenic differentiation of C2C12 cells by isoforms of the multiple homeodomain zinc finger transcription factor ATBF1. J Biol Chem 2001;276:25057-65.

303. Dong XY, Sun X, Guo P, et al. ATBF1 inhibits estrogen receptor (ER) function by selectively competing with AIB1 for binding to the ER in ER-positive breast cancer cells. J Biol Chem 2010;285:32801-9.

304. From AM, Maleszewski JJ, Rihal CS. Current status of endomyocardial biopsy. Mayo Clin Proc 2011;86:1095-102.

305. Ambrosi CM, Yamada KA, Nerbonne JM, Efimov IR. Gender differences in electrophysiological gene expression in failing and non-failing human hearts. PloS one 2013;8:e54635.

306. Zhang Y, Wang X, Xu X, Wang J, Liu X, Chen Y. Distinct microRNA expression signatures in human right atrial and ventricular myocardium. Mol Cell Biochem 2012;371:23-9.

307. Stollberger C, Ernst G, Finsterer J. Is the left atrial appendage our most lethal attachment? Eur J Cardio-Thorac 2000;18:625-6.

308. Hillis LD, Smith PK, Anderson JL, et al. 2011 ACCF/AHA Guideline for Coronary Artery Bypass Graft Surgery. A report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. Developed in collaboration with the American Association for Thoracic Surgery, Society of Cardiovascular Anesthesiologists, and Society of Thoracic Surgeons. J Am Coll Cardiol 2011;58:e123-210.

309. Abecasis GR, Auton A, Brooks LD, et al. An integrated map of genetic variation from 1,092 human genomes. Nature 2012;491:56-65.

310. Kahr PC, Piccini I, Fabritz L, et al. Systematic analysis of gene expression differences between left and right atria in different mouse strains and in human atrial tissue. PloS one 2011;6:e26389.

311. . at http://blast.ncbi.nlm.nih.gov/Blast.cgi.)

312. Herrmann S, Layh B, Ludwig A. Novel insights into the distribution of cardiac HCN channels: an expression study in the mouse heart. J Mol Cell Cardiol 2011;51:997-1006.

313. Reynard LN, Bui C, Syddall CM, Loughlin J. CpG methylation regulates allelic expression of GDF5 by modulating binding of SP1 and SP3 repressor proteins to the osteoarthritis susceptibility SNP rs143383. Hum Genet 2014;133:1059-73.

314. Brahler S, Kaistha A, Schmidt VJ, et al. Genetic deficit of SK3 and IK1 channels disrupts the endothelium-derived hyperpolarizing factor vasodilator pathway and causes hypertension. Circulation 2009;119:2323-32.

315. Newton-Cheh C, Johnson T, Gateva V, et al. Genome-wide association study identifies eight loci associated with blood pressure. Nat Genet 2009;41:666-76.

316. Levy D, Ehret GB, Rice K, et al. Genome-wide association study of blood pressure and hypertension. Nat Genet 2009;41:677-87.

317. Wain LV, Verwoert GC, O'Reilly PF, et al. Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. Nat Genet 2011;43:1005-11.

318. Kroigaard C, Dalsgaard T, Nielsen G, et al. Activation of endothelial and epithelial K(Ca) 2.3 calcium-activated potassium channels by NS309 relaxes human small pulmonary arteries and bronchioles. Br J Pharmacol 2012;167:37-47.

319. Rottlaender D, Motloch LJ, Schmidt D, et al. Clinical impact of atrial fibrillation in patients with pulmonary hypertension. PloS one 2012;7:e33902.

320. Owan TE, Hodge DO, Herges RM, Jacobsen SJ, Roger VL, Redfield MM. Trends in prevalence and outcome of heart failure with preserved ejection fraction. N Engl J Med 2006;355:251-9.

321. Bhatia RS, Tu JV, Lee DS, et al. Outcome of heart failure with preserved ejection fraction in a population-based study. N Engl J Med 2006;355:260-9.

322. Schmiegelow MD, Pedersen OD, Kober L, Seibaek M, Abildstrom SZ, Torp-Pedersen C. Incidence of Atrial Fibrillation in Patients with either Heart Failure or Acute Myocardial Infarction and Left Ventricular Dysfunction: A Cohort Study. Bmc Cardiovasc Disor 2011;11.

323. Liu C, Liu W, Palie J, Lu MF, Brown NA, Martin JF. Pitx2c patterns anterior myocardium and aortic arch vessels and is required for local cell movement into atrioventricular cushions. Development 2002;129:5081-91.

324. Tischkowitz MD, Hodgson SV. Fanconi anaemia. Journal of medical genetics 2003;40:1-10.

325. Pang Q, Christianson TA, Keeble W, Koretsky T, Bagby GC. The anti-apoptotic function of Hsp70 in the interferon-inducible double-stranded RNA-dependent protein kinase-mediated death signaling pathway requires the Fanconi anemia protein, FANCC. J Biol Chem 2002;277:49638-43.

326. Shyamsunder P, Vidyasekar P, Shukla AR, Mohan S, Verma RS. Lowered expression levels of a tumor suppressor gene - caveolin-1 within dysregulated gene networks of Fanconi anemia. Gene 2013;527:521-8.

327. Sanders P, Morton JB, Davidson NC, et al. Electrical remodeling of the atria in congestive heart failure: electrophysiological and electroanatomic mapping in humans. Circulation 2003;108:1461-8.

328. Kumagai K, Nakashima H, Urata H, Gondo N, Arakawa K, Saku K. Effects of angiotensin II type 1 receptor antagonist on electrical and structural remodeling in atrial fibrillation. J Am Coll Cardiol 2003;41:2197-204.

329. Healey JS, Morillo CA, Connolly SJ. Role of the renin-angiotensin-aldosterone system in atrial fibrillation and cardiac remodeling. Current opinion in cardiology 2005;20:31-7.

330. Zhang Y, Zhang P, Mu Y, et al. The role of renin-angiotensin system blockade therapy in the prevention of atrial fibrillation: a meta-analysis of randomized controlled trials. Clinical pharmacology and therapeutics 2010;88:521-31.

331. Arola AM, Sanchez X, Murphy RT, et al. Mutations in PDLIM3 and MYOZ1 encoding myocyte Z line proteins are infrequently found in idiopathic dilated cardiomyopathy. Molecular Genetics and Metabolism 2007;90:435-40.

332. Calkins H, Reynolds MR, Spector P, et al. Treatment of atrial fibrillation with antiarrhythmic drugs or radiofrequency ablation: two systematic literature reviews and meta-analyses. Circ Arrhythm Electrophysiol 2009;2:349-61.

333. Smemo S, Tena JJ, Kim KH, et al. Obesity-associated variants within FTO form long-range functional connections with IRX3. Nature 2014;507:371-5.

334. Green EM, Jiang Y, Joyner R, Weis K. A negative feedback loop at the nuclear periphery regulates GAL gene expression. Mol Biol Cell 2012;23:1367-75.

335. Zheng L, Liu H, Wang P, Song W, Sun X. Regulator of calcineurin 1 gene transcription is regulated by Nuclear factor-kappaB. Curr Alzheimer Res 2013.

336. Cargill M, Altshuler D, Ireland J, et al. Characterization of single-nucleotide polymorphisms in coding regions of human genes. Nat Genet 1999;22:231-8.

337. Traylor M, Farrall M, Holliday EG, et al. Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. Lancet Neurol 2012;11:951-62.

338. Francesconi M, Lehner B. The effects of genetic variation on gene expression dynamics during development. Nature 2014;505:208-11.

339. Burgermeister E, Tencer L, Liscovitch M. Peroxisome proliferator-activated receptor-gamma upregulates caveolin-1 and caveolin-2 expression in human carcinoma cells. Oncogene 2003;22:3888-900.

340. Gu J, Liu X, Wang X, et al. Beneficial effect of pioglitazone on the outcome of catheter ablation in patients with paroxysmal atrial fibrillation and type 2 diabetes mellitus. Europace 2011;13:1256-61.

341. Liu B, Wang J, Wang G. Beneficial Effects of Pioglitazone on Retardation of Persistent Atrial Fibrillation Progression in Diabetes Mellitus Patients. International heart journal 2014.

342. Kakuta H, Kurosaki E, Niimi T, et al. Distinct properties of telmisartan on agonistic activities for peroxisome proliferator-activated receptor gamma among clinically used angiotensin II receptor blockers: drug-target interaction analyses. J Pharmacol Exp Ther 2014;349:10-20.

343. Pan G, Zhou X, Zhao J. Effect of telmisartan on atrial fibrillation recurrences in patients with hypertension: a systematic review and meta-analysis. Cardiovascular therapeutics 2014;32:184-8.

Appendix 1

Assay details

A1 Assay details

Table A1.1 Real-time PCR assays

Transcript	Assay
ACTB	Hs99999903_m1
GAPDH	Hs03929097_g1
HPRT1	Hs02800695_m1
KCNN3	Hs01546820_m1
PRRX1	Hs00246567_m1
GORAB	Hs01552848_m1
PITX2a/b	Hs00165626_m1
PITX2c	Hs04183413_m1
CAV1	Hs00971716_m1
CAV2	Hs00184597_m1
C9orf3	Hs00262414_m1
FANCC	Hs00984545_m1
SYNPO2L	Hs00227591_m1
MYOZ1	Hs01122904_m1
HCN4	Hs00975492_m1
SYNE2	Hs00794881_m1
ESR2	Hs01100353_m1
ZFHX3	Hs00199344_m1
ZFHX3a	Hs00994898_m1

Assay details are given as Life Technologies assay codes.

Table A1.2 Genotyping assays

Multiplex	SNP	Forward primer	Reverse primer	Amplicon length	Extension primer
W1	rs117951282	ACGTTGGATGGGAAAAGTTAAGCACTCCC	ACGTTGGATGGCTACAAGTCCTCAACACTC	100	GCACTCCCACTTTCA
W1	rs1419110	ACGTTGGATGGCCCACTGGAAAATACCTTG	ACGTTGGATGCTTGGAGAAATGTGGACCTG	111	ACACTCAGCAGTGAC
W1	rs13376333	ACGTTGGATGATTGGGAGGTCATGGGATCG	ACGTTGGATGGGCTTTGCTCTGGAGAGAAC	99	GGTTCTGCTCCTTCCT
W1	rs16971384	ACGTTGGATGTCAGTCCATTGAACTGGAGG	ACGTTGGATGGTGGTTGCTGTCTTTCCCTG	100	GCCCCATGATGGAAGA
W1	rs1061122	ACGTTGGATGAGGAGTAGGTGCCAGCTCAG	ACGTTGGATGGAGTGCTGCTCAGTGTCAAA	100	ctCCAGCTCAGTCCCTG
W1	rs10852515	ACGTTGGATGTAATCACACACCTCGCACCG	ACGTTGGATGAATATCAGTAGCTCCTGCGG	100	TTGGTTTTTGGTTTGGT
W1	rs1051614	ACGTTGGATGGATGACTGGCGGATAGCCAT	ACGTTGGATGATGGCGCACACCAGCATCTC	91	gaaCTACGAGCGCATCCT
W1	rs4492614	ACGTTGGATGCTAGATGGTCTTACCCTCTC	ACGTTGGATGGTGTCATCTCACTCCCAATG	81	acTACCCTCTCCTCTCCTT
W1	rs4845663	ACGTTGGATGTTGCAGGTTCTCCGATAGTG	ACGTTGGATGGCATGACTGTAACTTCAAGG	99	cCTGATCATCACTCAGCCT
W1	rs6682291	ACGTTGGATGGTACGAAAAAAACTAGCCACG	ACGTTGGATGCTTTCCTCTTCCTCATTGGC	98	ctCCACGAATGAGCAGTTT
W1	rs7199343	ACGTTGGATGGCCCTGATAGGGTTTTAAGC	ACGTTGGATGCATGGATTCTGCTGGCACTC	92	GCTTTGGAATTTGGAAGTC
W1	rs2157786	ACGTTGGATGCACCCTATTCACCATTCTCC	ACGTTGGATGCTCTAGGTGAGGATGACTAC	95	tcCCATTCTCCCTGGTTAAA
W1	rs740178	ACGTTGGATGACCTCCTGGCAATGGGAGAC	ACGTTGGATGTCCTCCTTGTCTTCCTCAAC	90	ATGGGAGACCCCACTCTGGC
W1	rs76110919	ACGTTGGATGACTCCTCTCCCACCGCTTTC	ACGTTGGATGTGGAGGAAGGGTGGAGGAT	89	AGGGCCCCCCCTTCGTCCAAC
W1	rs7193343	ACGTTGGATGGAGGGGAAAGTTTGAACAGC	ACGTTGGATGAAATGTCGAGTCCTAATGGC	89	GAAAGTTTGAACAGCTTGTTT
W1	rs1218601	ACGTTGGATGCTTGGGATAACTCCTTCCTC	ACGTTGGATGGCCCCCCAATCTTGTAATTC	113	agAACTCCTTCCTCCACTCGCA
W1	rs16971474	ACGTTGGATGTAGGCAGACCCCATCATTTC	ACGTTGGATGCCCACAGGCAAGATCTTGTC	100	cGACCCCATCATTTCGAGATTA
W1	rs4788482	ACGTTGGATGAGGAACTGCTGGATTCTTGG	ACGTTGGATGTGTTCCTGCTGCGAGAGTTC	118	caTTAGCATGGGCCTTTCTCAA
W1	rs2228200	ACGTTGGATGAGAGCTTGCACTGGTATGAG	ACGTTGGATGCTGCACATGAACGTGGAGC	99	cTCACCGCCTTCCACTCGTCCTC
W1	rs2106261	ACGTTGGATGCACAGATAGAGCTCGTCCAG	ACGTTGGATGAGAGCAGTCTCTGGCACACT	100	ATAGAGCTCGTCCAGAGAATTGT
W1	rs6691316	ACGTTGGATGGGCTCAACTCATGCTCAATG	ACGTTGGATGAAGAGGTGCTCGATGCACAG	101	ggGAATGAGTGAATAAGCTGATC
W1	rs9940520	ACGTTGGATGCAGTTCAGACTCTCTGTTAG	ACGTTGGATGAAGCAACATAAGGCAGGCTC	82	CAGACTCTCTGTTAGAGGCCTACT
W1	rs1218565	ACGTTGGATGAACACCTCAGAATCCTTCCC	ACGTTGGATGGTGCAGGAACAGCAAATAGG	116	CACCTCAGAATCCTTCCCAACACAG

Multiplex	SNP	Forward primer	Reverse primer	Amplicon length	Extension primer
W1	rs7193297	ACGTTGGATGCATTCGTTGCAGGTGACCTC	ACGTTGGATGAATGAGCGCCTCGCGGAGA	100	GTGACCTCCTTGCTGGCGGGCTCGG
W1	rs16971464	ACGTTGGATGAGGAGTCACTCAGATGTTTG	ACGTTGGATGCTCATTTGACCTCTAACAATC	120	tgAACCTGTAAAATCGAAAATATGA
W1	rs4788679	ACGTTGGATGACCTACCTACAGCCTTATGG	ACGTTGGATGGAAGAAAGCCAGGTACGAAG	84	CCTACCTACAGCCTTATGGAACGACC
W1	rs2266943	ACGTTGGATGTTCCCAAAAGGCCATTCCAC	ACGTTGGATGTGGCTCTGTTTTGACTGGTG	111	ccTCAATTTCAAGACACCTTGCTCCCT
W1	rs1476646	ACGTTGGATGTTCAGTAGCAACTCAGAGCG	ACGTTGGATGCCCGGGAATACTTTCACATC	119	CAGAGCGTCGTGCTCTCTTGTCTTTGT
W1	rs2106258	ACGTTGGATGTTAGATCCTGGTTAGCGGAG	ACGTTGGATGAGCCTGAACAGAGCCAAATC	101	tTTAGATCCTGGTTAGCGGAGCCTTCG
W1	rs8052905	ACGTTGGATGACCAAAGTATTCCCCAGTCC	ACGTTGGATGCCACTACATGCAGCTTCTTG	109	GGGAAAGCAAGATCCTGAGAAAGAATA
W1	rs8056528	ACGTTGGATGAAGCTAAGTAGCCTAAGGAC	ACGTTGGATGCTTGTGAATTAGCGAGAACC	97	agTTTTGTGAAGCATCTTTGCTGCTTTT
W1	rs6695232	ACGTTGGATGACACCCCGATCCTATTTCAC	ACGTTGGATGTCTCACCTTACTTGCTGCTG	119	ctGAAGATATGGACATGATTTATGCCCC
W2	rs11640106	ACGTTGGATGAACAGACTGTCCTTCTCCAC	ACGTTGGATGAAATCTGAAGGCCCTGAGTG	113	CTCACCCTGCCCCCA
W2	rs925042	ACGTTGGATGGCCGCTTCTCGTTTTCAAAG	ACGTTGGATGGAAAGTGGAGATGGAGGGC	119	TTGGTCCTTGCACAA
W2	rs756720	ACGTTGGATGTGTTTTCCTGATCACCAGGG	ACGTTGGATGGCAGCATTGGACTAAATGGG	96	tATCCTCAGCCTTCCC
W2	rs12929452	ACGTTGGATGTATAAGGAACACCATCCTGC	ACGTTGGATGGCACAGGGTGTTATATTGGC	98	ACCATCCTGCCACTTA
W2	rs4788480	ACGTTGGATGGAAGAGTCTGCTTCAAGAGG	ACGTTGGATGTTTGCAGAACTGTGGCCCTC	101	CTGTATGCAGAACGGA
W2	rs8051826	ACGTTGGATGAACTTCCCTAGTATGGCCAG	ACGTTGGATGAGGGTTGGAACCCATGTAAC	96	tGAGGCTGGAAAACCA
W2	rs12373097	ACGTTGGATGGAGAGGGTGAAACGAGCATT	ACGTTGGATGTACAACAGACTTACTGCTTC	106	CATTTCTAGAGGCTCCC
W2	rs739414	ACGTTGGATGGAGAACGAGCTTCCCAAATC	ACGTTGGATGTTTCCGAAAGAGCTGTCGTC	84	tcCCCAAATCCAACTCCA
W2	rs12029542	ACGTTGGATGTACATCATGCCCAAGTAGCG	ACGTTGGATGCCCTTTGAGCCTCTGTTTTC	94	ATCCCAGGAATGCAAATA
W2	rs10908435	ACGTTGGATGAATTCCTAGCACAGTGCCTG	ACGTTGGATGTGCTCTCATGAGAGACTTGC	106	GTGAATAGCAGAGCTCCA
W2	rs11641701	ACGTTGGATGTGATGCTTGCTAGACACTCC	ACGTTGGATGTTCCAGCAGTCTGTAATCAC	120	cccaGGCCTGGCTCATCTT
W2	rs1106236	ACGTTGGATGCCTCATAGATGTGAAAACCC	ACGTTGGATGTACATTAGCAAGCGCCTCAG	107	ggaAAAGCGCTAGACAAGA
W2	rs1218574	ACGTTGGATGCTGAGATGGTGACTGAGTTC	ACGTTGGATGTTCCTGAGGTCTGCAGATCG	114	tctcGCATAACCACGTGCCC
W2	rs1131820	ACGTTGGATGCTAATGGACTTTCTGCTCCC	ACGTTGGATGGCTCGTAGGTCATGGCTATC	100	aAGCTCTTCGTGATCGACAA
W2	rs9930445	ACGTTGGATGTTTGCTTTGTTGAGGCACAG	ACGTTGGATGACCAGTGATGGCAATGTGTC	100	cGTTGAGGCACAGATTAGGT
W2	rs11264261	ACGTTGGATGAATTCTCCACCTCCATCCTG	ACGTTGGATGCTAACCTAGAGGCCTTGGTC	117	TCCATCCTGTACTTAGTAACC

Multiplex	SNP	Forward primer	Reverse primer	Amplicon length	Extension primer
W2	rs1106235	ACGTTGGATGATAGAGAGAGGAGCTACACC	ACGTTGGATGGCTCTGAGCTTTCTGTCTTC	119	aAACACCATGGAATAAATCCT
W2	rs4788683	ACGTTGGATGATTACACAGTGCGACTGTTC	ACGTTGGATGGGTGCAGTTATGACGAAGGG	104	tcccGTGCGACTGTTCATGTTT
W2	rs1858800	ACGTTGGATGGATCAGTATGTGGTGATGGG	ACGTTGGATGTCTAAAGCCCATGTCTCAGG	109	CACGATGTGTGAGACTGATAAC
W2	rs8057081	ACGTTGGATGTATTCCAGAGCTGCACCATC	ACGTTGGATGTCGAGACAAGCCACATGTCAC	98	tcGCCACCGACCCTTGTCACTAC
W2	rs10908429	ACGTTGGATGCCGTGACAACCTTACGTATC	ACGTTGGATGGACCAGAAATCTGAATTGGC	103	CTTACGTATCTGAAAATGTCTTC
W2	rs4845671	ACGTTGGATGTATGTGCAGAATCACCTGGC	ACGTTGGATGGGGAGACACAATTCAGTTC	111	agGGGTAATTTTGACACAAAAGG
W2	rs954785	ACGTTGGATGACTAGCCTGGAAGCTCTATG	ACGTTGGATGTTACACTGTTCCTGGCACTG	100	GGATCATATCTGTCCTGTTCATTA
W2	rs12740456	ACGTTGGATGTCTACCAGGTATTGTCTCCC	ACGTTGGATGCAAGCAAAGATGTCTGCAGG	89	aggaTCTCCCTACTACATCCCCATC
W2	rs9940321	ACGTTGGATGTCACATTGCTTATACCGCTC	ACGTTGGATGACTCAGATATCCAGGAGCAG	119	tcCCGCTCAAACTTATGACCAGTTC
W2	rs6426905	ACGTTGGATGAGAGGACTCACCTTCACCAG	ACGTTGGATGGAACCCTAATGAAAGCCCAG	110	cacaAGCTCCCTTGCTGCACTTTGT
W2	rs883319	ACGTTGGATGTGGCCTCCTAAGCACTTTTC	ACGTTGGATGTAGTTGCTATGAGGTCTGGG	113	gggAACTCTTTGAGTCTCACAATAG
W2	rs11264268	ACGTTGGATGGGCCCACTGGAGATACAAAG	ACGTTGGATGGCTCAGCAAATGAGAAGTGG	108	CTAATAAAGTAACTAATTCATTCACT
W2	rs8055406	ACGTTGGATGTTGGGCAAAAACAGACCCTA	ACGTTGGATGGTTTCTTTGAAGAAAGCATC	100	gcagAAATTACAGTGCACAGAGCTCC
W2	rs16971456	ACGTTGGATGGGATGATGGAGGTGGATTAG	ACGTTGGATGAGTCCTAAGAGAGCGTGGAG	115	ggcgGCTACATCTGTAAAAGAGTAAA
W2	rs9936884	ACGTTGGATGTTTGCGTCTCTGTCTCTCTC	ACGTTGGATGTCTGGCTTTTGTTACTTGGG	99	agagCACACTCTCACTCTATCCTGCCC
W2	rs12145386	ACGTTGGATGCTGTAGGTGACTTTATCATCC	ACGTTGGATGACAATCTCAGGGAGGGAATG	101	TCATCCTTTAATTTATAATTAGCAGTT
W2	rs4845391	ACGTTGGATGAGCCTACAGGGTTTAGCTTG	ACGTTGGATGGGAAAGTGGTCATTACAGGG	94	cttcGCTTGCAGGCTTCCATCAACCCCA
W2	rs6499594	ACGTTGGATGTGACCTCTCTTTACCCTTGC	ACGTTGGATGGACATTCCCTGATGAAGGAC	115	acAAAGCTTGCTGAAATTTACAATCCCT
W2	rs960182	ACGTTGGATGTTCCATGGCCTTCTAGCTTC	ACGTTGGATGGAGGCCAAATACAAAGAACC	100	cgcgTCCAGTGTTGCTGTCAAAAAGTCC
W2	rs9940310	ACGTTGGATGCACGTGATCAGTGAACACAG	ACGTTGGATGTGAGTGCTTCCTATCAAGGG	98	tgacCACAGGACGTGTGCATGTCAGTAC
W3	rs1218578	ACGTTGGATGAAGCATCACTCGTTCTGCTG	ACGTTGGATGAACTCCTCCTCCTGGTTCCT	105	GTTCTGCTGCCTTCC
W3	rs11863932	ACGTTGGATGAGAAGGCTCTCTAGGAACGG	ACGTTGGATGCCTGGAGATTGTGAGTCATC	100	TGAGACGTTGCTTCC
W3	rs2335407	ACGTTGGATGGGAAATAACAGAGAGGACCC	ACGTTGGATGCTAACACCCGGTGGATAAAC	98	GAACACGAGGCGTAG
W3	rs10494301	ACGTTGGATGTGGGCAGAAAATGAGATGGG	ACGTTGGATGATTCTTCAGCCTCCTCCTTC	110	GGGGAACAAGTTGCGT
W3	rs884664	ACGTTGGATGTACCTCCCATCGGTTAGTAG	ACGTTGGATGTGGTAAGAAGCGTCTTTCCC	101	CATAGCACCAACTGTCA

Multiplex	SNP	Forward primer	Reverse primer	Amplicon length	Extension primer
W3	rs1506982	ACGTTGGATGATAGTGGCTCGCTCAAATGC	ACGTTGGATGAAATCCTTTCCCCTGTCCTC	89	AATGCCGAAACTTGGCT
W3	rs6499600	ACGTTGGATGAGATTTCGCCTGCATGCCCT	ACGTTGGATGTCCCACCAGCACCAACCATC	106	cGCCTGCGGTTCATTTCC
W3	rs16971366	ACGTTGGATGAGATACCTGTAGGAGGTGAC	ACGTTGGATGAACCCTTCAGAAATAACCCC	99	tcCTGGGGATGAGGTCCA
W3	rs4788489	ACGTTGGATGGCAACTTAATAACCACTGACC	ACGTTGGATGTTCCGCATTAATGCTCTTCC	101	tacCCACTGACCACCTTCT
W3	rs8060701	ACGTTGGATGAAAACAATGGCAGGTTCCCG	ACGTTGGATGTGGAATCACATACGCAGAGG	109	ttGCAGGTTCCCGGCCACT
W3	rs1218551	ACGTTGGATGTTTTCTCACAGCTGTTAGGG	ACGTTGGATGTGCATGCTTCTCTTTGTGTG	102	TGAGATGAGATCATGCATA
W3	rs6681725	ACGTTGGATGTCCCTACATCAATCAACTGG	ACGTTGGATGGATGCTTTCAGACCCCTTTG	119	ACAGCTCACAATGTTTTCCA
W3	rs2256209	ACGTTGGATGTTCTCTTCTTCCAGAGCCAG	ACGTTGGATGGACTGCATGTGTATGTGTGG	105	GCAGTATTGGCTGTAAAACT
W3	rs4845394	ACGTTGGATGATGTCAGTTTAGGGCTTTCC	ACGTTGGATGTTCCATCTGTGGGTTAGGAG	120	gtagTTGATGCTGGGAGGTA
W3	rs8058014	ACGTTGGATGTAACTAGCCACAGACAAGAG	ACGTTGGATGGGGACAGTTGTAGAGCATTA	118	AGCCACAGACAAGAGAAATCT
W3	rs9925261	ACGTTGGATGCCACACTCTAGCTCATGTTG	ACGTTGGATGAATGGCTCTCAAAGTGGCAG	99	cccgCTTACACTGCCATGAATT
W3	rs756717	ACGTTGGATGTAAGAGGAGAACGGAGCAAG	ACGTTGGATGAAAGCCAATCTCCTCTCCAG	115	cGCCCACCAATTAATTGGTAAA
W3	rs7523074	ACGTTGGATGACCACCCTGCAGACTTGCTC	ACGTTGGATGGAGTGTGGAATAGAACAGGG	118	cccaCCGGCTGCCCTCTCTTTCA
W3	rs2335406	ACGTTGGATGGGCAGATTTACTCCCTCCTC	ACGTTGGATGTCAGGCTCTCAGCCTGTTAG	83	gCTCCTCATTCTAAGGAGAAAGT
W3	rs2040508	ACGTTGGATGTTGGGATCCCCTTCCAAATG	ACGTTGGATGTGGCTGGAGGACATCTTTAC	99	aAAATGTCAACAAACAAGAAAAG
W3	rs1995662	ACGTTGGATGCTGCTTATCCCTCCTGAATG	ACGTTGGATGGCTTTGTCCTGGATATGCAC	100	ttttGCTTGGCATCCATTTCAAAT
W3	rs7532286	ACGTTGGATGTCAACAAGGTTAGGCCATTG	ACGTTGGATGCTACTTGAGATTGTGCACTG	99	ggatGCCATTGTGGGTATATAAAT
W3	rs1548373	ACGTTGGATGCTGAGACTCAAACTTCACCC	ACGTTGGATGGTGAGAGGTTGAGCTTTGTG	96	gggaACAGGATTGAGACAAGTTAT
W3	rs10494300	ACGTTGGATGCCTTCCATGTAGGAATGCAA	ACGTTGGATGCATGCAATTCCAGTCAACCC	120	ccctATTATTACCATTGTTGCTGAT
W3	rs906274	ACGTTGGATGGGCCTAATTAATACTGCAGTC	ACGTTGGATGAATAGTCTGTACTGGCTGTG	91	AGTCATTATTGATTCAATCCTGAGA
W3	rs6426987	ACGTTGGATGTGCCTCTGTTGAGTCAAACC	ACGTTGGATGTACACTGTGCTCACCATCTG	101	caGGACTCTTCCCTTGAGGCCCTGAT
W3	rs719353	ACGTTGGATGGGATGTTAAGTCTCTGTGCC	ACGTTGGATGGGGTGACTTTGCTTCTGAAA	98	ggacGAAAATTTTATCTTGGGCATAT
W3	rs4845396	ACGTTGGATGTGACTCAAGGACCCTTGGC	ACGTTGGATGTCTGTTAGCAGGGTCATTGG	109	ccccGACCCTTGGCGCTTGCTTTTGCT
W3	rs906276	ACGTTGGATGACAACTGTCTGGGCTCTTTC	ACGTTGGATGGGAACAAAGGAGAAGCATCG	101	ctcaTTCTGTGGTTGCTCAACTCCATC
W3	rs4788689	ACGTTGGATGTCCAGGACTGGGTTTTACAC	ACGTTGGATGCATTCTTCAAGGTGCCATGC	95	ttcaAGGACTGGGTTTTACACACCATA

Multiplex	SNP	Forward primer	Reverse primer	Amplicon length	Extension primer
W3	rs4788671	ACGTTGGATGTTCAAGTGCTTAATAGTGCC	ACGTTGGATGCAGAACCTTCTGAACTGCTC	98	ggatTGCTTAATAGTGCCATGGGGCTC
W3	rs4788692	ACGTTGGATGACAGTGGGCCTTCTTTTCAG	ACGTTGGATGGAATGCATGAACCAGAACGC	120	gaacGAGGGCCCGTGGTGACCTGCAGAG
W4	rs4788696	ACGTTGGATGTCCTAAACTCCTTTTCCCCC	ACGTTGGATGAGCCATTTGCATTCCTCCAG	99	TTCCCCCCACAAAGC
W4	rs4999127	ACGTTGGATGAGTGAGAACTGACTCCAAGC	ACGTTGGATGCAGCTGTTCTCTTCAGCAGA	96	GAATTTGGCAGCCAC
W4	rs6426944	ACGTTGGATGCTTACCACATGAGCATTCCC	ACGTTGGATGAAAATGCAGAGTTGCACGGG	111	AGCATTCCCAGAGGA
W4	rs11264270	ACGTTGGATGTGGCCATGCCTGGTGTCTC	ACGTTGGATGGTGAGAACAACATCAGTGGG	117	aTGGGGTGACCAGAGC
W4	rs10128027	ACGTTGGATGTTCCTTCCCTCTAACTTCCC	ACGTTGGATGGTCATTGACATGGGATGCTG	116	CAACAGAACACACACCT
W4	rs11075958	ACGTTGGATGTGCACACCTTCTCTTGCCTG	ACGTTGGATGAGCTCTAAGAGCAGGGCAAC	117	ggGCCTGCAGCCAAAGC
W4	rs699444	ACGTTGGATGATCTGCAGGCAAATCTGGTG	ACGTTGGATGGCAGATTGGTCATTTGAGTG	94	ggATCTGGTGACAGCGG
W4	rs12087736	ACGTTGGATGCCAAGGATGCACACATCCAG	ACGTTGGATGAAAGCTGGCAGGACTAGAAC	99	CACACATCCAGTTTCATT
W4	rs6682000	ACGTTGGATGGGTTTAGACAACTGCCAGAG	ACGTTGGATGCCTGGAGTTGTTAATGATCG	113	gGCCAGAGTTTGCACTTG
W4	rs12445932	ACGTTGGATGTGACTGGTGCTCCTGACTTG	ACGTTGGATGTTGGTGAAAGGAGGTTCTGC	102	ccatCACGGATCCCCTCCC
W4	rs4788684	ACGTTGGATGTTCTTCTTTTGCCCGGGCCT	ACGTTGGATGGCCCCATAAGAATACCTAGC	118	cctttCCCGGGCCTCAATTT
W4	rs17719287	ACGTTGGATGTCCTGGGAAGAGGATTTGTC	ACGTTGGATGAATGTGCACAGCACTGTCCC	107	TTGTCTACATTCTACGCACA
W4	rs11584635	ACGTTGGATGAAGTCGGCTTTGAAATGCAG	ACGTTGGATGGCATTTCCCAGTGTTCCTTG	116	TCTTGTTTTGAAGTAACAGC
W4	rs11264249	ACGTTGGATGTTCAGGAGATAGTAGGTGGG	ACGTTGGATGTAGACTCTCAGGCTTAGCTC	101	GTTTTTGAGCAGGAAGTAAT
W4	rs12402320	ACGTTGGATGGGACAGGCAATCCTTATACC	ACGTTGGATGGCCCTTTAGAGGACATGAAG	102	cACCACTTGAAGATGGACACA
W4	rs16971465	ACGTTGGATGGGTGGTCAGTGGTTATTAAG	ACGTTGGATGTTTCTGCCATCACGGTGGAG	115	cGTCTCATTGTTTTCTATCACG
W4	rs4788488	ACGTTGGATGCCCTTCCCAAGTAACAAAAG	ACGTTGGATGGTGTTTGGAAAGTGGAGTCG	96	GCCAGATAAATCCATGGAAGAG
W4	rs6688473	ACGTTGGATGCATCCATCTCCAGAACTGAC	ACGTTGGATGGAGATGATGTGCGAAGAGAG	112	AGAACTGACTTTTGTCATTTTCT
W4	rs6690388	ACGTTGGATGCACAAGAGCAAGGGAATAAG	ACGTTGGATGGCATGCAGTAAGTGCTAATA	98	gaatGGAATAAGAACATGGCCTT
W4	rs11264274	ACGTTGGATGTTCCCAGTCAACCTTCAGGA	ACGTTGGATGAAGGGACAGGGTGGAGGAAG	119	ccccTCAGGACCGGCTCCGTTCTC
W4	rs4788697	ACGTTGGATGAAACACTCCGACATCCACTC	ACGTTGGATGGGAGGTGTTACAAAGAAAGG	101	actaGCCCTGGTATTTTCATGGCA
W4	rs12135902	ACGTTGGATGATTCCTCTGGACAGGCTTGG	ACGTTGGATGCCTGTTCATTTCTAGGGCAC	98	gtttGGTCAGAGGAGACCAACACA
W4	rs2798601	ACGTTGGATGCGCTCAGCTTCACCAAAACC	ACGTTGGATGATGGGAAAGACTGAGGCAAG	105	ggagGCTTCACCAAAACCACCTTGT

Multiplex	SNP	Forward primer	Reverse primer	Amplicon length	Extension primer
W4	rs951241	ACGTTGGATGCAAAGGACCGGGATGGTTTC	ACGTTGGATGCTGCTGTGGACCTGAGAAGT	103	gaACATGAGGAAAAACATCTGGATA
W4	rs7204751	ACGTTGGATGGGTCACTTCCCTCTCTCAAC	ACGTTGGATGGAAATGGAAAACAGTAGCTC	115	CCCCCACTTTTTTTTTTTTTTTGAA
W4	rs6698628	ACGTTGGATGGCATAACTGAGAAGAGCAAG	ACGTTGGATGTGGAGGGAAGTGCTGCATAG	100	tcagAACTGAGAAGAGCAAGACACCC
W4	rs11075954	ACGTTGGATGGCGTGTGAGAGGGTAATAGG	ACGTTGGATGTTCTGCTCTGGGATGGAAAG	111	AGGAAATAAAAATAAAAACCCACAATTA
W5	rs13336412	ACGTTGGATGGGATTGTCATTCCAGTGCTC	ACGTTGGATGTTCCTTAGTGGCAAGAGAAG	100	CCCTTTGGTCCCTCA
W5	rs17719419	ACGTTGGATGCTCTTTGGACAGTTCACAGG	ACGTTGGATGTGAATGCACAGCAGTGCCAG	100	TCTAGAGGTGCCCTG
W5	rs12128882	ACGTTGGATGGGTTAGAGATGTCCTGAGTG	ACGTTGGATGGCAGCTTTGACAGTACCTTG	99	GCACTGGCCCTTCTGT
W5	rs9921395	ACGTTGGATGTGCTGGTGATGGGAACTTTG	ACGTTGGATGCAACAATAACTTCACTGTGG	93	TCAATGAAGGTGGTGT
W5	rs4788668	ACGTTGGATGATCCCATTCGTCAGTGGCTC	ACGTTGGATGACAGAGTCTGCCAAGGAACC	106	GTGGCTCAGTTAGTTCC
W5	rs11264251	ACGTTGGATGGATGTCAGGAGCACACCAG	ACGTTGGATGCACTTCCCCATGAATTCAAT	117	GGGAAGTTGTGGGTTTA
W5	rs7404992	ACGTTGGATGTAGAGGGCTGGAGAGATGAC	ACGTTGGATGGCGGGAATTTCCATTCTTTG	99	CGTTCTCCACTGTGGCCC
W5	rs868108	ACGTTGGATGTTAGGCTGGTGAAAAAGCCC	ACGTTGGATGATAAGCAGTTGCTTGGTGGG	106	GTAGCGTTGAGTGCACTGC
W5	rs8055870	ACGTTGGATGGGAAGAGCATATTTCCTGTC	ACGTTGGATGAGCAGCTGGAGAGACAGTTC	98	TCTTTTATGCTCTAGTGACA
W5	rs7554577	ACGTTGGATGGCTCAACAGTGGATTTCCAG	ACGTTGGATGAGCATTCCTGATAACTGCCC	101	CAAGTCTTCAAGTCTTTACAAT
W5	rs4845390	ACGTTGGATGTCATAGCCTGGCTTGACAAC	ACGTTGGATGGTTTATCAGGAACCAGCCAG	118	TCACCTGCCCTCTATTAAATAAGC
W6	rs117951282	ACGTTGGATGGGAAAAGTTAAGCACTCCC	ACGTTGGATGGCTACAAGTCCTCAACACTC	100	GCACTCCCACTTTCA
W6	rs1419110	ACGTTGGATGGCCCACTGGAAAATACCTTG	ACGTTGGATGCTTGGAGAAATGTGGACCTG	111	ACACTCAGCAGTGAC
W6	rs13376333	ACGTTGGATGATTGGGAGGTCATGGGATCG	ACGTTGGATGGGCTTTGCTCTGGAGAGAAC	99	GGTTCTGCTCCTTCCT
W6	rs16971384	ACGTTGGATGTCAGTCCATTGAACTGGAGG	ACGTTGGATGGTGGTTGCTGTCTTTCCCTG	100	GCCCCATGATGGAAGA
W6	rs1061122	ACGTTGGATGAGGAGTAGGTGCCAGCTCAG	ACGTTGGATGGAGTGCTGCTCAGTGTCAAA	100	ctCCAGCTCAGTCCCTG
W6	rs10852515	ACGTTGGATGTAATCACACACCTCGCACCG	ACGTTGGATGAATATCAGTAGCTCCTGCGG	100	TTGGTTTTTGGTTTGGT
W6	rs1051614	ACGTTGGATGGATGACTGGCGGATAGCCAT	ACGTTGGATGATGGCGCACACCAGCATCTC	91	gaaCTACGAGCGCATCCT
W6	rs4492614	ACGTTGGATGCTAGATGGTCTTACCCTCTC	ACGTTGGATGGTGTCATCTCACTCCCAATG	81	acTACCCTCTCCTCTCCTT
W6	rs4845663	ACGTTGGATGTTGCAGGTTCTCCGATAGTG	ACGTTGGATGGCATGACTGTAACTTCAAGG	99	cCTGATCATCACTCAGCCT
W6	rs6682291	ACGTTGGATGGTACGAAAAAAACTAGCCACG	ACGTTGGATGCTTTCCTCTTCCTCATTGGC	98	ctCCACGAATGAGCAGTTT

Multiplex	SNP	Forward primer	Reverse primer	Amplicon length	Extension primer
W6	rs7199343	ACGTTGGATGGCCCTGATAGGGTTTTAAGC	ACGTTGGATGCATGGATTCTGCTGGCACTC	92	GCTTTGGAATTTGGAAGTC
W6	rs2157786	ACGTTGGATGCACCCTATTCACCATTCTCC	ACGTTGGATGCTCTAGGTGAGGATGACTAC	95	tcCCATTCTCCCTGGTTAAA
W6	rs740178	ACGTTGGATGACCTCCTGGCAATGGGAGAC	ACGTTGGATGTCCTCCTTGTCTTCCTCAAC	90	ATGGGAGACCCCACTCTGGC
W6	rs76110919	ACGTTGGATGACTCCTCTCCCACCGCTTTC	ACGTTGGATGTGGAGGAAGGGTGGAGGAT	89	AGGGCCCCCCCTTCGTCCAAC
W6	rs7193343	ACGTTGGATGGAGGGGAAAGTTTGAACAGC	ACGTTGGATGAAATGTCGAGTCCTAATGGC	89	GAAAGTTTGAACAGCTTGTTT
W6	rs1218601	ACGTTGGATGCTTGGGATAACTCCTTCCTC	ACGTTGGATGGCCCCCCAATCTTGTAATTC	113	agAACTCCTTCCTCCACTCGCA
W6	rs16971474	ACGTTGGATGTAGGCAGACCCCATCATTTC	ACGTTGGATGCCCACAGGCAAGATCTTGTC	100	cGACCCCATCATTTCGAGATTA
W6	rs4788482	ACGTTGGATGAGGAACTGCTGGATTCTTGG	ACGTTGGATGTGTTCCTGCTGCGAGAGTTC	118	caTTAGCATGGGCCTTTCTCAA
W6	rs2228200	ACGTTGGATGAGAGCTTGCACTGGTATGAG	ACGTTGGATGCTGCACATGAACGTGGAGC	99	cTCACCGCCTTCCACTCGTCCTC
W6	rs2106261	ACGTTGGATGCACAGATAGAGCTCGTCCAG	ACGTTGGATGAGAGCAGTCTCTGGCACACT	100	ATAGAGCTCGTCCAGAGAATTGT
W6	rs6691316	ACGTTGGATGGGCTCAACTCATGCTCAATG	ACGTTGGATGAAGAGGTGCTCGATGCACAG	101	ggGAATGAGTGAATAAGCTGATC
W6	rs9940520	ACGTTGGATGCAGTTCAGACTCTCTGTTAG	ACGTTGGATGAAGCAACATAAGGCAGGCTC	82	CAGACTCTCTGTTAGAGGCCTACT
W6	rs1218565	ACGTTGGATGAACACCTCAGAATCCTTCCC	ACGTTGGATGGTGCAGGAACAGCAAATAGG	116	CACCTCAGAATCCTTCCCAACACAG
W6	rs7193297	ACGTTGGATGCATTCGTTGCAGGTGACCTC	ACGTTGGATGAATGAGCGCCTCGCGGAGA	100	GTGACCTCCTTGCTGGCGGGCTCGG
W6	rs16971464	ACGTTGGATGAGGAGTCACTCAGATGTTTG	ACGTTGGATGCTCATTTGACCTCTAACAATC	120	tgAACCTGTAAAATCGAAAATATGA
W6	rs4788679	ACGTTGGATGACCTACCTACAGCCTTATGG	ACGTTGGATGGAAGAAAGCCAGGTACGAAG	84	CCTACCTACAGCCTTATGGAACGACC
W6	rs2266943	ACGTTGGATGTTCCCAAAAGGCCATTCCAC	ACGTTGGATGTGGCTCTGTTTTGACTGGTG	111	ccTCAATTTCAAGACACCTTGCTCCCT
W6	rs1476646	ACGTTGGATGTTCAGTAGCAACTCAGAGCG	ACGTTGGATGCCCGGGAATACTTTCACATC	119	CAGAGCGTCGTGCTCTCTTGTCTTTGT
W6	rs2106258	ACGTTGGATGTTAGATCCTGGTTAGCGGAG	ACGTTGGATGAGCCTGAACAGAGCCAAATC	101	tTTAGATCCTGGTTAGCGGAGCCTTCG
W6	rs8052905	ACGTTGGATGACCAAAGTATTCCCCAGTCC	ACGTTGGATGCCACTACATGCAGCTTCTTG	109	GGGAAAGCAAGATCCTGAGAAAGAATA
W6	rs8056528	ACGTTGGATGAAGCTAAGTAGCCTAAGGAC	ACGTTGGATGCTTGTGAATTAGCGAGAACC	97	agTTTTGTGAAGCATCTTTGCTGCTTTT
W6	rs6695232	ACGTTGGATGACACCCCGATCCTATTTCAC	ACGTTGGATGTCTCACCTTACTTGCTGCTG	119	ctGAAGATATGGACATGATTTATGCCCC
W7	rs11640106	ACGTTGGATGAACAGACTGTCCTTCTCCAC	ACGTTGGATGAAATCTGAAGGCCCTGAGTG	113	CTCACCCTGCCCCCA
W7	rs925042	ACGTTGGATGGCCGCTTCTCGTTTTCAAAG	ACGTTGGATGGAAAGTGGAGATGGAGGGC	119	TTGGTCCTTGCACAA
W7	rs756720	ACGTTGGATGTGTTTTCCTGATCACCAGGG	ACGTTGGATGGCAGCATTGGACTAAATGGG	96	tATCCTCAGCCTTCCC

Multiplex	SNP	Forward primer	Reverse primer	Amplicon length	Extension primer
W7	rs12929452	ACGTTGGATGTATAAGGAACACCATCCTGC	ACGTTGGATGGCACAGGGTGTTATATTGGC	98	ACCATCCTGCCACTTA
W7	rs4788480	ACGTTGGATGGAAGAGTCTGCTTCAAGAGG	ACGTTGGATGTTTGCAGAACTGTGGCCCTC	101	CTGTATGCAGAACGGA
W7	rs8051826	ACGTTGGATGAACTTCCCTAGTATGGCCAG	ACGTTGGATGAGGGTTGGAACCCATGTAAC	96	tGAGGCTGGAAAACCA
W7	rs12373097	ACGTTGGATGGAGAGGGTGAAACGAGCATT	ACGTTGGATGTACAACAGACTTACTGCTTC	106	CATTTCTAGAGGCTCCC
W7	rs739414	ACGTTGGATGGAGAACGAGCTTCCCAAATC	ACGTTGGATGTTTCCGAAAGAGCTGTCGTC	84	tcCCCAAATCCAACTCCA
W7	rs12029542	ACGTTGGATGTACATCATGCCCAAGTAGCG	ACGTTGGATGCCCTTTGAGCCTCTGTTTTC	94	ATCCCAGGAATGCAAATA
W7	rs10908435	ACGTTGGATGAATTCCTAGCACAGTGCCTG	ACGTTGGATGTGCTCTCATGAGAGACTTGC	106	GTGAATAGCAGAGCTCCA
W7	rs11641701	ACGTTGGATGTGATGCTTGCTAGACACTCC	ACGTTGGATGTTCCAGCAGTCTGTAATCAC	120	cccaGGCCTGGCTCATCTT
W7	rs1106236	ACGTTGGATGCCTCATAGATGTGAAAACCC	ACGTTGGATGTACATTAGCAAGCGCCTCAG	107	ggaAAAGCGCTAGACAAGA
W7	rs1218574	ACGTTGGATGCTGAGATGGTGACTGAGTTC	ACGTTGGATGTTCCTGAGGTCTGCAGATCG	114	tctcGCATAACCACGTGCCC
W7	rs1131820	ACGTTGGATGCTAATGGACTTTCTGCTCCC	ACGTTGGATGGCTCGTAGGTCATGGCTATC	100	aAGCTCTTCGTGATCGACAA
W7	rs9930445	ACGTTGGATGTTTGCTTTGTTGAGGCACAG	ACGTTGGATGACCAGTGATGGCAATGTGTC	100	cGTTGAGGCACAGATTAGGT
W7	rs11264261	ACGTTGGATGAATTCTCCACCTCCATCCTG	ACGTTGGATGCTAACCTAGAGGCCTTGGTC	117	TCCATCCTGTACTTAGTAACC
W7	rs1106235	ACGTTGGATGATAGAGAGAGGAGCTACACC	ACGTTGGATGGCTCTGAGCTTTCTGTCTTC	119	aAACACCATGGAATAAATCCT
W7	rs4788683	ACGTTGGATGATTACACAGTGCGACTGTTC	ACGTTGGATGGGTGCAGTTATGACGAAGGG	104	tcccGTGCGACTGTTCATGTTT
W7	rs1858800	ACGTTGGATGGATCAGTATGTGGTGATGGG	ACGTTGGATGTCTAAAGCCCATGTCTCAGG	109	CACGATGTGTGAGACTGATAAC
W7	rs8057081	ACGTTGGATGTATTCCAGAGCTGCACCATC	ACGTTGGATGTCGAGACAAGCCACATGTCAC	98	tcGCCACCGACCCTTGTCACTAC
W7	rs10908429	ACGTTGGATGCCGTGACAACCTTACGTATC	ACGTTGGATGGACCAGAAATCTGAATTGGC	103	CTTACGTATCTGAAAATGTCTTC
W7	rs4845671	ACGTTGGATGTATGTGCAGAATCACCTGGC	ACGTTGGATGGGGAGACACAATTCAGTTC	111	agGGGTAATTTTGACACAAAAGG
W7	rs954785	ACGTTGGATGACTAGCCTGGAAGCTCTATG	ACGTTGGATGTTACACTGTTCCTGGCACTG	100	GGATCATATCTGTCCTGTTCATTA
W7	rs12740456	ACGTTGGATGTCTACCAGGTATTGTCTCCC	ACGTTGGATGCAAGCAAAGATGTCTGCAGG	89	aggaTCTCCCTACTACATCCCCATC
W7	rs9940321	ACGTTGGATGTCACATTGCTTATACCGCTC	ACGTTGGATGACTCAGATATCCAGGAGCAG	119	tcCCGCTCAAACTTATGACCAGTTC
W7	rs6426905	ACGTTGGATGAGAGGACTCACCTTCACCAG	ACGTTGGATGGAACCCTAATGAAAGCCCAG	110	cacaAGCTCCCTTGCTGCACTTTGT
W7	rs883319	ACGTTGGATGTGGCCTCCTAAGCACTTTTC	ACGTTGGATGTAGTTGCTATGAGGTCTGGG	113	gggAACTCTTTGAGTCTCACAATAG
W7	rs11264268	ACGTTGGATGGGCCCACTGGAGATACAAAG	ACGTTGGATGGCTCAGCAAATGAGAAGTGG	108	CTAATAAAGTAACTAATTCATTCACT

Multiplex	SNP	Forward primer	Reverse primer	Amplicon length	Extension primer
W7	rs8055406	ACGTTGGATGTTGGGCAAAAACAGACCCTA	ACGTTGGATGGTTTCTTTGAAGAAAGCATC	100	gcagAAATTACAGTGCACAGAGCTCC
W7	rs16971456	ACGTTGGATGGGATGATGGAGGTGGATTAG	ACGTTGGATGAGTCCTAAGAGAGCGTGGAG	115	ggcgGCTACATCTGTAAAAGAGTAAA
W7	rs9936884	ACGTTGGATGTTTGCGTCTCTGTCTCTCTC	ACGTTGGATGTCTGGCTTTTGTTACTTGGG	99	agagCACACTCTCACTCTATCCTGCCC
W7	rs12145386	ACGTTGGATGCTGTAGGTGACTTTATCATCC	ACGTTGGATGACAATCTCAGGGAGGGAATG	101	TCATCCTTTAATTTATAATTAGCAGTT
W7	rs4845391	ACGTTGGATGAGCCTACAGGGTTTAGCTTG	ACGTTGGATGGGAAAGTGGTCATTACAGGG	94	cttcGCTTGCAGGCTTCCATCAACCCCA
W7	rs6499594	ACGTTGGATGTGACCTCTCTTTACCCTTGC	ACGTTGGATGGACATTCCCTGATGAAGGAC	115	acAAAGCTTGCTGAAATTTACAATCCCT
W7	rs960182	ACGTTGGATGTTCCATGGCCTTCTAGCTTC	ACGTTGGATGGAGGCCAAATACAAAGAACC	100	cgcgTCCAGTGTTGCTGTCAAAAAGTCC
W7	rs9940310	ACGTTGGATGCACGTGATCAGTGAACACAG	ACGTTGGATGTGAGTGCTTCCTATCAAGGG	98	tgacCACAGGACGTGTGCATGTCAGTAC
W8	rs1218578	ACGTTGGATGAAGCATCACTCGTTCTGCTG	ACGTTGGATGAACTCCTCCTCCTGGTTCCT	105	GTTCTGCTGCCTTCC
W8	rs11863932	ACGTTGGATGAGAAGGCTCTCTAGGAACGG	ACGTTGGATGCCTGGAGATTGTGAGTCATC	100	TGAGACGTTGCTTCC
W8	rs2335407	ACGTTGGATGGGAAATAACAGAGAGGACCC	ACGTTGGATGCTAACACCCGGTGGATAAAC	98	GAACACGAGGCGTAG
W8	rs10494301	ACGTTGGATGTGGGCAGAAAATGAGATGGG	ACGTTGGATGATTCTTCAGCCTCCTCCTTC	110	GGGGAACAAGTTGCGT
W8	rs884664	ACGTTGGATGTACCTCCCATCGGTTAGTAG	ACGTTGGATGTGGTAAGAAGCGTCTTTCCC	101	CATAGCACCAACTGTCA
W8	rs1506982	ACGTTGGATGATAGTGGCTCGCTCAAATGC	ACGTTGGATGAAATCCTTTCCCCTGTCCTC	89	AATGCCGAAACTTGGCT
W8	rs6499600	ACGTTGGATGAGATTTCGCCTGCATGCCCT	ACGTTGGATGTCCCACCAGCACCAACCATC	106	cGCCTGCGGTTCATTTCC
W8	rs16971366	ACGTTGGATGAGATACCTGTAGGAGGTGAC	ACGTTGGATGAACCCTTCAGAAATAACCCC	99	tcCTGGGGATGAGGTCCA
W8	rs4788489	ACGTTGGATGGCAACTTAATAACCACTGACC	ACGTTGGATGTTCCGCATTAATGCTCTTCC	101	tacCCACTGACCACCTTCT
W8	rs8060701	ACGTTGGATGAAAACAATGGCAGGTTCCCG	ACGTTGGATGTGGAATCACATACGCAGAGG	109	ttGCAGGTTCCCGGCCACT
W8	rs1218551	ACGTTGGATGTTTTCTCACAGCTGTTAGGG	ACGTTGGATGTGCATGCTTCTCTTTGTGTG	102	TGAGATGAGATCATGCATA
W8	rs6681725	ACGTTGGATGTCCCTACATCAATCAACTGG	ACGTTGGATGGATGCTTTCAGACCCCTTTG	119	ACAGCTCACAATGTTTTCCA
W8	rs2256209	ACGTTGGATGTTCTCTTCTTCCAGAGCCAG	ACGTTGGATGGACTGCATGTGTATGTGTGG	105	GCAGTATTGGCTGTAAAACT
W8	rs4845394	ACGTTGGATGATGTCAGTTTAGGGCTTTCC	ACGTTGGATGTTCCATCTGTGGGTTAGGAG	120	gtagTTGATGCTGGGAGGTA
W8	rs8058014	ACGTTGGATGTAACTAGCCACAGACAAGAG	ACGTTGGATGGGGACAGTTGTAGAGCATTA	118	AGCCACAGACAAGAGAAATCT
W8	rs9925261	ACGTTGGATGCCACACTCTAGCTCATGTTG	ACGTTGGATGAATGGCTCTCAAAGTGGCAG	99	cccgCTTACACTGCCATGAATT
W8	rs756717	ACGTTGGATGTAAGAGGAGAACGGAGCAAG	ACGTTGGATGAAAGCCAATCTCCTCTCCAG	115	cGCCCACCAATTAATTGGTAAA

Multiplex	SNP	Forward primer	Reverse primer	Amplicon length	Extension primer
W8	rs7523074	ACGTTGGATGACCACCCTGCAGACTTGCTC	ACGTTGGATGGAGTGTGGAATAGAACAGGG	118	cccaCCGGCTGCCCTCTCTTTCA
W8	rs2335406	ACGTTGGATGGGCAGATTTACTCCCTCCTC	ACGTTGGATGTCAGGCTCTCAGCCTGTTAG	83	gCTCCTCATTCTAAGGAGAAAGT
W8	rs2040508	ACGTTGGATGTTGGGATCCCCTTCCAAATG	ACGTTGGATGTGGCTGGAGGACATCTTTAC	99	aAAATGTCAACAAACAAGAAAAG
W8	rs1995662	ACGTTGGATGCTGCTTATCCCTCCTGAATG	ACGTTGGATGGCTTTGTCCTGGATATGCAC	100	ttttGCTTGGCATCCATTTCAAAT
W8	rs7532286	ACGTTGGATGTCAACAAGGTTAGGCCATTG	ACGTTGGATGCTACTTGAGATTGTGCACTG	99	ggatGCCATTGTGGGTATATAAAT
W8	rs1548373	ACGTTGGATGCTGAGACTCAAACTTCACCC	ACGTTGGATGGTGAGAGGTTGAGCTTTGTG	96	gggaACAGGATTGAGACAAGTTAT
W8	rs10494300	ACGTTGGATGCCTTCCATGTAGGAATGCAA	ACGTTGGATGCATGCAATTCCAGTCAACCC	120	ccctATTATTACCATTGTTGCTGAT
W8	rs906274	ACGTTGGATGGGCCTAATTAATACTGCAGTC	ACGTTGGATGAATAGTCTGTACTGGCTGTG	91	AGTCATTATTGATTCAATCCTGAGA
W8	rs6426987	ACGTTGGATGTGCCTCTGTTGAGTCAAACC	ACGTTGGATGTACACTGTGCTCACCATCTG	101	caGGACTCTTCCCTTGAGGCCCTGAT
W8	rs719353	ACGTTGGATGGGATGTTAAGTCTCTGTGCC	ACGTTGGATGGGGTGACTTTGCTTCTGAAA	98	ggacGAAAATTTTATCTTGGGCATAT
W8	rs4845396	ACGTTGGATGTGACTCAAGGACCCTTGGC	ACGTTGGATGTCTGTTAGCAGGGTCATTGG	109	ccccGACCCTTGGCGCTTGCTTTTGCT
W8	rs906276	ACGTTGGATGACAACTGTCTGGGCTCTTTC	ACGTTGGATGGGAACAAAGGAGAAGCATCG	101	ctcaTTCTGTGGTTGCTCAACTCCATC
W8	rs4788689	ACGTTGGATGTCCAGGACTGGGTTTTACAC	ACGTTGGATGCATTCTTCAAGGTGCCATGC	95	ttcaAGGACTGGGTTTTACACACCATA
W8	rs4788671	ACGTTGGATGTTCAAGTGCTTAATAGTGCC	ACGTTGGATGCAGAACCTTCTGAACTGCTC	98	ggatTGCTTAATAGTGCCATGGGGCTC
W8	rs4788692	ACGTTGGATGACAGTGGGCCTTCTTTTCAG	ACGTTGGATGGAATGCATGAACCAGAACGC	120	gaacGAGGGCCCGTGGTGACCTGCAGAG
W9	rs4788696	ACGTTGGATGTCCTAAACTCCTTTTCCCCC	ACGTTGGATGAGCCATTTGCATTCCTCCAG	99	TTCCCCCCACAAAGC
W9	rs4999127	ACGTTGGATGAGTGAGAACTGACTCCAAGC	ACGTTGGATGCAGCTGTTCTCTTCAGCAGA	96	GAATTTGGCAGCCAC
W9	rs6426944	ACGTTGGATGCTTACCACATGAGCATTCCC	ACGTTGGATGAAAATGCAGAGTTGCACGGG	111	AGCATTCCCAGAGGA
W9	rs11264270	ACGTTGGATGTGGCCATGCCTGGTGTCTC	ACGTTGGATGGTGAGAACAACATCAGTGGG	117	aTGGGGTGACCAGAGC
W9	rs10128027	ACGTTGGATGTTCCTTCCCTCTAACTTCCC	ACGTTGGATGGTCATTGACATGGGATGCTG	116	CAACAGAACACACACCT
W9	rs11075958	ACGTTGGATGTGCACACCTTCTCTTGCCTG	ACGTTGGATGAGCTCTAAGAGCAGGGCAAC	117	ggGCCTGCAGCCAAAGC
W9	rs699444	ACGTTGGATGATCTGCAGGCAAATCTGGTG	ACGTTGGATGGCAGATTGGTCATTTGAGTG	94	ggATCTGGTGACAGCGG
W9	rs12087736	ACGTTGGATGCCAAGGATGCACACATCCAG	ACGTTGGATGAAAGCTGGCAGGACTAGAAC	99	CACACATCCAGTTTCATT
W9	rs6682000	ACGTTGGATGGGTTTAGACAACTGCCAGAG	ACGTTGGATGCCTGGAGTTGTTAATGATCG	113	gGCCAGAGTTTGCACTTG
W9	rs12445932	ACGTTGGATGTGACTGGTGCTCCTGACTTG	ACGTTGGATGTTGGTGAAAGGAGGTTCTGC	102	ccatCACGGATCCCCTCCC

Multiplex	SNP	Forward primer	Reverse primer	Amplicon length	Extension primer
W9	rs4788684	ACGTTGGATGTTCTTCTTTTGCCCGGGCCT	ACGTTGGATGGCCCCATAAGAATACCTAGC	118	cctttCCCGGGCCTCAATTT
W9	rs17719287	ACGTTGGATGTCCTGGGAAGAGGATTTGTC	ACGTTGGATGAATGTGCACAGCACTGTCCC	107	TTGTCTACATTCTACGCACA
W9	rs11584635	ACGTTGGATGAAGTCGGCTTTGAAATGCAG	ACGTTGGATGGCATTTCCCAGTGTTCCTTG	116	TCTTGTTTTGAAGTAACAGC
W9	rs11264249	ACGTTGGATGTTCAGGAGATAGTAGGTGGG	ACGTTGGATGTAGACTCTCAGGCTTAGCTC	101	GTTTTTGAGCAGGAAGTAAT
W9	rs12402320	ACGTTGGATGGGACAGGCAATCCTTATACC	ACGTTGGATGGCCCTTTAGAGGACATGAAG	102	cACCACTTGAAGATGGACACA
W9	rs16971465	ACGTTGGATGGGTGGTCAGTGGTTATTAAG	ACGTTGGATGTTTCTGCCATCACGGTGGAG	115	cGTCTCATTGTTTTCTATCACG
W9	rs4788488	ACGTTGGATGCCCTTCCCAAGTAACAAAAG	ACGTTGGATGGTGTTTGGAAAGTGGAGTCG	96	GCCAGATAAATCCATGGAAGAG
W9	rs6688473	ACGTTGGATGCATCCATCTCCAGAACTGAC	ACGTTGGATGGAGATGATGTGCGAAGAGAG	112	AGAACTGACTTTTGTCATTTTCT
W9	rs6690388	ACGTTGGATGCACAAGAGCAAGGGAATAAG	ACGTTGGATGGCATGCAGTAAGTGCTAATA	98	gaatGGAATAAGAACATGGCCTT
W9	rs11264274	ACGTTGGATGTTCCCAGTCAACCTTCAGGA	ACGTTGGATGAAGGGACAGGGTGGAGGAAG	119	ccccTCAGGACCGGCTCCGTTCTC
W9	rs4788697	ACGTTGGATGAAACACTCCGACATCCACTC	ACGTTGGATGGGAGGTGTTACAAAGAAAGG	101	actaGCCCTGGTATTTTCATGGCA
W9	rs12135902	ACGTTGGATGATTCCTCTGGACAGGCTTGG	ACGTTGGATGCCTGTTCATTTCTAGGGCAC	98	gtttGGTCAGAGGAGACCAACACA
W9	rs2798601	ACGTTGGATGCGCTCAGCTTCACCAAAACC	ACGTTGGATGATGGGAAAGACTGAGGCAAG	105	ggagGCTTCACCAAAACCACCTTGT
W9	rs951241	ACGTTGGATGCAAAGGACCGGGATGGTTTC	ACGTTGGATGCTGCTGTGGACCTGAGAAGT	103	gaACATGAGGAAAAACATCTGGATA
W9	rs7204751	ACGTTGGATGGGTCACTTCCCTCTCTCAAC	ACGTTGGATGGAAATGGAAAACAGTAGCTC	115	CCCCCACTTTTTTTTTTTTTTTGAA
W9	rs6698628	ACGTTGGATGGCATAACTGAGAAGAGCAAG	ACGTTGGATGTGGAGGGAAGTGCTGCATAG	100	tcagAACTGAGAAGAGCAAGACACCC
W9	rs11075954	ACGTTGGATGGCGTGTGAGAGGGTAATAGG	ACGTTGGATGTTCTGCTCTGGGATGGAAAG	111	AGGAAATAAAAATAAAAACCCACAATTA
W10	rs13336412	ACGTTGGATGGGATTGTCATTCCAGTGCTC	ACGTTGGATGTTCCTTAGTGGCAAGAGAAG	100	CCCTTTGGTCCCTCA
W10	rs17719419	ACGTTGGATGCTCTTTGGACAGTTCACAGG	ACGTTGGATGTGAATGCACAGCAGTGCCAG	100	TCTAGAGGTGCCCTG
W10	rs12128882	ACGTTGGATGGGTTAGAGATGTCCTGAGTG	ACGTTGGATGGCAGCTTTGACAGTACCTTG	99	GCACTGGCCCTTCTGT
W10	rs9921395	ACGTTGGATGTGCTGGTGATGGGAACTTTG	ACGTTGGATGCAACAATAACTTCACTGTGG	93	TCAATGAAGGTGGTGT
W10	rs4788668	ACGTTGGATGATCCCATTCGTCAGTGGCTC	ACGTTGGATGACAGAGTCTGCCAAGGAACC	106	GTGGCTCAGTTAGTTCC
W10	rs11264251	ACGTTGGATGGATGTCAGGAGCACACCAG	ACGTTGGATGCACTTCCCCATGAATTCAAT	117	GGGAAGTTGTGGGTTTA
W10	rs7404992	ACGTTGGATGTAGAGGGCTGGAGAGATGAC	ACGTTGGATGGCGGGAATTTCCATTCTTTG	99	CGTTCTCCACTGTGGCCC
W10	rs868108	ACGTTGGATGTTAGGCTGGTGAAAAAGCCC	ACGTTGGATGATAAGCAGTTGCTTGGTGGG	106	GTAGCGTTGAGTGCACTGC

Multiplex	SNP	Forward primer	Reverse primer	Amplicon length	Extension primer
W10	rs8055870	ACGTTGGATGGGAAGAGCATATTTCCTGTC	ACGTTGGATGAGCAGCTGGAGAGACAGTTC	98	TCTTTTATGCTCTAGTGACA
W10	rs7554577	ACGTTGGATGGCTCAACAGTGGATTTCCAG	ACGTTGGATGAGCATTCCTGATAACTGCCC	101	CAAGTCTTCAAGTCTTTACAAT
W10	rs4845390	ACGTTGGATGTCATAGCCTGGCTTGACAAC	ACGTTGGATGGTTTATCAGGAACCAGCCAG	118	TCACCTGCCCTCTATTAAATAAGC
W11	rs3784807	ACGTTGGATGCAAGCTTTCCCCAGTCTGTG	ACGTTGGATGCTTGTTGTAGCTCACGGAAC	103	TCCCTAGCCCCAACC
W11	rs4746139	ACGTTGGATGTGACCTCCACCGCTTCTATC	ACGTTGGATGCTCAGGAGCCCCTGGGCCT	109	CTTCCTATCTGCGCC
W11	rs502612	ACGTTGGATGATACAACTTAAAGGCCAAGG	ACGTTGGATGGCTAGGTCTTACACTTTGCC	118	GGCCAAGGAGCCCAA
W11	rs60632610	ACGTTGGATGGGGACACTGTTCAAACTCCG	ACGTTGGATGTCCCCCTGATAGTGTGACCA	105	cCCAAGCTGAGCGATG
W11	rs11947581	ACGTTGGATGCAGCCTAGCTTACAGAAACC	ACGTTGGATGATTGAACCTGGAAGACTGGG	113	ACCCTCTGGATTTTCTC
W11	rs2623998	ACGTTGGATGCCTTGACAGCAAGCTTAGAG	ACGTTGGATGCTTTCTTGACCACACTTGGG	85	AAGCTTAGAGGTGTCTG
W11	rs34163229	ACGTTGGATGTTACAACCCACTGCTCTCTC	ACGTTGGATGTGATGCCCTGCTTGGGTGTT	118	CTCTCCCTAAGGCCCAAT
W11	rs11589602	ACGTTGGATGAATCCCGAGGGAGAATTCTG	ACGTTGGATGCAAGGCACAATGATTGGCTC	86	TCTGTTTTCTTGGCTTGT
W11	rs488156	ACGTTGGATGTAGCTAGTGGCAAGCTGTTC	ACGTTGGATGACTGCAGGTGATGAGCTGTG	94	ACTAGTGACCGAAGATGA
W11	rs3743496	ACGTTGGATGCCGTTAGCTGTAACTTGGAG	ACGTTGGATGCTTTTCTCCCAAGGTCGCAG	119	ctttGAGCTCCACTCTGCC
W11	rs2723296	ACGTTGGATGGAATATTGCGGTAAGGAGGC	ACGTTGGATGCTTGTAGTTTACGCAGGCAC	89	GGCCATTAGGCTGGAAACG
W11	rs913257	ACGTTGGATGCTCTAAACGAACCACTGGTC	ACGTTGGATGGTAGAAGCCGATGAAGAGAC	119	CCTTGATTCTACTTCTTGTT
W11	rs12498374	ACGTTGGATGTTTTCCCACTTAATCTAGGC	ACGTTGGATGTAAGCAATACAGGTCTGAGC	105	cctcGTGAATCAGCCCCATA
W11	rs2723333	ACGTTGGATGCTGCATATGTACGACCCTTG	ACGTTGGATGCAAAAACCCACCCAGATTC	120	GAAGAAATGGAGCCTATATC
W11	rs3853445	ACGTTGGATGGCTCACTGATAAGCCAGTTC	ACGTTGGATGACAACTTACTGCCACATGCC	108	GCATTTTCTTAGCCAAGATAC
W11	rs7674295	ACGTTGGATGATCCTCCTTGAATTTTCTG	ACGTTGGATGATAGAAACAGTGCAGTGAGG	100	cCCTTGAATTTTCTGAAGTGC
W11	rs542331	ACGTTGGATGGCCTATGGTGCTGATATCTC	ACGTTGGATGCTCAGGAGAGGTTATCACAC	106	ggccGAAAGTGAGCAGATGTG
W11	rs11943026	ACGTTGGATGCTAGTTGATCCTCACTTGGC	ACGTTGGATGAATGTCTTACTTGGGAAGCG	98	ccctTTGGCTCACTTACCAATT
W11	rs10824026	ACGTTGGATGCTATAGTCACCCTAGGTGAA	ACGTTGGATGTCCACCTCAGCAGAGGAGAC	96	AGTCACCCTAGGTGAAGTATTT
W11	rs6533526	ACGTTGGATGAAGCATGTTATACAGAAGCG	ACGTTGGATGACATACAGTGTGGCATTTAG	98	TACAGAAGCGATTAGGATTTTT
W11	rs12021752	ACGTTGGATGGCTCACTATGCACTCTCTTG	ACGTTGGATGTGAAGCTCTCTGAGGGCTA	119	TTCTCTAACCTTTATTTTGATTG
W11	rs12247028	ACGTTGGATGTGAGCGTGGTATCTTGAGTG	ACGTTGGATGTCTGTGGTATAAACACTGGG	101	gtTATCTTGAGTGGAAAGTAGAT

Multiplex	SNP	Forward primer	Reverse primer	Amplicon length	Extension primer
W11	rs2595110	ACGTTGGATGGCCAAATAACCAATTTCAAGG	ACGTTGGATGGCCTGCTTGGACTTTAATGG	113	agTTAACAGTCATTTTCTTTTCCC
W11	rs7164883	ACGTTGGATGCTAAACCACAGATCAACCCC	ACGTTGGATGGTGGTATCTTTCTCTGATCC	116	ctatCATGAATAGATCACGTGACC
W11	rs3826046	ACGTTGGATGTGCCTTACTAGGCCCAGAAG	ACGTTGGATGGGGTAATTACTTCTGGGCTC	101	gttgaCTGCAATAATGTCACTGAG
W11	rs4032971	ACGTTGGATGCTTGTGAACCTGTTAGACTC	ACGTTGGATGTGTCCTTAGATACGTGTTTG	118	gggaCTGTTAGACTCTGTTTCTTTG
W11	rs591715	ACGTTGGATGATACTGGTCTAGACTCCAGC	ACGTTGGATGCATTCTCTCACTTGGTAAGC	111	ggaggAGTATACTACTGGGACAGTT
W11	rs12027180	ACGTTGGATGCCTGCCACAGAAATTCCATC	ACGTTGGATGTAGGCTGCCAGGTGTTAAAG	96	cccctACAGAAATTCCATCCATGGAC
W11	rs374582	ACGTTGGATGACTCTCCCAAGAGAAGAGCC	ACGTTGGATGCAGCCTATGCAATGCAGTAG	119	cgCTAATAAATCCTTTAAGACCAACA
W11	rs41280400	ACGTTGGATGTCCAGATGCCCAAGTTTGAC	ACGTTGGATGGGGTTCGATTGAAAGAAGGC	116	aggtGGCCCTTGCTGAGTGAACCCCT
W11	rs11799764	ACGTTGGATGAGTATGTTCTAGTCTCCTAC	ACGTTGGATGGAGAATCCTAACAGCAAACC	97	tgCTATAACAAACATACCTTCATAATG
W11	rs16863425	ACGTTGGATGACTCTTGTACTAGAAGCTAC	ACGTTGGATGTGAAGGCCATGCTGCTTTAC	95	gggagGCTACTTATTACACTAAGTTGA
W11	rs441624	ACGTTGGATGAAGCAATTCCCTCCCTAGTG	ACGTTGGATGGAAAGCTTGAAGCCTTTCTC	98	cccctCCTTTGCAAATTTAATGTCATTC
W11	rs2429074	ACGTTGGATGAAGAGGACTCCAGTATTTGC	ACGTTGGATGGGCTGGGAGAGACATTTTTC	113	CTCCAGTATTTGCTTTAATTTATTTTTG
W11	rs11938968	ACGTTGGATGTCTCCCAGTGTGGCAAGTTC	ACGTTGGATGCTGCCCTGCGAAATATTTCT	117	aaatAGAATTCTAACTTTGAACCTAACC
W11	rs4307025	ACGTTGGATGAACACTGTGCAGTGGAATCG	ACGTTGGATGAAGGTTCAACCTTTTTTCC	116	AAAAATTATAATAAACTAGCATAAGGAC
W12	rs3812629	ACGTTGGATGTTACAAGACCCTGCCTCACG	ACGTTGGATGGAGTCTTAGGAGTCATAGGG	103	CACCTAAGACCCCCC
W12	rs17571707	ACGTTGGATGCTCAGCTTCTTAAGGCTCAG	ACGTTGGATGCAGCCCACTTTCATTCCTAC	103	GCTCAGGGATCAGGA
W12	rs478438	ACGTTGGATGCTTCATTTCCAGGCCACTAC	ACGTTGGATGTTCCCCAAGTTCCATGTGTG	99	ACTACAAGCCTTGCCC
W12	rs75911264	ACGTTGGATGAAGCATGTTATACAGAAGCG	ACGTTGGATGACATACAGTGTGGCATTTAG	98	cCTTGCGAGCAAGGGA
W12	rs12066968	ACGTTGGATGCTGTAGGAGTCACACTCTGG	ACGTTGGATGTGATGGTGTCTTGAACTGGC	95	aACCACTGCTGCTCTAT
W12	rs3820416	ACGTTGGATGGAAGAGTTCCCCCATTCATC	ACGTTGGATGTTGGGATGTACAGTTCCTTC	97	TGTCCCATTAGTTGCTG
W12	rs12038857	ACGTTGGATGTGTAGGCTTTAAAGTCAGGC	ACGTTGGATGCAGATAGGTTAAATTACCAG	97	aTGAAATCCAGGCTCTCT
W12	rs6533531	ACGTTGGATGAGAGTTCCAACATGACAGGG	ACGTTGGATGCTTTGAGCAAAACTGAGTCC	86	ACATGACAGGGAAACCAT
W12	rs6838973	ACGTTGGATGAGGAATGAAAGTGGGCTGAG	ACGTTGGATGTCTCTTCCGATTGGAAACAG	102	GGAGGAATCTGTGGAGTA
W12	rs3866823	ACGTTGGATGGGCCTGATCTTTAGACCTTC	ACGTTGGATGTACTTCGTGTTTGGAATGGC	101	aAAAAATTGGCCTCAACCT
W12	rs13126975	ACGTTGGATGGATCTAATAGTGGAGGGTGG	ACGTTGGATGATTCACAGTGGCCCTATATG	112	tcGTGGAGGGTGGATAGGG

Multiplex	SNP	Forward primer	Reverse primer	Amplicon length	Extension primer
W12	rs4033102	ACGTTGGATGCTATGACTTCTGCTAGTCTC	ACGTTGGATGAGGAAAAACACCATGTTGAG	86	TCTGCTAGTCTCAGGGAATA
W12	rs4656799	ACGTTGGATGGGGAGACCAGTTTTACAGAG	ACGTTGGATGAGCCATCACCTGTTCTTCAC	99	TGGGACTTGAATTTTATCAG
W12	rs17513835	ACGTTGGATGGGATAAAAACCGAGAGTAAG	ACGTTGGATGAGGAAGCTCAAAGGCAAATC	116	ggggTGTGCCATGTAGTCTTG
W12	rs1947187	ACGTTGGATGTCACCTTTGGAGACATCAAC	ACGTTGGATGGAATATGTTTGGATTTCTTTC	109	TCTTTATAGCACTGTTTCCAAG
W12	rs12440104	ACGTTGGATGTCTCTCCTTTCCTCGTCGG	ACGTTGGATGTATTCTTCACGATGGCCTCC	94	ctttCTCGTCGGAGGCTTATCT
W12	rs8030574	ACGTTGGATGACATGGCATAGACATTTGGG	ACGTTGGATGAGTGGGGCCCTTCTTTCCT	87	gtttGGTCTTTCTAAGGGTACT
W12	rs10222783	ACGTTGGATGCTAGGACTTGGGTAGGAAAT	ACGTTGGATGCCTGCCACTGTGTTTTCCTT	119	ggagATATGAAATGAGCCAGGA
W12	rs11857639	ACGTTGGATGATAAGGCCGGATGCCAAGAG	ACGTTGGATGAGCTGGCATATGGGTACTGG	111	ctcttCTCTCTACCCCATGACCCA
W12	rs8040516	ACGTTGGATGGTTTGAGTCTTTTCCTGGGC	ACGTTGGATGGCTATTGTCAGAAACTGCCC	119	tttgTTCTCACTCCTTTCTGAATG
W12	rs2200733	ACGTTGGATGCCCCAAACTTTCTGGAAAAT	ACGTTGGATGTGGTGGTACTTGGGTTTTG	119	agATTACCTGTTCTAATTTTCTCT
W12	rs3853444	ACGTTGGATGGAGAGCTGACTTGTAATATG	ACGTTGGATGCCGGCAACTGTGAAATCATA	120	GAGCTGACTTGTAATATGTATATA
W12	rs16997168	ACGTTGGATGCTTGGGCCTTCTGGTGATTA	ACGTTGGATGAGCTCATAAAACTCTCCCTC	86	TCTGGTGATTATTTCTTCAACTCAG
W12	rs529004	ACGTTGGATGTGACTGCTGGACCCCAGAG	ACGTTGGATGAAGGAAGGGCCCAGCTCATA	109	ccccAGCCAGTGCGCTCCAAACTGCC
W12	rs7080456	ACGTTGGATGGCACATGTGTTTGTGCGTGA	ACGTTGGATGACACCCATACGCACAAAACC	113	ccctcTGTGCGTGATTATTTGTCTCA
W12	rs3866836	ACGTTGGATGCAGCCGCGGCATTAGATTTT	ACGTTGGATGCAGCAGAGACCCCTCACAT	91	CGGCATTAGATTTTCATAGGAGCATA
W12	rs12129225	ACGTTGGATGCCTGGGGAGTGTATAATAG	ACGTTGGATGAGACTGAGCATTGACATCCC	86	gctgtGGAGTGTATAATAGTCTCTCA
W12	rs12038255	ACGTTGGATGATCACTATGTCATTTTCAC	ACGTTGGATGAAGGCCCTAGGCAGTACAAT	116	cTTTCTTTCTTTCCAATTTTTTAATGA
W12	rs35177144	ACGTTGGATGCCTGGTTATTTTCTGCTGTC	ACGTTGGATGGGTTTAACTGTGATTAGGAG	116	ggagGTCTTTTGTTTTTCTGGTGTGTG
W12	rs659580	ACGTTGGATGTCCGCGGCGGACTCTCTTTA	ACGTTGGATGCACCCCGCTAAGACCTTTTC	117	gtcctCGGCGGACTCTCTTTACTTCGGC
W12	rs1448799	ACGTTGGATGTCTTAAGCCCATCAGAGCTA	ACGTTGGATGTTCAGAAGGGTTTATTGCTC	110	tttcgTAAGCCCATCAGAGCTAAAATCA
W12	rs570881	ACGTTGGATGGCTTCAAAGTAATCATCTG	ACGTTGGATGCCTTTAAACAGATCTTAGTG	86	AATCATCTGTAAAATAAAAAGTTTGAAA
W13	rs4745718	ACGTTGGATGGCAGGGCCACAGGCTTGAA	ACGTTGGATGAAGGTGGAGGAAGATCATGC	109	AGGCTTGAAAGCCCC
W13	rs2680344	ACGTTGGATGGGCAGTACTTCGATCAGAG	ACGTTGGATGACTATGCTGGCCAAACAGAG	104	TGCACCAAATACAGCA
W13	rs581241	ACGTTGGATGCCTAAGTTTATAACAGGGTG	ACGTTGGATGTGGCCATTCATGCATCACAC	88	AACAGGGTGAGCAAAT
W13	rs11638230	ACGTTGGATGTCACTGCCATTACTCACCTC	ACGTTGGATGAAGGCTCATAACCCACACAG	94	CTCCCCAACCCTGACTG

Multiplex	SNP	Forward primer	Reverse primer	Amplicon length	Extension primer
W13	rs2623997	ACGTTGGATGGAAGGTTGCCTTCATCAGAG	ACGTTGGATGAGGATGCAACTGTGTTGCCG	97	gtATTGCCGGTCTCCAC
W13	rs619456	ACGTTGGATGCCACAGGTAAGTTGGACAGC	ACGTTGGATGGATTTCAATTAAGGTAAGCAG	112	GCAGAAGGCACATACTT
W13	rs2739200	ACGTTGGATGGCGACGAGAAACGAGGCAG	ACGTTGGATGTCCCCTCTCCTTTCGCTCT	108	AGGGAAGCAGATGCCAGC
W13	rs6843082	ACGTTGGATGTTCTGGTGTCCTGGGATTTG	ACGTTGGATGACACTGTTGGTGATGAGTGG	119	catcGCCTCAGAGCTGAAA
W13	rs976568	ACGTTGGATGGTGTGAAGGTGTGTGTCGC	ACGTTGGATGTGTTCGACCCAAGTCGCTG	120	gggCAGAGGGAGGAAGTCT
W13	rs7552922	ACGTTGGATGAGGTCAGGGCATGATCTACT	ACGTTGGATGTTTGGCCACATGAAGTTCTG	106	GGGCATGATCTACTGCAAAA
W13	rs58473244	ACGTTGGATGTGTCTCCATCTTTGTTGGTC	ACGTTGGATGTCCAGTGAGCAGATGTTGTG	103	ctgTGTTGGTCATGGTAAGG
W13	rs7679158	ACGTTGGATGCACCTGGACTTTGGTAATCG	ACGTTGGATGGGAGATTAACCTGACAGTAG	103	ggtagCTTCCTGCCCCAAACA
W13	rs17570669	ACGTTGGATGCTTTGTGATCCAGTGTAGGC	ACGTTGGATGAGGGTAAAGGGTCTTTACTC	105	GATAATGTTCCTAGAAGTTGT
W13	rs17042171	ACGTTGGATGCCCAAGGCAACATGAGGAAT	ACGTTGGATGTAGACAGGGTGCCTGAGGAC	119	AATTAAATTACCTCACTGGAAA
W13	rs513287	ACGTTGGATGTGATGGTGGAGATGGTTTGC	ACGTTGGATGTCATGGGGTATGGTTGTTGC	114	cccctATGGTTTGCTAAATGTCC
W13	rs6817202	ACGTTGGATGCACAAATCTGTAAACTGGGC	ACGTTGGATGCAATTCAACCCAGACAAGCC	107	ctcgGTAAACTGGGCAGTTATTT
W13	rs561873	ACGTTGGATGTCTTCGGTATCCTCTGTGAC	ACGTTGGATGGGGGAAATTGATAGAGTAG	120	GCAAAATTCTGAGTAGTGTATAT
W13	rs7535322	ACGTTGGATGGCCTTTGAGAGAAGTCTCTT	ACGTTGGATGGGAGCTCACTTCTTATGTAA	120	gggTGAGAGAAGTCTCTTGAAATA
W13	rs10033464	ACGTTGGATGCTGAGGAATTCTAAATGAC	ACGTTGGATGAACTCAGAGCTTGATGAAA	109	TTTTTACATTGTTAGAGTCAAGAAA
W14	rs12905211	ACGTTGGATGAAAGAGTTGCAAAGAGCCCC	ACGTTGGATGAGGAAAGAAGGGCCCCACTG	103	GTGGGGAGAGGAGCC
W14	rs2421494	ACGTTGGATGGCAAGATCCCAGCAGATTGT	ACGTTGGATGTTCCCAATCCCTTCCTGCTG	109	AGCAGATTGTAATTTACAGC
W14	rs7540713	ACGTTGGATGGAGAAGGAACAACTTCTTGC	ACGTTGGATGATGACAGTAAGTCTCAGCAG	107	TTACTTGTTTGAGTTCTCTAC
W14	rs3903239	ACGTTGGATGAAGATGAAAACAATGACCC	ACGTTGGATGGTTCACTACCAATCTTTTCC	108	AAATACAATGGACAGGATTAA

Table A1.3 Allelic expression ratio assays

Multiplex	Transcribed SNP	Froward primer	Reverse primer	Amplicon length	Extension Primer
W1	rs10852515	ACGTTGGATGTAATCACACACCTCGCACCG	ACGTTGGATGAATATCAGTAGCTCCTGCGG	100	TTGGTTTTTGGTTTGGT
W1	rs1051614	ACGTTGGATGGATGACTGGCGGATAGCCAT	ACGTTGGATGATGGCGCACACCAGCATCTC	91	gaaCTACGAGCGCATCCT
W1	rs740178	ACGTTGGATGACCTCCTGGCAATGGGAGAC	ACGTTGGATGTCCTCCTTGTCTTCCTCAAC	90	ATGGGAGACCCCACTCTGGC
W2	rs3743496	ACGTTGGATGCCGTTAGCTGTAACTTGGAG	ACGTTGGATGCTTTTCTCCCAAGGTCGCAG	119	ctttGAGCTCCACTCTGCC
W2	rs41280400	ACGTTGGATGTCCAGATGCCCAAGTTTGAC	ACGTTGGATGGGGTTCGATTGAAAGAAGGC	116	aggtGGCCCTTGCTGAGTGAACCCCT
W2	rs4746139	ACGTTGGATGTGACCTCCACCGCTTCTATC	ACGTTGGATGCTCAGGAGCCCCTGGGCCT	109	CTTCCTATCTGCGCC
W2	rs60632610	ACGTTGGATGGGGACACTGTTCAAACTCCG	ACGTTGGATGTCCCCCTGATAGTGTGACCA	105	cCCAAGCTGAGCGATG
W2	rs6533526	ACGTTGGATGAAGCATGTTATACAGAAGCG	ACGTTGGATGACATACAGTGTGGCATTTAG	98	TACAGAAGCGATTAGGATTTTT
W2	rs7164883	ACGTTGGATGCTAAACCACAGATCAACCCC	ACGTTGGATGGTGGTATCTTTCTCTGATCC	116	ctatCATGAATAGATCACGTGACC
W2	rs913257	ACGTTGGATGCTCTAAACGAACCACTGGTC	ACGTTGGATGGTAGAAGCCGATGAAGAGAC	119	CCTTGATTCTACTTCTTGTT
W3	rs3812629	ACGTTGGATGTTACAAGACCCTGCCTCACG	ACGTTGGATGGAGTCTTAGGAGTCATAGGG	103	CACCTAAGACCCCCC
W3	rs3820416	ACGTTGGATGGAAGAGTTCCCCCATTCATC	ACGTTGGATGTTGGGATGTACAGTTCCTTC	97	TGTCCCATTAGTTGCTG
W3	rs529004	ACGTTGGATGTGACTGCTGGACCCCAGAG	ACGTTGGATGAAGGAAGGGCCCAGCTCATA	109	ccccAGCCAGTGCGCTCCAAACTGCC
W3	rs75911264	ACGTTGGATGAAGCATGTTATACAGAAGCG	ACGTTGGATGACATACAGTGTGGCATTTAG	98	cCTTGCGAGCAAGGGA
W4	rs2739200	ACGTTGGATGGCGACGAGAAACGAGGCAG	ACGTTGGATGTCCCCTCTCCTTTCGCTCT	108	AGGGAAGCAGATGCCAGC
W4	rs58473244	ACGTTGGATGTGTCTCCATCTTTGTTGGTC	ACGTTGGATGTCCAGTGAGCAGATGTTGTG	103	ctgTGTTGGTCATGGTAAGG
W4	rs7552922	ACGTTGGATGAGGTCAGGGCATGATCTACT	ACGTTGGATGTTTGGCCACATGAAGTTCTG	106	GGGCATGATCTACTGCAAAA
W5	rs4657	ACGTTGGATGATGTGGCTGCTAAAGCCATC	ACGTTGGATGATCTGCCTTTTGTGTGTGCC	118	CACAGCCCTGTTCAC
W5	rs76455830	ACGTTGGATGGCACCTCCTTAAAGCGATAC	ACGTTGGATGGTGATTTATTATGACTTCTC	115	ACGTTTTCCTCAGCAG
W5	rs4986938	ACGTTGGATGTGAGGTGAACTGGCCCACA	ACGTTGGATGACAGCAGAAAGATGAAGCCC	102	GGCCCACAGAGGTCACA
W5	rs7229	ACGTTGGATGATGCACAAATAACCACATCC	ACGTTGGATGTTCAGTACCTAGCTCTGCTC	100	AACCACATCCATGCAATA
W5	rs8940	ACGTTGGATGAACATCTGTCACACTCTTCC	ACGTTGGATGGCCTTTTGTAAAGACCTGCC	90	CACACTCTTCCATATTGTCT
W5	rs928554	ACGTTGGATGGCAGTGACCCTCTAATCAAC	ACGTTGGATGTCATGGATTACAATGATCCC	112	ATCTTGGGTAACATTTTCACTTCA
W5	rs1049334	ACGTTGGATGGGAACACAATGTTGAGCCAC	ACGTTGGATGAATCACGCTTTCCTGAATCC	87	gtttAGCCACTAAACCACCC

Multiplex	Transcribed SNP	Froward primer	Reverse primer	Amplicon length	Extension Primer
W5	rs1049337	ACGTTGGATGCAGGAGCTTTGGACCTAATC	ACGTTGGATGTTTATTACTGCCTCCTCCCC	89	GGACCTAATCCAAGCATC
W5	rs1152582	ACGTTGGATGAATGAGCAGTGGTGTCCATC	ACGTTGGATGATCAGCTAAGGGTGCCTATG	115	ATATATTATAGAAGCAAGCGAG
W5	rs1256049	ACGTTGGATGCTTGCTTTCCCCAGGCTTTG	ACGTTGGATGTAACACCTCCATCCAACAGC	99	ccccGCCTGTTCGACCAAGT
W5	rs4657	ACGTTGGATGATCTGCCTTTTGTGTGTGCC	ACGTTGGATGATGTGGCTGCTAAAGCCATC	118	tggtAAAGAGAAGCACCAAGA
W5	rs7161192	ACGTTGGATGCAGCTCATGGATCAGACTTC	ACGTTGGATGTGTTTCGCTTCCTCACTGAC	118	ATCAGACTTCTGGTTTGGTA
W5	rs9920	ACGTTGGATGAGCAACTCGCTTTAGGTCAG	ACGTTGGATGAGCTCAGAGTCAGAAACACG	110	ggaCCTGAAGACCAAAATTAGAATA

Appendix 2

Associations between genotype and gene expression

A2 Associations between genotype and gene expression

Table A2.1 Associations at chromosome 1q21 in the NE cohort in blood

SNP	Importance	Position	*KCNN3* aerQTL		*KCNN3* eQTL	
			Beta	p value	Beta	p value
rs1995662		154679567	-0.02	8.66E-01	-0.09	2.02E-01
rs6695232		154681950	0.11	1.78E-01	0.10	1.59E-02
rs10128027		154691445	-0.01	8.65E-01	0.02	7.20E-01
rs4845663		154692088	-0.02	7.55E-01	-0.03	5.44E-01
rs7523074		154692269	0.06	6.17E-01	-0.10	8.30E-02
rs7554577		154692883	-0.03	7.28E-01	0.12	5.58E-03
rs11264249		154693181	-0.04	6.46E-01	-0.06	1.54E-01
rs11264251		154693538	0.01	9.23E-01	-0.01	7.65E-01
rs12087736		154695678	0.09	5.13E-01	-0.11	9.97E-02
rs10494301		154701218	-0.14	1.01E-01	-0.04	3.44E-01
rs12402320		154701406	0.08	3.62E-01	-0.01	8.13E-01
rs884664		154703725	0.07	4.03E-01	0.04	4.20E-01
rs6426905		154706620	-0.02	8.93E-01	0.03	7.40E-01
rs960182		154710756	-0.02	7.85E-01	0.01	8.05E-01
rs10908429		154710904	0.05	5.41E-01	0.07	1.30E-01
rs6682000		154711226	-0.03	7.50E-01	0.07	1.11E-01
rs12145386		154713499	0.00	9.99E-01	0.02	7.74E-01
rs4999127		154714006	-0.01	9.01E-01	-0.09	1.30E-01

SNP	Importance	Position	*KCNN3* aerQTL		*KCNN3* eQTL	
			Beta	p value	Beta	p value
rs868108		154720289	-0.13	1.65E-01	0.10	2.94E-02
rs6698628		154721780	-0.19	4.78E-01	0.10	3.05E-01
rs906276		154724392	-0.09	4.09E-01	-0.09	1.43E-01
rs4845390		154726374	-0.13	6.99E-02	0.06	1.27E-01
rs10494300		154726546	0.15	5.98E-02	-0.07	1.04E-01
rs6682291		154732776	0.25	4.13E-02	0.02	8.15E-01
rs4845391		154739104	-0.06	4.54E-01	-0.07	9.68E-02
rs1051614	*KCNN3* Transcribed	154744807	0.13	2.17E-02	-0.10	2.52E-02
rs1131820		154744852	0.17	1.88E-02	-0.10	3.80E-02
rs7532286		154750816	0.00	9.51E-01	**0.15**	**4.21E-04**
rs1506982		154753697	0.20	1.76E-01	0.03	7.29E-01
rs4845671		154755126	-0.05	6.01E-01	-0.07	1.21E-01
rs12135902		154756104	0.01	9.71E-01	-0.19	1.65E-02
rs11264261		154756443	0.15	4.13E-02	-0.08	8.46E-02
rs883319		154758422	-0.10	3.25E-01	-0.10	5.32E-02
rs11584635		154760965	-0.22	9.95E-03	0.10	3.05E-02
rs17719287		154766495	0.14	3.13E-01	0.09	2.19E-01
rs1419110		154768689	-0.09	2.26E-01	0.00	9.48E-01
rs4492614		154769561	0.13	1.54E-01	-0.07	1.90E-01
rs17719419		154771916	0.04	7.90E-01	-0.23	1.10E-03
rs6690388		154774447	-0.23	1.44E-02	0.12	1.77E-02
rs6426944		154776412	-0.19	4.64E-02	0.04	3.55E-01
rs2256209		154779544	0.20	2.01E-01	-0.10	2.04E-01
rs2798601		154779685	-0.33	1.69E-03	0.11	5.06E-02

SNP	Importance	Position	*KCNN3* aerQTL		*KCNN3* eQTL	
			Beta	p value	Beta	p value
rs6688473		154783773	0.15	1.74E-01	-0.12	9.61E-03
rs2335406		154786686	0.00	9.58E-01	0.00	9.97E-01
rs1106236		154787136	-0.03	6.95E-01	**0.15**	**4.76E-04**
rs1106235		154787213	-0.06	5.28E-01	0.09	9.18E-02
rs1218601		154791676	-0.04	6.41E-01	0.11	2.23E-02
rs951241		154793798	0.03	7.83E-01	-0.14	1.32E-02
rs12029542		154794897	0.06	5.90E-01	-0.07	1.43E-01
rs11264268		154796520	0.09	3.55E-01	0.09	7.08E-02
rs6681725		154797062	-0.08	5.12E-01	-0.12	5.58E-02
rs1218551		154801173	0.02	8.39E-01	-0.11	1.46E-02
rs11264270		154802379	-0.02	9.07E-01	0.12	7.13E-02
rs4845394		154808287	-0.04	6.10E-01	0.03	5.12E-01
rs1218578		154810030	-0.10	1.90E-01	0.02	6.32E-01
rs11264274		154811127	0.06	4.78E-01	-0.06	1.86E-01
rs12128882		154811435	-0.02	8.22E-01	0.00	9.96E-01
rs1218574		154811677	-0.08	4.51E-01	0.11	7.08E-02
rs12740456		154814197	0.08	4.62E-01	-0.14	1.82E-02
rs13376333	AF GWAS	154814353	0.10	2.89E-01	**-0.17**	**2.11E-04**
rs6426987		154815257	0.08	3.85E-01	**-0.15**	**5.32E-04**
rs1218565		154817687	-0.05	5.18E-01	0.10	1.68E-02
rs4845396		154828409	-0.08	3.66E-01	**0.16**	**9.74E-05**
rs1061122		154832290	0.18	7.04E-02	-0.10	4.04E-02
rs76110919		154842095	-	-	-	-
rs2335407		154843347	-0.14	7.80E-02	**0.21**	**3.12E-07**

Results of association testing at the 1q21 locus. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. The significance threshold after correction for multiple testing at this locus was 7.57e-04. Significant associations are **in bold**.

Table A2.2 Associations at chromosome 7q31 in blood

SNP	Importance	Position	*CAV1* aerQTL		*CAV1* eQTL		*CAV2* aerQTL	
			Beta	p value	Beta	p value	Beta	p value
rs17138749		116133098	**0.10**	**2.42E-03**	-0.07	3.71E-01	**-0.43**	**1.60E-08**
rs8940	*CAV2* Transcribed	116146074	0.03	4.39E-01	-0.07	4.21E-01	**0.14**	**1.90E-04**
rs6466578		116150869	0.04	3.17E-01	0.08	3.39E-01	0.12	3.43E-01
rs3919515		116151784	**0.07**	**8.18E-04**	-0.03	6.16E-01	0.11	1.86E-02
rs17588172		116154015	**0.08**	**2.70E-03**	-0.08	1.99E-01	0.06	2.08E-01
rs4236601	Glaucoma GWAS	116162729	0.00	1.00E+00	-0.05	5.37E-01	0.12	3.14E-03
rs4730748		116167595	0.03	4.39E-01	-0.09	3.30E-01	**0.13**	**3.13E-04**
rs959173		116182054	0.05	2.06E-01	0.02	7.99E-01	0.12	2.47E-01
rs3807989	AF GWAS	116186241	**0.10**	**4.15E-05**	-0.03	6.03E-01	0.04	1.00E+00
rs11773845	AF, PR interval GWAS	116191301	**0.10**	**4.15E-05**	-0.03	6.42E-01	0.04	1.00E+00
rs1049314		116199695	0.01	1.00E+00	-0.05	5.60E-01	0.13	2.26E-02
rs8713		116199797	0.02	5.27E-01	-0.02	7.81E-01	0.11	4.18E-02
rs9920		116200092	-0.01	1.00E+00	-0.02	8.66E-01	-0.13	1.91E-01
rs1049334	*CAV1* Transcribed	116200380	**0.23**	**7.50E-09**	-0.04	6.85E-01	**-0.56**	**1.17E-07**
rs1049337	*CAV1* Transcribed	116200587	**-0.08**	**3.73E-06**	0.04	5.21E-01	-0.08	3.43E-01
rs1860588		116206486	0.07	1.21E-01	0.04	6.45E-01	0.12	4.30E-01

The significance threshold after correction for multiple testing was 3.12E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. The major alleles of the SNPs rs3807989 and rs11773845 are associated with increased AF risk. Significant associations are **in bold**.

Table A2.3 Associations at chromosome 9q22 in blood

SNP	Importance	Position	*C9orf3* aerQTL		*C9orf3* eQTL		*FANCC* aerQTL	
			Beta	p value	Beta	p value	Beta	p value
rs7848139		97495467	-0.04	7.68E-02	-0.04	1.13E-01	0.02	1.87E-01
rs356122		97565461	0.09	2.94E-03	0.10	6.05E-03	0.03	3.90E-01
rs4744381		97601089						
rs1530331		97620602	0.00	9.26E-01	-0.05	3.02E-01	0.08	5.70E-02
rs10993391		97675840	**0.24**	**4.14E-07**	**0.21**	**3.57E-06**	-0.02	6.90E-01
rs7853195		97685298	-0.01	5.82E-01	-0.01	6.67E-01	-0.03	5.13E-02
rs76455830		97695714			-0.04	9.13E-01		
rs10821415	AF GWAS	97713459	-0.02	5.08E-01	0.00	8.78E-01	-0.03	5.13E-02
rs10993413		97729136	-0.06	1.92E-01	-0.01	8.33E-01	0.05	1.43E-01
rs3802458		97741274	0.01	8.75E-01	-0.22	1.81E-02	0.03	6.53E-01
rs3802457	PCOS GWAS	97741336	0.03	5.61E-01	-0.07	3.17E-01	0.02	6.91E-01
rs7048941		97809674	0.12	5.70E-04	0.08	3.52E-02	0.00	9.87E-01
rs4744437		97832830	0.00	9.68E-01	-0.01	7.51E-01	0.03	4.69E-01
rs7033633		97838781	0.00	8.15E-01	-0.02	5.01E-01	-0.02	4.40E-01
rs17679141		97842479	0.04	1.05E-01	0.07	1.35E-02	-0.05	5.65E-03
rs4657	*C9orf3* Transcribed	97849090	0.00	7.77E-01	-0.04	1.36E-01	-0.01	7.43E-01
rs4647554	*FANCC* Transcribed	97862701	-0.04	8.12E-02	-0.02	4.69E-01	0.03	1.80E-02

The significance threshold after correction for multiple testing was 2.94E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold**.

Table A2.4 Associations at chromosome 14q23in blood

SNP	Importance	Position	*SYNE2* aerQTL		*SYNE2* eQTL		*ESR2* aerQTL	
			Beta	p value	Beta	p value	Beta	p value
rs4459477		64346190	0.07	3.20E-02	-0.01	1.06E-01	0.01	5.27E-01
rs17750888		64355420	-0.01	3.17E-01	0.01	1.17E-01	0.01	1.80E-01
rs12588807		64363070	0.08	1.57E-03	-0.03	3.62E-01	0.00	1.00E+00
rs2275018		64375985	-0.04	2.01E-02	0.03	5.87E-01	0.01	2.37E-01
rs7153036		64377548	-0.03	9.43E-02	0.03	6.49E-01	0.01	1.57E-01
rs12879919		64382350	0.04	1.07E-01	-0.02	3.50E-01	0.00	1.00E+00
rs11158521		64401255	-0.03	2.85E-02	0.04	9.64E-01	0.00	6.55E-01
rs2184292		64408135	-0.04	1.28E-02	0.03	7.07E-01	0.00	1.00E+00
rs7153680		64474227	0.00	1.00E+00	0.03	6.24E-01	0.00	1.00E+00
rs4566057		64506541	**-0.09**	**1.26E-03**	-0.02	2.68E-01	0.01	5.27E-01
rs1890908		64519035	-0.02	2.37E-01	0.01	1.02E-01	0.01	4.39E-01
rs10137972		64557734	-0.01	5.27E-01	0.01	9.09E-02	-0.02	2.06E-01
rs17101651		64561324	-0.08	3.20E-02	-0.01	9.90E-02	0.00	1.00E+00
rs8022428		64563613	-0.05	4.55E-02	-0.01	1.01E-01	0.00	1.00E+00
rs8007972		64564527	0.02	1.00E+00	-0.05	9.24E-01	0.01	5.27E-01
rs8016917		64569951	**-0.12**	**1.57E-03**	0.04	3.38E-01	-0.05	6.52E-03
rs10142318		64604281	**-0.11**	**1.19E-06**	-0.01	1.18E-01	0.00	1.00E+00
rs7144688		64613473	**-0.13**	**1.39E-12**	-0.02	3.54E-01	0.00	1.00E+00
rs11158533		64622988	**-0.09**	**2.73E-05**	0.03	6.20E-01	0.00	6.55E-01
rs17751454		64625181	-0.08	2.01E-02	0.07	1.60E+00	0.00	1.00E+00
rs3866743		64633322	**-0.18**	**1.83E-04**	0.04	3.80E-01	-0.04	8.33E-02
rs7161192	*SYNE2* Transcribed	64637147	**-0.10**	**1.19E-06**	0.04	8.10E-01	0.00	6.55E-01

SNP	Importance	Position	*SYNE2* aerQTL		*SYNE2* eQTL		*ESR2* aerQTL	
			Beta	p value	Beta	p value	Beta	p value
rs2256191		64656855	**-0.10**	**1.45E-11**	0.03	6.38E-01	0.01	5.27E-01
rs1152591	AF GWAS	64680848	**-0.11**	**7.52E-24**	0.00	5.72E-02	0.00	1.00E+00
rs7145919		64687601	-0.10	8.15E-03	-0.09	1.35E+00	-0.01	6.55E-01
rs8020646		64691320	**-0.14**	**5.13E-05**	0.01	7.68E-02	0.02	2.06E-01
rs12434245		64691853	-0.10	8.15E-03	-0.09	1.35E+00	-0.01	6.55E-01
rs1048315		64692465	**-0.14**	**1.80E-06**	-0.02	2.56E-01	0.02	2.37E-01
rs1152582	*SYNE2* Transcribed	64692630	**0.12**	**6.19E-37**	-0.01	1.29E-01	0.01	3.17E-01
rs7229		64692825	**0.12**	**2.83E-35**	-0.01	2.65E-01	0.00	4.39E-01
rs2772163		64693385	**-0.10**	**2.68E-09**	0.05	1.32E+00	0.00	1.00E+00
rs928554	*ESR2* Transcribed	64694195	**0.12**	**3.98E-32**	0.00	3.99E-02	0.00	5.27E-01
rs4986938	*ESR2* Transcribed	64699816	**-0.12**	**1.99E-20**	0.04	9.85E-01	-0.01	1.38E-01
rs1256064		64700739	**-0.12**	**4.15E-05**	-0.03	2.99E-01	0.02	1.57E-01
rs1256063		64702217	**0.10**	**4.15E-05**	-0.05	5.50E-01	-0.01	6.55E-01
rs10144225		64704994	**-0.14**	**1.18E-05**	0.01	4.90E-02	0.00	1.00E+00
rs1256049	*ESR2* Transcribed	64724051	-0.09	1.57E-01	-0.08	6.22E-01	0.01	5.27E-01

The significance threshold after correction for multiple testing was 1.35E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold**.

Table A2.5 Associations with expression at chromosome 16q22 in blood

SNP	Importance	Position	*ZFHX3* aerQTL		*ZFHX3* eQTL		*ZFHX3a* aerQTL		*ZFHX3a* eQTL	
			Beta	p value	Beta	p value	Beta	p value	Beta	p value
rs117951282		72820328	-	-	-	-	-	-	-	-
rs699444		72827758	0.03	5.22E-02	0.02	7.79E-01	0.15	2.67E-02	0.04	5.54E-01
rs740178	*ZFHX3* Transcribed	72832135	0.03	2.86E-02	-0.01	8.70E-01	0.18	5.49E-03	0.10	1.42E-01
rs12929452		72838680	-0.04	1.19E-01	-0.01	8.20E-01	-0.16	1.70E-03	-0.01	8.14E-01
rs2266943		72854307	0.03	2.96E-02	0.00	9.34E-01	0.17	7.39E-03	0.03	6.40E-01
rs8058014		72857544	0.03	4.67E-02	0.01	8.00E-01	0.19	3.53E-03	0.09	1.65E-01
rs1476646		72898348	-0.05	8.63E-03	-0.05	2.47E-01	-0.19	1.55E-04	-0.06	2.07E-01
rs4788668		72901874	0.03	1.06E-01	0.02	6.72E-01	0.14	2.63E-03	0.02	7.70E-01
rs4788671		72915085	-0.04	6.22E-02	0.01	7.96E-01	**-0.18**	**6.08E-04**	-0.04	3.46E-01
rs6499594		72917601	0.06	1.33E-02	0.10	1.22E-01	0.11	1.06E-01	0.16	2.31E-02
rs16971366		72919582	-0.02	4.58E-01	0.04	4.71E-01	**-0.21**	**6.49E-05**	0.03	6.36E-01
rs9940310		72925982	0.06	2.42E-02	0.13	7.95E-02	0.20	1.51E-02	0.20	1.88E-02
rs9925261		72927533	0.03	1.23E-01	0.06	1.45E-01	-0.13	8.54E-03	0.08	1.07E-01
rs16971384		72931085	0.02	2.58E-01	0.03	5.45E-01	**-0.17**	**2.08E-04**	0.05	3.69E-01
rs4788482		72937079	-0.06	6.34E-02	-0.07	3.22E-01	**-0.25**	**5.08E-08**	-0.09	2.33E-01
rs4788679		72956544	0.01	6.94E-01	-0.01	8.57E-01	-0.19	4.26E-05	0.03	6.64E-01
rs9936884		72960230	0.02	3.85E-01	-0.05	5.13E-01	-0.07	4.72E-01	0.01	8.89E-01
rs4788488		72960283	0.01	4.69E-01	-0.02	6.10E-01	**-0.21**	**8.73E-06**	0.02	6.62E-01
rs8055870		72972090	-0.02	3.81E-01	0.06	1.32E-01	0.13	1.95E-02	0.01	8.77E-01
rs6499600		72979374	0.01	8.13E-01	-0.06	1.55E-01	**-0.20**	**3.94E-05**	-0.02	7.40E-01
rs13336412		72981949	-0.01	6.99E-01	0.01	8.68E-01	**0.20**	**3.10E-04**	-0.02	6.61E-01
rs2228200		72984668	0.03	4.42E-01	0.10	2.42E-01	0.12	4.29E-01	0.07	4.82E-01

SNP	Importance	Position	*ZFHX3* aerQTL		*ZFHX3* eQTL		*ZFHX3a* aerQTL		*ZFHX3a* eQTL	
			Beta	p value	Beta	p value	Beta	p value	Beta	p value
rs2106258		72990553	-0.02	4.80E-01	-0.06	2.02E-01	**-0.26**	**3.61E-12**	-0.02	6.96E-01
rs2157786		72991286	-0.01	6.94E-01	-0.06	1.88E-01	**-0.24**	**2.70E-09**	-0.05	3.37E-01
rs10852515	*ZFHX3a* Transcribed	72991660	-0.04	1.77E-01	-0.10	1.13E-01	**-0.25**	**7.47E-15**	-0.15	3.68E-02
rs7193297		72993831	0.04	1.27E-01	0.03	4.04E-01	0.21	1.20E-03	-0.01	8.98E-01
rs7404992		72994419	-0.01	8.49E-01	-0.06	4.77E-01	-0.13	1.84E-01	-0.10	3.06E-01
rs756717		72996162	-0.02	4.61E-01	0.06	1.56E-01	0.06	2.60E-01	0.11	2.35E-02
rs4788683		72997747	0.01	4.60E-01	-0.01	9.03E-01	-0.05	3.47E-01	-0.03	4.93E-01
rs9921395		73001957	-0.04	7.91E-02	-0.03	5.41E-01	-0.09	9.19E-02	-0.07	1.30E-01
rs12445932		73004432	-0.03	2.31E-01	0.01	8.13E-01	-0.16	1.19E-02	-0.03	6.57E-01
rs7199343	Kawasaki GWAS	73009024	-0.02	2.71E-01	-0.06	1.54E-01	-0.02	6.91E-01	-0.12	1.40E-02
rs11075954		73012164	0.02	2.44E-01	0.05	2.66E-01	-0.09	6.86E-02	0.08	1.03E-01
rs2040508		73012685	-0.01	5.97E-01	-0.08	1.17E-01	-0.02	8.31E-01	-0.07	2.61E-01
rs16971456		73013036	-0.03	3.67E-01	-0.02	7.54E-01	0.18	3.22E-02	0.03	6.40E-01
rs9930445		73013482	-0.03	1.71E-01	-0.03	5.20E-01	0.08	1.90E-01	-0.02	6.60E-01
rs4788684		73013633	0.03	8.34E-02	0.09	2.82E-02	0.02	7.12E-01	0.10	5.11E-02
rs16971464		73016143	0.05	1.77E-01	0.15	6.07E-02	0.28	1.78E-02	0.17	8.25E-02
rs16971465		73017061	0.01	7.14E-01	0.05	4.23E-01	0.16	8.41E-02	0.10	2.19E-01
rs4788489		73017118	0.01	8.13E-01	0.00	9.89E-01	-0.05	3.29E-01	-0.01	8.38E-01
rs16971474		73019004	-0.02	5.46E-01	-0.03	5.23E-01	0.00	9.49E-01	0.00	9.50E-01
rs11640106		73020116	-0.02	4.14E-01	-0.02	6.61E-01	0.03	5.69E-01	-0.07	1.93E-01
rs1858800		73024276	0.02	3.18E-01	0.04	3.12E-01	-0.10	7.52E-02	0.02	6.34E-01
rs756720		73028921	0.00	9.95E-01	0.07	1.70E-01	0.10	1.39E-01	0.06	3.26E-01
rs7193343	AF GWAS	73029160	-0.02	4.84E-01	0.02	6.50E-01	-0.05	4.35E-01	0.02	7.64E-01
rs11075958		73033869	-0.01	6.23E-01	0.05	2.72E-01	0.13	1.25E-01	0.01	8.02E-01

SNP	Importance	Position	*ZFHX3* aerQTL		*ZFHX3* eQTL		*ZFHX3a* aerQTL		*ZFHX3a* eQTL	
			Beta	**p value**	**Beta**	**p value**	**Beta**	**p value**	**Beta**	**p value**
rs8056528		73036633	-0.01	7.83E-01	-0.05	2.87E-01	0.06	3.91E-01	-0.02	6.67E-01
rs719353		73042551	0.01	7.48E-01	-0.04	3.06E-01	0.00	9.49E-01	-0.02	6.46E-01
rs4788689		73049830	-0.03	2.62E-01	-0.10	1.02E-01	-0.03	7.29E-01	-0.11	1.20E-01
rs2106261	AF GWAS	73051620	-0.01	7.10E-01	-0.02	6.66E-01	0.01	8.85E-01	-0.02	8.04E-01
rs11863932		73053579	-0.05	2.94E-01	-0.02	8.37E-01	-0.13	8.88E-02	0.04	6.67E-01
rs1548373		73059861	0.02	3.68E-01	-0.01	7.39E-01	0.09	1.45E-01	0.01	9.20E-01
rs4788692		73065656	-0.06	7.31E-02	-0.09	2.08E-01	0.05	6.25E-01	-0.14	1.05E-01
rs12373097		73068515	-0.02	5.79E-01	0.07	1.66E-01	-0.01	8.46E-01	0.07	2.59E-01
rs8057081		73068977	-0.01	7.74E-01	-0.03	4.90E-01	0.00	9.86E-01	-0.01	8.64E-01
rs4788696		73070310	0.00	9.45E-01	-0.01	8.99E-01	-0.12	3.41E-02	0.08	2.27E-01
rs8060701		73073289	**0.17**	**1.29E-04**	-0.04	6.15E-01	0.09	3.91E-01	-0.04	6.69E-01
rs9940321		73073808	0.03	2.87E-01	-0.03	5.27E-01	0.02	7.24E-01	-0.06	2.33E-01
rs9940520		73074012	0.11	3.63E-03	-0.09	1.50E-01	0.20	2.89E-02	-0.07	3.21E-01
rs11641701		73079212	0.01	7.58E-01	-0.01	8.30E-01	-0.12	2.17E-02	-0.05	3.18E-01
rs7204751		73079683	0.05	1.24E-01	0.02	8.05E-01	-0.09	2.49E-01	0.02	7.90E-01
rs4788697		73087494	-0.02	2.94E-01	-0.05	2.98E-01	-0.10	6.68E-02	-0.09	8.47E-02
rs8052905		73097663	0.03	5.60E-01	-0.04	6.91E-01	0.14	2.83E-01	-0.17	1.56E-01
rs739414		73097956	0.00	9.88E-01	-0.01	8.04E-01	-0.10	8.17E-02	0.02	7.80E-01
rs8051826		73102456	-0.01	6.36E-01	-0.03	6.30E-01	0.11	2.99E-01	-0.12	1.68E-01

The significance threshold after correction for multiple testing was 7.69E-04. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold**.

Table A2.6 Comparison of aerQTLs between the NE and SA cohorts for *ZFHX3α* in blood

SNP	Importance	Position	NE Cohort		SA Cohort		Combined data set	
			Beta	p value	Beta	p value	Beta	p value
rs117951282		72820328	-	-	0.36	2.09E-01	0.36	1.82E-01
rs699444		72827758	0.15	2.67E-02	0.06	8.06E-01	0.14	4.00E-02
rs740178	Transcribed	72832135	0.18	5.49E-03	0.00	1.00E+00	0.13	2.50E-02
rs12929452		72838680	-0.16	1.70E-03	0.22	6.93E-02	-0.09	7.64E-02
rs2266943		72854307	0.17	7.39E-03	-0.02	8.28E-01	0.09	8.24E-02
rs8058014		72857544	0.19	3.53E-03	-0.02	8.73E-01	0.13	2.87E-02
rs1476646		72898348	**-0.19**	**1.55E-04**	0.24	2.07E-02	-0.09	6.78E-02
rs4788668		72901874	0.14	2.63E-03	0.01	9.04E-01	0.09	2.00E-02
rs4788671		72915085	-0.18	6.08E-04	0.21	2.33E-02	-0.06	1.81E-01
rs6499594		72917601	0.11	1.06E-01	-0.11	1.39E-01	0.00	9.63E-01
rs16971366		72919582	**-0.21**	**6.49E-05**	-0.03	8.17E-01	-0.17	8.36E-04
rs9940310		72925982	0.20	1.51E-02	-0.13	1.17E-01	0.02	7.70E-01
rs9925261		72927533	-0.13	8.54E-03	-0.18	3.84E-02	-0.14	1.00E-03
rs16971384		72931085	**-0.17**	**2.08E-04**	0.01	9.16E-01	-0.11	7.25E-03
rs4788482		72937079	**-0.25**	**5.08E-08**	-0.13	3.96E-01	**-0.23**	**7.26E-07**
rs4788679		72956544	**-0.19**	**4.26E-05**	-0.24	5.57E-02	**-0.20**	**1.59E-05**
rs9936884		72960230	-0.07	4.72E-01	0.04	5.51E-01	0.01	9.07E-01
rs4788488		72960283	**-0.21**	**8.73E-06**	0.01	8.43E-01	-0.13	2.04E-03
rs8055870		72972090	0.13	1.95E-02	0.18	8.18E-02	0.14	4.38E-03
rs6499600		72979374	**-0.20**	**3.94E-05**	-0.16	1.78E-02	**-0.18**	**4.88E-06**
rs13336412		72981949	**0.20**	**3.10E-04**	0.08	2.71E-01	0.15	9.51E-04
rs2228200		72984668	0.12	4.29E-01	0.42	1.35E-01	0.19	1.50E-01

SNP	Importance	Position	NE Cohort		SA Cohort		Combined data set	
			Beta	p value	Beta	p value	Beta	p value
rs2106258		72990553	**-0.26**	**3.61E-12**	**-0.34**	**1.36E-11**	**-0.29**	**7.09E-22**
rs2157786		72991286	**-0.24**	**2.70E-09**	**-0.34**	**1.25E-09**	**-0.27**	**5.41E-17**
rs10852515	Transcribed	72991660	**-0.25**	**7.47E-15**	**-0.34**	**1.73E-18**	**-0.29**	**7.06E-31**
rs7193297		72993831	0.21	1.20E-03	-0.18	1.47E-02	0.02	7.13E-01
rs7404992		72994419	-0.13	1.84E-01	-0.16	4.24E-02	-0.15	1.30E-02
rs756717		72996162	0.06	2.60E-01	0.08	3.47E-01	0.07	1.47E-01
rs4788683		72997747	-0.05	3.47E-01	-0.06	3.98E-01	-0.06	2.06E-01
rs9921395		73001957	-0.09	9.19E-02	0.19	1.24E-02	0.02	7.05E-01
rs12445932		73004432	-0.16	1.19E-02	0.13	2.92E-01	-0.08	1.60E-01
rs7199343	Kawasaki GWAS	73009024	-0.02	6.91E-01	0.34	1.63E-03	0.07	1.89E-01
rs11075954		73012164	-0.09	6.86E-02	-0.20	1.27E-03	**-0.14**	**3.70E-04**
rs2040508		73012685	-0.02	8.31E-01	-0.10	1.89E-01	-0.06	2.46E-01
rs16971456		73013036	0.18	3.22E-02	0.13	1.47E-01	0.15	1.21E-02
rs9930445		73013482	0.08	1.90E-01	0.17	7.62E-03	0.13	3.48E-03
rs4788684		73013633	0.02	7.12E-01	-0.20	8.60E-04	-0.08	3.68E-02
rs16971464		73016143	0.28	1.78E-02	-0.21	5.98E-01	0.23	5.43E-02
rs16971465		73017061	0.16	8.41E-02	0.00	9.91E-01	0.07	2.87E-01
rs4788489		73017118	-0.05	3.29E-01	-0.19	4.30E-03	-0.11	7.79E-03
rs16971474		73019004	0.00	9.49E-01	0.17	3.25E-02	0.08	1.16E-01
rs11640106		73020116	0.03	5.69E-01	0.26	2.84E-03	0.10	2.99E-02
rs1858800		73024276	-0.10	7.52E-02	-0.17	7.02E-02	-0.12	1.34E-02
rs756720		73028921	0.10	1.39E-01	0.20	3.97E-02	0.14	1.41E-02
rs7193343	AF GWAS	73029160	-0.05	4.35E-01	-0.04	6.29E-01	-0.05	3.70E-01
rs11075958		73033869	0.13	1.25E-01	0.48	2.63E-03	0.21	4.94E-03

SNP	Importance	Position	NE Cohort		SA Cohort		Combined data set	
			Beta	p value	Beta	p value	Beta	p value
rs8056528		73036633	0.06	3.91E-01	0.05	5.19E-01	0.05	2.88E-01
rs719353		73042551	0.00	9.49E-01	0.01	9.23E-01	0.00	1.00E+00
rs4788689		73049830	-0.03	7.29E-01	0.14	1.22E-01	0.05	3.72E-01
rs2106261	AF GWAS	73051620	0.01	8.85E-01	0.15	1.06E-01	0.07	2.31E-01
rs11863932		73053579	-0.13	8.88E-02	0.11	4.62E-01	-0.07	3.21E-01
rs1548373		73059861	0.09	1.45E-01	0.00	9.98E-01	0.05	3.00E-01
rs4788692		73065656	0.05	6.25E-01	-0.10	1.88E-01	-0.05	3.82E-01
rs12373097		73068515	-0.01	8.46E-01	-0.09	2.39E-01	-0.05	3.16E-01
rs8057081		73068977	0.00	9.86E-01	0.05	5.25E-01	0.02	6.58E-01
rs4788696		73070310	-0.12	3.41E-02	-0.25	1.68E-01	-0.13	1.76E-02
rs8060701		73073289	0.09	3.91E-01	-0.04	6.36E-01	0.00	9.52E-01
rs9940321		73073808	0.02	7.24E-01	0.04	5.94E-01	0.03	5.22E-01
rs9940520		73074012	0.20	2.89E-02	0.01	9.44E-01	0.08	1.90E-01
rs11641701		73079212	-0.12	2.17E-02	0.02	8.08E-01	-0.06	1.43E-01
rs7204751		73079683	-0.09	2.49E-01	-0.05	5.05E-01	-0.07	2.09E-01
rs4788697		73087494	-0.10	6.68E-02	-0.01	8.77E-01	-0.06	1.59E-01
rs8052905		73097663	0.14	2.83E-01	0.08	3.13E-01	0.09	1.51E-01
rs739414		73097956	-0.10	8.17E-02	-0.15	5.47E-02	-0.12	9.91E-03
rs8051826		73102456	0.11	2.99E-01	0.24	1.55E-02	0.18	9.97E-03

Effect sizes and p-values are shown for each of the NE and SA cohorts and for the combined data set. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. The significance threshold after correction for multiple testing was 7.69E-04. Significant associations are **in bold**.

Table A2.7 Comparison of aerQTLs between the NE and SA cohorts for *KCNN3* in blood

SNP	Importance	Position	NE Cohort		SA Cohort		Combined data set	
			Beta	p value	Beta	p value	Beta	p value
rs1995662		154679567	-0.02	8.66E-01	-0.04	7.42E-01	-0.02	7.40E-01
rs6695232		154681950	0.11	1.78E-01	0.02	7.91E-01	0.07	2.06E-01
rs10128027		154691445	-0.01	8.65E-01	0.08	3.43E-01	0.02	6.79E-01
rs4845663		154692088	-0.02	7.55E-01	0.06	4.96E-01	0.01	9.24E-01
rs7523074		154692269	0.06	6.17E-01	-0.16	1.13E-01	-0.05	5.38E-01
rs7554577		154692883	-0.03	7.28E-01	-0.13	9.63E-02	-0.07	2.04E-01
rs11264249		154693181	-0.04	6.46E-01	0.09	2.81E-01	0.01	8.12E-01
rs11264251		154693538	0.01	9.23E-01	0.07	3.83E-01	0.03	5.55E-01
rs12087736		154695678	0.09	5.13E-01	-0.12	3.42E-01	0.00	9.67E-01
rs10494301		154701218	-0.14	1.01E-01	0.03	8.17E-01	-0.09	1.62E-01
rs12402320		154701406	0.08	3.62E-01	-0.06	5.13E-01	0.02	7.19E-01
rs884664		154703725	0.07	4.03E-01	0.04	6.69E-01	0.06	3.51E-01
rs6426905		154706620	-0.02	8.93E-01	-0.27	1.32E-03	-0.19	1.39E-02
rs960182		154710756	-0.02	7.85E-01	-0.05	5.76E-01	-0.03	5.88E-01
rs10908429		154710904	0.05	5.41E-01	0.06	5.34E-01	0.05	3.88E-01
rs6682000		154711226	-0.03	7.50E-01	0.04	6.12E-01	0.00	9.63E-01
rs12145386		154713499	0.00	9.99E-01	-0.15	9.47E-02	-0.07	3.02E-01
rs4999127		154714006	-0.01	9.01E-01	0.03	8.18E-01	0.00	9.89E-01
rs868108		154720289	-0.13	1.65E-01	-0.07	4.44E-01	-0.10	1.17E-01
rs6698628		154721780	-0.19	4.78E-01	0.26	2.54E-02	0.17	1.52E-01
rs906276		154724392	-0.09	4.09E-01	-0.10	4.37E-01	-0.10	2.59E-01
rs4845390		154726374	-0.13	6.99E-02	0.01	9.25E-01	-0.07	1.57E-01
rs10494300		154726546	0.15	5.98E-02	0.06	5.51E-01	0.12	5.23E-02

SNP	Importance	Position	NE Cohort		SA Cohort		Combined data set	
			Beta	p value	Beta	p value	Beta	p value
rs6682291		154732776	0.25	4.13E-02	-0.04	7.04E-01	0.11	1.95E-01
rs4845391		154739104	-0.06	4.54E-01	-0.03	7.07E-01	-0.05	4.10E-01
rs1051614		154744807	0.13	2.17E-02	0.09	9.61E-02	0.11	4.60E-03
rs1131820		154744852	0.17	1.88E-02	0.15	2.31E-02	0.16	1.21E-03
rs7532286		154750816	0.00	9.51E-01	-0.14	6.59E-02	-0.06	2.46E-01
rs1506982		154753697	0.20	1.76E-01	-0.01	9.48E-01	0.12	2.67E-01
rs4845671		154755126	-0.05	6.01E-01	-0.03	6.49E-01	-0.04	4.93E-01
rs12135902		154756104	0.01	9.71E-01	-0.25	4.80E-01	-0.03	8.42E-01
rs11264261		154756443	0.15	4.13E-02	0.14	6.19E-02	0.15	5.86E-03
rs883319		154758422	-0.10	3.25E-01	-0.05	6.44E-01	-0.08	2.80E-01
rs11584635		154760965	-0.22	9.95E-03	-0.29	8.12E-03	**-0.24**	**2.94E-04**
rs17719287		154766495	0.14	3.13E-01	-0.13	5.65E-01	0.08	4.83E-01
rs1419110		154768689	-0.09	2.26E-01	-0.10	2.87E-01	-0.09	1.07E-01
rs4492614		154769561	0.13	1.54E-01	0.06	6.32E-01	0.11	1.33E-01
rs17719419		154771916	0.04	7.90E-01	-0.04	8.77E-01	0.02	8.53E-01
rs6690388		154774447	-0.23	1.44E-02	-0.10	3.20E-01	-0.18	9.58E-03
rs6426944		154776412	-0.19	4.64E-02	-0.12	1.36E-01	-0.16	1.39E-02
rs2256209		154779544	0.20	2.01E-01	-0.06	5.81E-01	0.05	6.13E-01
rs2798601		154779685	-0.33	1.69E-03	-0.19	5.99E-02	**-0.26**	**2.99E-04**
rs6688473		154783773	0.15	1.74E-01	0.01	8.89E-01	0.08	2.78E-01
rs2335406		154786686	0.00	9.58E-01	-0.05	6.07E-01	-0.02	7.42E-01
rs1106236		154787136	-0.03	6.95E-01	-0.04	5.53E-01	-0.04	5.03E-01
rs1106235		154787213	-0.06	5.28E-01	0.24	3.65E-02	0.04	5.62E-01
rs1218601		154791676	-0.04	6.41E-01	-0.04	6.86E-01	-0.04	5.37E-01
rs951241		154793798	0.03	7.83E-01	0.09	2.92E-01	0.06	3.68E-01

SNP	Importance	Position	NE Cohort		SA Cohort		Combined data set	
			Beta	p value	Beta	p value	Beta	p value
rs12029542		154794897	0.06	5.90E-01	0.14	9.25E-02	0.10	1.34E-01
rs11264268		154796520	0.09	3.55E-01	0.02	8.31E-01	0.06	3.73E-01
rs6681725		154797062	-0.08	5.12E-01	0.16	4.68E-02	0.07	3.13E-01
rs1218551		154801173	0.02	8.39E-01	-0.14	2.03E-01	-0.03	6.51E-01
rs11264270		154802379	-0.02	9.07E-01	-0.13	2.74E-01	-0.07	4.47E-01
rs4845394		154808287	-0.04	6.10E-01	-0.03	7.43E-01	-0.04	5.38E-01
rs1218578		154810030	-0.10	1.90E-01	-0.07	3.45E-01	-0.09	1.07E-01
rs11264274		154811127	0.06	4.78E-01	0.13	2.22E-01	0.08	2.10E-01
rs12128882		154811435	-0.02	8.22E-01	0.04	6.05E-01	0.01	8.80E-01
rs1218574		154811677	-0.08	4.51E-01	0.08	3.12E-01	0.01	8.47E-01
rs12740456		154814197	0.08	4.62E-01	0.07	7.57E-01	0.08	4.06E-01
rs13376333		154814353	0.10	2.89E-01	-0.04	5.89E-01	0.03	6.40E-01
rs6426987		154815257	0.08	3.85E-01	-0.01	8.86E-01	0.04	5.64E-01
rs1218565		154817687	-0.05	5.18E-01	-0.11	2.07E-01	-0.07	2.11E-01
rs4845396		154828409	-0.08	3.66E-01	-0.13	5.50E-02	-0.11	5.31E-02
rs1061122		154832290	0.18	7.04E-02	0.08	3.81E-01	0.13	4.91E-02
rs76110919		154842095						
rs2335407		154843347	-0.14	7.80E-02	-0.05	5.41E-01	-0.10	7.88E-02

Effect sizes and p-values are shown for each of the NE and SA cohorts and for the combined data set. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. The significance threshold after correction for multiple testing at this locus was 7.57e-04. Significant associations are **in bold.**

Table A2.8 Associations between SNPs at 1q21 and 14q23 and CAV1 expression in blood

SNP	Importance	Chromosome	Position	*CAV1* eQTL	
				Beta	p value
rs1995662		1	154679567	-0.05	5.01E-01
rs6695232		1	154681950	0.07	1.14E-01
rs10128027		1	154691445	-0.03	4.80E-01
rs4845663		1	154692088	0.05	2.77E-01
rs7523074		1	154692269	-0.03	6.41E-01
rs7554577		1	154692883	0.00	9.23E-01
rs11264249		1	154693181	-0.02	6.81E-01
rs11264251		1	154693538	0.03	5.51E-01
rs12087736		1	154695678	-0.08	2.59E-01
rs10494301		1	154701218	0.04	4.39E-01
rs12402320		1	154701406	-0.04	4.23E-01
rs884664		1	154703725	-0.05	3.31E-01
rs6426905		1	154706620	-0.03	7.43E-01
rs960182		1	154710756	0.00	9.29E-01
rs10908429		1	154710904	-0.04	4.13E-01
rs6682000		1	154711226	0.09	4.25E-02
rs12145386		1	154713499	-0.09	1.43E-01
rs4999127		1	154714006	-0.08	1.94E-01
rs868108		1	154720289	0.08	1.02E-01
rs6698628		1	154721780	-0.07	5.54E-01
rs906276		1	154724392	-0.06	3.83E-01
rs4845390		1	154726374	0.03	5.51E-01
rs10494300		1	154726546	0.03	5.21E-01

SNP	Importance	Chromosome	Position	*CAV1* eQTL	
				Beta	p value
rs6682291		1	154732776	0.07	4.00E-01
rs4845391		1	154739104	-0.04	4.17E-01
rs1051614	Transcribed	1	154744807	**-0.18**	**1.98E-04**
rs1131820		1	154744852	-0.07	1.93E-01
rs7532286		1	154750816	0.06	1.67E-01
rs1506982		1	154753697	0.16	5.28E-02
rs4845671		1	154755126	0.08	7.61E-02
rs12135902		1	154756104	0.03	7.45E-01
rs11264261		1	154756443	-0.12	1.92E-02
rs883319		1	154758422	0.01	8.72E-01
rs11584635		1	154760965	0.05	3.12E-01
rs17719287		1	154766495	-0.07	4.27E-01
rs1419110		1	154768689	-0.02	6.34E-01
rs4492614		1	154769561	-0.12	4.62E-02
rs17719419		1	154771916	-0.02	8.07E-01
rs6690388		1	154774447	-0.03	5.46E-01
rs6426944		1	154776412	-0.04	3.64E-01
rs2256209		1	154779544	-0.08	2.96E-01
rs2798601		1	154779685	-0.02	7.80E-01
rs6688473		1	154783773	-0.03	5.68E-01
rs2335406		1	154786686	0.01	8.16E-01
rs1106236		1	154787136	-0.01	7.48E-01
rs1106235		1	154787213	-0.07	2.57E-01
rs1218601		1	154791676	-0.01	8.94E-01
rs951241		1	154793798	0.05	4.38E-01

SNP	Importance	Chromosome	Position	*CAV1* eQTL	
				Beta	p value
rs12029542		1	154794897	-0.01	8.90E-01
rs11264268		1	154796520	0.03	5.04E-01
rs6681725		1	154797062	0.03	6.27E-01
rs1218551		1	154801173	-0.04	4.44E-01
rs11264270		1	154802379	0.01	8.70E-01
rs4845394		1	154808287	-0.01	7.39E-01
rs1218578		1	154810030	0.04	4.13E-01
rs11264274		1	154811127	-0.04	3.72E-01
rs12128882		1	154811435	-0.06	2.16E-01
rs1218574		1	154811677	-0.08	2.42E-01
rs12740456		1	154814197	0.03	6.81E-01
rs13376333	AF GWAS	1	154814353	0.02	7.51E-01
rs6426987		1	154815257	0.01	8.67E-01
rs1218565		1	154817687	0.02	6.25E-01
rs4845396		1	154828409	0.02	7.15E-01
rs1061122		1	154832290	-0.02	7.28E-01
rs76110919		1	154842095	NA	
rs2335407		1	154843347	0.06	1.90E-01
rs4459477		14	64346190	-0.04	6.72E-01
rs17750888		14	64355420	-0.09	1.50E-01
rs12588807		14	64363070	-0.01	8.88E-01
rs2275018		14	64375985	-0.06	3.99E-01
rs7153036		14	64377548	0.00	9.59E-01
rs12879919		14	64382350	-0.08	3.59E-01
rs11158521		14	64401255	-0.12	6.10E-02

SNP	Importance	Chromosome	Position	*CAV1* eQTL	
				Beta	p value
rs2184292		14	64408135	-0.09	1.91E-01
rs7153680		14	64474227	-0.11	9.21E-02
rs4566057		14	64506541	-0.11	2.57E-01
rs1890908		14	64519035	0.08	5.22E-01
rs10137972		14	64557734	-0.08	5.29E-01
rs17101651		14	64561324	-0.17	1.70E-01
rs8022428		14	64563613	0.01	8.52E-01
rs8007972		14	64564527	0.13	1.43E-01
rs8016917		14	64569951	-0.12	3.41E-01
rs10142318		14	64604281	-0.13	1.74E-01
rs7144688		14	64613473	0.02	7.77E-01
rs11158533		14	64622988	0.22	9.35E-04
rs17751454		14	64625181	0.14	7.29E-02
rs3866743		14	64633322	-0.09	4.66E-01
rs7161192		14	64637147	0.19	4.23E-03
rs2256191		14	64656855	0.10	1.39E-01
rs1152591	AF GWAS	14	64680848	0.20	1.26E-03
rs7145919		14	64687601	0.31	4.89E-03
rs8020646		14	64691320	-0.15	1.73E-01
rs12434245		14	64691853	0.31	4.89E-03
rs1048315		14	64692465	-0.07	4.49E-01
rs1152582	Transcribed	14	64692630	-0.11	5.27E-02
rs7229		14	64692825	-0.09	1.20E-01
rs2772163		14	64693385	0.12	9.24E-02
rs928554		14	64694195	-0.14	1.75E-02

SNP	Importance	Chromosome	Position	*CAV1* eQTL	
				Beta	**p value**
rs4986938		14	64699816	0.07	3.12E-01
rs1256064		14	64700739	-0.05	5.96E-01
rs1256063		14	64702217	0.09	4.23E-01
rs10144225		14	64704994	-0.15	1.53E-01
rs1256049		14	64724051	0.16	3.85E-01

Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. The significance threshold after correction for multiple testing for trans-acting associations was 2.48E-04. Significant associations are **in bold.**

Table A2.9 Associations with total expression of *KCNN3* in chromosome band 1q21 in tissue

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs1995662		154679567	-0.02	8.07E-01	0.21	2.46E-01	-0.37	2.37E-02
rs6695232		154681950	-0.02	6.26E-01	-0.10	4.31E-01	0.19	4.97E-02
rs10128027		154691445	-0.02	6.59E-01	0.11	3.17E-01	0.09	3.58E-01
rs4845663		154692088	0.03	5.96E-01	-0.08	4.83E-01	0.16	1.04E-01
rs7554577		154692883	0.02	6.41E-01	-0.07	4.05E-01	0.10	2.95E-01
rs11264249		154693181	0.04	4.31E-01	-0.07	4.28E-01	0.06	5.29E-01
rs11264251		154693538	-0.01	8.93E-01	0.01	8.94E-01	0.06	5.67E-01
rs12087736		154695678	0.11	1.27E-01	-0.04	7.87E-01	-0.05	7.45E-01
rs10494301		154701218	0.06	2.70E-01	-0.12	4.81E-01	-0.04	7.53E-01
rs12402320		154701406	-0.05	2.89E-01	0.01	9.66E-01	-0.07	5.14E-01
rs884664		154703725	-0.01	9.08E-01	-0.09	5.56E-01	0.13	2.13E-01
rs925042		154704595	0.01	8.75E-01	0.03	7.57E-01	-0.14	1.92E-01
rs6426905		154706620	0.16	6.23E-02	-0.14	5.65E-01	0.12	4.63E-01
rs960182		154710756	-0.04	3.87E-01	0.16	7.34E-02	-0.13	2.07E-01
rs10908429		154710904	-0.06	3.12E-01	-0.10	3.71E-01	0.05	7.25E-01
rs6682000		154711226	0.03	5.95E-01	0.13	4.24E-01	-0.15	1.51E-01
rs12145386		154713499	0.04	4.89E-01	0.11	3.37E-01	0.12	3.92E-01
rs4999127		154714006	0.05	4.47E-01	-0.38	2.04E-02	0.12	3.90E-01
rs6698628		154721780	-0.18	1.44E-01	0.01	9.65E-01	-0.20	5.55E-01
rs906276		154724392	-0.01	8.51E-01	0.01	9.36E-01	-0.03	8.45E-01
rs4845390		154726374	-0.12	8.21E-03	-0.14	2.32E-01	0.00	9.98E-01
rs10494300		154726546	0.12	1.13E-02	0.09	4.16E-01	-0.09	4.22E-01
rs10908435		154729394	0.08	1.11E-01	0.01	8.92E-01	-0.06	6.02E-01
rs6682291		154732776	0.14	1.54E-01	0.09	7.31E-01	0.21	3.44E-01

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs4845391		154739104	0.03	5.97E-01	-0.16	2.47E-01	-0.07	5.34E-01
rs1051614	*KCNN3* Transcribed	154744807	0.04	4.25E-01	0.29	2.53E-02	-0.03	7.66E-01
rs1131820		154744852	0.08	1.70E-01	0.18	1.78E-01	0.04	7.28E-01
rs7532286		154750816	-0.02	6.05E-01	0.01	9.65E-01	0.11	2.39E-01
rs1506982		154753697	0.02	8.34E-01	-0.02	9.38E-01	0.01	9.69E-01
rs4845671		154755126	0.01	8.70E-01	-0.29	2.53E-02	-0.09	3.95E-01
rs11264261		154756443	0.06	2.65E-01	0.29	2.53E-02	0.00	9.88E-01
rs883319		154758422	-0.06	2.67E-01	-0.26	8.26E-02	-0.19	1.36E-01
rs11584635		154760965	-0.08	1.22E-01	-0.17	2.35E-01	-0.04	7.42E-01
rs17719287		154766495	0.01	9.56E-01	0.09	7.31E-01	0.27	2.63E-01
rs1419110		154768689	0.02	6.73E-01	-0.17	2.45E-01	-0.06	5.41E-01
rs4492614		154769561	-0.08	2.62E-01	0.26	5.17E-02	0.02	8.98E-01
rs17719419		154771916	0.13	2.72E-01	0.09	7.31E-01	0.20	5.02E-01
rs6690388		154774447	-0.09	1.47E-01	-0.07	5.37E-01	0.05	7.08E-01
rs6426944		154776412	-0.13	1.49E-02	0.01	9.32E-01	0.19	7.39E-02
rs2256209		154779544	-0.14	5.81E-02	0.16	2.98E-01	0.19	2.25E-01
rs2798601		154779685	-0.07	2.82E-01	-0.03	7.65E-01	0.13	2.46E-01
rs6688473		154783773	-0.05	4.89E-01	0.10	3.67E-01	0.22	1.26E-01
rs2335406		154786686	0.06	1.58E-01	0.17	1.56E-01	-0.09	3.41E-01
rs1106236		154787136	-0.03	5.03E-01	-0.09	4.53E-01	-0.03	7.47E-01
rs1106235		154787213	-0.02	7.06E-01	-0.10	2.11E-01	-0.15	2.06E-01
rs1218601		154791676	-0.06	2.54E-01	-0.19	5.25E-02	-0.07	5.38E-01
rs951241		154793798	0.02	7.75E-01	0.11	3.17E-01	0.19	2.21E-01
rs12029542		154794897	-0.05	4.01E-01	0.20	4.28E-02	0.18	1.45E-01
rs11264268		154796520	0.05	3.74E-01	-0.02	8.86E-01	-0.09	4.73E-01
rs6681725		154797062	0.05	4.52E-01	-0.09	5.93E-01	0.04	8.17E-01

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs1218551		154801173	-0.01	7.68E-01	0.13	1.84E-01	0.07	5.18E-01
rs11264270		154802379	-0.16	7.72E-03	0.21	2.09E-01	0.03	8.56E-01
rs4845394		154808287	0.03	5.51E-01	-0.07	4.51E-01	0.01	9.05E-01
rs1218578		154810030	0.04	3.75E-01	-0.08	4.33E-01	0.05	6.26E-01
rs12128882		154811435	0.04	4.62E-01	0.20	1.89E-01	0.07	5.41E-01
rs1218574		154811677	-0.04	5.17E-01	0.01	9.66E-01	-0.08	4.89E-01
rs12740456		154814197	0.06	3.50E-01	-0.10	5.87E-01	-0.26	4.26E-02
rs13376333	AF GWAS	154814353	0.08	9.43E-02	-0.05	7.38E-01	0.02	8.51E-01
rs6426987		154815257	0.05	2.74E-01	0.00	9.99E-01	0.05	6.16E-01
rs1218565		154817687	0.01	8.89E-01	0.00	9.81E-01	0.02	7.88E-01
rs4845396		154828409	0.08	8.63E-02	-0.12	1.68E-01	-0.05	6.06E-01
rs1061122		154832290	0.10	6.66E-02	-0.10	3.32E-01	0.07	6.14E-01
rs76110919		154842095						
rs2335407		154843347	-0.05	3.40E-01	0.12	1.83E-01	-0.12	3.03E-01

The significance threshold after correction for multiple testing was 7.81E-04. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.10 Associations with allelic expression ratio of *KCNN3* in chromosome band 1q21 in tissue

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs1995662		154679567	-0.10	4.39E-01	-0.32	3.57E-01	-0.43	6.50E-03
rs6695232		154681950	0.10	1.12E-01	-0.01	9.32E-01	0.05	6.26E-01
rs10128027		154691445	0.02	7.59E-01	0.05	7.83E-01	-0.16	1.06E-01
rs4845663		154692088	0.10	1.21E-01	-0.09	6.68E-01	0.02	8.17E-01
rs7554577		154692883	0.08	2.56E-01	0.07	7.94E-01	0.12	1.85E-01
rs11264249		154693181	0.12	8.49E-02	0.07	7.94E-01	0.15	1.09E-01
rs11264251		154693538	-0.04	6.38E-01	0.07	7.33E-01	-0.15	1.35E-01
rs12087736		154695678	0.11	2.41E-01	0.30	2.05E-01	0.12	3.39E-01
rs10494301		154701218	0.08	2.29E-01	0.09	5.30E-01	-0.02	8.18E-01
rs12402320		154701406	-0.12	1.58E-01	-0.03	9.14E-01	0.01	9.19E-01
rs884664		154703725	0.06	3.89E-01	-0.15	4.04E-01	0.09	4.39E-01
rs925042		154704595	-0.16	2.52E-02	-0.09	7.29E-01	-0.07	4.89E-01
rs6426905		154706620	0.33	3.20E-03	-0.36	2.95E-01	0.23	1.57E-01
rs960182		154710756	-0.04	5.14E-01	-0.07	7.46E-01	-0.04	6.25E-01
rs10908429		154710904	0.13	9.11E-02	0.19	2.12E-01	0.12	2.42E-01
rs6682000		154711226	0.09	2.13E-01	0.20	1.88E-01	0.19	6.16E-02
rs12145386		154713499	-0.24	1.45E-02	0.27	4.49E-01	-0.03	8.06E-01
rs4999127		154714006	0.24	2.65E-02			0.12	3.82E-01
rs6698628		154721780	0.08	7.34E-01				
rs906276		154724392	0.23	9.58E-03	0.33	1.61E-01	0.21	1.54E-01
rs4845390		154726374	**0.25**	**9.08E-06**	0.33	1.10E-02	0.20	1.64E-02
rs10494300		154726546	-0.16	8.84E-03	-0.31	4.60E-02	-0.16	5.61E-02
rs10908435		154729394	-0.03	6.52E-01	-0.09	6.68E-01	-0.01	9.11E-01

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs6682291		154732776	-0.31	1.19E-03	-0.46	1.68E-01	-0.44	5.14E-03
rs4845391		154739104	0.03	6.15E-01	0.03	8.99E-01	0.10	2.54E-01
rs1051614	KCNN3 Transcribed	154744807	**-0.31**	**8.38E-23**	**-0.35**	**7.73E-07**	**-0.30**	**9.80E-10**
rs1131820		154744852	**-0.30**	**4.39E-09**	-0.35	8.49E-04	**-0.31**	**1.25E-05**
rs7532286		154750816	0.17	3.49E-03	0.22	6.75E-02	0.17	7.33E-02
rs1506982		154753697	-0.13	5.11E-01			-0.14	4.17E-01
rs4845671		154755126	0.18	1.15E-02			0.24	1.13E-02
rs11264261		154756443	**-0.27**	**3.58E-07**	-0.19	1.52E-01	-0.25	9.59E-04
rs883319		154758422	0.11	1.73E-01	-0.27	4.49E-01	0.26	1.43E-02
rs11584635		154760965	**0.33**	**3.39E-08**	0.32	8.35E-02	0.27	1.06E-03
rs17719287		154766495	-0.31	9.47E-03	-0.46	1.68E-01	-0.37	2.30E-02
rs1419110		154768689	**-0.21**	**6.08E-04**	-0.43	5.42E-02	-0.14	1.41E-01
rs4492614		154769561	-0.01	9.19E-01	0.11	5.76E-01	-0.08	6.06E-01
rs17719419		154771916	-0.09	6.19E-01	0.46	1.68E-01	-0.25	2.87E-01
rs6690388		154774447	0.22	1.25E-02	0.32	1.72E-01	0.25	5.86E-02
rs6426944		154776412	0.05	5.14E-01	0.00	9.83E-01	-0.02	8.58E-01
rs2256209		154779544	-0.03	7.52E-01	-0.33	1.61E-01	-0.06	6.38E-01
rs2798601		154779685	0.13	2.23E-01	0.32	1.72E-01	0.04	8.04E-01
rs6688473		154783773	-0.06	4.64E-01	-0.07	7.44E-01	-0.18	1.16E-01
rs2335406		154786686	-0.08	2.09E-01	-0.04	8.23E-01	0.07	4.85E-01
rs1106236		154787136	0.11	8.43E-02	0.34	2.69E-02	0.00	9.99E-01
rs1106235		154787213	0.23	1.22E-02	0.32	3.57E-01	0.28	5.69E-02
rs1218601		154791676	0.15	7.71E-02	0.09	6.58E-01	0.18	8.75E-02
rs951241		154793798	0.08	4.10E-01	0.46	1.68E-01	-0.08	6.53E-01
rs12029542		154794897	0.06	4.71E-01	0.13	5.29E-01	-0.03	7.84E-01

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs11264268		154796520	-0.10	1.99E-01	-0.38	3.11E-02	-0.10	3.69E-01
rs6681725		154797062	-0.10	3.91E-01	0.36	1.18E-01	-0.26	1.18E-01
rs1218551		154801173	0.05	4.51E-01	0.22	1.86E-01	0.04	6.08E-01
rs11264270		154802379	-0.09	3.34E-01	-0.33	1.61E-01	-0.33	4.16E-02
rs4845394		154808287	0.12	5.76E-02	-0.22	1.48E-01	0.06	5.48E-01
rs1218578		154810030	0.07	3.54E-01	0.21	2.12E-01	0.07	4.56E-01
rs12128882		154811435	-0.04	5.28E-01	0.36	4.93E-02	0.01	9.62E-01
rs1218574		154811677	-0.04	6.79E-01	0.32	3.57E-01	-0.11	4.62E-01
rs12740456		154814197	-0.05	6.01E-01	0.39	9.26E-02	-0.03	8.37E-01
rs13376333	AF GWAS	154814353	-0.01	9.27E-01	0.38	9.62E-04	0.07	5.16E-01
rs6426987		154815257	-0.03	6.55E-01	0.37	9.85E-03	0.05	6.72E-01
rs1218565		154817687	-0.03	6.24E-01	0.36	2.61E-03	0.04	7.17E-01
rs4845396		154828409	-0.09	1.29E-01	-0.03	8.57E-01	0.00	9.90E-01
rs1061122		154832290	0.09	2.53E-01	-0.10	5.23E-01	0.07	5.07E-01
rs76110919		154842095						
rs2335407		154843347	0.01	8.27E-01	-0.07	6.76E-01	-0.09	3.06E-01

The significance threshold after correction for multiple testing was 7.81E-04. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.11 Associations with total expression of *PRRX1* in chromosome band 1q24 in tissue

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs12021752		170510520	-0.15	4.78E-02	-0.17	5.40E-01	0.15	1.46E-01
rs16863425		170519884	-0.14	4.30E-02	-0.17	5.40E-01	0.10	2.71E-01
rs913257	*GORAB* Transcribed	170521376	-0.08	1.49E-01	-0.29	2.46E-01	0.02	8.07E-01
rs7552922	*GORAB* Transcribed	170521650	-0.13	1.47E-01	0.61	9.67E-02	0.14	2.82E-01
rs12038857		170535636	-0.01	8.65E-01	-0.17	7.09E-01	-0.22	5.13E-02
rs7535322		170554924	-0.04	6.10E-01	0.06	8.35E-01	-0.04	7.03E-01
rs12027180		170563205	0.13	4.11E-02	-0.07	8.49E-01	-0.05	6.17E-01
rs3903239	AF GWAS	170569317	-0.11	7.23E-02	0.43	3.33E-01	-0.09	2.91E-01
rs619456		170585832	0.12	4.49E-02	-0.50	6.64E-02	-0.09	2.56E-01
rs12129225		170615833	0.01	8.38E-01	-0.09	7.55E-01	0.11	1.87E-01
rs12038255		170618461	0.14	9.10E-02	0.11	6.97E-01	-0.12	3.86E-01
rs659580		170633896	-0.01	8.32E-01	0.55	1.10E-01	-0.06	5.32E-01
rs502612		170641803	-0.10	1.01E-01	0.43	3.33E-01	-0.05	4.70E-01
rs591715		170642509	0.13	7.55E-02	-0.07	8.49E-01	-0.02	8.49E-01
rs12066968		170652434	-0.08	4.75E-01	0.07	8.76E-01	0.21	1.41E-01
rs513287		170664237	0.01	9.17E-01	0.14	6.00E-01	0.05	5.51E-01
rs7540713		170674903	0.05	4.77E-01	-0.33	3.52E-01	-0.11	3.08E-01
rs58473244		170705898	0.00		0.00		0.00	
rs3820416	*PRRX1* Transcribed	170707675	-0.08	2.17E-01	0.18	4.12E-01	0.09	3.18E-01
rs4656799		170709195	-0.02	7.18E-01	0.09	6.67E-01	0.01	8.74E-01
rs11799764		170714230	0.05	4.94E-01	-0.20	4.65E-01	-0.15	2.07E-01
rs2421494		170715531	-0.07	3.35E-01	-0.39	3.76E-01	0.07	4.83E-01

rs11589602	170716202	0.09	2.06E-01	-0.21	4.32E-01	0.00	9.96E-01

The significance threshold after correction for multiple testing was 2.17E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.12 Associations with allelic expression ratio of *PRRX1* in chromosome band 1q24 in tissue

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs12021752		170510520	-0.14	2.72E-01			-0.14	4.37E-01
rs16863425		170519884	-0.10	4.42E-01			-0.10	5.96E-01
rs913257	GORAB Transcribed	170521376	-0.18	1.81E-01			-0.21	2.75E-01
rs7552922	GORAB Transcribed	170521650	0.08	6.73E-01			-0.07	8.10E-01
rs12038857		170535636	-0.15	3.52E-01			-0.12	5.56E-01
rs7535322		170554924	-0.01	9.58E-01			0.04	8.54E-01
rs12027180		170563205	**0.65**	**8.60E-06**			0.68	4.34E-04
rs3903239	AF GWAS	170569317	-0.37	2.53E-03			-0.47	1.12E-02
rs619456		170585832	**0.52**	**1.88E-04**			0.54	7.66E-03
rs12129225		170615833	0.02	9.15E-01			-0.09	6.41E-01
rs12038255		170618461	0.21	1.15E-01			0.37	1.92E-02
rs659580		170633896	0.01	9.40E-01			-0.12	6.91E-01
rs502612		170641803	**-0.40**	**1.31E-03**			-0.39	3.29E-02
rs591715		170642509	**0.71**	**1.93E-05**			**0.69**	**8.55E-04**
rs12066968		170652434	0.61	9.09E-03			0.69	3.85E-03
rs513287		170664237	**-0.66**	**1.22E-12**			**-0.68**	**2.13E-08**
rs7540713		170674903	**0.52**	**1.37E-03**			0.52	1.78E-02
rs58473244		170705898						
rs3820416	PRRX1 Transcribed	170707675	**-0.68**	**2.50E-36**			**-0.69**	**3.25E-25**
rs4656799		170709195	**-0.68**	**3.25E-20**			**-0.69**	**2.90E-11**
rs11799764		170714230	**0.69**	**1.51E-03**			**0.65**	**5.12E-02**
rs2421494		170715531	**-0.68**	**1.24E-12**			**-0.69**	**5.74E-09**
rs11589602		170716202	0.49	1.42E-02			0.46	1.27E-01

The significance threshold after correction for multiple testing was 2.17E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.13 Associations with total expression of *GORAB* in chromosome band 1q24 in tissue

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs12021752		170510520	-0.08	1.09E-01	0.08	3.82E-01	0.01	8.70E-01
rs16863425		170519884	-0.06	1.89E-01	0.08	3.82E-01	-0.03	6.72E-01
rs913257	*GORAB* Transcribed	170521376	-0.08	6.41E-02	0.04	6.46E-01	0.02	6.83E-01
rs7552922	*GORAB* Transcribed	170521650	0.02	7.16E-01	0.03	7.99E-01	-0.09	3.10E-01
rs12038857		170535636	-0.04	4.20E-01	-0.03	8.35E-01	-0.02	7.77E-01
rs7535322		170554924	-0.02	6.55E-01	0.06	5.25E-01	0.00	9.58E-01
rs12027180		170563205	0.11	8.92E-03	0.09	3.05E-01	-0.06	3.03E-01
rs3903239	AF GWAS	170569317	-0.04	3.36E-01	0.26	3.11E-02	0.13	1.09E-02
rs619456		170585832	0.09	2.80E-02	-0.14	1.25E-01	-0.06	2.50E-01
rs12129225		170615833	-0.08	5.72E-02	-0.19	1.30E-02	-0.08	1.48E-01
rs12038255		170618461	0.24	3.02E-05	0.03	7.30E-01	0.03	7.17E-01
rs659580		170633896	-0.09	5.17E-02	0.18	1.10E-01	0.11	7.50E-02
rs502612		170641803	0.01	8.45E-01	0.30	2.32E-02	0.11	1.83E-02
rs591715		170642509	0.10	4.13E-02	0.19	1.08E-01	-0.07	3.23E-01
rs12066968		170652434	-0.02	7.69E-01	0.18	2.41E-01	0.01	9.15E-01
rs513287		170664237	-0.02	6.34E-01	-0.13	1.73E-01	-0.04	4.54E-01
rs7540713		170674903	0.05	3.06E-01	-0.11	4.02E-01	0.05	5.15E-01
rs58473244		170705898	0.00		0.00		0.00	
rs3820416	*PRRX1* Transcribed	170707675	-0.05	2.50E-01	-0.10	2.98E-01	0.05	4.26E-01
rs4656799		170709195	0.01	7.49E-01	-0.18	3.63E-03	0.09	9.25E-02
rs11799764		170714230	0.03	5.43E-01	-0.14	9.84E-02	0.02	8.23E-01
rs2421494		170715531	-0.01	8.75E-01	0.05	7.68E-01	0.05	4.65E-01

rs11589602	170716202	0.05	3.24E-01	0.02	8.43E-01	-0.11	6.78E-02

The significance threshold after correction for multiple testing was 2.17E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.14 Associations with total expression of *GORAB* in chromosome band 1q24 in tissue

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs12021752		170510520	**0.19**	**3.52E-08**	-0.35	8.33E-02	**0.23**	**2.17E-03**
rs16863425		170519884	**0.17**	**9.87E-08**	-0.35	8.33E-02	**0.19**	**2.80E-04**
rs913257	GORAB Transcribed	170521376	**0.23**	**8.02E-14**	-0.38	3.36E-03	**0.31**	**1.71E-08**
rs7552922	GORAB Transcribed	170521650	**-0.09**	**6.60E-04**	0.00	1.00E+00	0.05	2.06E-01
rs12038857		170535636	**0.15**	**1.40E-03**	-0.19	1.38E-01	0.09	2.37E-01
rs7535322		170554924	0.08	2.01E-02	-0.24	2.01E-02	0.12	4.04E-02
rs12027180		170563205	**-0.12**	**1.57E-03**	0.40	1.21E-01	-0.06	3.71E-01
rs3903239	AF GWAS	170569317	0.06	8.33E-02	**-0.17**	**2.17E-03**	0.04	4.39E-01
rs619456		170585832	-0.11	2.70E-03	0.40	1.21E-01	-0.11	9.43E-02
rs12129225		170615833	0.06	1.80E-01	0.16	5.84E-03	-0.01	1.00E+00
rs12038255		170618461	-0.06	1.60E-02	0.24	2.85E-02	-0.05	2.37E-01
rs659580		170633896	0.01	1.00E+00	**-0.17**	**2.17E-03**	0.03	5.27E-01
rs502612		170641803	0.07	2.85E-02	**-0.14**	**4.29E-04**	0.07	2.06E-01
rs591715		170642509	**-0.15**	**1.83E-04**	0.18	1.80E-01	-0.08	1.80E-01
rs12066968		170652434	0.10	7.36E-02	-0.19	1.38E-01	-0.02	1.00E+00
rs513287		170664237	-0.02	5.27E-01	0.12	1.80E-01	0.01	6.55E-01
rs7540713		170674903	**-0.12**	**1.13E-03**	0.19	1.21E-01	-0.05	4.39E-01
rs58473244		170705898						
rs3820416	PRRX1 Transcribed	170707675	0.03	3.71E-01	-0.25	5.27E-01	0.03	5.27E-01
rs4656799		170709195	-0.02	6.55E-01	0.10	4.39E-01	0.08	1.80E-01
rs11799764		170714230	-0.07	1.21E-01	0.12	3.71E-01	-0.01	1.00E+00
rs2421494		170715531	0.08	6.52E-02			0.01	1.00E+00
rs11589602		170716202	-0.06	3.17E-01	-0.15	2.37E-01	-0.03	6.55E-01

The significance threshold after correction for multiple testing was 2.17E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.15 Associations with total expression of *PITX2a/b* in chromosome band 4q25 in tissue

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs75911264		111538810						
rs6533526		111538827						
rs2595110		111545323	-0.09	5.63E-01	-0.09	6.76E-01	0.26	3.43E-01
rs976568		111550721	0.16	3.42E-01	0.23	2.53E-01	-0.40	2.72E-01
rs2739200	*PITX2* Transcribed	111558411	0.12	4.67E-01	0.31	1.87E-01	-0.48	1.84E-01
rs12498374		111584419	0.17	3.54E-01	0.05	8.29E-01	-0.71	6.19E-02
rs2723296		111604483	0.16	3.30E-01	0.31	1.87E-01	-0.50	1.68E-01
rs1448799		111622782	0.42	1.79E-02	0.09	6.89E-01	0.19	5.73E-01
rs16997168		111629039	**0.66**	**8.96E-04**	0.16	4.77E-01	-0.87	6.17E-02
rs10222783		111634826	0.08	7.35E-01	-0.42	2.45E-01	0.05	9.14E-01
rs4307025		111657503	0.44	1.13E-02	0.23	3.13E-01	0.11	7.28E-01
rs2723333		111699091	0.04	8.57E-01	0.47	1.78E-01	0.41	2.49E-01
rs17042171	AF GWAS	111708287	**0.70**	**6.48E-04**	0.21	3.30E-01	-0.31	4.85E-01
rs2200733	AF GWAS	111710169	**0.68**	**1.42E-03**	0.14	2.80E-01	-0.44	3.71E-01
rs13126975		111717525	-0.18	4.39E-01	-0.04	9.43E-01	0.15	6.44E-01
rs6843082	AF GWAS	111718067	**0.58**	**1.35E-03**	0.23	2.73E-01	0.04	9.19E-01
rs10033464	AF GWAS	111720761	0.09	7.64E-01	0.13	7.90E-01	0.38	4.04E-01
rs6533531		111731965	0.32	3.88E-02	0.19	4.93E-01	0.23	4.83E-01
rs3853444		111734136	0.17	2.88E-01	-0.48	7.22E-02	0.59	2.18E-02
rs17570669		111736882	-0.66	8.02E-02			0.24	7.62E-01
rs7674295		111741438	-0.46	2.47E-02	0.13	6.85E-01	-0.37	3.45E-01
rs11938968		111742752	-0.24	1.15E-01	0.24	1.84E-01	-0.04	8.67E-01

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs3866836		111745640	0.15	3.89E-01	-0.16	4.88E-01	-0.29	3.82E-01
rs4032971		111751723	0.26	1.70E-01	-0.23	2.79E-01	-0.18	6.43E-01
rs3853445		111761487	-0.14	4.55E-01	-0.15	6.10E-01	0.45	1.74E-01
rs17571707		111765413	-0.31	3.10E-01	0.19	6.02E-01	-0.50	3.26E-01
rs6838973		111765495	-0.38	1.12E-02	-0.09	6.87E-01	0.26	3.98E-01
rs3866823		111782436	-0.02	8.87E-01	-0.26	1.71E-01	-0.07	8.08E-01
rs4033102		111784205	-0.15	7.41E-01			-0.01	9.88E-01
rs542331		111798258	-0.16	4.26E-01	-0.19	5.58E-01	0.41	3.13E-01
rs6817202		111799254	0.13	4.85E-01	-0.23	2.79E-01	-0.45	1.82E-01
rs7679158		111812327	-0.26	1.04E-01	-0.15	6.13E-01	0.16	6.22E-01
rs11947581		111812531	0.08	6.53E-01	-0.12	5.79E-01	-0.52	8.33E-02
rs570881		111816849						
rs561873		111829967	-0.23	4.78E-01	-0.23	5.37E-01	-0.28	6.58E-01
rs581241		111836832	-0.44	6.45E-02	0.00	9.96E-01	0.80	6.87E-02
rs441624		111844449	0.49	1.66E-01	1.01	1.39E-02	0.29	6.54E-01
rs374582		111849422	0.16	2.76E-01	0.19	3.67E-01	0.08	7.43E-01
rs11943026		111872144	-0.30	1.30E-01	0.00	9.96E-01	0.50	2.64E-01
rs17513835		111894848	0.18	5.16E-01	-0.21	3.62E-01	0.05	9.27E-01

The significance threshold after correction for multiple testing was 1.35E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.16 Associations with total expression of *PITX2c* in chromosome band 4q25 in tissue

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs75911264		111538810						
rs6533526		111538827						
rs2595110		111545323			-0.09	4.96E-01	-0.41	3.58E-01
rs976568		111550721			-0.03	8.47E-01	-0.36	4.54E-01
rs2739200	*PITX2* Transcribed	111558411			0.02	8.76E-01	-0.36	4.54E-01
rs12498374		111584419			0.05	7.23E-01	-0.91	4.37E-02
rs2723296		111604483			0.01	9.17E-01	-0.40	4.00E-01
rs1448799		111622782			0.13	3.76E-01	0.58	2.12E-01
rs16997168		111629039			-0.01	9.24E-01	-0.43	4.04E-01
rs10222783		111634826			0.03	8.94E-01	0.59	3.68E-01
rs4307025		111657503			-0.10	4.75E-01	0.57	1.17E-01
rs2723333		111699091			-0.15	5.09E-01	0.28	5.87E-01
rs17042171	AF GWAS	111708287			-0.09	5.25E-01	0.09	8.70E-01
rs2200733	AF GWAS	111710169			-0.12	3.27E-01	-0.02	9.76E-01
rs13126975		111717525			0.11	7.12E-01	-0.82	1.31E-01
rs6843082	AF GWAS	111718067			-0.09	4.84E-01	0.09	8.70E-01
rs10033464	AF GWAS	111720761			-0.05	8.67E-01	0.00	
rs6533531		111731965			-0.09	5.68E-01	0.79	7.86E-02
rs3853444		111734136			0.12	4.89E-01	0.30	3.38E-01
rs17570669		111736882			-0.02	9.44E-01	0.00	
rs7674295		111741438			-0.24	2.07E-01	0.17	8.01E-01
rs11938968		111742752			-0.12	2.42E-01	0.44	1.60E-01

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs3866836		111745640			0.32	7.71E-03	-0.50	1.72E-01
rs4032971		111751723			0.31	5.29E-03	-0.41	4.04E-01
rs3853445		111761487			-0.06	7.46E-01	0.48	3.05E-01
rs17571707		111765413			-0.39	6.65E-02	0.16	8.61E-01
rs6838973		111765495			-0.01	9.23E-01	0.09	8.57E-01
rs3866823		111782436			0.27	9.87E-03	-0.44	2.07E-01
rs4033102		111784205			-0.02	9.44E-01	0.00	
rs542331		111798258			0.04	8.60E-01	0.05	9.14E-01
rs6817202		111799254			0.31	5.29E-03	-0.51	1.57E-01
rs7679158		111812327			-0.04	8.33E-01	0.04	9.11E-01
rs11947581		111812531			0.17	2.02E-01	-0.51	1.57E-01
rs570881		111816849			0.00		0.00	
rs561873		111829967			-0.02	9.43E-01	-0.72	2.65E-01
rs581241		111836832			-0.03	8.62E-01	0.42	4.14E-01
rs441624		111844449			-0.14	6.54E-01	0.33	7.17E-01
rs374582		111849422			-0.11	3.50E-01	0.33	2.79E-01
rs11943026		111872144			-0.03	8.62E-01	0.65	2.43E-01
rs17513835		111894848			0.00	9.91E-01	0.05	9.28E-01

The significance threshold after correction for multiple testing was 1.35E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.17 Associations with allelic expression ratio of *PITX2* in chromosome band 4q25 in tissue

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs75911264		111538810						
rs6533526		111538827						
rs2595110		111545323	-0.08	6.37E-01	-0.25	3.12E-01	0.35	2.60E-01
rs976568		111550721	-0.04	7.70E-01	-0.01	9.65E-01	-0.23	3.05E-01
rs2739200	PITX2 Transcribed	111558411	0.00	9.80E-01	-0.01	9.65E-01	-0.26	2.11E-01
rs12498374		111584419	0.04	7.95E-01	-0.12	6.51E-01	-0.32	1.31E-01
rs2723296		111604483	0.00	9.80E-01	-0.20	4.36E-01	-0.26	2.11E-01
rs1448799		111622782	0.08	6.85E-01	0.44	4.59E-03	0.05	8.60E-01
rs16997168		111629039	0.06	7.93E-01	0.44	4.59E-03	0.05	8.60E-01
rs10222783		111634826	0.25	6.23E-01				
rs4307025		111657503	0.03	8.92E-01	0.32	1.48E-01	0.06	8.48E-01
rs2723333		111699091	-0.47	4.70E-02	-0.52	7.78E-02	0.27	6.77E-01
rs17042171	AF GWAS	111708287	-0.02	9.39E-01	0.52	7.78E-02	0.06	8.48E-01
rs2200733	AF GWAS	111710169	0.23	5.20E-01	0.52	7.78E-02	0.06	8.48E-01
rs13126975		111717525	0.22	5.31E-01			0.23	6.04E-01
rs6843082	AF GWAS	111718067	0.10	6.41E-01	0.52	7.78E-02	0.06	8.48E-01
rs10033464	AF GWAS	111720761	0.30	4.03E-01				
rs6533531		111731965	0.13	5.30E-01	0.32	1.48E-01	0.06	8.48E-01
rs3853444		111734136	0.14	5.02E-01			0.48	1.73E-01
rs17570669		111736882	-0.06	9.11E-01	0.36	2.87E-01		
rs7674295		111741438	-0.21	3.46E-01	0.13	7.28E-01	0.36	3.18E-01
rs11938968		111742752	-0.21	2.97E-01	0.13	7.28E-01	0.56	5.77E-02
rs3866836		111745640	0.03	9.06E-01			-0.36	4.26E-01

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs4032971		111751723	0.03	9.06E-01			-0.25	5.00E-01
rs3853445		111761487	0.12	5.93E-01	0.36	2.87E-01	-0.75	7.08E-02
rs17571707		111765413	0.06	9.11E-01	0.13	7.28E-01		
rs6838973		111765495	-0.13	4.44E-01			-0.15	7.39E-01
rs3866823		111782436	-0.03	8.71E-01	0.36	2.87E-01	-0.56	5.77E-02
rs4033102		111784205			0.36	2.87E-01		
rs542331		111798258	-0.08	6.77E-01			-0.75	7.08E-02
rs6817202		111799254	0.06	7.75E-01			-0.36	4.26E-01
rs7679158		111812327	-0.17	3.59E-01	-0.36	2.87E-01	0.06	8.89E-01
rs11947581		111812531	0.02	9.57E-01	0.13	7.28E-01	-0.50	4.29E-01
rs570881		111816849						
rs561873		111829967	-0.56	4.09E-02			-0.94	1.19E-01
rs581241		111836832	0.15	5.63E-01	-0.36	2.87E-01	0.82	4.59E-02
rs441624		111844449	-0.01	9.57E-01	-0.52	7.78E-02	0.60	1.62E-01
rs374582		111849422	0.18	3.35E-01	-0.13	7.28E-01	0.33	2.94E-01
rs11943026		111872144	0.11	6.41E-01	-0.36	2.87E-01	0.82	4.59E-02
rs17513835		111894848	-0.56	4.09E-02			-0.94	1.19E-01

Table A2.18 Associations with total expression of *CAV1* in chromosome band 7q31 in tissue

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
	CAV2 Transcribed							
rs6466578		116150869	-0.03	5.95E-01	-0.15	4.31E-02	0.13	8.19E-02
rs3919515		116151784	0.08	4.87E-02	-0.07	3.80E-01	0.08	2.29E-01
rs959173		116182054	0.02	7.15E-01	-0.11	2.11E-01	0.12	9.80E-02
rs3807989	AF GWAS	116186241	0.11	1.33E-02	0.09	2.34E-01	0.07	3.38E-01
rs11773845	AF, PR interval GWAS	116191301	0.11	7.73E-03	0.06	3.56E-01	0.09	2.03E-01
rs8713		116199797	-0.07	1.08E-01	0.09	1.48E-01	-0.03	7.25E-01
rs9920	*CAV1* Transcribed	116200092	**0.20**	**4.04E-03**	0.02	8.24E-01	0.07	4.29E-01
rs1049334	*CAV1* Transcribed	116200380	0.06	4.66E-01	0.10	1.80E-01	-0.11	3.15E-01
rs1049337	*CAV1* Transcribed	116200587	-0.09	3.32E-02	-0.01	8.85E-01	0.06	3.73E-01
rs1860588		116206486	-0.03	7.07E-01	-0.11	2.11E-01	0.12	1.61E-01

The significance threshold after correction for multiple testing was 4.55E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.19 Associations with total expression of *CAV2* in chromosome band 7q31 in tissue

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
	CAV2 Transcribed							
rs6466578		116150869	0.00	9.70E-01	-0.18	5.64E-02	0.17	3.79E-02
rs3919515		116151784	0.08	1.20E-02	-0.09	5.59E-01	0.13	7.40E-02
rs959173		116182054	0.04	3.78E-01	-0.18	6.43E-02	0.17	2.99E-02
rs3807989	AF GWAS	116186241	0.10	6.17E-03	-0.04	6.50E-01	0.09	2.34E-01
rs11773845	AF, PR interval GWAS	116191301	**0.10**	**3.66E-03**	-0.07	5.71E-01	0.10	1.58E-01
rs8713		116199797	-0.02	5.18E-01	-0.01	9.14E-01	0.00	9.66E-01
rs9920	*CAV1* Transcribed	116200092	**0.19**	**7.29E-04**	0.06	7.62E-01	0.09	3.33E-01
rs1049334	*CAV1* Transcribed	116200380	-0.03	6.75E-01	0.02	8.65E-01	-0.21	8.05E-02
rs1049337	*CAV1* Transcribed	116200587	-0.07	3.71E-02	0.07	6.03E-01	0.03	7.08E-01
rs1860588		116206486	0.01	8.97E-01	-0.18	6.43E-02	0.17	5.95E-02

The significance threshold after correction for multiple testing was 4.55E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.20 Associations with allelic expression ratio of *CAV1* in chromosome band 7q31 in tissue

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs8940	CAV2 Transcribed	116146074	0.05	3.71E-01	0.42	1.57E-01	0.18	4.04E-02
rs6466578		116150869	0.05	1.00E+00	-0.01	1.00E+00	0.37	1.28E-02
rs3919515		116151784	**0.17**	**3.79E-07**	0.06	1.80E-01	**0.34**	**3.47E-04**
rs959173		116182054	0.09	1.57E-01	-0.01	1.00E+00	0.34	1.14E-02
rs3807989	AF GWAS	116186241	**0.15**	**5.89E-08**	0.06	1.80E-01	0.06	3.59E-02
rs11773845	AF, PR interval GWAS	116191301	**0.15**	**5.89E-08**	0.06	1.80E-01	0.06	3.59E-02
rs8713		116199797	-0.03	2.73E-01	0.12	7.29E-03	0.01	1.00E+00
rs9920	CAV1 Transcribed	116200092	-0.06	5.22E-03			-0.11	1.80E-01
rs1049334	CAV1 Transcribed	116200380	**0.14**	**2.29E-10**	0.04	2.73E-01	0.04	1.21E-01
rs1049337	CAV1 Transcribed	116200587	**-0.29**	**5.07E-19**	-0.03	6.55E-01	**-0.34**	**1.49E-07**
rs1860588		116206486	0.08	2.37E-01	-0.01	1.00E+00	0.30	4.55E-02

The significance threshold after correction for multiple testing was 4.55E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.21 Associations with allelic expression ratio of *CAV2* in chromosome band 7q31 in tissue

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs8940	CAV2 Transcribed	116146074	**-0.11**	**2.52E-08**			**-0.11**	**2.91E-04**
rs6466578		116150869	0.11	1.50E-01			0.04	6.05E-01
rs3919515		116151784	**-0.11**	**2.56E-04**			-0.09	2.65E-02
rs959173		116182054	0.11	1.50E-01			0.05	4.56E-01
rs3807989	AF GWAS	116186241	-0.05	2.97E-01			-0.04	4.51E-01
rs11773845	AF, PR interval GWAS	116191301	-0.05	2.97E-01			-0.04	4.51E-01
rs8713		116199797	**-0.10**	**1.61E-03**			-0.05	2.37E-01
rs9920	CAV1 Transcribed	116200092	0.07	1.00E+00			0.19	6.01E-02
rs1049334	CAV1 Transcribed	116200380	0.15	2.44E-01				
rs1049337	CAV1 Transcribed	116200587	0.09	2.26E-02			0.07	1.28E-01
rs1860588		116206486	0.08	4.66E-01			0.03	6.93E-01

The significance threshold after correction for multiple testing was 4.55E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.22 Associations with total expression of *C9orf3* in chromosome band 9q22 in tissue

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs356122		97565461	-0.02	6.79E-01	0.23	2.76E-01	-0.13	1.01E-01
rs4744381		97601089	0.00		0.00		0.00	
rs1530331		97620602	-0.01	8.66E-01	0.00		-0.07	4.72E-01
rs10993391		97675840	0.16	8.41E-02	0.23	5.73E-01	-0.07	5.25E-01
rs7853195		97685298	0.00	9.26E-01	-0.18	2.18E-01	-0.09	1.87E-01
rs10821415	AF GWAS	97713459	0.01	9.09E-01	-0.19	3.12E-01	-0.09	2.13E-01
rs10993413		97729136	0.02	7.78E-01	-0.01	9.83E-01	0.12	2.33E-01
rs3802458		97741274	0.07	7.57E-01	0.00		0.78	1.39E-02
rs3802457	PCOS GWAS	97741336	-0.07	6.47E-01	-0.14	6.92E-01	0.29	2.21E-01
rs7048941		97809674	0.05	4.70E-01	0.20	5.60E-01	0.06	5.50E-01
rs4744437		97832830	0.06	5.12E-01	0.00		0.02	8.22E-01
rs7033633		97838781	-0.10	9.13E-02	-0.26	1.81E-01	-0.03	6.94E-01
rs17679141		97842479	0.06	2.10E-01	0.00	9.83E-01	-0.09	1.87E-01
rs4657	*C9orf3* Transcribed	97849090	-0.05	3.05E-01	-0.11	5.07E-01	0.01	9.35E-01
rs4647554	*FANCC* Transcribed	97862701	-0.03	5.80E-01	0.14	6.07E-01	0.09	1.83E-01

The significance threshold after correction for multiple testing was 3.33E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.23 Associations with total expression of *FANCC* in chromosome band 9q22 in tissue

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs356122		97565461	0.00	9.38E-01	-0.13	2.80E-01	0.09	5.46E-01
rs4744381		97601089	0.00		0.00		0.00	
rs1530331		97620602	0.01	8.58E-01	0.00		-0.02	9.11E-01
rs10993391		97675840	0.11	1.35E-01	0.03	8.70E-01	-0.21	3.08E-01
rs7853195		97685298	0.04	3.18E-01	0.05	6.75E-01	0.21	6.56E-02
rs10821415	AF GWAS	97713459	0.05	1.98E-01	0.06	5.65E-01	0.26	2.60E-02
rs10993413		97729136	0.10	1.74E-01	-0.43	2.59E-02	0.13	4.06E-01
rs3802458		97741274	0.08	7.07E-01	0.00		1.00	6.87E-02
rs3802457	PCOS GWAS	97741336	-0.01	9.44E-01	0.02	9.38E-01	0.59	1.36E-01
rs7048941		97809674	0.08	2.00E-01	0.09	7.09E-01	-0.01	9.43E-01
rs4744437		97832830	0.08	3.18E-01	0.00		0.02	9.24E-01
rs7033633		97838781	-0.11	2.94E-02	0.10	4.66E-01	0.19	1.27E-01
rs17679141		97842479	0.09	3.27E-02	0.01	9.38E-01	0.03	7.89E-01
rs4657	*C9orf3* Transcribed	97849090	-0.06	1.33E-01	0.12	2.65E-01	0.09	3.91E-01
rs4647554	*FANCC* Transcribed	97862701	-0.05	2.58E-01	0.04	8.41E-01	-0.16	1.43E-01

The significance threshold after correction for multiple testing was 3.33E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.14 Associations with allelic expression ratio of *C9orf3* in chromosome band 9q22 in tissue

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs356122		97565461	-0.05	1.71E-01	-0.07	4.18E-01	0.01	8.64E-01
rs4744381		97601089						
rs1530331		97620602	-0.05	2.80E-01			-0.05	6.47E-01
rs10993391		97675840	0.18	1.02E-02			0.08	2.03E-01
rs7853195		97685298	**-0.10**	**1.51E-05**	-0.09	4.39E-02	0.01	7.51E-01
rs10821415	AF GWAS	97713459	**-0.10**	**1.51E-05**	-0.09	4.39E-02	0.01	7.51E-01
rs10993413		97729136	0.10	2.55E-02			0.04	6.89E-01
rs3802458		97741274						
rs3802457	PCOS GWAS	97741336	0.11	2.08E-01			0.13	2.34E-01
rs7048941		97809674	-0.01	8.62E-01			0.10	2.12E-01
rs4744437		97832830	-0.11	3.30E-03			-0.07	3.48E-01
rs7033633		97838781	**-0.09**	**5.07E-08**	-0.09	6.45E-03	0.00	9.89E-01
rs17679141		97842479	0.07	1.88E-02	0.10	2.67E-01	-0.02	6.15E-01
rs4657	C9orf3 Transcribed	97849090	**-0.10**	**7.69E-12**	**-0.09**	**1.62E-05**	-0.01	7.42E-01
rs4647554	FANCC Transcribed	97862701	**0.11**	**4.67E-09**	0.09	6.45E-03	0.03	3.80E-01

The significance threshold after correction for multiple testing was 3.33E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.25 Associations with allelic expression ratio of *FANCC* in chromosome band 9q22 in tissue

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs356122		97565461	-0.02	3.64E-01	-0.09	1.09E-01	-0.01	7.17E-01
rs4744381		97601089						
rs1530331		97620602	0.04	1.72E-01			-0.05	1.30E-01
rs10993391		97675840	-0.05	1.58E-01	-0.06	4.84E-01	0.02	3.94E-01
rs7853195		97685298	**-0.05**	**4.14E-07**	-0.05	7.35E-02	-0.03	5.83E-03
rs10821415	AF GWAS	97713459	**-0.06**	**8.92E-08**	-0.05	7.35E-02	**-0.04**	**1.88E-03**
rs10993413		97729136	0.06	1.32E-02			-0.04	7.42E-02
rs3802458		97741274						
rs3802457	PCOS GWAS	97741336	0.05	5.61E-01				
rs7048941		97809674	-0.01	7.46E-01	-0.06	4.84E-01	0.04	1.93E-01
rs4744437		97832830	0.05	9.42E-02			-0.09	3.93E-02
rs7033633		97838781	-0.03	3.53E-02	-0.08	9.33E-03	-0.01	7.31E-01
rs17679141		97842479	-0.06	1.06E-06	0.00	9.57E-01	-0.04	8.88E-03
rs4657	C9orf3 Transcribed	97849090	-0.01	2.91E-01	-0.09	8.12E-03	-0.01	4.47E-01
rs4647554	FANCC Transcribed	97862701	**0.04**	**7.65E-05**	**0.07**	**2.97E-03**	0.03	8.69E-03

The significance threshold after correction for multiple testing was 3.33E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.26 Associations with total expression of *MYOZ1* in chromosome band 10q22 in tissue

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs7080456		75391217	**-1.08**	**2.38E-03**			0.31	1.13E-01
rs41280400	*MYOZ* Transcribed	75391823	-0.80	5.09E-02			0.34	4.49E-01
rs4745718		75402788	**-0.62**	**5.54E-03**	-2.12	2.46E-02	-0.01	9.31E-01
rs34163229	*SYNPO2L* Transcribed	75406912	**-1.36**	**9.49E-15**	**-2.85**	**8.35E-04**	0.13	3.45E-01
rs3812629	*SYNPO2L* Transcribed	75407290	**-1.37**	**4.68E-15**	**-2.85**	**8.35E-04**	0.13	3.38E-01
rs4746139	*SYNPO2L* Transcribed	75407649	**-1.37**	**8.56E-15**	**-2.85**	**8.35E-04**	0.12	4.02E-01
rs12247028		75410052	**-0.78**	**1.38E-08**	-1.47	4.93E-02	0.16	1.37E-01
rs60632610	*SYNPO2L* Transcribed	75415677	**-1.37**	**1.12E-14**	**-2.85**	**8.35E-04**	0.11	4.24E-01
rs10824026	AF GWAS	75421208	**-1.37**	**5.52E-15**	**-2.85**	**8.35E-04**	0.11	4.17E-01

Table A2.27 Associations with total expression of *SYNPO2L* in chromosome band 10q22 in tissue

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs7080456		75391217	-0.06	6.90E-01			1.51	2.22E-02
rs41280400	*MYOZ* Transcribed	75391823	0.02	9.27E-01			0.05	9.72E-01
rs4745718		75402788	0.03	7.59E-01	-0.49	1.12E-01	0.80	1.59E-01
rs34163229	*SYNPO2L* Transcribed	75406912	0.06	5.10E-01	-0.43	2.13E-01	0.82	1.17E-01
rs3812629	*SYNPO2L* Transcribed	75407290	0.06	5.21E-01	-0.43	2.13E-01	0.83	1.02E-01
rs4746139	*SYNPO2L* Transcribed	75407649	0.06	5.08E-01	-0.43	2.13E-01	0.66	1.94E-01
rs12247028		75410052	-0.01	8.38E-01	-0.12	6.56E-01	0.82	6.45E-02
rs60632610	*SYNPO2L* Transcribed	75415677	0.06	5.16E-01	-0.43	2.13E-01	0.82	1.17E-01
rs10824026	AF GWAS	75421208	0.05	5.28E-01	-0.43	2.13E-01	0.83	1.02E-01

The significance threshold after correction for multiple testing was 5.56E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.28 Associations with allelic expression ratio of *MYOZ1* in chromosome band 10q22 in tissue

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs7080456		75391217	**1.09**	**5.24E-04**				
rs41280400	MYOZ Transcribed	75391823	**1.09**	**5.24E-04**				
rs4745718		75402788	1.72	1.00E+00				
rs34163229	SYNPO2L Transcribed	75406912	**1.34**	**5.66E-06**				
rs3812629	SYNPO2L Transcribed	75407290	**1.34**	**5.66E-06**				
rs4746139	SYNPO2L Transcribed	75407649	**1.34**	**5.66E-06**				
rs12247028		75410052	1.35	1.61E-02				
rs60632610	SYNPO2L Transcribed	75415677	**1.34**	**5.66E-06**				
rs10824026	AF GWAS	75421208	**1.34**	**5.66E-06**				

Table A2.29 Associations with allelic expression ratio of *SYNPO2L* in chromosome band 10q22 in tissue

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs7080456		75391217	-0.04	4.39E-01	-0.09	6.55E-01	0.13	1.38E-01
rs41280400	MYOZ Transcribed	75391823	-0.02	1.00E+00	-0.10	1.00E+00	0.25	1.38E-01
rs4745718		75402788	**-0.13**	**3.03E-05**	-0.09	3.17E-01	0.08	1.00E+00
rs34163229	SYNPO2L Transcribed	75406912	**-0.14**	**6.53E-08**			0.10	2.01E-02
rs3812629	SYNPO2L Transcribed	75407290	**-0.14**	**6.53E-08**			0.10	2.01E-02
rs4746139	SYNPO2L Transcribed	75407649	**-0.14**	**6.53E-08**			0.10	2.01E-02
rs12247028		75410052	**-0.11**	**4.68E-03**	-0.11	5.13E-02	0.14	9.12E-03
rs60632610	SYNPO2L Transcribed	75415677	0.16	9.43E-02			0.10	2.85E-02
rs10824026	AF GWAS	75421208	0.16	9.43E-02			0.10	2.85E-02

The significance threshold after correction for multiple testing was 5.56E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.30 Associations with total expression of *SYNE2* in chromosome band 14q23 in tissue

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs4459477		64346190	0.04	6.06E-01			-0.15	3.35E-01
rs17750888		64355420	0.00	9.70E-01	-0.12	6.48E-01	-0.12	1.19E-01
rs12588807		64363070	0.03	7.96E-01			-0.06	7.24E-01
rs2275018		64375985	-0.01	9.06E-01	-0.01	9.60E-01	-0.16	5.43E-02
rs7153036		64377548	0.06	2.83E-01	-0.07	7.26E-01	-0.09	2.96E-01
rs12879919		64382350	-0.04	5.69E-01	-0.12	7.82E-01	0.09	4.04E-01
rs11158521		64401255	0.02	6.59E-01	0.10	7.47E-01	-0.08	2.69E-01
rs2184292		64408135	0.04	4.41E-01	0.44	1.19E-01	-0.14	1.08E-01
rs7153680		64474227	0.03	4.96E-01	0.47	5.97E-02	-0.11	1.37E-01
rs4566057		64506541	-0.03	6.71E-01	-0.42	1.98E-01	0.10	4.02E-01
rs1890908		64519035	0.00	9.72E-01	0.00		0.16	3.37E-01
rs10137972		64557734	0.05	6.68E-01	0.13	7.65E-01	-0.02	9.24E-01
rs17101651		64561324	-0.08	4.35E-01	-0.12	7.82E-01	-0.07	6.88E-01
rs8022428		64563613	0.03	5.92E-01	-0.30	3.26E-01	-0.01	8.76E-01
rs8007972		64564527	0.02	7.58E-01			0.00	9.67E-01
rs8016917		64569951	-0.16	2.81E-01			0.07	7.25E-01
rs10142318		64604281	0.00	9.57E-01	-0.38	2.31E-01	0.07	4.64E-01
rs7144688		64613473	-0.08	2.05E-01	-0.52	1.66E-02	0.13	1.41E-01
rs11158533		64622988	0.01	8.45E-01	0.40	2.03E-01	-0.06	4.11E-01
rs17751454		64625181	0.08	1.79E-01	0.42	1.07E-01	-0.10	2.55E-01
rs3866743		64633322	-0.02	9.20E-01			0.09	6.89E-01
rs7161192	*SYNE2* Transcribed	64637147	0.02	7.13E-01	0.23	6.24E-01	-0.10	1.84E-01
rs2256191		64656855	0.09	1.05E-01	0.34	6.59E-02	-0.08	3.06E-01

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs1152591	AF GWAS	64680848	-0.10	6.05E-02	-0.16	4.77E-01	0.01	9.21E-01
rs7145919		64687601	-0.02	8.25E-01	-0.21	5.34E-01	0.00	9.69E-01
rs12434245		64691853	-0.01	8.79E-01	-0.21	5.34E-01	0.01	9.12E-01
rs1152582	*SYNE2* Transcribed	64692630	-0.08	1.35E-01	-0.20	4.31E-01	0.08	2.81E-01
rs928554	*ESR2* Transcribed	64694195	-0.06	2.39E-01	-0.23	2.42E-01	0.05	4.70E-01
rs4986938	*ESR2* Transcribed	64699816	0.11	3.91E-02	**0.47**	**1.39E-03**	-0.13	8.49E-02
rs1256064		64700739	-0.07	3.46E-01	-0.35	2.09E-01	0.07	5.17E-01
rs1256063		64702217	-0.31	2.35E-03			-0.18	4.34E-01
rs10144225		64704994	0.10	1.94E-01	-0.45	1.62E-01	0.01	9.21E-01
rs1256049	*ESR2* Transcribed	64724051	-0.43	6.57E-03	-0.17	7.07E-01	-0.08	7.26E-01

The significance threshold after correction for multiple testing was 1.52E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.31 Associations with total expression of *ESR2* in chromosome band 14q23 in tissue

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs4459477		64346190	-0.06	6.37E-01			-0.55	3.19E-01
rs17750888		64355420	0.08	3.09E-01	-0.14	6.64E-01	-0.31	2.72E-01
rs12588807		64363070	-0.08	6.16E-01			-0.62	3.08E-01
rs2275018		64375985	0.11	2.02E-01	-0.12	6.67E-01	-0.17	5.38E-01
rs7153036		64377548	0.05	5.90E-01	0.31	1.97E-01	0.00	9.90E-01
rs12879919		64382350	-0.14	2.00E-01	-0.11	7.95E-01	-0.48	2.14E-01
rs11158521		64401255	0.02	8.39E-01	-0.28	3.19E-01	-0.09	7.38E-01
rs2184292		64408135	0.09	3.19E-01	0.18	5.84E-01	0.06	8.35E-01
rs7153680		64474227	0.02	8.12E-01	0.13	6.16E-01	-0.15	5.53E-01
rs4566057		64506541	-0.01	9.08E-01	-0.78	8.65E-02	-0.41	3.44E-01
rs1890908		64519035	0.03	8.27E-01			-0.56	3.84E-01
rs10137972		64557734	-0.03	8.61E-01	0.30	6.03E-01	0.14	8.04E-01
rs17101651		64561324	0.09	5.83E-01	-0.03	9.55E-01	-0.23	6.73E-01
rs8022428		64563613	-0.01	8.88E-01	-0.47	2.67E-01	-0.16	6.43E-01
rs8007972		64564527	0.05	6.75E-01	-0.17	7.70E-01	-0.41	3.13E-01
rs8016917		64569951	-0.02	9.28E-01			0.08	9.12E-01
rs10142318		64604281	0.09	4.27E-01	-0.46	2.73E-01	0.05	8.98E-01
rs7144688		64613473	-0.08	4.44E-01	-0.31	3.78E-01	0.67	6.20E-02
rs11158533		64622988	-0.04	5.81E-01	-0.15	6.20E-01	0.05	8.58E-01
rs17751454		64625181	-0.02	8.60E-01	0.16	6.24E-01	0.04	8.89E-01
rs3866743		64633322	0.20	4.76E-01			0.81	3.57E-01
rs7161192	*SYNE2* Transcribed	64637147	-0.06	4.33E-01	-0.45	4.90E-01	0.24	3.68E-01
rs2256191		64656855	0.07	4.38E-01	-0.26	2.98E-01	0.23	3.53E-01

rs1152591	AF GWAS	64680848	0.00	9.69E-01	0.31	2.56E-01	-0.61	2.23E-02
rs7145919		64687601	-0.14	3.32E-01	-0.54	1.87E-01	0.75	8.05E-02
rs12434245		64691853	-0.13	3.54E-01	-0.54	1.87E-01	0.78	6.56E-02
rs1152582	*SYNE2* Transcribed	64692630	0.02	7.90E-01	0.51	6.89E-02	-0.60	2.93E-02
rs928554	*ESR2* Transcribed	64694195	0.00	9.75E-01	0.43	5.94E-02	-0.54	3.72E-02
rs4986938	*ESR2* Transcribed	64699816	-0.04	6.16E-01	-0.09	7.26E-01	0.32	2.18E-01
rs1256064		64700739	0.12	2.92E-01	-0.36	3.27E-01	-0.18	6.58E-01
rs1256063		64702217	-0.10	5.32E-01	-0.09	7.84E-01	-0.78	2.96E-01
rs10144225		64704994	0.25	2.98E-02	-0.49	3.10E-01	-0.05	8.96E-01
rs1256049	*ESR2* Transcribed	64724051	-0.25	3.23E-01	-0.04	9.49E-01	-1.03	2.49E-01

The significance threshold after correction for multiple testing was 1.52E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.32 Associations with allelic expression ratio of *SYNE2* in chromosome band 14q23 in tissue

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs4459477		64346190	-0.03	3.17E-01			0.03	6.55E-01
rs17750888		64355420	-0.02	3.71E-01	0.05	2.37E-01	-0.02	6.55E-01
rs12588807		64363070	0.00	1.00E+00			0.07	3.17E-01
rs2275018		64375985	0.02	3.17E-01	0.06	2.06E-01	0.01	6.55E-01
rs7153036		64377548	0.03	3.17E-01	0.07	2.06E-01	-0.03	5.27E-01
rs12879919		64382350	-0.03	1.80E-01	-0.06	5.27E-01	-0.02	6.55E-01
rs11158521		64401255	-0.01	6.55E-01	0.03	5.27E-01	-0.06	1.57E-01
rs2184292		64408135	0.03	1.80E-01	0.09	3.59E-02	-0.02	6.55E-01
rs7153680		64474227	0.03	2.06E-01	0.09	3.59E-02	0.01	6.55E-01
rs4566057		64506541	-0.12	2.42E-03	0.04	6.55E-01	-0.19	7.29E-03
rs1890908		64519035	-0.09	6.52E-03			-0.19	3.36E-03
rs10137972		64557734	0.01	6.55E-01	0.15	4.04E-02	-0.08	4.39E-01
rs17101651		64561324	0.01	1.00E+00	0.05	6.55E-01	-0.02	6.55E-01
rs8022428		64563613	-0.05	1.07E-01	0.09	8.33E-02	-0.10	3.59E-02
rs8007972		64564527	-0.06	2.26E-02	-0.21	5.27E-01	-0.05	3.71E-01
rs8016917		64569951	0.12	5.13E-02			-0.16	1.38E-01
rs10142318		64604281	-0.08	5.22E-03	0.03	1.00E+00	**-0.15**	**1.26E-03**
rs7144688		64613473	**-0.08**	**3.74E-05**	0.08	8.33E-02	-0.10	3.01E-03
rs11158533		64622988	0.00	1.00E+00	0.10	1.43E-02	0.01	1.00E+00
rs17751454		64625181	0.04	1.07E-01	0.10	1.21E-01	0.04	5.27E-01
rs3866743		64633322	0.11	3.17E-01			-0.18	2.73E-01
rs7161192	SYNE2 Transcribed	64637147	0.00	1.00E+00	0.10	1.43E-02	0.03	6.55E-01
rs2256191		64656855	-0.02	2.73E-01	0.06	2.37E-01	-0.07	4.55E-02

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs1152591	AF GWAS	64680848	0.03	5.13E-02	-0.09	6.52E-03	0.09	1.95E-03
rs7145919		64687601	-0.05	1.57E-01	0.14	4.55E-02	-0.06	3.17E-01
rs12434245		64691853	-0.05	1.57E-01	0.14	4.55E-02	-0.06	3.17E-01
rs1152582	SYNE2 Transcribed	64692630	**0.05**	**5.13E-05**	-0.08	8.15E-03	**0.12**	**1.07E-06**
rs928554	ESR2 Transcribed	64694195	**0.05**	**3.86E-04**	-0.08	8.15E-03	**0.11**	**2.45E-05**
rs4986938	ESR2 Transcribed	64699816	-0.02	3.17E-01	0.08	1.38E-01	-0.10	9.12E-03
rs1256064		64700739	**-0.11**	**8.60E-06**	0.04	5.27E-01	-0.10	2.85E-02
rs1256063		64702217	-0.01	6.55E-01			0.10	2.06E-01
rs10144225		64704994	**-0.10**	**1.48E-04**	0.03	1.00E+00	-0.11	2.53E-02
rs1256049	ESR2 Transcribed	64724051	-0.11	1.21E-01	0.05	6.55E-01	0.04	6.55E-01

The significance threshold after correction for multiple testing was 1.52E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.33 Associations with allelic expression ratio of *ESR2* in chromosome band 14q23 in tissue

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs4459477		64346190	0.00	1.00E+00			0.04	5.27E-01
rs17750888		64355420	0.01	5.27E-01	0.05	1.07E-01	0.01	1.00E+00
rs12588807		64363070	0.03	4.39E-01			0.01	1.00E+00
rs2275018		64375985	-0.01	6.55E-01	0.05	1.07E-01	-0.02	6.55E-01
rs7153036		64377548	-0.01	6.55E-01	0.02	1.00E+00	0.01	1.00E+00
rs12879919		64382350	0.04	1.21E-01	-0.04	6.55E-01	0.05	4.39E-01
rs11158521		64401255	0.03	2.37E-01	0.04	3.17E-01	0.01	1.00E+00
rs2184292		64408135	0.00	1.00E+00	0.08	4.68E-03	0.00	1.00E+00
rs7153680		64474227	0.01	6.55E-01	0.08	4.68E-03	0.01	1.00E+00
rs4566057		64506541	-0.03	5.27E-01	0.04	6.55E-01	0.10	2.06E-01
rs1890908		64519035	0.06	8.33E-02			0.04	6.55E-01
rs10137972		64557734	-0.09	3.20E-02	0.02	1.00E+00	0.09	5.27E-01
rs17101651		64561324	**-0.12**	**1.33E-04**	0.04	6.55E-01	0.07	4.39E-01
rs8022428		64563613	-0.07	2.01E-02	0.03	6.55E-01	0.05	3.71E-01
rs8007972		64564527	0.02	5.27E-01			0.02	6.55E-01
rs8016917		64569951	0.07	1.80E-01			-0.06	5.27E-01
rs10142318		64604281	0.05	6.52E-02	0.01	1.00E+00	-0.04	4.39E-01
rs7144688		64613473	0.00	1.00E+00	0.02	6.55E-01	-0.03	4.39E-01
rs11158533		64622988	-0.01	6.55E-01	0.07	4.68E-03	-0.02	5.27E-01
rs17751454		64625181	0.04	8.33E-02	0.09	1.02E-03	0.02	6.55E-01
rs3866743		64633322	0.09	1.80E-01			-0.06	5.27E-01
rs7161192	SYNE2 Transcribed	64637147	-0.01	6.55E-01	0.07	4.68E-03	-0.02	6.55E-01
rs2256191		64656855	0.03	8.33E-02	0.06	1.21E-01	0.03	5.27E-01

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs1152591	AF GWAS	64680848	-0.01	5.27E-01	-0.07	7.29E-03	0.00	1.00E+00
rs7145919		64687601	-0.08	4.04E-02	0.02	1.00E+00	-0.07	2.06E-01
rs12434245		64691853	-0.08	4.04E-02	0.02	1.00E+00	-0.07	2.06E-01
rs1152582	SYNE2 Transcribed	64692630	-0.02	1.57E-01	-0.05	3.59E-02	0.01	6.55E-01
rs928554	ESR2 Transcribed	64694195	-0.02	1.21E-01	-0.05	3.59E-02	0.02	5.27E-01
rs4986938	ESR2 Transcribed	64699816	0.04	2.26E-02	0.08	1.43E-02	0.03	3.71E-01
rs1256064		64700739	0.03	2.37E-01	0.02	1.00E+00	-0.05	3.71E-01
rs1256063		64702217	-0.03	3.71E-01			-0.11	2.06E-01
rs10144225		64704994	0.06	2.01E-02	0.01	1.00E+00	-0.04	5.27E-01
rs1256049	ESR2 Transcribed	64724051	-0.09	4.55E-02	0.04	6.55E-01	0.00	1.00E+00

The significance threshold after correction for multiple testing was 1.52E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.34 Associations with total expression of *HCN4* in chromosome band 15q24.

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs8040516		73598548	0.03	7.36E-01	-0.58	5.23E-01		
rs3743496	*HCN4* Transcribed	73614408	0.04	6.12E-01	-0.41	3.29E-01		
rs35177144	*HCN4* Transcribed	73614704	0.04	6.22E-01	-0.41	3.29E-01		
rs529004	*HCN4* Transcribed	73614834	0.20	2.61E-01	0.42	5.49E-01		
rs488156		73625895	0.08	2.85E-01	0.51	8.28E-02		
rs2623998		73628086	-0.05	5.88E-01	0.40	4.75E-01		
rs12905211		73628168	-0.02	8.36E-01	-0.58	1.21E-01		
rs8030574		73628214	0.08	3.67E-01	0.55	2.95E-01		
rs2623997		73628714	0.06	4.13E-01	0.38	5.00E-01		
rs478438		73630686	-0.27	1.33E-02	-0.86	7.42E-02		
rs3784807		73631830	0.10	2.60E-01	0.37	4.25E-01		
rs11857639		73637772	-0.12	3.91E-01	0.39	6.46E-01		
rs3826046		73650824	**0.47**	**4.24E-04**	0.40	4.75E-01		
rs7164883	AF GWAS	73652174	-0.10	2.88E-01	-0.58	1.20E-01		
rs12440104		73653009	0.11	1.84E-01	0.25	5.33E-01		
rs2680344		73653485	-0.09	2.85E-01	-0.40	3.36E-01		
rs11638230		73656261	0.10	4.41E-01	-0.74	3.73E-01		

The significance threshold after correction for multiple testing was 2.94E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.35 Associations with allelic expression ratio of *HCN4* in chromosome band 15q24.

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs8040516		73598548	**-0.08**	**1.83E-04**	**-0.03**	**1.64E-04**		
rs3743496	HCN4 Transcribed	73614408	**-0.08**	**1.62E-06**	**-0.03**	**1.48E-04**		
rs35177144	HCN4 Transcribed	73614704	**-0.08**	**1.62E-06**	**-0.03**	**1.48E-04**		
rs529004	HCN4 Transcribed	73614834	0.11	5.22E-03	0.27	1.00E+00		
rs488156		73625895	**0.11**	**2.26E-07**	0.03	2.06E-01		
rs2623998		73628086	0.08	1.13E-03	0.03	5.27E-01		
rs12905211		73628168	**0.09**	**1.83E-04**	0.03	5.78E-02		
rs8030574		73628214	0.08	2.26E-02	0.03	6.55E-01		
rs2623997		73628714	**-0.07**	**9.11E-04**	**-0.03**	**1.48E-04**		
rs478438		73630686	0.04	3.17E-01	0.03	1.60E-02		
rs3784807		73631830	-0.06	9.12E-03	-0.03	4.55E-02		
rs11857639		73637772	0.09	2.26E-02				
rs3826046		73650824	-0.03	5.27E-01	-0.03	4.39E-01		
rs7164883	AF GWAS	73652174	0.02	6.55E-01	0.04	3.71E-01		
rs12440104		73653009	-0.06	2.26E-02	-0.03	1.21E-01		
rs2680344		73653485	-0.01	1.00E+00	0.02	1.00E+00		
rs11638230		73656261	-0.03	6.55E-01	-0.03	4.39E-01		

The significance threshold after correction for multiple testing was 2.94E-03. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.36 Associations with total expression of *ZFHX3* in chromosome band 16q22 in tissue

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs117951282		72820328	-0.32	2.25E-01				
rs699444		72827758	0.04	4.11E-01	-0.06	8.22E-01	0.05	4.68E-01
rs740178	*ZFHX3* Transcribed	72832135	0.02	6.77E-01	0.05	6.50E-01	0.15	3.15E-02
rs12929452		72838680	-0.05	1.41E-01	0.05	6.12E-01	-0.22	1.80E-04
rs2266943		72854307	0.02	5.84E-01	0.14	4.21E-01	0.16	2.02E-02
rs8058014		72857544	0.02	6.23E-01	0.05	6.50E-01	0.15	3.57E-02
rs1476646		72898348	-0.02	5.34E-01	0.01	9.09E-01	-0.10	7.20E-02
rs4788668		72901874	-0.02	5.72E-01	0.08	4.34E-01	0.10	9.82E-02
rs4788671		72915085	-0.01	7.14E-01	-0.06	5.61E-01	-0.08	2.02E-01
rs6499594		72917601	0.04	4.58E-01	-0.48	4.31E-02	0.05	6.13E-01
rs16971366		72919582	-0.04	3.16E-01	0.07	4.14E-01	-0.08	2.10E-01
rs9940310		72925982	0.02	7.46E-01	-0.48	4.31E-02	0.05	6.88E-01
rs9925261		72927533	-0.03	4.34E-01	0.03	7.78E-01	-0.04	5.70E-01
rs16971384		72931085	-0.04	2.67E-01	0.07	3.89E-01	-0.08	2.19E-01
rs4788482		72937079	-0.01	8.79E-01	-0.06	7.85E-01	-0.14	2.07E-01
rs4788679		72956544	-0.01	8.74E-01	-0.02	8.19E-01	-0.08	3.25E-01
rs9936884		72960230	-0.01	8.26E-01	-0.06	8.22E-01	0.20	5.61E-02
rs4788488		72960283	-0.01	7.61E-01	-0.03	7.61E-01	0.01	8.57E-01
rs8055870		72972090	0.03	4.47E-01	0.00	9.97E-01	0.02	6.95E-01
rs6499600		72979374	0.01	7.49E-01	-0.01	9.05E-01	-0.04	4.37E-01
rs13336412		72981949	0.01	7.30E-01	0.20	2.26E-02	0.02	6.92E-01
rs2228200		72984668	0.08	5.18E-01			0.14	5.69E-01
rs2106258		72990553	0.00	9.16E-01	-0.09	3.73E-01	-0.15	4.63E-02

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs2157786		72991286	-0.03	3.79E-01	-0.07	4.58E-01	-0.02	8.13E-01
rs10852515	*ZFHX3a* Transcribed	72991660	-0.02	6.40E-01	-0.02	9.17E-01	-0.13	1.50E-01
rs7193297		72993831	0.02	5.38E-01	0.15	6.60E-02	0.03	6.26E-01
rs7404992		72994419	-0.05	4.53E-01			0.10	3.70E-01
rs4788683		72997747	0.08	2.76E-02	0.04	6.59E-01	0.03	5.72E-01
rs9921395		73001957	-0.02	5.44E-01	0.11	2.82E-01	0.01	9.14E-01
rs12445932		73004432	0.05	1.60E-01	0.02	7.98E-01	0.05	4.19E-01
rs7199343	Kawasaki GWAS	73009024	-0.07	5.80E-02	0.10	2.52E-01	0.06	3.37E-01
rs11075954		73012164	0.05	1.17E-01	-0.12	2.93E-01	-0.02	7.33E-01
rs2040508		73012685	-0.06	1.06E-01	0.00	9.93E-01	0.04	5.67E-01
rs16971456		73013036	0.02	7.75E-01	0.02	9.29E-01	-0.04	6.37E-01
rs9930445		73013482	-0.03	3.30E-01	0.06	4.48E-01	-0.01	8.61E-01
rs4788684		73013633	-0.03	4.43E-01	-0.04	7.25E-01	-0.02	7.95E-01
rs16971464		73016143	0.07	2.72E-01	-0.16	2.80E-01	-0.01	9.50E-01
rs16971465		73017061	-0.01	8.01E-01	0.02	9.09E-01	0.02	8.39E-01
rs4788489		73017118	-0.02	5.01E-01	-0.06	5.43E-01	0.02	7.60E-01
rs16971474		73019004	0.07	1.24E-01	-0.03	8.75E-01	-0.08	2.67E-01
rs11640106		73020116	-0.07	3.72E-02	0.12	3.03E-01	0.04	4.35E-01
rs1858800		73024276	0.06	9.26E-02	-0.18	2.38E-01	0.00	9.87E-01
rs756720		73028921	0.00	9.85E-01	0.00	9.99E-01	0.00	9.41E-01
rs7193343	AF GWAS	73029160	0.03	5.39E-01	0.02	9.29E-01	-0.01	8.60E-01
rs11075958		73033869	-0.01	8.76E-01	0.01	8.78E-01	-0.04	5.41E-01
rs8056528		73036633	0.00	9.92E-01	-0.02	8.29E-01	0.03	6.84E-01
rs719353		73042551	0.00	9.58E-01	-0.02	8.51E-01	0.01	9.09E-01
rs4788689		73049830	-0.02	7.55E-01	0.06	7.61E-01	-0.01	9.39E-01

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs2106261	AF GWAS	73051620	0.03	5.44E-01	0.02	9.29E-01	0.02	8.20E-01
rs11863932		73053579	0.11	2.17E-01			-0.09	5.06E-01
rs1548373		73059861	0.01	7.83E-01	-0.05	6.93E-01	0.05	3.97E-01
rs4788692		73065656	-0.10	1.78E-01	0.06	7.61E-01	-0.01	9.20E-01
rs12373097		73068515	-0.01	7.91E-01	-0.18	2.87E-01	0.08	2.97E-01
rs8057081		73068977	0.02	6.99E-01	0.02	9.29E-01	-0.04	6.37E-01
rs4788696		73070310	0.02	6.12E-01	-0.05	6.64E-01	0.01	9.28E-01
rs8060701		73073289	-0.01	9.15E-01			-0.08	4.66E-01
rs9940321		73073808	0.06	1.28E-01	-0.20	2.28E-01	0.03	6.45E-01
rs9940520		73074012	-0.07	1.62E-01	0.06	8.22E-01	-0.04	5.93E-01
rs11641701		73079212	0.08	1.93E-02	-0.07	6.49E-01	0.02	6.94E-01
rs4788697		73087494	0.08	3.44E-02	-0.18	2.28E-01	-0.07	2.92E-01
rs8052905		73097663	-0.12	8.74E-02	0.06	8.22E-01	0.00	9.80E-01
rs739414		73097956	0.02	6.32E-01	0.20	2.41E-02	0.04	5.06E-01
rs8051826		73102456	-0.02	8.20E-01	0.07	6.82E-01	0.05	5.94E-01

The significance threshold after correction for multiple testing was 7.90E-04. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.37 Associations with total expression of *ZFHX3a* in chromosome band 16q22 in tissue

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs117951282		72820328	-0.12	7.37E-01				
rs699444		72827758	0.05	3.67E-01	-0.16	6.92E-01	0.24	1.86E-02
rs740178	*ZFHX3* Transcribed	72832135	0.05	3.45E-01	0.14	4.37E-01	0.26	9.11E-03
rs12929452		72838680	-0.03	5.51E-01	0.04	7.71E-01	-0.24	5.15E-03
rs2266943		72854307	0.06	2.82E-01	0.28	2.64E-01	0.27	5.39E-03
rs8058014		72857544	0.05	3.21E-01	0.14	4.37E-01	0.26	7.69E-03
rs1476646		72898348	0.01	8.88E-01	-0.02	8.88E-01	-0.16	3.49E-02
rs4788668		72901874	-0.04	4.13E-01	0.15	3.36E-01	0.16	6.93E-02
rs4788671		72915085	0.02	5.83E-01	-0.10	5.18E-01	-0.17	4.79E-02
rs6499594		72917601	0.11	1.67E-01	-0.47	2.19E-01	0.04	8.10E-01
rs16971366		72919582	-0.02	6.39E-01	0.17	1.62E-01	-0.05	5.95E-01
rs9940310		72925982	0.09	2.58E-01	-0.47	2.19E-01	0.00	9.83E-01
rs9925261		72927533	0.02	7.46E-01	0.14	2.89E-01	0.05	5.61E-01
rs16971384		72931085	-0.02	6.54E-01	0.17	1.52E-01	0.03	7.41E-01
rs4788482		72937079	0.09	2.46E-01	0.14	6.66E-01	-0.01	9.59E-01
rs4788679		72956544	0.04	4.61E-01	0.07	5.63E-01	0.08	4.97E-01
rs9936884		72960230	-0.06	4.90E-01	-0.16	6.92E-01	0.19	1.92E-01
rs4788488		72960283	0.03	5.89E-01	0.06	6.51E-01	0.11	2.46E-01
rs8055870		72972090	0.03	5.25E-01	-0.10	4.06E-01	-0.08	3.36E-01
rs6499600		72979374	0.03	4.94E-01	0.15	3.43E-01	0.03	6.87E-01
rs13336412		72981949	0.02	6.31E-01	0.25	6.84E-02	-0.03	7.17E-01
rs2228200		72984668	0.01	9.44E-01			0.26	4.54E-01
rs2106258		72990553	0.07	1.69E-01	-0.01	9.26E-01	0.03	7.73E-01

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs2157786		72991286	-0.02	7.09E-01	0.02	9.07E-01	0.13	1.69E-01
rs10852515	*ZFHX3a* Transcribed	72991660	0.03	5.91E-01	0.19	5.22E-01	0.08	5.33E-01
rs7193297		72993831	0.02	5.76E-01	0.22	8.63E-02	-0.02	8.22E-01
rs7404992		72994419	0.03	7.77E-01			0.15	3.31E-01
rs4788683		72997747	0.13	3.16E-03	0.08	5.13E-01	0.06	5.09E-01
rs9921395		73001957	-0.07	1.15E-01	0.19	1.83E-01	0.03	6.88E-01
rs12445932		73004432	0.03	5.77E-01	0.01	9.26E-01	-0.05	6.23E-01
rs7199343	Kawasaki GWAS	73009024	-0.05	2.62E-01	0.15	2.39E-01	0.07	4.15E-01
rs11075954		73012164	0.06	1.52E-01	-0.17	3.23E-01	-0.01	9.31E-01
rs2040508		73012685	-0.08	1.09E-01	0.05	7.74E-01	0.14	1.45E-01
rs16971456		73013036	0.03	6.35E-01	-0.10	8.02E-01	-0.07	5.87E-01
rs9930445		73013482	0.02	6.27E-01	0.06	6.09E-01	-0.04	6.34E-01
rs4788684		73013633	-0.05	2.31E-01	0.00	9.90E-01	-0.01	9.36E-01
rs16971464		73016143	-0.12	1.86E-01	-0.13	6.07E-01	-0.04	7.96E-01
rs16971465		73017061	-0.10	1.71E-01	0.22	4.45E-01	-0.04	7.89E-01
rs4788489		73017118	-0.03	4.81E-01	-0.10	4.79E-01	0.04	5.84E-01
rs16971474		73019004	0.12	4.66E-02	0.00	9.98E-01	-0.03	8.04E-01
rs11640106		73020116	-0.10	2.62E-02	0.20	2.57E-01	0.04	5.68E-01
rs1858800		73024276	0.09	6.30E-02	-0.32	1.50E-01	-0.03	6.83E-01
rs756720		73028921	0.05	2.49E-01	0.01	9.34E-01	0.01	8.92E-01
rs7193343	AF GWAS	73029160	0.02	7.99E-01	-0.10	8.02E-01	-0.02	8.41E-01
rs11075958		73033869	0.06	2.77E-01	0.01	9.55E-01	-0.09	3.82E-01
rs8056528		73036633	-0.05	2.86E-01	0.03	8.28E-01	0.07	4.20E-01
rs719353		73042551	-0.04	4.07E-01	0.02	8.98E-01	0.05	5.65E-01
rs4788689		73049830	0.04	5.22E-01	0.13	6.72E-01	0.14	2.80E-01

			RAA eQTL		LAA eQTL		IMA eQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs2106261	AF GWAS	73051620	0.02	7.74E-01	-0.10	8.02E-01	0.01	9.47E-01
rs11863932		73053579	0.15	1.71E-01			0.00	9.90E-01
rs1548373		73059861	-0.07	1.33E-01	-0.02	9.19E-01	0.02	8.47E-01
rs4788692		73065656	-0.08	3.81E-01	0.13	6.72E-01	0.21	3.12E-01
rs12373097		73068515	-0.02	6.94E-01	-0.42	7.58E-02	0.06	5.63E-01
rs8057081		73068977	0.03	5.76E-01	-0.10	8.02E-01	-0.02	8.81E-01
rs4788696		73070310	-0.05	3.39E-01	0.00	9.94E-01	0.10	3.63E-01
rs8060701		73073289	0.06	5.20E-01			0.02	9.07E-01
rs9940321		73073808	0.06	2.15E-01	-0.09	7.24E-01	-0.09	3.47E-01
rs9940520		73074012	-0.06	3.35E-01	0.24	5.48E-01	-0.10	3.60E-01
rs11641701		73079212	0.06	1.48E-01	-0.18	4.26E-01	0.09	2.70E-01
rs4788697		73087494	0.12	1.47E-02	-0.07	7.51E-01	-0.13	2.06E-01
rs8052905		73097663	-0.23	9.67E-03	0.24	5.48E-01	-0.17	2.23E-01
rs739414		73097956	0.05	3.30E-01	0.16	2.77E-01	0.17	3.39E-02
rs8051826		73102456	0.07	4.56E-01	0.19	4.47E-01	0.06	6.51E-01

The significance threshold after correction for multiple testing was 7.90E-04. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.38 Associations with allelic expression ratio of *ZFHX3* in chromosome band 16q22 in tissue

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs117951282		72820328	NA	1.00E+00			NA	1.00E+00
rs699444		72827758	0.02	2.10E-01			-0.04	4.01E-03
rs740178	ZFHX3 Transcribed	72832135	0.01	3.85E-01			-0.03	1.28E-02
rs12929452		72838680	-0.01	7.56E-01			0.03	1.47E-01
rs2266943		72854307	0.01	6.56E-01			-0.03	1.15E-02
rs8058014		72857544	0.01	5.77E-01			-0.03	1.15E-02
rs1476646		72898348	-0.02	3.21E-01			0.02	4.05E-01
rs4788668		72901874	0.01	6.27E-01			-0.03	6.80E-02
rs4788671		72915085	-0.03	1.44E-01			0.01	5.84E-01
rs6499594		72917601	-0.03	2.05E-01			0.01	7.37E-01
rs16971366		72919582	-0.02	1.96E-01			-0.03	8.63E-02
rs9940310		72925982	-0.03	2.74E-01			0.00	9.27E-01
rs9925261		72927533	-0.02	3.14E-01			-0.01	4.09E-01
rs16971384		72931085	-0.01	5.02E-01			-0.02	2.22E-01
rs4788482		72937079	-0.05	3.18E-02			0.00	9.75E-01
rs4788679		72956544	-0.03	1.26E-01			0.02	3.76E-01
rs9936884		72960230	0.05	1.64E-01			-0.06	5.78E-02
rs4788488		72960283	-0.01	4.16E-01			-0.02	3.45E-01
rs8055870		72972090	0.01	6.75E-01			0.00	9.11E-01
rs6499600		72979374	-0.01	5.29E-01			0.02	4.61E-01
rs13336412		72981949	0.02	3.31E-01			-0.02	3.35E-01
rs2228200		72984668	-0.06	4.06E-01				
rs2106258		72990553	-0.04	3.15E-02			0.00	9.00E-01

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs2157786		72991286	-0.03	1.05E-01			-0.01	7.67E-01
rs10852515	ZFHX3a Transcribed	72991660	-0.05	2.62E-02			0.01	7.86E-01
rs7193297		72993831	0.02	2.39E-01			-0.05	3.84E-02
rs7404992		72994419	0.03	4.06E-01			-0.06	5.78E-02
rs4788683		72997747	0.00	9.98E-01			-0.04	1.39E-02
rs9921395		73001957	-0.02	2.42E-01			0.00	9.66E-01
rs12445932		73004432	-0.02	2.47E-01			-0.02	1.91E-01
rs7199343	Kawasaki GWAS	73009024	-0.01	7.06E-01			-0.01	5.21E-01
rs11075954		73012164	0.03	1.01E-01			0.00	9.97E-01
rs2040508		73012685	-0.01	6.41E-01			0.00	9.57E-01
rs16971456		73013036	-0.02	5.21E-01			0.01	6.20E-01
rs9930445		73013482	-0.01	4.97E-01			-0.03	1.78E-01
rs4788684		73013633	0.03	1.01E-01			0.00	9.88E-01
rs16971464		73016143	-0.05	5.38E-01			0.03	5.13E-01
rs16971465		73017061	-0.05	2.79E-01			-0.03	4.42E-01
rs4788489		73017118	0.02	3.89E-01			0.03	3.87E-01
rs16971474		73019004	-0.03	3.49E-01			0.03	2.34E-01
rs11640106		73020116	-0.01	5.89E-01			0.00	9.21E-01
rs1858800		73024276	0.02	4.10E-01			0.03	1.24E-01
rs756720		73028921	0.02	4.15E-01			-0.04	7.67E-02
rs7193343	AF GWAS	73029160	-0.03	6.78E-02			0.02	3.72E-01
rs11075958		73033869	0.00	8.66E-01			-0.05	7.47E-02
rs8056528		73036633	0.00	9.41E-01			0.00	9.09E-01
rs719353		73042551	-0.03	1.17E-01			0.01	4.06E-01
rs4788689		73049830	-0.01	8.35E-01			0.03	3.84E-01

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs2106261	AF GWAS	73051620	-0.04	3.84E-02			0.01	5.09E-01
rs11863932		73053579	-0.07	2.05E-01			-0.01	8.68E-01
rs1548373		73059861	-0.03	6.00E-02			0.01	7.85E-01
rs4788692		73065656	0.01	8.42E-01			0.05	2.89E-01
rs12373097		73068515	0.04	8.94E-02			0.02	4.48E-01
rs8057081		73068977	-0.03	1.19E-01			0.00	9.95E-01
rs4788696		73070310	-0.02	3.56E-01			0.02	4.95E-01
rs8060701		73073289	0.04	3.15E-01			-0.05	1.61E-01
rs9940321		73073808	-0.02	3.48E-01			-0.02	2.62E-01
rs9940520		73074012	0.03	3.28E-01			-0.04	1.83E-01
rs11641701		73079212	-0.01	5.96E-01			0.01	4.44E-01
rs4788697		73087494	-0.02	3.64E-01			0.00	8.46E-01
rs8052905		73097663	0.01	8.58E-01			0.02	5.75E-01
rs739414		73097956	0.00	8.65E-01			0.00	9.55E-01
rs8051826		73102456	0.01	8.45E-01			-0.06	1.36E-01

The significance threshold after correction for multiple testing was 7.90E-04. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**

Table A2.39 Associations with allelic expression ratio of *ZFHX3a* in chromosome band 16q22 in tissue

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs117951282		72820328	NA	1.00E+00			NA	1.00E+00
rs699444		72827758	0.08	2.25E-01			0.01	8.83E-01
rs740178	ZFHX3 Transcribed	72832135	0.10	7.29E-02			0.01	8.35E-01
rs12929452		72838680	0.03	5.05E-01			0.02	7.53E-01
rs2266943		72854307	0.12	4.24E-02			0.01	8.35E-01
rs8058014		72857544	0.10	7.29E-02			0.01	8.53E-01
rs1476646		72898348	-0.02	7.10E-01			-0.06	3.91E-01
rs4788668		72901874	0.04	3.38E-01			0.09	2.00E-01
rs4788671		72915085	-0.01	7.43E-01			-0.08	2.65E-01
rs6499594		72917601	-0.01	8.64E-01			-0.16	2.07E-01
rs16971366		72919582	0.13	3.30E-03			0.09	7.59E-02
rs9940310		72925982	-0.05	6.08E-01			-0.24	1.85E-01
rs9925261		72927533	0.06	1.30E-01			0.08	1.46E-01
rs16971384		72931085	0.08	6.12E-02			0.10	7.68E-02
rs4788482		72937079	0.09	1.03E-02			0.12	1.37E-02
rs4788679		72956544	0.08	5.17E-02			0.12	1.53E-02
rs9936884		72960230	-0.18	8.38E-03			-0.03	7.44E-01
rs4788488		72960283	0.03	4.93E-01			0.12	1.88E-02
rs8055870		72972090	-0.06	2.80E-01			-0.15	5.97E-02
rs6499600		72979374	0.05	3.14E-01			0.12	6.34E-02
rs13336412		72981949	-0.12	1.48E-02			-0.06	4.14E-01
rs2228200		72984668						
rs2106258		72990553	0.06	8.99E-02			0.11	8.33E-03

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs2157786		72991286	0.04	2.37E-01			0.14	2.91E-03
rs10852515	ZFHX3a Transcribed	72991660	0.07	4.02E-02			0.11	1.16E-02
rs7193297		72993831	-0.15	1.41E-02			-0.11	3.06E-01
rs7404992		72994419	-0.13	8.89E-02			-0.03	7.44E-01
rs4788683		72997747	0.04	3.78E-01			0.07	1.75E-01
rs9921395		73001957	0.03	5.26E-01			0.11	4.45E-02
rs12445932		73004432	0.06	2.27E-01			0.10	7.65E-02
rs7199343	Kawasaki GWAS	73009024	-0.04	4.90E-01			0.00	9.72E-01
rs11075954		73012164	0.04	3.99E-01			0.02	7.22E-01
rs2040508		73012685	0.00	9.98E-01			0.00	9.79E-01
rs16971456		73013036	-0.05	5.50E-01			-0.06	6.02E-01
rs9930445		73013482	-0.02	6.82E-01			-0.10	2.61E-01
rs4788684		73013633	0.07	1.10E-01			-0.07	2.63E-01
rs16971464		73016143	0.04	7.22E-01				
rs16971465		73017061	0.02	8.34E-01				
rs4788489		73017118	0.04	3.56E-01			-0.05	5.78E-01
rs16971474		73019004	-0.06	4.04E-01			0.10	2.70E-01
rs11640106		73020116	-0.04	3.68E-01			0.00	9.91E-01
rs1858800		73024276	0.05	3.83E-01			-0.03	7.53E-01
rs756720		73028921	0.04	5.02E-01			-0.06	3.31E-01
rs7193343	AF GWAS	73029160	0.00	9.97E-01			-0.01	9.07E-01
rs11075958		73033869	-0.05	4.29E-01			-0.08	3.21E-01
rs8056528		73036633	0.03	6.05E-01			0.00	9.72E-01
rs719353		73042551	0.02	6.96E-01			0.00	9.54E-01
rs4788689		73049830	0.02	7.79E-01			0.02	8.08E-01

			RAA aerQTL		LAA aerQTL		IMA aerQTL	
SNP	Importance	Position	Beta	p value	Beta	p value	Beta	p value
rs2106261	AF GWAS	73051620	0.03	6.38E-01			-0.01	9.07E-01
rs11863932		73053579	-0.01	8.93E-01			0.15	6.07E-02
rs1548373		73059861	0.04	4.97E-01			-0.02	7.67E-01
rs4788692		73065656	-0.13	7.28E-02			-0.01	9.09E-01
rs12373097		73068515	0.03	6.36E-01			-0.15	1.51E-01
rs8057081		73068977	0.08	1.49E-01			-0.02	7.27E-01
rs4788696		73070310	-0.01	8.01E-01			0.12	5.97E-02
rs8060701		73073289	-0.03	7.43E-01			0.08	3.11E-01
rs9940321		73073808	0.00	9.85E-01			0.07	1.88E-01
rs9940520		73074012	-0.03	7.43E-01			0.08	3.11E-01
rs11641701		73079212	-0.04	3.82E-01			0.07	2.49E-01
rs4788697		73087494	0.05	2.82E-01			0.01	9.23E-01
rs8052905		73097663	-0.12	3.56E-01			0.09	4.71E-01
rs739414		73097956	-0.05	4.60E-01			-0.02	8.15E-01
rs8051826		73102456	-0.05	4.55E-01			0.01	8.94E-01

The significance threshold after correction for multiple testing was 7.90E-04. Effects are presented as log effect sizes (**Beta**) of the minor allele referenced to the major allele. Significant associations are **in bold.**