Integrated transcriptional profiling and linkage analysis for disease gene identification in a rat model of metabolic syndrome

Fiona Maciver

Thesis submitted to the University of London
for the degree of Doctor of Philosophy

September 2006

Physiological Genomics and Medicine Group
Division of Clinical Sciences Centre
Imperial College Faculty of Medicine
Hammersmith Hospital
Du Cane Road
London W12 0NN

Abstract

Development of the metabolic syndrome is determined by a complex interaction of environmental and genetic factors and as with other complex human traits, disease gene identification is problematic due to a lack of statistical power and genetic heterogeneity in patients. Identifying genes that underlie metabolic syndrome risk factors will allow better understanding and development of treatments which would decrease disease health burden. The spontaneously hypertensive rat (SHR) is a widely studied model of insulin resistance and hypertension and, as in humans, the hypertensive phenotype clusters with metabolic traits. SHR was crossed with normotensive Brown Norway rats to develop the BXH/HXB recombinant inbred (RI) strain panel.

In these studies this RI panel is being used to identify genomic regions influencing gene expression, termed expression quantitative trait loci (eQTL). Affymetrix microarray expression profiles from two tissues (fat and kidney) in the RI and parental strains were generated. Transcript abundance for each gene on the microarray was treated as a quantitative trait and these eQTL were defined as: *cis*-acting, and *trans*-acting. This work examines a subset of *cis*-acting eQTL ($P<10^{-4}$) to look for polymorphisms in the putative promoters of these genes. These polymorphisms may cause the differential expression detected in these strains and underlie the phenotypic variation seen.

In silico comparative mapping was carried out to investigate the applicability of the detected rat *cis*-acting eQTL genes to human disease. This used human QTL for metabolic and cardiovascular phenotypes to identify rat *cis*-acting eQTL and their human orthologs as potential candidate genes. This analysis detected several genes with known involvement in human metabolic syndrome phenotypes and a large number of novel candidate genes that merit testing in human populations.

The results demonstrate an integrated genome-wide expression profiling with linkage analysis approach and its use in identification of candidate genes which may underlie complex traits.

Acknowledgements

I gratefully acknowledge the support of the MRC for funding this studentship.

I would like to thank everyone in the team and the microarray centre team past and present with extra special thanks to Penny for making it through my PhD in the team and her critical reading of my thesis and Helen for all the support and laughing at me when I cry. I also thank Professor Tim Aitman for his supervision during my thesis.

I have lots of other people to thank for their support. Too many to mention all by name but particularly my friends from home you keep me grounded and are always there for me and my American family for late night stressed out phone calls. The Holts for letting me invade your home and seeing me through a bad time, especially Jacquie for all the times when I just needed to be and you let me be whatever that was (and all the wine of course).

My Mum, Dad and Joan who I love very much, I thank all three of you for knowing me so well and all the support both emotional and financial throughout, life, my university days and now PhD time. You knew when I was stressed even when I was being stubbornly able to cope. I also want to thank Mum for putting up with all the times I needed to hit out at someone as she is the usual target and takes the stressed out abuse when it's really not deserved.

Finally, I want to thank Nick I wasn't expecting to find my lobster but here you are making me smile, looking after me, putting up with all the tears and I love you. Also thanks Nick for helping me see it wasn't just me in the thesis boat - we were all sinking together! For arriving into my world I hadn't seen you coming and didn't think I'd let you in when you got here but it just feels right.

I don't think my acknowledgements would be me without some cheesy music reference so to all of you who helped and were there for me you are the wind beneath my wings – you know who you are.

That's enough with the mushy stuff now – please enjoy the rest of the thesis a lot of blood sweat and mainly tears went into it and not just mine. This is for all of you!!

Preface

The work presented in **Chapters 3** and **5** has been published:

N Hübner, CA Wallace, H Zimdahl, E Petretto, H Schulz, **F Maciver**, M Mueller, O Hummel, J Monti, V Zidek, A Musilova, V Kren, H Causton, L Game, G Born, S Schmidt, A Muller, SA Cook, TW Kurtz, J Whittaker, M Pravenec, TJ Aitman. Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. *Nature Genetics* 2005, 37: 243-253

Presentation of the work in **Chapter 3** at the Genetic Systems in the Rat meeting in Copenhagen, Denmark in September 2004 and at the Science, Education and Technology for Britain poster competition at the House of Commons, UK in September 2004.

Statement of contribution

I carried out all optimisation protocols described in **Chapter 2 and section 3.2.1** including which RNA extraction protocol to use and testing the Bioanalyser. I generated and assessed for quality expression profiles for 66 of the 130 rat fat tissue samples with Dr Caroline Wallace carrying out the remaining 64.

Two methods of data analysis were compared and I carried out the initial comparisons with the first data set of 92 samples including 6 parental strain samples (**Chapter 3.2.2**). The false discovery rate (FDR) and reproducibility experiments were carried out in collaboration with Dr Enrico Petretto in the group.

I carried out the Map Manager QTX analysis and sequencing on the first 47 genes determined to be differentially expressed (**3.2.3**). I selected the genes to be tested for linkage using fold change and *P*-value to determine differential expression, carried out Map Manager analysis and mapped the genes to the genome. I normalised the final data set of 130 expression profiles using RMA and Dr Enrico Petretto performed the QTL Reaper and FDR analysis. Collaboratively with Drs Caroline Wallace and Enrico Petretto and Michael Muller I designed an algorithm to remove redundancy from the linkage data set (**3.2.3, page 91**).

I carried out all parts of the target probe set sequencing and the tests to determine if the polymorphisms were significant that are described in **Chapter 3.2.4**.

I selected, characterised and sequenced all of the genes described in **Chapter 4.**

For the work described in **Chapter 5**, which is *in-silco* comparative mapping of eQTL genes to the human genome, I collated all known SHR and human physiological QTL for metabolic and cardiovascular disease phenotypes. I then designed a method of comparative mapping of the human pQTL to identify human genes from rat eQTL which lie within disease regions. Nicholas J Dickens of the physiological genomics and medicine group designed a program to automate the process.

Table of Contents

ABSTRACT **2**

ACKNOWLEDGEMENTS **3**

PREFACE **4**

LIST OF FIGURES **10**

LIST OF TABLES **11**

ABBREVIATIONS **12**

CHAPTER 1: INTRODUCTION **14**

1.1 GENETIC DISSECTION OF COMPLEX TRAITS 14

1.1.1 Identifying complex trait genes *15*

1.1.1.1 Candidate gene approach 16

1.1.1.2 Genome wide screen – QTL analysis 17

1.1.1.3 Linkage Analysis and Association Studies 17

1.2 METABOLIC SYNDROME, CARDIOVASCULAR DISEASE AND TYPE 2 DIABETES 19

1.2.1 Hypertension *21*

1.2.2 Insulin Resistance *26*

1.2.3 Dyslipidaemia *28*

1.2.4 Obesity *29*

1.3 ANIMAL MODELS 30

1.3.1 Rat Models of Metabolic Syndrome *31*

1.3.1.1 Dahl Salt Sensitive Rat 32

1.3.1.2 Fructose fed rats 33

1.3.1.3 Lyon hypertensive rat 33

1.3.1.4 Spontaneously hypertensive rat 33

1.3.1.5 Stroke prone spontaneously hypertensive rat 34

1.4 THE SPONTANEOUSLY HYPERTENSIVE RAT AS A MODEL OF METABOLIC SYNDROME 35

1.4.1 Hypertension in SHR *36*

1.4.2 Insulin resistance and dyslipidemia in SHR *36*

1.4.3 SHR QTL for metabolic syndrome *39*

1.5 STRATEGIES FOR ANALYSING QTL IN RAT MODELS 41

1.5.1 Recombinant inbred strains (RI strains) *42*

1.5.2 Congenic strains *44*

1.5.3 Positional candidates *44*

1.5.4 Differential gene expression *45*

1.6 GENERATION OF THE BXH/HXB RECOMBINANT INBRED STRAINS 45

1.7 MICROARRAY EXPRESSION PROFILING 47

1.7.1 DNA microarrays *47*

1.7.2 Integrated linkage analysis and gene expression approach to disease gene identification *48*

1.8 COMPARATIVE GENOMICS 51

1.9 AIMS OF THIS WORK 52

CHAPTER 2: MATERIALS AND METHODS 54

2.1 STOCK SOLUTIONS 54

2.1.1 Solutions for DNA and RNA analysis *54*

2.2 PREPARATION OF DNA 55

2.2.1 Extraction of DNA from rat tissue *55*

2.3 PREPARATION OF RNA AND CDNA 56

2.3.1 Extraction of RNA from rat tissue to select and optimise procedures *56*

2.3.2 First strand cDNA synthesis *57*

2.4 AMPLIFICATION OF DNA (PCR) 58

2.4.1 Standard protocol for PCR *58*

2.5 GEL ELECTROPHORESIS 59

2.5.1 Agarose gel electrophoresis *59*

2.5.2 Denaturing gel electrophoresis *60*

2.6 AGILENT BIOANALYSER 61

2.7 PREPARATION OF RNA, CDNA AND CRNA FOR AFFYMETRIX GENECHIP™ STUDIES 62

2.7.1 Extraction and purification of RNA from rat fat tissue *62*

2.7.2 Double stranded cDNA synthesis and purification *62*

2.7.3 cRNA synthesis and purification *63*

2.7.4 Hybridisation of cRNA to Affymetrix GeneChips 64

2.7.5 Expression data from RI strain kidneys 65

2.8 SEQUENCING OF PCR PRODUCTS 65

2.9 LINKAGE ANALYSIS AND DEFINITION OF EQTL 66

2.9.1 Mapping of eQTL 66

2.9.1.1 Using Map Manager QTX 66

2.9.1.2 Mapping eQTL across the whole genome using QTL Reaper 67

2.9.3 Definition of cis- and trans-acting eQTL 68

2.10 *IN SILICO* ANALYSIS 68

2.11 DESCRIPTION OF THE GENERATION AND ANALYSIS OF DATA SETS USED IN THIS THESIS 70

CHAPTER 3: MICROARRAY DATA GENERATION, QUALITY CONTROL, NORMALISATION, VALIDATION AND LINKAGE ANALYSIS 72

3.1 INTRODUCTION 72

3.1.1 Gene expression studies 72

3.1.2 Data extraction and normalisation methods 73

3.1.3 Quality control indices 74

3.2 RESULTS 75

3.2.1 Generation and quality control of microarray data sets 75

3.2.2 Comparison of data extraction and normalisation methods 81

3.2.3 Linkage analysis 88

3.2.4 Target probe set sequencing 101

3.3 DISCUSSION 116

3.4 CONCLUSIONS 122

CHAPTER 4: GENERATION AND CHARACTERISATION OF A DATA SET OF DIFFERENTIALLY EXPRESSED *CIS*-ACTING EQTL GENES 123

4.1 INTRODUCTION 123

4.1.1 Control of gene expression by cis elements 123

4.2 RESULTS 124

4.2.1 Selection and description of genes to be sequenced 124

4.2.2 Sequencing polymorphisms 129

4.2.3 Trans gene sequencing *139*

4.3 DISCUSSION 140

4.4 CONCLUSION 145

CHAPTER 5: *IN SILICO* STUDIES: PQTL DETERMINATION AND COMPARATIVE MAPPING 146

5.1 INTRODUCTION 146

5.1.1 QTL analysis *147*

5.1.2 Comparative mapping *147*

5.2 RESULTS 149

5.2.1 Collation of QTL in SHR and humans *149*

5.2.2 Identification of cis-acting human candidate genes *160*

5.3 DISCUSSION 170

5.4 CONCLUSION 175

CHAPTER 6: GENERAL DISCUSSION 177

6.1 CONCLUDING REMARKS 184

REFERENCES 185

APPENDIX I MICROARRAY SAMPLE PREPARATION RECORD 200

APPENDIX II PHYSIOLOGICAL SHR QTLS MAPPED IN PREVIOUS GENOME SCREENS. 202

APPENDIX III *TRANS* GENE PRIMERS 209

List of Figures

Figure 1.1 Screenshot from VC map showing all rat QTL 40

Figure 1.2 Generation of the BXH/HXB RI strains 43

Figure 1.3 Control of expression QTL (eQTL) 50

Figure 2.1 Flowchart describing the preliminary data sets 70

Figure 2.2 Flowchart describing the final data set 71

Figure 3.1 Comparison of Formaldehyde Gels and the Bioanalyser assay 77

Figure 3.2 Bioanalyser traces for total RNA and cRNA 79

Figure 3.3 Comparison of t-test data from two normalisation methods 83

Figure 3.4 P-value against fold change in both data sets 86

Figure 3.5 Example graphs showing the FDR and reproducibility in the RI strains following different normalisations 87

Figure 3.6 Chromosome map of markers significantly linked to differentially expressed genes using Map Manager QTX 93

Figure 3.7 Location of *cis*-acting eQTL and previously mapped SHR pQTL 96

Figure 3.8 Location of *trans*-acting eQTL and previously mapped SHR pQTL 97

Figure 3.9 Chromosome map of the initial 47 eQTL linkages reclassified according to *cis/trans* definition 100

Figure 4.1 Sequencher contig view of the *Akr1c6* gene 134

Figure 4.2 Sequence electropherogram of a polymorphism in *Ark1c6* gene 136

Figure 4.3 Sequence electropherogram of polymorphisms in *Aox1* gene 137

Figure 4.4 Genomic PCR of fragment 3 in SHR and BN in *Fmo3* gene 138

Figure 5.1 Synteny map derived from Rat Genome Database VC map 148

Figure 5.2 Original Database Schema for eQTL explorer 153

Figure 5.3 Screenshot from eQTL explorer 155

List of Tables

Table 3.1 Comparison of RNA yield using RNA Bee and Trizol 76

Table 3.2 Differentially expressed genes in fat tissue 82

Table 3.3 Fold Change of fat tissue genes 85

Table 3.4 47 genes initially tested in Map Manager QTX 89

Table 3.5 Number and *cis*/trans classification of non redundant eQTL detected in fat and kidney datasets at different genome-wide thresholds of significance 95

Table 3.6 Initial MapManager linkages reclassified 99

Table 3.7 cis-acting eQTL in fat tissue mapped at P≤10-4 102

Table 3.8 *cis*-acting eQTL in kidney tissue mapped at $P \leq 10^{-4}$ 108

Table 3.9 Significance of probe set sequence variants 115

Table 4.1 All cis eQTL at P <10-4 and with fold change >2 in both kidney and fat tissue 126

Table 4.2 All of the candidate probesets that were sequenced 130

Table 4.3 Polymorphisms detected in the *cis*-acting eQTL genes 131

Table 4.4 *trans* genes selected to test overall polymorphism rate 141

Table 5.1 Physiological SHR QTLs mapped in previous genome screens 150

Table 5.2 Human pQTL list 156

Table 5.3 Gene List generated by compartive mapping 162

Table 5.4 eQTLs deleted from Table 5.3 169

Abbreviations

AGT	Angiotensinogen
ALL	Acute lymphoblastiod leukaemia
AML	Acute myeloid leukaemia
ATP III	Adult treatment panel III
BN	Brown Norway
bp	Base pair
BP	Blood pressure
cDNA	Complementary deoxyribonucleic acid
cM	CentiMorgan
cRNA	Complementary ribonucleic acid
cSNP	Coding single nucleotide polymorphism
CVD	Cardiovascular disease
Dahl R	Dahl salt resistant
Dahl S	Dahl salt sensitive
DEPC	Diethyl pyrocarbonate
DNA	Deoxyribonucleic acid
eQTL	Expression quantitative trait loci
EST	Expressed sequence tags
FCHL	Familial combined hyperlipidaemia
FDR	False discovery rate
g; mg, μg, ng	Grams; milli-, micro-, nano-
HDL	High density lipoprotein
Hg	Mercury
IBD	Identity by descent
IBS	Identity by state
IRS	Insulin receptor substrate
Kb	Kilobase pair
l; ml, μl	Litre; milli-, micro
LCFA	Long chain fatty acid
LDL	Low density lipoprotein
LH	Lyon hypertensive

LOD	Logarithm of odds
LRS	Likelihood ratio statistic
MAS5.0	Microarray analysis suite 5.0
Mb	Megabase pair
MDC	Max-Delbruck Center
MM	Mismatch probe
mRNA	Messenger RNA
NEFA	Non-esterifeid fatty acids
PCR	Polymerase chain reaction
PM	Perfect match probe
pQTL	Physiological quantitative trait loci
QT	Quantitative trait
QTL	Quantitative trait loci
qRT-PCR	Quantitative reverse transcriptase PCR
RAE230A	Rat Array expression 230 A
RGD	Rat genome database
RI	Recombinant inbred
RMA	Robust Multichip Average
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RT-PCR	Reverse transcriptase PCR
SDP	Strain distribution pattern
SHR	Spontaneously hypertensive rat
SHRSP	Stroke-prone spontaneously hypertensive rat
SNP	Single nucleotide polymorphism
T2D	Type 2 diabetes
TBE	Tris-borate EDTA
TE	Tris-EDTA
Tm	Melting temperature
Tris	Tris-[hydroxymethyl]methylamine
UTR	Untranslated region
VLDL	Very low density lipoprotein
WHO	World health organisation
WKY	Wistar Kyoto rat

Chapter 1: Introduction

1.1 Genetic dissection of complex traits

Genetically complex traits, such as hypertension and the metabolic syndrome, are those diseases without a clear Mendelian pattern of inheritance; they tend to be highly prevalent and more common than monogenic (Mendelian) disorders (such as cystic fibrosis). The genetic dissection of inherited traits which lead to increased susceptibility to these multifactorial diseases will be invaluable for their prevention and treatment. However, identification of genes underlying complex traits is inherently more difficult than in a monogenic disorder, where one mutated gene leads to disease. This is because in a complex trait the disease susceptibility and progression is influenced by multiple genes and modulated by age, gender and environmental factors. In addition, the genetic variants that underlie complex traits are likely to be present in both affected and unaffected individuals; to be associated with relatively modest increases in individual risk; and to interact in a non-linear way with other susceptibility factors contributing to disease. Falconer and Mackay (Falconer and Mackay 1996) extended this theory by postulating a threshold of risk whereby the inheritance of multiple susceptibility genes puts individuals above the mean risk of the population, with a proportion of individuals exceeding the threshold and developing disease.

Complex trait susceptibility genes have many influences which confound their genetic dissection including; genetic heterogeneity, epistasis, low penetrance, variable expression and pleiotropy (Lander and Schork 1994; Risch 2000). Genetic heterogeneity can be divided into locus and allelic heterogeneity. Locus

heterogeneity means that the disease may be related to multiple genes at different locations in the genome, each acting independently to increase or decrease susceptibility to disease. Allelic heterogeneity occurs when different alleles of the same gene act on disease susceptibility and this can also occur in monogenic disorders. Epistasis refers to the interactions between genes which leads to a greater or lesser phenotypic effect than expected from the sum of the effects any of the genes individually. Penetrance describes the extent to which the phenotype of a gene is expressed. Low penetrance, though difficult to estimate in a complex trait since it requires knowledge of potential genes and environmental factors; occurs when the disease phenotype is not expressed even though the disease causing allele is present. Pleiotropy occurs when a single gene mutation results in multiple phenotypic effects which often appear unrelated. In addition, apparent disease in some individuals or families may be due to the presence of phenocopies; that is environmentally produced phenotypes that mimic genetically produced phenotypes. The interaction between disease genes and environment can also occur so that in a particular environment a gene carrier may be spared development of the disease entirely.

1.1.1 Identifying complex trait genes

Methods for mapping and identifying complex trait susceptibility genes fall broadly into two categories: either candidate gene identification or a genome wide screen (also called quantitative trait analysis) but ultimately all result in identification and study of a candidate gene. The traditional candidate gene approach is carried out using knowledge of the biology of known genes and testing of allele frequencies in these genes in case versus control association studies. Quantitative trait analysis originally used linkage analysis in pedigrees to identify regions of the genome associated with

disease and have more recently utilised whole genome tests of association with genome wide linkage disequilibrium. A genome wide screen followed by fine mapping of candidate genes can provide the most powerful method to detect genes for a particular disease. In this case the genomic screen is performed first to identify regions of the genome potentially containing susceptibility genes. Once these regions are identified, candidate genes in these areas can be preferentially investigated. This positional cloning approach successfully identified the apolipoprotein E4 allele as an Alzheimer's disease gene (Pericak-Vance et al., 1991; Saunders et al., 1993) as one of the first examples of a complex trait disease gene.

1.1.1.1 Candidate gene approach

Candidate gene studies use knowledge of the biology and pathophysiology of a disease to enable potential candidates to be selected based on their known or presumed biological function. These selected genes are analysed to detect polymorphisms (or alleles) which are then tested for association with disease; that is does one allele occur more frequently in patients than in normal controls. Historically this process required vast amounts of sequencing firstly to detect polymorphisms and then genotyping to determine if they are disease associated which was laborious and time consuming and often led to results which were difficult to interpret (Hirschhorn and Altshuler 2002). However with the completion of the human genome and availability of a wealth of single nucleotide polymorphism (SNP) data and the HapMap project (Hirschhorn and Daly 2005; Morton 2005) SNPs can be identified, *in silico,* and tested for association with disease. The improved sequence and SNP data have made it easier and more cost effective to pursue this type of approach. In addition, the candidate gene approach can be highly successful when readily

identifiable genes are known to be associated with variation in biologically relevant processes that have been well characterised.

1.1.1.2 Genome wide screen – QTL analysis

The main alternative to the candidate gene approach is to carry out genome wide screens to identify quantitative trait loci (QTL); regions of the genome which contribute to disease. Polymorphic genetic markers spanning the entire genome, at preferentially no more than 10cM intervals, are tested for linkage to the phenotype. Traditionally microsatellites, which are simple short repeats of mono-, di-, tri- or tetra-nucleotide repeats that are scattered in non-coding DNA have been the genetic markers of choice. These highly polymorphic repeats can be easily genotyped using PCR. However due to the availability of increasing amounts of SNP data the genetic maps are becoming more dense and this allows the linked regions of the genome to be more closely defined. For this type of analysis to succeed the flanking markers and the disease gene must co-segregate (be inherited together) in subjects with the disease more frequently than expected by chance or in control subjects.

1.1.1.3 Linkage Analysis and Association Studies

Both the candidate gene and QTL analyses utilise linkage analysis and association studies. Linkage is a genetic relationship between loci (eg flanking markers and disease gene) while association is a statistical observation of a relationship between alleles and/or phenotypes. Linkage occurs when alleles from two or more genes are coinherited, the closer alleles are genetically the more often they will be inherited together. Linkage disequilibrium occurs when the combinations of alleles (or markers and disease) occur more or less frequently in a population than would be expected based on the frequencies of the individual alleles.

Linkage analysis measures the co-segregation of markers and phenotype to identify regions of the genome which influence a trait. In human complex traits non-parametric approaches to linkage analysis are more frequently used as they are model free and thus more powerful. These methods identify the sharing of alleles identical by descent (IBD) or identical by state (IBS) at marker loci in families. The principle behind linkage analysis is that two relatives affected with the same disease will share alleles IBD near the disease gene(s) more often than expected under the random segregation of alleles.

Association studies, in a simple description, assess whether common allelic variants occur together more often than expected by chance. The goal of these studies is to find linkage disequilibrium (allelic association) between markers and the disease. Linkage disequilibrium is a function of the distance between the marker and the disease gene and the genetic history of the population under study and is variable at different locations in the genome.

Linkage and association are complementary to each other and a strategy often employed is to carry out linkage analysis across the genome to localise a candidate region and then narrow this region by linkage disequilibrium mapping and perform association studies on polymorphisms in individual genes in significant regions.

In comparison to monogenic disease, and taking into consideration the prevalence of complex traits, relatively few underlying genes have been positionally cloned. One of the first examples, *ApoE4*, was mentioned in **Section 1.1.1.2** and at least two susceptibility genes have been identified in type 2 diabetes, peroxisome proliferator-

activated receptor-gamma (*PPARγ*) and *CALPAIN-10*. A common *PPARγ* allele is associated with a decreased risk (Altshuler et al. 2000) while *CALPAIN-10* showed association with type 2 diabetes in both Mexican Americans and a Finnish population (Horikawa et al. 2000).

The recent advances in genome resources such as availability of annotated sequences and databases, for example ensembl (www.ensembl.org) and NCBI (www.ncbi.nlm.nih.gov), with the development of technologies such as DNA microarrays for the global analysis of gene expression have started to facilitate and expedite the identification of complex trait genes (Glazier et al. 2002; Farrall 2004; Darvasi 2005; Peltonen et al. 2006; Wu and Lin 2006).

1.2 Metabolic Syndrome, Cardiovascular Disease and Type 2 Diabetes

Cardiovascular diseases (CVD), such as coronary artery disease or atherosclerosis, and type 2 diabetes (T2D) often result from the complex interaction of environmental and genetic factors and are the major cause of morbidity and mortality in westernised societies. The occurrence of T2D and CVD are determined by multiple risk factors including hypertension, insulin resistance, dyslipidemia and obesity, which tend to cluster together in individuals more frequently than expected by chance (Reaven 1988a; Reaven 1993). For example, as many as 50% of patients with essential hypertension are also found to be insulin resistant (i.e. have impaired insulin action on glucose and fat metabolism) leading to an increased risk of cardiovascular disease. This clustering of risk factors has been drawn together under the term Metabolic Syndrome and is thought to affect 15-25% of people in industrialised countries

(Alberti and Zimmet 1998; 2002). Metabolic syndrome is associated with a 2-4 fold increase in CVD morbidity and mortality (Isomaa et al. 2001; Lakka et al. 2002)and 5-9 fold increase in developing T2D (Balkau et al. 2003; Sattar et al. 2003). The increasing prevalence of metabolic syndrome is due to the dramatic rise in obesity; it worsens with age and often culminates in T2D if not CVD. Environmental and other factors that augment the primary risk of developing these highly prevalent disorders include age, race, sex, diet, and cigarette smoking (Thompson and Wilson 1992). There are different sets of criteria to define the metabolic syndrome and they are not totally concordant. The World Health Organisation (WHO) definition requires the patient to be insulin resistant or have impaired glucose tolerance and at least 2 other risk factors from hypertension, dyslipidaemia, obesity and micoalbuminuria (Alberti and Zimmet 1998). The other main criteria are set out by the Adult Treatment Panel III and define metabolic syndrome as having any 3 of the following: abdominal obesity, atherogenic dyslipidemia, high blood pressure, glucose intolerance and prothrombic and proinflamatory states (2002).

Many cardiovascular diseases are multifactorial, resulting from the complex interaction of environmental and genetic factors. Many of the genetic variants predisposing to cardiovascular disease have been proposed to be common in a "normal" population (Todd 2001). The simultaneous presence of several predisposing minor risk alleles, occasionally accompanied by more major, highly deleterious ones, may lead to the addition or multiplication of deleterious allelic effects on the phenotype. This constitutes an individual "risk allele accumulation", which translates into a high individual cardiovascular risk.

Identifying genes that underlie risk factors for metabolic syndrome will allow a better understanding and development of new treatments which would decrease the health burden of these diseases. Following is a discussion of some of the risk factors and the underlying genetics that have been elucidated. The risk factors themselves are complex traits and therefore also under multigenic and environmental control.

1.2.1 Hypertension

Several decades of population research from the North American Framingham Study and elsewhere have demonstrated that elevated systemic blood pressure (hypertension) is a common and powerful contributor to the major cardiovascular diseases associated with accelerated atherogenesis (Kannel 2000).

Hypertension has been defined by the WHO and the International Society of Hypertension as a systolic blood pressure exceeding 140 mm Hg and diastolic pressure in excess of 90 mm Hg. Although these criteria are somewhat arbitrary, Professor Geoffrey Rose suggested that the best clinical definition of hypertension should be "the level of blood pressure above which the benefits ....of action exceed those of inaction" (Rose 1980).

Hypertension can be divided into two categories, primary (or essential) and secondary hypertension. Secondary hypertension refers to all cases of hypertension with a known cause (either as part of a syndrome or monogenic disorder) and renal disease accounts for over 80% of cases. Essential hypertension has no single identifiable cause and accounts for over 95% of all hypertension (Sharma et al. 2000). In essential hypertension, cardiac output is initially normal and elevated blood pressure is

maintained by an increased peripheral resistance in the small arteries and arterioles. This increased peripheral resistance is due to structural changes in the vasculature, such as narrowing of the arterioles caused by constriction, growth, and proliferation of arteriolar smooth muscle cells (Beevers and MacGregor 1999). Patients with essential hypertension are also prone to developing atherosclerosis due to lipid-rich deposits and inflammation in the arterial wall of cerebral and coronary arteries. In response to the increase in peripheral resistance, there is an increased left ventricular load on the heart in order to maintain stroke volume and therefore cardiac output. This increased load on the heart can result in the development of left ventricular hypertrophy where the heart is increased in size. Ultimately, the function of this chamber deteriorates, the cavity dilates, and the symptoms and signs of heart failure appear.

Pickering first proposed that essential hypertension was multifactorial, with hypertensive subjects inheriting an aggregate of hypertension-determining genes (Pickering 1961). As stated in **Section 1.2**, within individual patients, hypertension often clusters with other risk factors for cardiovascular disease such as hyperlipidaemia, obesity, glucose intolerance and insulin resistance (Reaven 1988a). However, the association of hypertension with these metabolic abnormalities is not uniform and the mechanisms responsible for these associations remain speculative. For example, hyperinsulinaemia has been shown to promote sodium and water retention (DeFronzo et al. 1975; Baum 1987) and to stimulate the sympathetic nervous system (Reaven et al. 1996). However, studies in patients with insulinomas show that isolated hyperinsulinaemia is not alone sufficient to cause hypertension (Sawicki et al. 1992).

Blood pressure must be tightly regulated to permit uninterrupted perfusion of all vital organs and is determined by cardiac output and vascular resistance to blood flow. This simple definition belies the integrated actions of the cardiovascular, renal, neural and endocrine systems in the regulation of blood pressure. Included in these systems are baroreceptors that sense acute changes in pressure in vessels; natriuretic peptides produced by the brain and heart in response to increased pressure in these organs; the renin-angiotensin-aldosterone system, which influences vascular volume homeostasis and vascular tone; the kinin-kallikrein system, which affects vascular tone and renal salt handling; the adrenergic receptor system which influences heart rate, cardiac contraction, and vascular tone; and factors produced by blood vessels that cause vasodilation, such as nitric oxide, or contraction, such as endothelin.

The genetics of hypertension have been studied and estimates from epidemiological studies suggest that between 30-50% of the blood pressure variance in the normal population can be accounted for by genetic differences (Hingorini 2000). Race is an important determinant of blood pressure, for example, in the US, African Americans tend to have higher levels of blood pressure than non- African Americans, and overall hypertension-related mortality rates are higher among African Americans (Gillum 1996). There are also examples taken from non industrialised populations where low blood pressure is more common, such as the Yanomamo Indians in the Amazon (Oliver et al. 1975) and in a study of bushmen in Botswana it was observed that blood pressure tended to fall with age (Truswell et al. 1972). Similar observations have been made in the Xhosa people of South Africa (Sever et al. 1980) and inhabitants of the Solomon Islands (Page et al. 1974). Therefore, care must be taken when selecting

individuals for studying hypertension as the genetic effects vary from population to population and this will affect results of a genome screen.

An environmental influence on blood pressure which has been considered is salt intake; possibly in light of the fact that all the acquired and monogenic forms of hypertension alter blood pressure by changing renal salt balance (Lifton et al. 2001). Epidemiological studies demonstrate that the amount of salt consumed in different communities is a major factor in determining blood pressure differences between these communities (MacGregor 1985; Elliott et al. 1996). The relationship between sodium intake and sodium excretion is of major importance in hypertension as it is this mechanism that determines extracellular volume. It has long been recognised that total-body sodium and water balance is regulated in part by renal arterial perfusion pressure (Starling 1909). Guyton proposed that blood pressure and sodium homeostasis are related through the mechanism of pressure natriuresis i.e. when renal arterial perfusion pressure increases, renal sodium output increases and blood volume contracts by an amount sufficient to return arterial blood pressure to its baseline (Guyton et al. 1972). Several other factors such as alcohol intake, nutrition affecting blood lipids and stress have been shown to influence blood pressure (Timio et al. 1997; Kannel 2000).

In most populations, both systolic and diastolic blood pressure exhibit a continuous unimodal distribution, a feature which is characteristic of a complex trait. Evidence suggests that while environmental factors exert a major influence on the average level of blood pressure in any given population, it is genetic factors which determine the position of an individual within the blood pressure distribution (Harrap 1994).

Candidate gene and genome wide screens have both been carried out to identify genes involved in hypertension. Candidate gene studies focussed on the molecules with a known involvement in the vasoconstriction and dilation of blood vessels and in the renin-angiotensin system which plays a central role in blood pressure regulation, where it functions to regulate sodium homeostasis and modulate vascular tone and structure (Sealey and Laragh 1990). One example of a candidate gene for hypertension is angiotensinogen (AGT) which was first studied because a strong correlation was detected between AGT concentration and blood pressure (Walker et al. 1979). The role of AGT in hypertension has been extensively studied and is still unclear; a recent review (Dickson and Sigmund 2006) summaries the 500 or more association studies carried out and concludes that some variants in this gene are likely to confer a small risk for hypertension. A mouse model featuring increased expression of the angiotensinogen gene exists and demonstrates elevated blood pressure which supports this genes involvement in development of hypertension (Takahashi and Smithies 1999).

The *angiotensin converting enzyme* (*ACE*) is another widely studied human hypertension candidate gene and certain allelic genotypes do associate with disease; however the results remain largely equivocal (Lindpaintner et al. 1996; Montgomery et al. 1997).

There have been several genome wide scans carried out for essential hypertension which pinpoint many regions of the genome, some of which contain good candidate genes. This will be discussed further in **Chapter 5** where all human QTL from genome wide screens have been gathered and used in a comparative mapping study.

1.2.2 Insulin Resistance

Insulin resistance can be operationally defined as a state when insulin action on both glucose and fat metabolism is impaired. This can be detected in an individual tissue or at the whole body level and is often accompanied by compensatory hyperinsulinaemia and hyperglycaemia. Insulin resistance is one of the central metabolic defects of type 2 diabetes and predisposes an individual to other features of the metabolic syndrome such as dyslipidemia and hypertension (Reaven 1995). In addition, insulin resistance is a common accompaniment of obesity and can also be secondary to it. Metabolic studies of type 2 diabetes, familial combined hyperlipiaemia (FCHL), hypertension and obesity have shown that insulin action is defective on both glucose and fatty acid metabolism in these conditions (Chen et al. 1987; Ferrannini et al. 1987; Aitman et al. 1997b; Kooner et al. 1998). Catecholamine-mediated lipolysis, the release of fatty acids from fat stores, is also defective in some of these disorders (Reynisdottir et al. 1994; Reynisdottir et al. 1995; Bougneres et al. 1997). These metabolic studies have suggested the presence of a primary defect in insulin and catecholamine action at the level of the adipocyte in several of the insulin resistance syndromes (Arner 1997; Aitman et al. 1997a).

Rare monogenic forms of insulin resistance occur and in most cases have been attributed to mutations in the insulin receptor gene including homologous deletions, mutations in the tyrosine kinase domain and nonsense mutations resulting in truncated protein (Wertheimer et al. 1993; Desbois-Mouthon et al. 1996; Jospe et al. 1996). However, the common forms of insulin resistance are believed to be under the influence of many genetic and environmental factors. As with hypertension, candidate gene investigations and genome screens have been carried out and again the

genome screens are presented in **Chapter 5**. One candidate gene that has been examined is the insulin receptor which, as outlined above, accounts for many monogenic forms of insulin resistance and is thought to account for <2% of complex insulin resistance and T2D (Taylor and Moller 1993).

Insulin receptor substrate 1(IRS1) binds to the phosphotyrosines on the insulin receptor and is itself phosphorylated and activated following binding of insulin to the insulin receptor. This gene was therefore identified and studied as another good candidate and a Gly^{972}→Arg polymorphism confers lower insulin sensitivity to obese carriers (Clausen et al. 1995). This polymorphism is more prevalent in patients within the top 2 percentile for insulin resistance and dyslipidemia in Caucasian T2D subjects (Zhang et al. 1996). A Pro^{12}→Ala polymorphism has been investigated in another candidate gene *PPARγ*. PPARγ, is a transcription factor which regulates adipocyte differentiation (Tontonoz et al. 1994) and plays an important role in lipid metabolism, glucose homeostasis, insulin sensitivity, obesity, diabetes, foam cell formation and atherogenesis and it is the target molecule for the thiazolidinedione group of drugs which enhance insulin sensitivity (Lehmann et al. 1995). It was hypothesised that the Pro^{12}→Ala substitution may cause a reduction in the transcriptional activity of *PPARγ* and thus may be associated with increased insulin sensitivity and decreased risk of type 2 diabetes (Hara et al. 2000). Another study failed to confirm this but agrees that genetic variability of *PPARγ* affects body weight control and lipid homeostasis in humans (Meirhaeghe et al. 2000). The genetics of type 2 diabetes has been reviewed recently and the above are just a few examples of studies carried out (McCarthy 2004).

1.2.3 Dyslipidaemia

Lipoprotein particles transport lipids, cholesterol, triglycerides and phospholipids around the body in the bloodstream. Serum cholesterol and triglyceride levels are determined by several metabolic pathways, ligands and receptors and involve several tissues including the liver, adipose, the small intestine and skeletal muscle. Chylomicrons transport dietary lipids like triglycerides from the intestines to the liver. Very low density lipoproteins (VLDL), low density lipoproteins (LDL) and high density lipoproteins (HDL) transport endogenous lipids. VLDL and LDL transport cholesterol from the liver to the tissues and LDL is considered 'bad' cholesterol as it can build up which leads to narrowing and inflammation (clogging) of the arteries (atherosclerosis). HDL transports excess cholesterol from tissues back to the liver for removal; this is a very important mechanism as tissues do not have a mechanism of breaking down cholesterol.

Genetic linkage and sequence studies of the lipoprotein particles have elucidated mutations and often the mechanism of action which leads to dysfunction. This led to the development of statins, (HMG-CoA-reductase inhibitors) which are lipid lowering drugs. Many of the initial insights into disorders of lipid metabolism came from the study of monogenic hypercholesterolaemias. Familial hypercholesterolaemia (FH) is the most common form (Goldstein et al. 1973); patients have high LDL cholesterol caused by mutation(s) in the LDL receptor gene. Other monogenic forms detected mutations in the *apoA1*, *apoB-100*, *PSCK9* and *ABCA1* genes. The most common inherited form of hyperlipidaemia, FCHL, is found in approximately 30% of myocardial infarction survivors and may affect up to one in 50 of the general population (Goldstein et al. 1973; 1998). FCHL has hypertriglyceridaemia associated

with hypercholesterolaemia and is most often due to an increase in VLDL and LDL. The lipid abnormalities seen in FCHL may occur as part of the metabolic syndrome however patients with FCHL must have a blood relative with primary hyperlipidaemia. At least 3 chromosomal locations have been investigated in multiple genome wide scans for candidate genes underlying this disease and the *LCAT* gene and *APOA1/C3/A4/A5* gene cluster have been shown to be important in FCHL (Shoulders et al. 2004).

1.2.4 Obesity

Obesity, particularly central (abdominal) obesity is a major and important feature of the metabolic syndrome and is characterised by an increased fat mass which occurs when the intake of food exceeds the energy requirement of the body (Spiegelman and Flier 1996). The dramatic escalation of obesity in the western world in the past 30 years is due, at least in part, to an increasingly sedentary lifestyle and dietary changes including increase in dietary fat and carbohydrate.

Abdominal obesity, as opposed to whole body obesity, is proposed to be a better indicator of CVD risk overall (Rimm et al. 1995; Lakka et al. 2002), therefore it has been suggested that waist circumference or waist to hip ratio is a better measure of the cardiovascular complications of obesity than body mass index (BMI) (Carey et al. 1996). The components of the metabolic syndrome occur more frequently in obese people but this is thought not to be due to the obesity per se, but to the fact that increased adiposity increases the likelihood that an individual will be insulin resistant (Abbasi et al. 2002; McLaughlin et al. 2003). An increase in free fatty acids (which make up triglycerides) leads to hyperinsulinaemia as there is a reduction in hepatic

insulin sensitivity which results in increased glucose output and decreased insulin clearance. Availability of more fatty acids also provides substrates for triglyceride synthesis which results in hypertriglyceridaemia and lastly may also inhibit glucose uptake by skeletal muscle.

Adipose tissue comprises the body's largest endocrine organ as it not only stores excess body energy in the form of fat but also secretes many systemically functioning adipokines, for example TNFα, resistin and adiponectin. Clinical studies have demonstrated that a low blood level of adiponectin is associated with atherosclerotic CVD, T2D, hypertension and dyslipidemia (Zoccali et al. 2002; Pischon et al. 2004; Schulze et al. 2005). Also genetic analyses of SNPs within this gene have identified some variants which may determine plasma concentration or affect the ability of adiponectin to multimerise (Kondo et al. 2002; Stumvoll et al. 2002; Kishida et al. 2003; Ohashi et al. 2004).

Recent reviews of the genetics of obesity (Bell et al. 2005) present the overview of the outcomes of the many genome wide screens for obesity carried out in humans.

1.3 Animal models

The features of the metabolic syndrome and thus CVD and T2D cluster together and can cause and/or affect each other. This is common in genetically complex traits, that is, it is difficult to separate out the different parts of the trait to enable the genetic dissection of the disease (Merikangas and Risch 2003). Animal models of human metabolic and cardiovascular disease have been widely used to elucidate the mechanisms underlying these disorders, since they have several advantages over

human studies for gene identification. Firstly, the use of inbred strains eliminates genetic heterogeneity between affected individuals thereby limiting the number of alleles involved to those that differ between strains. Secondly, rare alleles can be fixed in inbred strains by a programme of selective breeding. Thirdly, breeding protocols can be manipulated to generate experimental crosses that are most informative for linkage analysis. Fourthly, all offspring from many crosses can be combined into a single large group for analysis as a single pedigree (Risch 2000). Fifthly, very detailed phenotyping not possible in humans such as terminal phenotyping can be carried out and finally the environment can be controlled which minimises or eliminates the variability between individual animals. All of the above make linkage analysis more practicable in animal models and result in more robust QTL with higher statistical significance. However positional cloning of the genes underlying these QTL is still a significant challenge, mainly because of the lack of a one-to-one correspondence between genotype and phenotype.

1.3.1 Rat Models of Metabolic Syndrome

Inbred rat strains have been studied for over 50 years creating a wealth of genetic and phenotypic data for use in studies of the onset, development and progression of complex diseases. There is an immense amount of literature available about the rat as a model organism including information on physiological mechanisms including blood pressure regulation, making this model species particularly appropriate for disease-based research. For example, between 1975 and 2006 PubMed lists over 14,700 publications on the spontaneously hypertensive rat, the model used in this thesis. The publication of the rat genome sequence has greatly ameliorated the

previous paucity of genome sequence data in the rat compared to the mouse and provided valuable insights into mammalian evolution (Gibbs et al. 2004).

One drawback of using rat, or other rodent models, is that they generally do not develop atherosclerosis, that is the atheroma of the coronary arteries which leads to atherosclerotic plaque rupture, and MI and this is the main pathophysiological endpoint associated with coronary artery disease (CAD) (Rekhter 2002; Rekhter 2002). This is due to the fact that unlike humans, rats and mice have a naturally high level of HDL, the molecule which clears cholesterol from the body by transporting it back to the liver, along with relatively low levels of the atherogenic lipoprotein LDL.

This section outlines a number of animal models which may be amenable for study as models of the metabolic syndrome.

1.3.1.1 Dahl Salt Sensitive Rat

The Dahl salt sensitive (S) rat was generated from a colony of outbred Sprague-Dawley rats selected for sensitivity to salt-induced hypertension (Dahl et al. 1962; Rapp et al. 1984). A salt-resistant strain (R) was also maintained. Both strains were bred on the basis of their blood pressure after being fed a high salt (8% NaCl) diet. Dahl S rats develop severe hypertension and marked vascular and renal lesions after 3-4 weeks on a high salt diet and all die within 8 weeks (Rapp and Dene 1985). The strain is maintained on a low salt diet (0.3% NaCl) and although they develop marked hypertension, there is a slower rise in blood pressure (Rapp and Dene 1985). Kidney transplantation experiments between Dahl S and Dahl R rats suggest that the kidney plays a central role in blood pressure regulation in these strains (Dahl and Heine

1975). At least 14 QTL regions influencing blood pressure have been identified in Dahl S rats (Garrett et al. 1998; Rapp 2000).

1.3.1.2 Fructose fed rats

Environmental manipulations have been shown to induce both increased blood pressure and insulin resistance. The best example of this is the fructose-fed rat. To determine if hypertension could be induced in normal rats by feeding them a fructose-enriched diet, Sprague Dawley rats were fed a diet containing 66% fructose as a percentage of total calories for 2 weeks (Hwang et al. 1987). Systolic blood pressure increased in the fructose-fed rats, with no change in the control group (Hwang et al. 1987). In addition, hyperinsulinaemia and hypertriglyceridemia were associated with hypertension in this model (Hwang et al. 1987). Fructose fed rats are now a well established, environmentally induced model of increased blood pressure and insulin resistance.

1.3.1.3 Lyon hypertensive rat

Lyon hypertensive (LH) rats were bred from a colony of Sprague-Dawley rats selected for high systolic blood pressure (Vincent et al. 1984). A normotensive control (LN) and a hypotensive control (LL) were also bred at the same time (Vincent et al. 1984). This strain exhibits mild hypertension, increased body weight, increased plasma triglycerides and insulin when compared to the LN controls (Vincent et al. 1993).

1.3.1.4 Spontaneously hypertensive rat

The spontaneously hypertensive rat (SHR) was selectively bred for high blood pressure in the absence of dietary or environmental stimuli from outbred Wistar Kyoto rats. A male with mild hypertension was mated with a female with high blood

pressure followed by repeated brother x sister mating with continued selection for high blood pressure (Okamoto et al. 1972; Yamori 1984). SHR has been widely studied as a model of both hypertension (Yamori 1984; Pravenec et al. 1995; Munroe and Caulfield 2000; Rapp 2000) and insulin resistance (Reaven et al. 1989b; Aitman et al. 1998; Sechi 1999). Defective insulin-mediated glucose uptake has been observed in isolated adipocytes and at the whole body level (Reaven et al. 1989b; Hulman et al. 1991; Hulman et al. 1993). A reduction in sensitivity to catecholamine action has been observed in isolated adipocytes (Chiappe de Cingolani 1988). SHR also show hypertriglyceridemia, hyperinsulinaemia (Qu and Donnelly 1997) and abdominal obesity (Aitman et al. 1997b).

An inbred Wistar Kyoto strain, WKY, is often used as a normotensive control although its suitability has been questioned as it differs from SHR at 50% of genetic marker loci (St Lezin et al. 1992). Additional normotensive strains, for example Brown Norway (BN), are often used in comparative studies (Rapp 2000) including the work presented in this thesis. The SHR strain will be discussed in more detail in **Section 1.4**.

1.3.1.5 Stroke prone spontaneously hypertensive rat

Prior to SHR becoming fully inbred, a substrain with exceptionally high blood pressure was found to be more susceptible to stroke than other SHR substrains. This substrain was subsequently selectively bred for a high incidence of stroke [Okamoto, 1974]. Male systolic blood pressure is 190-250 mm Hg in SHRSP compared with 125-140 mm Hg in the normotensive control WKY (Jacob et al. 1991). Eighty-two percent of male SHRSP over 100 days of age and 58% of females over 150 days have cerebral haemorrhage or infarction (Yamori 1984). QTL regions conferring

sensitivity to stroke have been identified on chromosomes 1, 4 and 5 in an SHRSP x SHR cross (Rubattu et al. 1996; Jeffs et al. 1997), while at least nine blood pressure QTLs have been identified in SHRSP x WKY crosses (Rapp 2000). This strain also exhibits insulin resistance with abnormal insulin action on glucose and fatty acid metabolism which is distinct from the *Cd36* deficiency seen in SHR (Collison et al. 2000).

1.4 The Spontaneously Hypertensive Rat as a Model of Metabolic Syndrome

The spontaneously hypertensive rat (SHR) is a widely studied model of Metabolic Syndrome. In SHR, as in many humans with hypertension, the hypertensive phenotype clusters with metabolic traits such as insulin resistance and dyslipidemia, which are associated with increased risk of cardiovascular disease. As mentioned in **Section 1.3.1.4,** SHR was selectively bred for high blood pressure from outbred Wistar Kyoto rats and has been shown to have both hypertension and insulin resistance (Yamori 1984; Reaven et al. 1989b; Pravenec et al. 1995). Within this strain insulin resistance has been observed at the whole body level and in isolated adipocytes (Reaven et al. 1989b; Hulman et al. 1991; Hulman et al. 1993; Rao 1993; Paternostro et al. 1995). Also, a reduction in the sensitivity of adipocytes to catecholamine action on the promotion of lipolysis has been observed (Nelson et al. 1987; Chiappe de Cingolani 1988; Reaven et al. 1989b). These adipocyte phenotypes have also been observed in FCHL, Metabolic Syndrome and upper body obesity (Reynisdottir et al. 1994; Reynisdottir et al. 1995). These cellular defects in SHR adipocytes suggest a similarly altered fat cell metabolism to that seen in human insulin resistance.

1.4.1 Hypertension in SHR

The SHR strain was bred with continuous selection for high blood pressure over 5 generations and develops hypertension spontaneously, without dietary or environmental stimuli, at the age of 5-15 weeks (Yamori 1984). Systolic blood pressure reaches 170 mm Hg by 10 weeks of age (Tanase et al. 1982) and plateaus at approx. 200 mm Hg at 20-28 weeks (Yamori 1984). Hypertension in SHR is associated with increased peripheral vascular resistance and a normal cardiac output (Yamori 1984) while levels of plasma renin and angiotensin II are not elevated (Campbell et al. 1995). SHR rats have fewer glomeruli than WKY rats which results in a reduced glomerular volume consistent with the view that the kidney plays an important role in hypertension. As in humans, environmental factors including excess sodium intake, stress (Kvetnansky et al. 1979) and social alterations (Horie et al. 1991) have been shown to influence development of hypertension in SHR.

1.4.2 Insulin resistance and dyslipidemia in SHR

Initial comparisons of insulin metabolism in SHR and control rats using an oral glucose tolerance test showed indirect evidence of resistance to insulin-mediated glucose uptake in SHR (Mondon and Reaven 1988). The presence of insulin resistance and hyperinsulinaemia in SHR led to direct experiments to measure insulin-mediated glucose uptake and the ability of insulin to suppress catecholamine-mediated lipolysis in isolated adipocytes (Reaven et al. 1989b). Adipocytes were isolated from SHR rats, derived from the NIH colony, and normotensive Wistar Kyoto (WKY) control rats. Glucose transport was measured in isolated adipocytes incubated in the presence of differing concentrations of insulin and [^{14}-C] D-glucose. Lipolysis was also measured in isolated adipocytes incubated in the presence of

isoproterenol (or isoprenaline) and increasing concentrations of insulin. Glucose uptake was significantly lower in adipocytes from SHR rats when incubated in the absence of insulin and at every insulin concentration (Reaven et al. 1989b). Catecholamine-mediated lipolysis was lower in adipocytes isolated from SHR both in the absence of insulin and at every insulin concentration; however, the relative ability of insulin to suppress lipolysis was increased in adipocytes from SHR (Reaven et al. 1989b).

This work demonstrated insulin resistance in SHR adipocytes. Further work also showed insulin resistance in cardiac muscle (Paternostro et al. 1995). Measurements of basal and insulin-mediated glucose uptake in perfused hearts from SHR and WKY by ^{31}P NMR spectroscopy showed a 3-fold increase in basal glucose uptake in SHR but an 18% decrease in insulin-mediated glucose uptake (Paternostro et al. 1995).

Insulin resistance has been measured in SHR at the whole body level using the euglycaemic-hyperinsulinaemic clamp in conscious and anaesthetised rats (Hulman et al. 1991; Hulman et al. 1993). Total insulin-mediated glucose uptake was found to be significantly lower in SHR than in WKY, indicated by the reduced glucose infusion rate required to maintain euglycaemia in SHR (Hulman et al. 1993).

An important consideration when measuring insulin-mediated glucose uptake during clamp studies was raised by Rao (Rao 1993). During euglycaemic clamp studies, insulin-mediated glucose uptake and hepatic glucose output was shown not to differ significantly between WKY and SHR (Rao 1993). This had led to the suggestion that SHR was not insulin resistant (Buchanan et al. 1992). However, steady-state plasma

insulin levels in SHR were found to be double that of WKY, indicating that SHR had a marked defect in insulin clearance (Rao 1993). Additional studies were performed in order to produce equivalent steady-state plasma insulin levels in the two strains and, under these conditions; a marked difference was seen in both insulin-mediated glucose uptake and insulin-mediated suppression of hepatic glucose output between SHR and WKY. SHR was shown to have a 50% reduction in insulin-mediated glucose uptake and impaired suppression of hepatic glucose production (Rao 1993).

Abnormalities in lipid metabolism in SHR were observed in the late 1970s (Iritani et al. 1977). When comparing SHR to control WKY rats, plasma cholesterol was found to be decreased, and plasma triglycerides and free fatty acids were found to be increased (Iritani et al. 1977). A study of the responsiveness of isolated adipocytes from SHR and WKY to the effect of norepinephrine on lipolysis demonstrated that the response was lower in SHR adipocytes compared to WKY, but there was no significant change in the basal rate of lipolysis (Chiappe de Cingolani 1988).

Glucose and fat metabolism have been examined together by assaying insulin and catecholamine action in isolated adipocytes from SHR and normotensive WKY rats (Aitman et al. 1997b). Reteroperitoneal adipocytes from SHR showed a reduction in sensitivity to insulin measured by a reduction in insulin-mediated glucose uptake and a reduction in sensitivity to isoproterenol measured by isoproterenol-mediated non-esterified fatty acids (NEFA) secretion (Aitman et al. 1997b).

SHR also show the hypertriglyceridemia, hyperinsulinaemia and abdominal obesity, common phenotypes of human metabolic syndrome. Blood samples from SHR and

WKY rats were analysed for levels of serum glucose, insulin and triglyceride concentrations (Qu and Donnelly 1997). SHR rats were shown to have higher fasting insulin, higher triglycerides and higher NEFA concentration (Qu and Donnelly 1997). In addition, SHR have been observed to have increased abdominal fat when compared to the Brown Norway (BN) normotensive control (Aitman et al. 1997b).

1.4.3 SHR QTL for metabolic syndrome

The genetics of metabolic syndrome, and the underlying risk factors, hypertension, insulin resistance and dyslipidemia have been widely studied by genome wide linkage screens. As of October 2004 more than 60 QTL for hypertension and metabolic phenotypes had been identified in SHR (Rat Genome Database, http://rgd.mcw.edu), with this number being added to all the time. These QTL will be discussed further in **Chapter 5** and at the time of writing the number in RGD had increased to more than 200. QTL have also been mapped using a recombinant inbred (RI) strain set that was derived from the hypertensive SHR/Ola and the normotensive BN.Lx/Cub (**Section 1.6**).

A review by Rapp (Rapp 2000) collated all the hypertension QTL in the rat, at that time 24 QTL across 17 of the 20 autosomes and both sex chromosomes. The virtual comparative (VC) map tool at RGD can now be used to visualise all rat QTL and an example of one chromosome is shown in **Figure 1.1**. This demonstrates the increased amount of

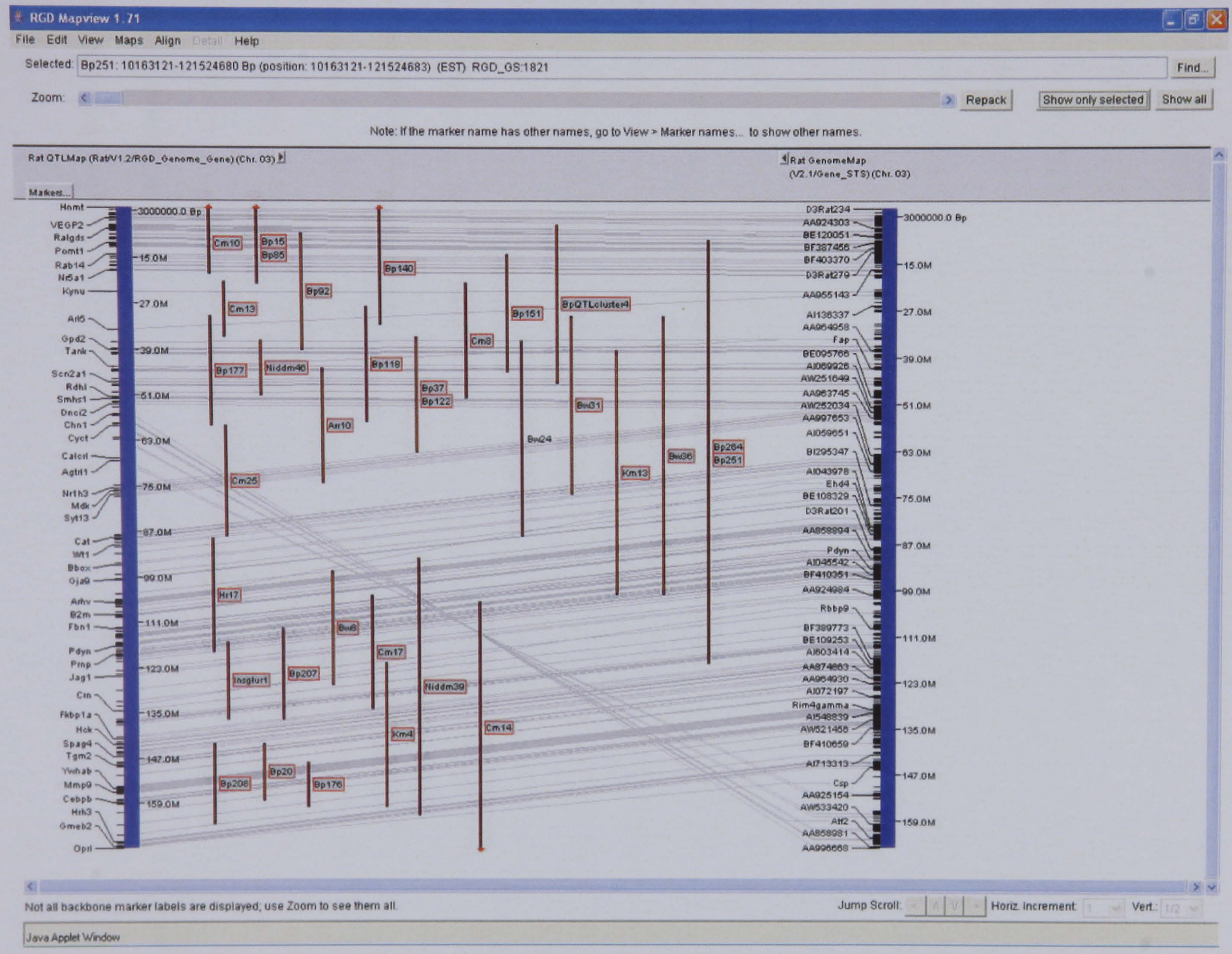


Figure 1.1 Screenshot from VC map showing all rat QTL

This is a screen shot taken from RGD (http://rgd.mcw.edu/) showing all rat blood pressure QTL on chromosome 3 and demonstrates the wealth of available QTL information.

available information and yet still only a handful of QTL genes have been positionally cloned. Elucidation of the molecular basis of these QTL in SHR may throw insights into the related human disorders; however the underlying genes have yet to be identified in most cases.

1.5 Strategies for analysing QTL in rat models

Section 1.3 describes the advantages animal models have over human studies for disease gene identification in complex traits. In animal models, as in human QTL studies, linkage analysis across the whole genome is carried out to identify regions which influence the trait of interest. An experimental cross is set up between two strains, for example one with hypertension and a normotensive control. This means all the animals in the F1 generation are heterozygous at all loci for each of the progenitor (parent) strains. The F1 generation is intercrossed to produce a large F2 population and more than 300 animals are required to achieve statistical significance. The F2 animals are then phenotyped for quantitative traits, such as blood pressure and genotyped across all the markers in the genome to determine which parental alleles is/are present at each marker. This is an example of linkage analysis in a segregating population. Linkage analysis identifies chromosomal regions, containing a gene (or genes) influencing complex traits such as blood pressure, that are frequently too large for positional cloning. The limiting factors in linkage analysis are the modest effects genes often have on the overall phenotype in a complex trait and the limited number of chromosomal recombination events that can be observed in segregating populations of practical size, e.g. F2 generations (Rapp 2000). Therefore approaches have been developed either to more precisely map or narrow the interval or to select genes from within the region based on known relevant biology and or gene expression patterns.

1.5.1 Recombinant inbred strains (RI strains)

Recombinant Inbred (RI) strain panels are one of the most powerful resources developed for disease gene mapping in rats and mice (Silver 1995). Like all inbred strains, RI strains are fixed to homozygosity at essentially all loci and are therefore genetically stable and a renewable resource. Each strain within an RI panel has undergone different recombination events during inbreeding and therefore has a unique allele distribution throughout the genome consisting of blocks of parental alleles, in this case BN and SHR (**Figure 1.2**). The comparison of this distribution of alleles across all the RI strains in a panel generates a unique strain distribution pattern (SDP), which can be correlated with a particular phenotype or other genetic markers, as can be seen for the hypothetical gene xyz in **Figure 1.2**. The advantage of RI strains is that once developed they are a permanent resource, as genetically identical animals can be propagated as required by repeated brother-sister matings. This allows the accumulation of phenotype and mapping data over several years or even decades. Since all RI strains within an RI panel are derived from the same parental strains they are effectively a permanent F2 generation as each strain is a genetic mix of parental alleles and therefore a segregating population in terms of linkage information.

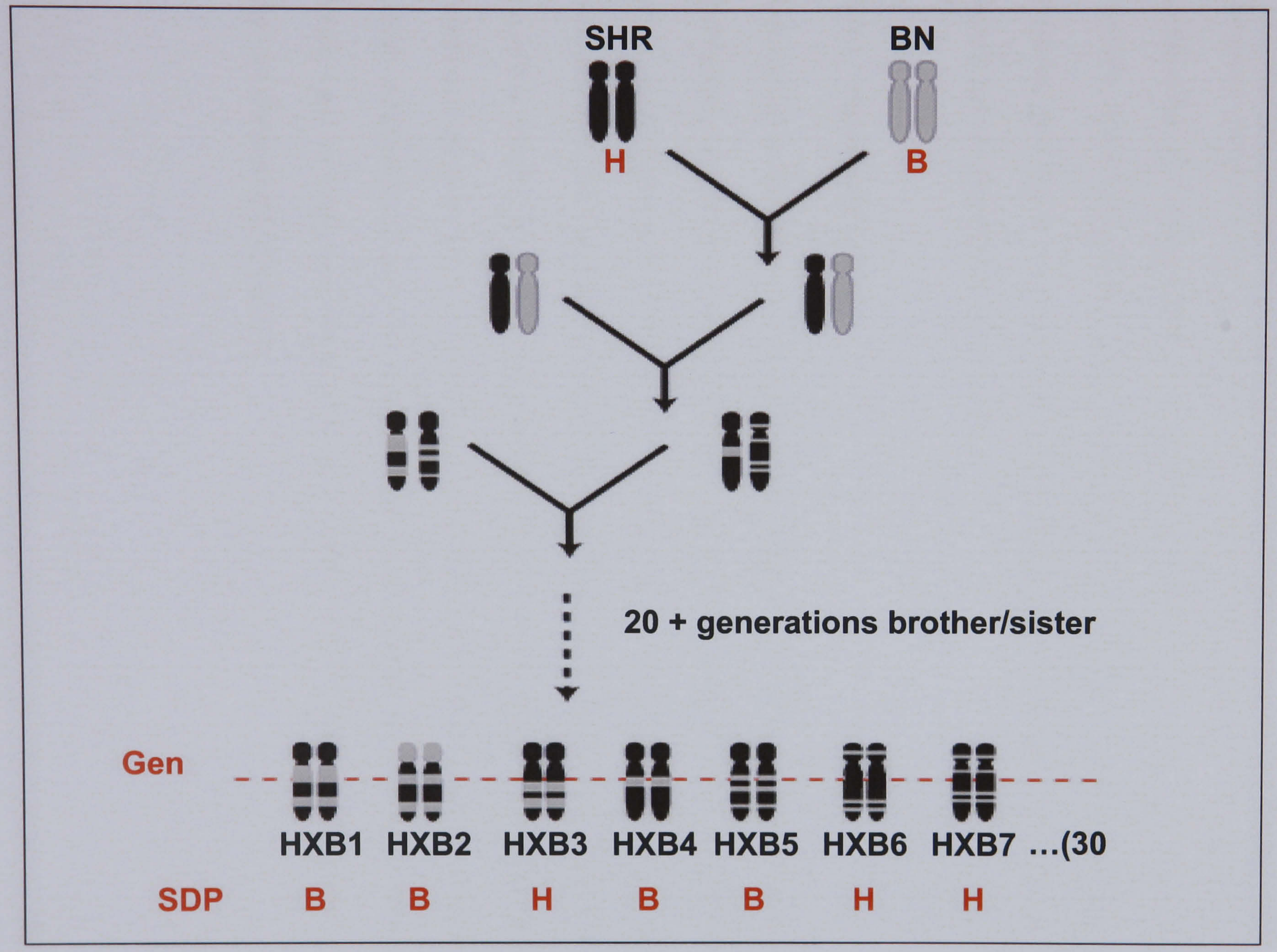


Figure 1.2 Generation of the BXH/HXB RI strains

SHR and BN strains were crossed and the resulting F1 generation intercrossed. Randomly selected pairs of F2 animals were inbred to homozygosity over 20 or more generations. As a result each RI strain has undergone different recombination events during its evolution and therefore has a unique allele distribution of BN (**B**) and SHR (**H**) alleles in the genome. This is termed the strain distribution pattern (**SDP**) and is shown above for the hypothetical xyz gene. The SDP can be compared at each marker to build a linkage map and detect linkage to QT.

1.5.2 Congenic strains

A further advantage of using model systems is the ability to generate congenic strains. In a congenic strain, a specific region of a chromosome is replaced with the corresponding chromosomal region from a different strain. Strain differences on this chromosomal segment can therefore be studied in isolation from strain differences on other chromosomes. Once a congenic strain is constructed that includes a QTL, congenic substrains can be created with smaller and smaller donor fragments to localise the QTL more precisely. Congenic strains have been created for several of the QTL identified in SHR. This is a very powerful technique, which facilitates fine mapping and increases the likelihood of complex gene identification and one that clearly can not be used in human genetics

1.5.3 Positional candidates

Subsequent to identification of a chromosomal region containing a gene (or genes) influencing the complex trait being studied e.g. blood pressure, biological candidate genes in the region can be analyzed for sequence variation between rat strains. The variation detected can then be assessed for association with the phenotype and the causative variant identified. If no known biological candidate exists in the region then analysis of all genes can be carried out. This is facilitated by use of congenic strains which significantly reduce the size of a QTL thus lowering the number of genes which must be tested and studying the gene expression of the genes to determine if they are differentially expressed.

1.5.4 Differential gene expression

It is not uncommon for QTL regions to be as large as 30cM or more, containing many hundreds of genes, the majority of which are uncharacterised and have little sequence polymorphism data available. Therefore, an alternative strategy for identifying candidate genes is to use the analysis of gene expression in affected and unaffected strains. The target gene of interest may be identified by characteristic changes in expression level across a series of samples of varying genotype, for example a mutant gene may show a reduced expression level in samples with the abnormal allele. The availability of expression data may also aid gene identification by revealing pathways related to the target gene and thereby facilitating the identification of candidate genes. A gene expression study using a microarray-based approach identified around 250 genes that were differentially expressed between SHR and BN in adipose tissue. One gene was also found to have a reduced hybridisation signal in SHR when compared with a chromosome 4 congenic strain. This gene was *Cd36* and further analysis confirmed that this gene underlies the chromosome 4 insulin resistance QTL in SHR (Aitman et al. 1999).

1.6 Generation of the BXH/HXB Recombinant inbred Strains

During the 1980's SHR (SHR/Ola) was intercrossed with normotensive Brown Norway (BN/Lx) rats to develop a panel of BXH/HXB recombinant inbred (RI) strains (Pravenec et al. 1989; Printz et al. 2003). To date more than 1000 genetic markers have been mapped on this RI panel forming a high-resolution genetic map for studies in these strains (Pravenec et al. 1996). The framework map contains around 850 markers with an average distance between markers of 2.7cM. Over 45

independent physiological and metabolic phenotypes have already been mapped within this RI panel and new phenotypes are continuously being added.

Mean arterial pressure and left ventricular heart weight are two of the phenotypes that have been measured in this RI strain set (Pravenec et al. 1995) and QTL for blood pressure were found on chromosomes 2, 4 and 19. Systolic blood pressure linked to chromosome 19, diastolic blood pressure linked to chromosome 2 and mean arterial pressure linked to the Il6 marker on chromosome 4. For each phenotype, the SHR allele was associated with an increase in blood pressure. An additional QTL for left ventricular heart weight (corrected for body weight) was identified on chromosome 17, with the SHR allele associated with a lower left ventricular mass than the BN allele (Pravenec et al. 1995). Dyslipidemia was also mapped in the RI strains to chromosomes 4, 8 and 20 (Bottger et al. 1996). Lipid phenotypes of total cholesterol and phospholipid, VLDL cholesterol and phospholipids, LDL/HDL1 cholesterol and phospholipids, IDL cholesterol and phospholipids and HDL2 phospholipids were measured. The SHR/Ola strain exhibited lower levels of total cholesterol and phospholipids, lower levels of VLDL cholesterol and phospholipids and higher levels of HDL2 phospholipids than the BN/Lx strain, when both were fed a high fat/high cholesterol diet. In addition, association was seen in the RI strains between basal levels of HDL2 phospholipids and the Il6 marker on rat chromosome 4 ($P<0.0005$), with animals carrying the SHR allele having lower levels of HDL2 phospholipid (Bottger et al. 1996). Finally QTL on chromosomes 4, 12 and 16 associated with defective insulin-mediated glucose uptake were mapped and a significant linkage was seen in the RI strains to chromosome 7 for basal glucose uptake (Aitman et al. 1997b).

1.7 Microarray Expression profiling

1.7.1 DNA microarrays

DNA microarrays have been developed to analyse the expression of numerous genes simultaneously. There are two main types: 1) cDNA clones or their PCR-amplified products are spotted on the slide and 2) oligonucleotide sequences representing selected genes, are spotted or synthesised on the array. Technologies developed for computer chips allow thousands of genes to be spotted on a single slide allowing the interrogation of the entire genome in a single experiment. The two types are manufactured differently but the experimental analysis principles are very similar (Harrington et al. 2000), both involve reverse transcription of RNA from a cell or tissue type and hybridisation of the resultant cDNA or cRNA onto the array.

Microarrays have been utilised in a number of ways and one of the first successful applications to human disease was in the study of cancer classification in acute leukaemias for class discovery and class prediction (Golub et al. 1999). Since the 1960s, the acute leukaemias have been classified as either acute lymphoblastic leukaemia (ALL) or acute myeloid leukaemia (AML) although no single test is sufficient for diagnosis. Distinguishing ALL from AML is critical for successful treatment. The authors identified 1100 genes with an expression pattern that was strongly correlated with the AML-ALL class distinction. The expression levels of a subset of 50 genes most strongly correlated with class distinction were used to create a class predictor capable of assigning a new sample to one of the two classes. The list of informative genes also provided new markers of acute leukaemia subtype. The class discovery procedure could distinguish between AML and ALL without previous

knowledge of these classes and identified two subclasses of ALL: T-lineage ALL and B-lineage ALL.

This type of analysis also generates potentially useful data concerning sets of genes that are differentially regulated between normal and disease cells. This may help understand disease pathophysiology and identify other genes that act in the same way, or in compensatory pathways, as the primary defect.

1.7.2 Integrated linkage analysis and gene expression approach to disease gene identification

A novel approach using integrated linkage analysis and gene expression profiling to identify genes involved in complex traits was proposed in 2001 (Jansen and Nap 2001). This strategy, called "genetical genomics", uses the heritability of gene expression profiles and this bridges the gap between genotype and phenotype since transcript abundance is an intermediate phenotype between genomic sequence variation and phenotype. In this approach gene expression data is treated as a quantitative phenotype much as in conventional linkage analysis traits such as blood pressure or glucose levels would be treated quantitatively. Microarray based expression profiles are generated from all individuals in a segregating population and then mapped to the genome by the same methods of linkage analysis used to map physiological phenotypes. The BXH/HXB RI panel is an example of a segregating population in which replicates can be generated by study of multiple animals with identical genotype from within each strain. The example described in **Section 1.7.1** used microarrays and linkage analysis to identify the *Cd36* gene and the difference of the method described here is that this is on a whole genome scale.

Just as linkage analysis and mapping of a physiological phenotype results in a QTL (or QTLs) for that phenotype, mapping the expression profile of a particular gene onto the genome results in QTL(s) for expression of that gene. The expression profile acts as a phenotype and the resulting QTL can be termed an expression QTL (eQTL) which is distinct from a physiological QTL or pQTL, where a phenotype such as blood pressure is mapped.

This genetical genomic approach allows the genetic determinants of gene expression to be mapped by linkage analysis to discrete regions of the genome and will indicate what proportion of the variation in gene expression maps to the same chromosomal region as the gene itself (*cis*-acting eQTL) or to another region of the genome (*trans*-acting eQTL) (**Figure 1.3**). In the case of a *trans*-acting eQTL, the region of the genome to which the eQTL maps is regulating (*trans*-regulating region) expression and the gene itself is being regulated (*trans*-regulated gene). This approach not only allows *cis* and *trans*-acting factors to be identified but can also be used to identify gene-gene interactions and thus pathways can be determined. *Cis*-acting eQTL are caused by genomic sequence variants that reside within or close to the gene while for a *trans*-regulated gene the polymorphism that cause the differential expression will reside elsewhere in the genome, at the site of the eQTL (*trans*-regulating region). These eQTL may be considered as postional candidates for previously mapped QTL when they reside within a pQTL (**Figure 1.3**).

The first study to use this integrated whole genome approach was carried out using a simple eukaryote, *S. cerevisiae*. Brem *et al* (Brem et al. 2002) performed genetic

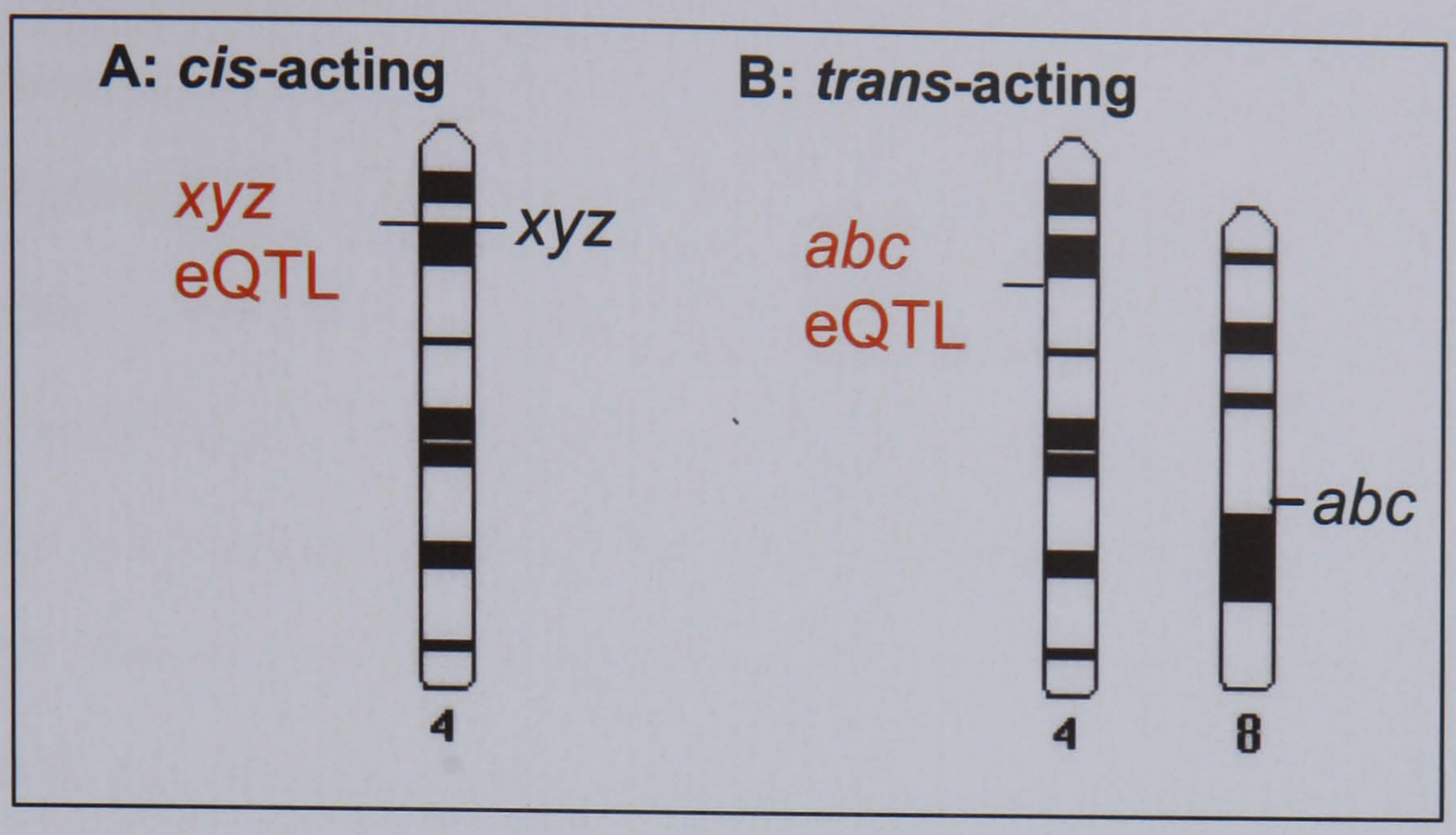


Figure 1.3 Control of expression QTL (eQTL)
eQTL are shown in red and genes in black for the hypothetical *xyz* and *abc* genes. **A.** Expression of the *xyz* gene is linked to the *xyz* gene itself and is therefore *cis*-acting, while **B.** the eQTL for gene *abc* is linked to a different region of the genome as the *abc* gene and is therefore regulated in *trans*.

linkage analysis of gene expression profiles in the offspring of a yeast cross and were able to define co-ordinated patterns of gene expression. They measured expression of all 6125 genes and detected over 1500 as differentially expressed between the parental yeast strains. Expression was also measured in 40 haploid offspring and 570 eQTL were identified of which 308 were in the genes differentially expressed between the parents. They concluded that differential gene expression is highly heritable and that there is a high rate of *cis*-acting alleles and a small number of *trans*-regulating factors with widespread effects. Eight *trans*-acting loci were found to act on five or more regions, each of which controlled expression of up to 100 genes of related function.

In a follow up paper Yvert et al (Yvert et al. 2003) used this data in a cluster analysis and determined that one third of all genes could be clustered. Each cluster when annotated was found to contain genes involved in similar processes. In addition, linkage of these clusters to the region of the genome controlling the expression of the genes in the cluster was carried out to identify *trans*-acting factors.

Schadt and colleagues (Schadt et al. 2003) were the first to utilise this approach in a mammalian system, using a mouse F2 population. From a total of approximately twenty-three thousand genes, 7,861 were determined to be differentially expressed in the parents or 10% of the F2 animals. Also without filtering for differential expression 4,339 eQTL were detected in 3,701 genes. Again genes determined to be controlled in *cis* resulted in higher LOD scores than those eQTL which were *trans*-regulated. They also discovered "eQTL hotspots", regions of the genome containing more eQTL than would be expected by chance, which could be indicative of possible regulatory elements which control expression of many genes. These regulatory elements could explain the correlation of expression patterns of genes on a particular pathway. Finally using cluster analysis this study identified both a gene expression pattern strongly associated with obesity and two obesity subtypes within the group, that were under the control of distinct loci. This is similar to the cancer classification decribed in **Section 1.7.1.**

Overall these studies demonstrate the power of this approach in that parental differences in gene expression were shown to be highly heritable and that known pathways can be detected.

1.8 Comparative Genomics

Comparative genomics or comparative mapping is a powerful approach to identify systematically genes and the important regulatory elements in human by looking at evolutionary conservation. It exploits the similarities and differences between the sequences of genomes of different organisms to determine which nucleotides have remained the same over long periods of time, that is, since the animals split in

evolutionary history. This provides important insights into gene structure and function and can thereby be used to develop strategies to combat human disease based on the analysis of the genomes of model organisms.

1.9 Aims of this work

The present study aims to use transcriptional profiling in a segregating population to integrate the substantial body of genetic and phenotypic data available for the BXH/HXB RI strains by generating and mapping gene expression profiles across the 30 strains (and SHR and BN/Lx parentals) using DNA microarray technology.

Following generation of gene expression data for the RI strains, expression profiles are mapped as a quantitative phenotypes. The clustering and map location data that will be derived will provide a new description of the metabolic pathways that mediate cardiovascular and metabolic traits represented within these RI strains and will suggest new regulatory control points and interactions between these pathways.

The specific aims presented in this thesis are to:

1. generate microarray data for the BXH/HXB recombinant inbred strains and SHR and BN/Lx parental strains.
2. use linkage analysis of the microarray data set to identify differentially expressed genes that are controlled in *cis* and in *trans*.
3. prioritise genes as candidates for sequencing in the promoter region to identify variants which may underlie the differential expression and thus biological phenotype in *cis*-regulated eQTL.

4. carry out comparative mapping of *cis* eQTL to identify potential candidate genes for human cardiovascular and metabolic phenotypes.

Chapter 3 describes the generation of the microarray data sets and the linkage analysis carried out. It also discusses some of the quality control and preliminary analyses performed on the data to assess the optimum methods.

Chapter 4 examines a subset of the *cis*-acting eQTL to look for polymorphisms in the putative promoters of these genes. These polymorphisms may cause the differential expression detected in these strains and underlie the phenotypic variation seen.

Chapter 5 describes a comparative mapping study carried out using human QTL for metabolic and cardiovascular phenotypes to identify rat *cis*-acting eQTL and their human orthologs as potential candidate genes.

Chapter 2: Materials and methods

2.1 Stock solutions

2.1.1 Solutions for DNA and RNA analysis

10 x MOPS	41.8g MOPS, 16.6ml DEPC-treated 3M sodium acetate (pH 5.2), 20ml DEPC-treated 0.5M EDTA in 800ml of ddH_2O. Adjust pH to 7.0 with 5M NaOH. Adjust volume to 1l with ddH_2O. Filter-sterilise and store in the dark.
10 x TBE	108g Tris base, 55g boric acid, 40ml 0.5M EDTA (pH 8) in 1l of ddH_20. Stir until the solutes have dissolved.
DEPC water	1ml of diethyl pyrocarbonate (DEPC) per 500ml of water. Shake and leave overnight before autoclaving.
EDTA	186g Na_2EDTA in 800ml of water. Adjust pH to 8.0 with 5M NaOH. Adjust volume to 1l with ddH_2O.
Loading buffer	0.25% w/v bromophenol blue, 0.25% w/v xylene cyanol, 30% w/v glycerol in water.
Lysis buffer	100mM Tris-HCl (pH 8.5), 5mM EDTA, 0.2% SDS, 200mM NaCl in water.
Proteinase K	20mg/ml in H_2O, stored at -20°C.
RNA loading buffer	200µl 2.5% bromophenol blue, 200µl 2.5% xylene cyanol, 400µl 0.5M EDTA in 9.2ml deionised formamide. Aliquot and store at -20°C.
RNA premix	2.5ml 10 x MOPS, 4.4ml 37% formaldehyde, 12.5ml deionised formamide. Store at -20°C in single-use aliquots.

TE buffers	Various combinations of Tris-HCl and EDTA. Typically
$T_{10}E_{0.1}$	(pH 7.5); 10mM Tris-HCl, 0.1mM EDTA (pH 7.5).

2.2 Preparation of DNA

2.2.1 Extraction of DNA from rat tissue

The DNA extraction protocol was modified from Laird *et al.* (Laird et al. 1991).

Genomic DNA was extracted from rat liver tissue. After CO_2 narcosis, animals were sacrificed by cervical dislocation. Tissues were removed quickly, placed in bijous and immediately placed into liquid nitrogen. For long-term storage, tissues were kept at -70°C.

A small piece of tissue was removed from -70°C storage and added to 750µl of lysis buffer plus proteinase K at 100µg/ml in an eppendorf. Lysis was carried out overnight, at 55°C in a gyratory water bath. Following lysis, the samples were vortexed then centrifuged in a bench-top microfuge (Eppendorf, Cambridge, UK) at 10,000 rpm for 10 minutes. The supernatant was placed in a fresh tube, 750µl of isopropanol was added and the sample mixed until precipitation was complete. The DNA was removed using a sterile plastic loop and resuspended in 500µl of TE.

DNA yield and purity was determined by absorption at 260nm and 280nm UV using the Gene Quant spectrophotometer (Amersham Pharmacia Biotech UK Ltd, Little Chalfont, UK). An OD unit of 1.0 indicated a concentration of 50µg/ml of double stranded DNA. The 260nm/280nm ratio was expected to fall between 1.7 and 2.0.

2.3 Preparation of RNA and cDNA

2.3.1 Extraction of RNA from rat tissue to select and optimise procedures

RNA was extracted from rat peritoneal fat and whole kidney tissue. After CO_2 narcosis, animals were sacrificed by cervical dislocation. Tissues were removed quickly, to avoid RNA degradation, placed in bijous and immediately placed into liquid nitrogen. For long-term storage, tissues were kept at -70°C.

Two methods were tested for RNA extractions:-

1) RNA was extracted from tissue using the RNAzol^{TM} B reagent (BioGenesis, Poole, UK). Frozen tissue was homogenised in 2ml of RNAzol^{TM} B per 100mg of tissue, 0.1 volumes of chloroform was added and the samples shaken vigorously. After incubation on ice for 5 minutes, the suspension was centrifuged at 12,000*g* for 15 minutes. The aqueous layer was transferred to a fresh tube, 1 volume of isopropanol was added and the sample incubated at 4°C for 15 minutes. The sample was centrifuged at 12,000*g* for 15 minutes and the supernatant removed. The RNA pellet was washed once with 75% ethanol and resuspended in DEPC-treated 1mM EDTA, pH 7. The RNA sample was stored at -70°C in single-use aliquots.

2) RNA was extracted from tissue using the Trizol^{TM} reagent (Invitrogen, Paisley, UK). Frozen tissue was homogenised in 5ml of Trizol^{TM} per 100mg of tissue and incubated at room temperature for 5mins. Next, 0.6mls per 100mg of tissue of chloroform was added and the samples shaken vigorously. After incubation at room temperature for 2-3 minutes, the suspension was centrifuged at 12,000*g* for 15

minutes. The aqueous layer was transferred to a fresh tube, 1 volume of isopropanol was added and the sample incubated at room temperature for 10 minutes. The sample was centrifuged at 12,000*g* for 10 minutes and the supernatant removed. The RNA pellet was washed once with 4mls of 75% ethanol and resuspended in DNase/RNase free water (Sigma-Aldrich, Irvine, UK). The RNA sample was stored at -70°C in single-use aliquots.

RNA yield was determined through absorption at 260nm UV and visualised initially on a 1% agarose gel with formaldehyde and ultimately using the Agilent bioanalyser. An appropriate dilution of the RNA was made and the OD at 260nm and 280nm UV was read using the Gene Quant spectrophotometer (Amersham Pharmacia Biotech UK Ltd, Little Chalfont, UK). An OD unit of 1.0 indicated a concentration of 40µg/ml of RNA. The 260nm/280nm ratio was used to assess the purity of the RNA sample and was expected to be in the range of 1.7 to 2.0.

2.3.2 First strand cDNA synthesis

Prior to first strand cDNA synthesis, total RNA was treated with Amplification Grade DNaseI (Invitrogen, Paisley, UK) which was inactivated by addition of a final concentration of 2.5mM EDTA and incubation at 65°C for 10 minutes. The DNase digested RNA was reversed transcribed into cDNA with the 1st Strand Synthesis Kit according to manufacturers instructions (Roche, Lewes, UK). Synthesis was carried out in a 20µl reaction volume with 1 x reaction buffer, $MgCl_2$ (5mM), dNTPs (1mM), random primer (3.2µg), RNase inhibitor (50U) and AMV reverse transcriptase (>20U) with 8.2µl of DNase digested RNA. The sample was incubated at 25°C for 10 minutes, followed by 42°C for 60 minutes and 99°C for 5 minutes. The cDNA was

purified through a G50 Spin Column (Amersham Pharmacia Biotech UK Ltd, Little Chalfont, UK) and stored at -20°C.

2.4 Amplification of DNA (PCR)

2.4.1 Standard protocol for PCR

PCR primers were designed both by hand and using Primer 3 (http://frodo.wi.mit.edu/cgi-bin/primer3/prime_3www.cgi). Primer pairs were checked to ensure non-complementary 3' ends and designed to ensure that the primers had a similar melting temperature (Tm). Tm was calculated as:

$$2(A + T) + 4(G + C).$$

Oligonucleotide primers were synthesised by SigmaGenosys Biotechnology (Cambridge, UK). When Primer3 was used the rodent and simple mispriming library was selected and the default settings of primer length 20 with Tm of 60°C were used and they were also checked by eye and using the above formula.

PCR reactions were performed in 10µl volumes and contained a final concentration of 1 x reaction buffer, 0.2mM dNTPs, 0.66µM of each oligonucleotide primer and 0.5U of *Taq* polymerase (Bioline, London, UK). Optimised concentrations of $MgCl_2$ were used with the final concentration in the range of 1.0 - 2.0mM. Genomic DNA was used at a final concentration of 5ng/µl. Stock solutions were stored at -20°C. PCR reactions containing water instead of DNA as a 'water blank', were also performed to act as negative controls.

PCR was carried out in 96-well plates or individual 0.2ml tubes (ABGene, Epsom, UK) without an oil overlay using a Tetrad thermal cycler (MJ Research, Rayne, UK). 96-well plates were sealed with heat-sealing foil and a Thermo-Sealer (ABGene, Epsom, UK). Standard amplification conditions were 94°C for 3 minutes, followed by 30-40 cycles consisting of denaturation at 94°C for 30 seconds, an optimised annealing temperature for 45 seconds, extension at 72°C for 1 minute and final extension at 72°C for 8 minutes. If Amplitaq Gold (Applied Biosystems, Warrington, UK) was used, the initial denaturation step was increased to 12 minutes.

Each primer pair was optimised for $MgCl_2$ concentration in the range of 1-2.5mM. Most primers were used at 1.5mM $MgCl_2$. Each primer was also optimised for annealing temperature in the range of 50°C to 65°C, according to the Tm of the primers.

2.5 Gel electrophoresis

2.5.1 Agarose gel electrophoresis

Agarose gel electrophoresis was carried out using horizontal submarine agarose gels. For agarose mini-gels, 10cm wide mini Sub^{TM} cells were used (Bio-Rad, Hemel Hempstead, UK) with 1 x TBE used as gel buffer and running buffer. Gels were made by dissolving agarose powder (Invitrogen, Paisley, UK) in gel buffer and boiling in a microwave. Agarose was cooled to 60°C before pouring. Agarose concentrations ranged between 1% and 4%. For agarose gels of concentration greater than 3%, standard agarose and NuSieve agarose (Flowgen, Ashby de la Zouch, UK)

were combined in a ratio of 1:3. Ethidium bromide was added to the gel and running buffer using 2µl of 5mg/ml stock per 100ml of gel or running buffer. DNA samples were loaded with 1/3 volume of loading buffer and electrophoresis carried out at 120 volts for 30 - 45 minutes. ϕx174/*Hae*III (AB Gene, Epsom, UK) and λ/*Hin*dIII fragments (Invitrogen, Paisley, UK) were used as DNA size markers.

DNA was visualised using a UV transilluminator (Camlab Ltd, Cambridge, UK) and photographed using polaroid film taken on a Polaroid MP4 land camera (Kodak, Hemel Hempstead, UK) with a red filter.

2.5.2 Denaturing gel electrophoresis

Formaldehyde gel electrophoresis was carried out to separate total RNA. Separate chemicals, designated for RNA work, were used for all these experiments. 1 x MOPS was used as running buffer for 15cm wide Sub™ cell (Bio-Rad, Hemel Hempstead, UK) formaldehyde gels. Formaldehyde gels were made by dissolving agarose powder (Invitrogen, Paisley, UK) in DEPC-treated water and boiling in a microwave. Agarose was cooled to 60°C and 10 x MOPS and formaldehyde were added. The gel was poured and allowed to set for 40 minutes. Five micrograms of RNA was combined with RNA premix, RNA loading buffer and ethidium bromide (0.5µg/µl). Samples were incubated at 55°C for 15 minutes and quenched on ice. Electrophoresis was carried out at 70 volts for 3 hours. 0.24-9.5kb RNA ladder (Invitrogen, Paisley, UK) was used as RNA size standard.

Following electrophoresis, RNA was visualised using a UV transilluminator (Camlab Ltd, Cambridge, UK) and photographed using polaroid film taken on a Polaroid MP4 land camera (Kodak, Hemel Hempstead, UK) with a red filter.

2.6 Agilent Bioanalyser

To assess RNA quality, 1µl of RNA (25-500ng) from each sample was tested using the Agilent Bioanalyser RNA 6000 Nano assay. Each chip can assess 12 samples simultaneously. Before use, the Bioanalyser electrodes were decontaminated by applying one electrode cleaner with 350µl of RNAZap (Ambion, Huntington, UK) to the electrodes for one minute followed by one electrode cleaner containing 350µl of water for 10 seconds. The electrodes were then allowed to dry. Filtered gel matrix was prepared by pipetting 550µl of RNA matrix into a spin filter and centrifuging at 1500*g* for 10 minutes. This was aliquoted into 65µl aliquots for use. Immediately prior to running the Bioanalyser assay 1µl of vortexed RNA dye concentrate was added to a 65µl aliquot of filtered gel matrix and centrifuged at 13,000*g* for 10 minutes. 9µl of this gel dye mix was used to prepare the RNA 6000 nanochip in the chip priming station. Once the chip was primed another 9µl was pipetted into each well of the chip marked with a G. 5µl of RNA 6000 Nano Marker was added to the remaining 13 wells of the chip. 1µl of each sample and 1µl of RNA 6000 Ladder were denatured at 70°C for 2 minutes and placed on ice before being added to the wells on the chip. The chip was then vortexed for 1 minute and run on the *Eukaryotic total RNA Nano* assay.

2.7 Preparation of RNA, cDNA and cRNA for Affymetrix GeneChip™ studies

2.7.1 Extraction and purification of RNA from rat fat tissue

One peritoneal fat pad from each animal in the study was added to 3ml Trizol™ for homogenisation. RNA extraction was carried out as detailed in **Section 2.3.1** (Method 2).

RNA samples were purified using the RNeasy Mini Kit (Qiagen, Crawley, UK). RNA was extracted and purified from 4 animals of each rat RI strain and 5 from each parental strain. RNA samples were concentrated to 2µg/µl by ethanol precipitation as follows. First the sample was added to an 1.5ml eppendorf tube containing 3 x volume of 100% ethanol and 1/10 volume of sodium acetate (pH 4.6). Tubes were vortexed and allowed to stand at room temperature for 15 minutes. The samples were then centrifuged in a bench-top microfuge (Eppendorf, Cambridge, UK) at maximum speed for 20 minutes. The supernatant was removed, and the pellet washed with 70% ethanol. Following vortexing, centrifugation and removal of the supernatant, the pellet was air-dried at room temperature. Pellets were resuspended in the appropriate volume of DNase/RNase free water to ensure the RNA concentration was 2µg/µl.

2.7.2 Double stranded cDNA synthesis and purification

Ten micrograms of total RNA was incubated at 70°C for 10 minutes with 100pmol T7-$(T)_{24}$ primer and Poly A^+ controls. After incubation and quenching on ice, first strand synthesis was carried out in a 20µl reaction volume with 1 x First strand buffer, 0.1M DTT, 10mM dNTPs and 400U Superscript II reverse transcriptase (Invitrogen,

Groningen, The Netherlands). Samples were incubated at 42°C for 1 hour and quenched on ice before proceeding to second strand synthesis.

Second strand cDNA synthesis was performed on each sample with 1 x Second strand buffer, 10mM dNTPs, 10U DNA ligase, 40U DNA Polymerase I and 2U RNase H (Invitrogen, Groningen, The Netherlands) to a final volume of 150µl. Samples were incubated at 16°C for 2 hours.

The double stranded cDNA was purified using GeneChip ® Sample Cleanup module (Affymetrix, Santa Clara) according to the manufacturer's instructions.

2.7.3 cRNA synthesis and purification

cRNA was generated to allow the incorporation of biotin-labelled nucleotides for detection of microarray hybridisation. *In vitro* transcription was performed on double stranded cDNA preparations using the Bioarray High Yield RNA Transcript Labelling Kit (Enzo Diagnostics Inc., New York, USA) which contains biotin-labelled nucleotides. Samples were incubated at 37°C for 5 hours, with gentle agitation at hourly intervals.

cRNA samples were purified using the GeneChip® Sample Cleanup module according to the manufacturer's instructions.

cRNA yield and quality was determined through absorption at 260nm UV and using the Agilent Bioanalyser as described in **Section 2.6**. An appropriate dilution of the cRNA was made and the OD at 260nm and 280nm UV was read using the Gene

Quant spectrophotometer (Amersham Pharmacia Biotech UK Ltd, Little Chalfont, UK). An OD unit of 1.0 indicated a concentration of 40μg/ml of cRNA. cRNA yields were expected to be in the range of 40 to 110μg. The 260nm/280nm ratio was expected to be in the range of 1.8 to 2.1.

2.7.4 Hybridisation of cRNA to Affymetrix GeneChips

Biotin-labelled cRNA samples were given to the Imperial College Microarray Centre for fragmentation and hybridisation to Test chips and GeneChips™ according to the protocol provided by Affymetrix (High Wycombe, UK). After hybridisation of the biotinylated cRNA, streptavidin-phycoerythrin (SAPE) and a secondary biotinylated goat anti-streptavidin antibody is added to the hybridisation mix. Briefly, the SAPE binds to the biotinylated cRNA and is subsequently bound by the goat anti-streptavidin antibody which is also biotinylated. The addition of a second dose of SAPE amplifies the hybridisation signal. In these studies, all arrays were analysed after antibody amplification. 10μg of cRNA was hybridised to a Test chip to assess protocols (e.g. background signal, signal to noise ratio, percentage of present calls) before the RI strain and parental samples were applied to the RAE230A GeneChip™. 15μg cRNA was hybridised to each GeneChip™. Initial analysis of the GeneChips™ was carried out using Microarray Analysis Suite 5.0 (Affymetrix, Santa Clara, CA, USA). This software gives present, marginal or absent calls for each gene or EST and generates values for the average difference in signal between perfect match and mismatch probes.

2.7.5 Expression data from RI strain kidneys

A second data set was generated in the RI strains by Norbert Hübners' group in the Max-Delbrück Centre, Berlin using the kidney tissue from the same animals used in the adipose tissue experiments in London and also hybridised to Affymetrix R230A GeneChips. This student acknowledges permission to include this data set in this thesis as it provides a comparison for the fat data set and a more complete data set overall since both tissues are involved in the metabolic syndrome.

2.8 Sequencing of PCR products

Template was prepared from a 50µl PCR reaction. PCR was performed as per the standard PCR protocol (**Section 2.4.1**), and all volumes were scaled up in proportion. 5µl of the PCR product was resolved on an agarose gel to check that a single, intense band had been amplified. The remaining 45µl was purified through a Centricon 100 column (Millipore, Watford, UK) according to manufacturer's instructions. If a single band could not be obtained, the bands were excised from the gel and the DNA was purified using the QiaQuick Gel Purification System (Qiagen Ltd, Crawley, UK).

Template reactions were prepared containing 30-90ng of purified PCR product and 3.2pmol of primer, with water added to a final volume of 10µl. Template reactions were given to the local sequencing service for cycle sequencing, purification and electrophoresis according to the protocols given below.

Cycle sequencing was performed on a Perkin Elmer 9700 Thermal Cycler using the BigDye v3.1 Sequencing Kit (Applied Biosystems, Warrington, UK). Five

microlitres of template reaction was added to 0.9µl of BigDye, 2µl of water and 2µl 5X buffer. Reactions were cycled with an initial denaturation step at 96°C for 1 minute, followed by 35 cycles consisting of 96°C for 10 seconds, 52°C for 15 seconds and 60°C for 4 minutes.

Cycle sequencing products were purified using EDTA-EtOH chemistry clean up. Five microlitres of 125mM EDTA and 30µl of 100% Ethanol is added to 10µl of sequencing reaction and incubated at room temperature for 15mins before centrifugation at 3,000*g* for 30mins. The pellets are washed once with 100% ethanol and then twice with 70% ethanol. The DNA is resuspended in 10-12µl of HiDi Formamide. Sequencing was performed on an ABI 3730 DNA Analyser (Applied Biosystems, Warrington, UK) for 4.5 hours. DNA sequences were analysed using both Sequence Navigator and Sequencher software (Applied Biosystems, Warrington, UK).

2.9 Linkage Analysis and definition of eQTL

2.9.1 Mapping of eQTL

2.9.1.1 Using Map Manager QTX

Expression QTL were mapped initially probe set by probe set using the Map Manager QTX program (Manly et al. 2001). This program generates likelihood ratio statistics for one quantitative trait (QT) to all the markers in the genome. Therefore each probe set (which is in essence a QT) needs to be added individually. Genes were prioritised based on the fact that they were differentially expressed in the parental strains since it was not practical to input all 15,000 probe set values from the microarray manually.

At the time no outlier detection could be carried out as not enough replicates were available. Mean expression values for each RI strain were derived and entered into the Map Manager program to generate a list of significantly linked markers for that probe set.

2.9.1.2 Mapping eQTL across the whole genome using QTL Reaper

Mean expression values from the four replicates for each RI strain and in the fat and kidney tissue were derived after application of the Nalimov outlier test at $P < 10^{-3}$. This test was carried out by Dr Enrico Petretto of the Physiological Genomics and Medicine group. Genome-wide linkage analysis for each of the 15,923 expression traits and the 1,011 genetic markers was carried out using the QTL Reaper program (© 2002 Kenneth F. Manly) which generates a likelihood ratio statistic (LRS) as a measure of the significance of a possible eQTL. QTL Reaper establishes genome-wide significance by permutation test, estimating an empirical genome-wide probability for observing a given LRS score by chance (Churchill and Doerge 1994). For each probe set permutations are carried out until a LRS greater than that for the real data is observed, or 1,000,000 permutations have been done, conditional on at least 1,000 permutations being performed. The permutation procedures for the QTL Reaper program correct for individual expression phenotypes for multiple testing across genetic markers to give a genome-wide corrected *P*-value, but not for multiple testing across the 15,923 expression measurements. The number of falsely discovered linkages at a given genome-wide significance level was estimated by Dr Enrico Petretto. The q-value (Storey 2002), which is defined as the minimum positive FDR for a fixed significance threshold was calculated.

2.9.3 Definition of *cis*- and *trans*-acting eQTL

Following the QTL Reaper analysis the eQTL defined were classified as being regulated in *cis*, that is by the gene itself or in *trans*, by another gene. *cis*-acting eQTL, are defined in this thesis as when the most significantly linked marker maps 10Mb up or down stream of the physical location of the probe set. eQTL where the marker was more distant than 10Mb or on another chromosome were defined as being regulated in *trans*. The analysis to define all eQTL as either *cis* or *trans* was automated using an algorithm devised by Michael Muller in the group following discussion of the method and decision on the size of the *cis* window with this student and Drs Caroline Wallace and Enrico Petretto.

2.10 *In silico* analysis

Probe set details such as sequence and mapping information were obtained from the Affymetrix NetAffx analysis centre (http://www.netaffx.com). Physical locations of probe sets were obtained from Affymetrix, or by mapping to the genome using BLAST both at Ensembl (http://www.ensembl.org/Rattus_norvegicus/blastview) and NCBI (http://www.ncbi.nlm.nih.gov/BLAST/). Finally, during the time this work was being carried out both the Ensembl and UCSC (http://genome.ucsc.edu/) genome browsers mapped all Affymetrix probe sets to the genome assemblies and these were used to locate probe sets as this data became available.

Previously mapped SHR physiological QTL were identified using the Rat Genome Database (http://rgd.mcw.edu/). These were then checked against the reference sequences to ensure reported *P*-values and peaks of linkage were correct. In addition, a literature search was carried out using Pubmed

(http://www.ncbi.nlm.nih.gov/entrez/) and Google (http://www.google.com/) to ensure all SHR pQTL were recorded.

Human QTL from genome wide scans for type 2 diabetes, insulin resistance, dyslipidemia, hypertension and cardiovascular phenotypes were collated. RGD was again used as a basis for this and the information from there was checked against the orginal references and a literature search, again using Pubmed, was carried out to identify other QTL not already included in RGD.

2.11 Description of the generation and analysis of data sets used in this thesis

Figure 2.1 Flowchart describing the preliminary data sets

The preliminary data sets were used to generate **Figure 3.6** and **Figure 3.9**. There were six gene expression profiles from the parental strains normalised by both MAS5.0 and RMA algorithms. These data sets were used to compare the normalisation methods. RMA was selected and used to select genes that were differentially expressed in the parental strains. These 47 genes were then tested for linkage in the available 97 expression profiles RI strains using Map Manager QTX. The most significantly linked marker for each probe set was selected and these are shown in **Figure 3.6** and **Figure 3.9**. Twenty of these were selected for sequencing in the target region.

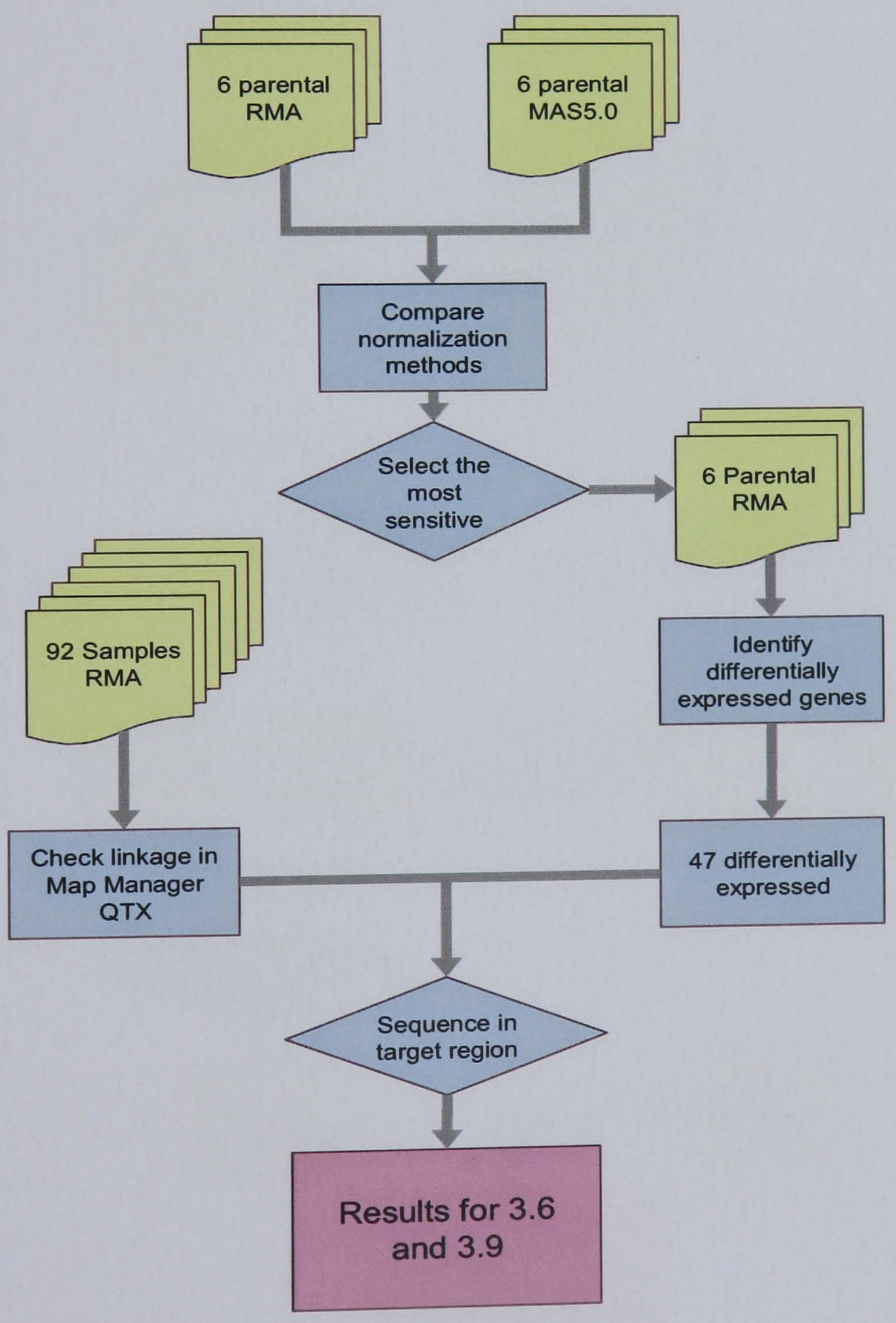
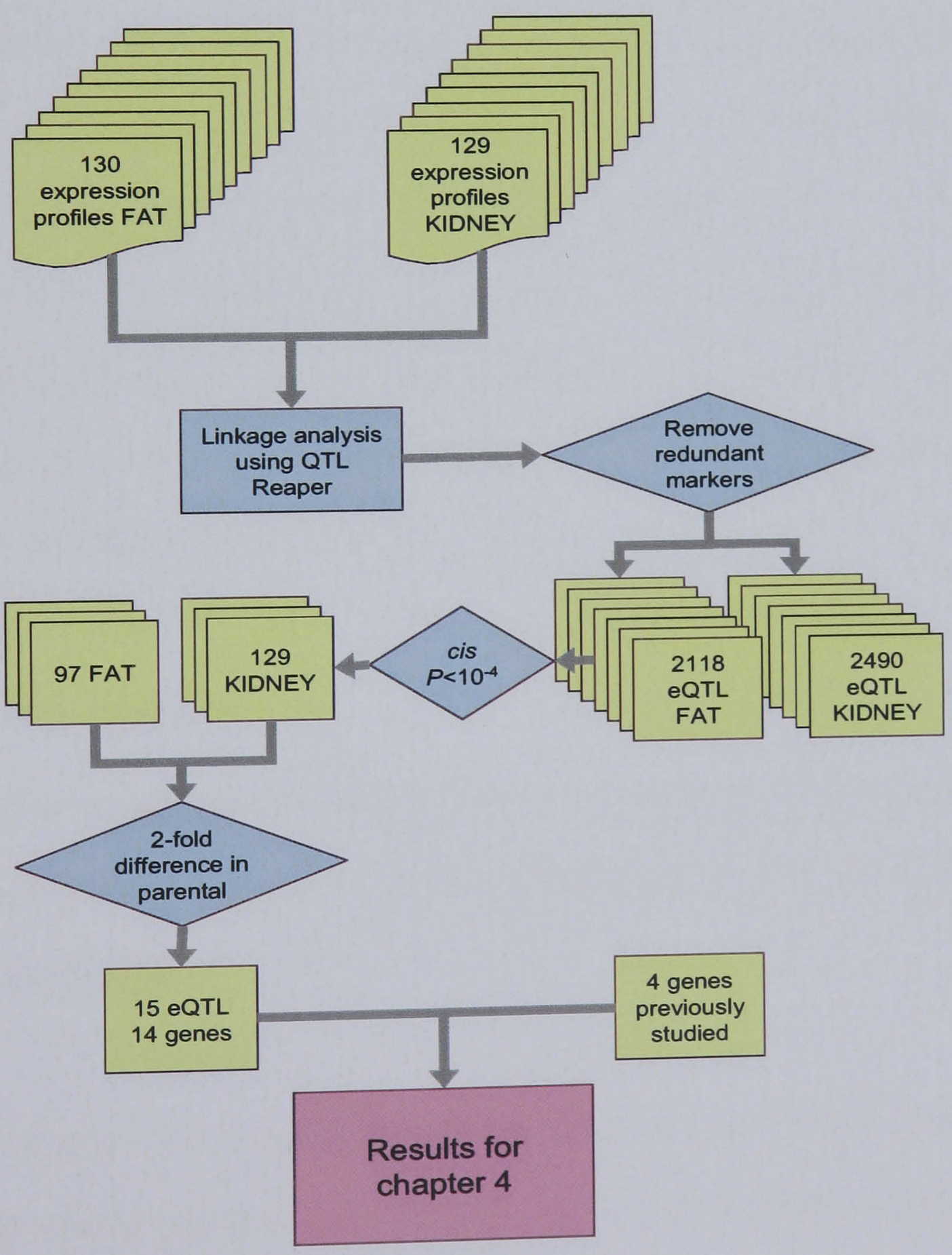


Figure 2.2 Flowchart describing the final data set

The final data set was analysed in **Chapter 4**. There were 130 gene expression profiles from fat tissue and 129 from kidney tissue, each of these includes 10 from the parental strains and the rest are from the RI strains. Linkage analysis was carried out in the RI strain data using QTL Reaper. Redundant markers that had the same genetic locations and the same SDP were removed. There were 2118 eQTL in fat and 2490 in kidney, the location of the marker relative to its linked probe set were calculated and those in *cis* and with $P<10^{-4}$ were selected for study. This left 97 in fat and 129 in kidney, this data was then reduced by selecting genes with a two-fold differential expression in the parental strains. Twenty-two eQTL were discarded because they did not have sufficient mapping information. The final data set sequenced in **Chapter 4** was comprised of the 15 eQTL that represented 14 genes and an additional 4 genes that had been studied previously and matched either the fold change or *P*-value criteria.



Chapter 3: Microarray data generation, quality control, normalisation, validation and linkage analysis

3.1 Introduction

The overall aim of the work in this chapter was to generate, validate and normalise microarray data from 30 RI strain and the BN and SHR parentals and to carry out linkage analysis of gene expression profiles to localise and ultimately identify genes which may play a causal role in SHR metabolic phenotypes.

The study design here was to extract fat and kidney RNA from 4 animals per RI strain and 5 animals in the BN and SHR parental strains and to hybridise them to Affymetrix GeneChips. This is a large data set and was generated in 2 centres: the kidney data by Dr Norbert Hubner's group at the Max-Delbruck centre (MDC) in Berlin. This student generated the data from 66 of the 130 tissue samples with the remainder being generated by Dr Caroline Wallace of the Physiological Genomics and Medicine group (**Appendix I**).

3.1.1 Gene expression studies

As discussed in the **Introduction section 1.7** the recent development of microarray technology has allowed the study of expression of thousands of genes and thus the entire genome simultaneously.

Rat Expression Analysis 230 (RAE 230) A Genechip® Affymetrix oligonucleotide microarrays were used to generate expression values in the SHR and BN parents and

all 30 RI strains. This is one of two arrays in the RAE 230 set and represents approximately 16,000 genes in the rat genome and contains genes which were represented on the previous Rat U34A GeneChip. A probe set is designed for each gene containing 11 probe pairs which are used to determine if the gene is expressed in the RNA sample. There are two types of probe in a probe pair: one perfect match (PM) and one mismatch (MM) both of 25 nucleotides in length (Lipshutz et al. 1999). The perfect match probe is exactly complementary to the RNA target sequence of the transcript while the mismatch probe differs by a single base in the central (13th) position and is used to detect non-specific hybridisation and therefore determine background measurements. The probe pairs are distributed across the chip and the information from each probe pair is combined (summarised) to derive an expression measure for the gene.

3.1.2 Data extraction and normalisation methods

There are several methods to summarise probe intensities from each probe in a probe set. Two methods were compared in the work presented here; the Affymetrix Microarray Analysis Suite 5.0 (MAS5.0) (Affymetrix, Santa Clara, CA, USA) and the Robust Multi-array Analysis (RMA) method developed by Irizarry (Irizarry et al. 2003). Both methods use the "* .cel" files to extract expression values for each gene in the sample, these files contain the captured image of the scanned array and the raw intensities for probe sets. The first method used: the MAS5.0 software gives present (P), marginal (M) or absent (A) calls for each gene or EST and generates values for the average difference in signal between perfect match and mismatch probes. The average difference is calculated from the combined background adjusted PM and MM values of the probe set. The data was extracted using the MAS5.0 software

(Affymetrix 2002) and then normalised using Microsoft Excel by scaling the raw average difference to the median of the medians of all the GeneChips in the analysis. This will be referred to as median scaling.

The second method RMA, uses raw probe set data ("* .cel" files) and simultaneously normalises the data using a quantile normalisation. This eliminates systematic differences between GeneChips, without significantly altering the relative intensity of probes within a GeneChip. The data is transformed to the $\log_2$ scale and a median-polish procedure is used to combine multiple probes into a single measure of expression for each gene on each array. The quantile normalisation aims to make the distribution of probe intensities for each array in the data set equal.

3.1.3 Quality control indices

During the derivation of this data set the quality of the RNA, cRNA and hybridisation were tested. Since the data was being generated by two people it was important to ensure the data was consistent and therefore could be analysed and compared together. Similar quality control was applied to the kidney data set at the MDC in Berlin. The Agilent Bioanalyser can be used to assess the quality of both the extracted RNA and the cRNA synthesized before the sample is hybridised to the microarray. The yield of cRNA was also determined and only in vitro transcription reactions achieving 35µg of cRNA or more were hybridised. Affymetrix MAS 5.0 software was used to extract the data from the microarrays to assess the quality of each hybridisation reaction based on the Raw Q and percentage P value. Raw Q is a measure of quality control which considers the pixel to pixel variation of probe cells on a GeneChip array. A maximum threshold Raw Q value of 4 was used in this study

following preliminary experiments using Affymetrix Test arrays and fat tissue samples from peritoneal fat pads of rats of similar age to that used in this experiment. This threshold was based on a guide from Affymetrix and also our observation that generally cRNA from fat samples resulted in a slightly higher Q than samples from other tissues. Any sample which produced a Raw Q value of 4 or over was excluded from the study and the cRNA synthesis repeated. Percentage P is the percentage of genes on the array which are called present by the MAS5.0 software that is genes that are detected in the hybridised RNA sample.

3.2 Results

3.2.1 Generation and quality control of microarray data sets

Two complete data sets were generated with a total of 130 and 129 microarrays in fat and kidney respectively. This student generated 66 of the 130 microarrays from the fat data set which was combined and normalised with the other 64 to facilitate effective analysis of the data and comparison with the kidney data set. This student acknowledges Dr Caroline Wallace for permission to include this data since this allows statistical analysis and interpretation of the whole data set. A table of samples prepared by the student and Dr Wallace is given in **Appendix I**

Prior to carrying out RNA extractions, cRNA synthesis and hybridisations on the tissues from the RI strains a number of test samples were used to determine which protocols were optimal. These samples were peritoneal fat pads from age matched rats to allow determination of which RNA extraction protocol yielded more RNA. RNABee and Trizol protocols were compared and with both fat and kidney tissue

Trizol consistently gave a higher yield of RNA with better A260/280 ratios. **Table 3.1** shows representative samples from this analysis. Normally a formaldehyde gel is used to assess both RNA and cRNA quality, however this uses 2.5μg of RNA while the Agilent Bioanalyser only uses between 25-500ng of the RNA sample. The fat pads used in this study were from 6 week old animals and varied in size depending on which strains they originated from. Consequently they yielded small amounts of RNA and often just enough to carry out cRNA synthesis (10μg), therefore the technique using the lowest amount of RNA used to determine quality was preferable. Thus a comparison was carried out between the bioanalyser and a formaldehyde gel to determine if they gave the same results. The same six test RNA and six test cRNA samples were run on both a formaldehyde gel and a bioanlyser trace. **Figure 3.1** shows gel images from both methods, the bioanalyser generates an artificial gel image following the trace. The 6 lanes on left of each gel in **Figure 3.1** contain the same total RNA samples with the expected bands corresponding to 28S and 18S ribosomal RNA detected. The last 6 lanes on each image show cRNA sample traces from the same samples. cRNA should result in a smear on electrophoresis as it represents the heterogeneous population of mRNA present in a tissue. The lack of any distinct bands is an indicator of quality. Since the methods gave corresponding results for quality and the Bioanalyser requires significantly less sample it was decided to use it to assess the RNA and cRNA quality.

	A260	A280	Ratio	yield (μg)
Fat RNA Bee	0.25	0.128	1.98	500
Kidney RNA Bee	0.349	0.11	1.61	698
Fat Trizol	0.34	0.164	2.07	680
Kidney Trizol	0.539	0.169	1.582	1078

Table 3.1 Comparison of RNA yield using RNA Bee and Trizol

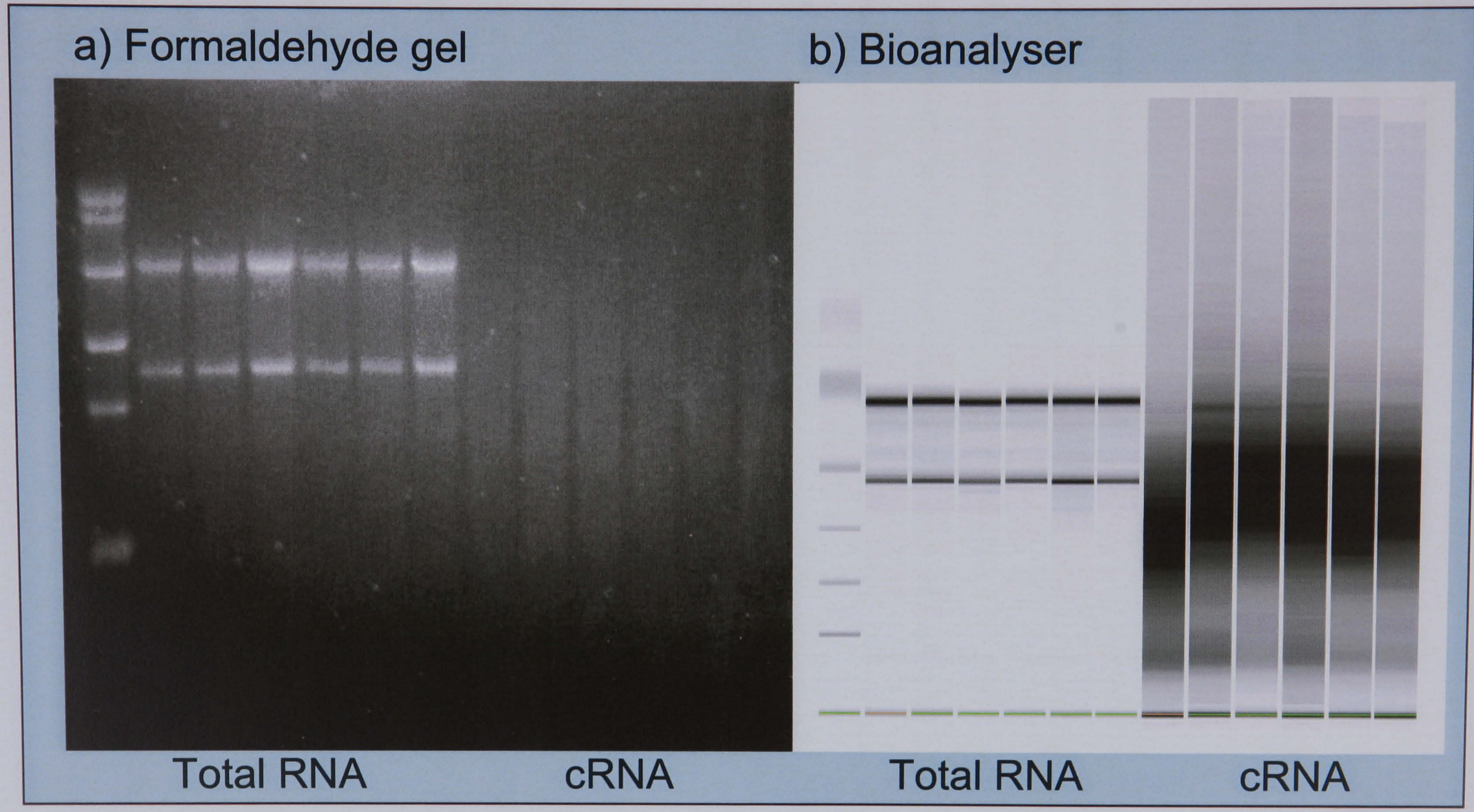


Figure 3.1 Comparison of Formaldehyde Gels and the Bioanalyser assay:
a) A formaldehyde gel showing total RNA and cRNA samples, b) A Bioanlyser trace image generated from running the same samples on the Bioanalyser *Eukaryotic total RNA Nano* assay. The total RNA traces both show the 18S and 28S ribosomal peaks and the cRNA traces show a smear as expected.

Hybridisations using test 3 arrays were also carried out on these preliminary fat tissue samples to ensure the cRNA synthesized would be suitable for RAE230A GeneChips and to determine expected Raw Q and percentage P values in fat tissue samples. Ten samples were checked on test arrays and this showed that fat tissue samples give a higher Raw Q than the recommended value of 3 and therefore when the study samples were hybridised a higher threshold of 4 was permitted in fat tissue.

Following these protocol assessments RNA was extracted from peritoneal fat tissue from 4 animals per RI strain and 5 animals for each parental strain (**Methods 2.3.1**) and was concentrated to 2μg/μl by ethanol precipitation. cRNA was generated (**Methods 2.7**) and those samples passing quality control were hybridised to the microarray. Those samples which did not pass the quality control were not hybridised and more cRNA generated either from the remaining RNA or from new RNA extracted from the other fat pad of the same animal.

The yield and purity of the RNA and cRNA samples were determined using the A260nm/280nm ratio. Only samples which had a 260nm/280nm ratio of between 1.8 and 2.1 were hybridised to the GeneChips. The quality of the RNA and cRNA was also assessed using the Agilent Bioanalyser and example traces are shown in **Figure 3.2**. **Figure 3.2A** shows traces from two total RNA samples, the trace in blue is indicative of good quality RNA with two peaks of approximately equal heights corresponding to 18S and 28S ribosomal RNA which should be present in equal quantities in a good total RNA sample. The trace in green which has no distinct peaks and a definite leftward shift is typical of a degraded RNA sample. All RNA samples were assessed and only those showing good quality traces were used to make cRNA.



Figure 3.2 Bioanalyser traces for total RNA and cRNA
Traces from samples of good quality are shown in blue and those from degraded samples in green. **A.** Total RNA traces. The blue trace has two peaks corresponding to 18S and 28S and the green trace has no distinct peaks and is indicative of degraded RNA and would be excluded from study. **B.** cRNA trace comparison. The green peak is indicative of a poor cRNA synthesis reaction or degraded cRNA and is shown by a leftwards shift of the peak

In the case of samples which failed quality the RNA was re-extracted from the other fat pad or from the fat pads of a new animal from the same strain. Following cRNA synthesis and clean-up the samples were again run on the Bioanalyser and representative traces are shown in **Figure 3.2B**. Again good quality cRNA is shown in blue, there are no distinct peaks detected and the small shoulder on the left of the curve is expected. This is because the cRNA represents only the mRNA population and should not contain the highly abundant ribosomal RNA molecules which make up the majority of a total RNA sample. A poor quality cRNA trace is shown in green in **Figure 3.2B** where the curve is again shifted leftwards meaning the cRNA contains many small molecules indicating the synthesis reaction was incomplete. Only reactions which when quantified achieved at least 35µg of cRNA were assessed on the bioanalyser and subsequently hybridised to Affymetrix GeneChips. All RNA and cRNA bioanalyser traces from the samples hybridised to the microarrays are on the **Data CD** inserted in the back cover of this thesis.

Affymetrix MAS 5.0 software was used to extract data from the microarrays and to determine the quality of each hybridisation reaction. Samples which had a Raw Q of less than 4 and percentage of genes called present of between 55% and 65% were included in the data set. If these parameters were not achieved cRNA was resynthesized either from previously extracted RNA, or from newly extracted RNA from the second fat pad of the same animal or from fat pads of a new animal.

Following the quality control checks the complete fat tissue data set has 120 chips from the RI strains, that is 4 per strain and 5 from each parental strain making a total of 130 hybridisations.

3.2.2 Comparison of data extraction and normalisation methods

A comparison of two commonly used data extraction and normalisation methods was carried out on the data generated in order to determine which method would be used to extract and normalise the data for the linkage analysis to identify genes which may be involved in SHR pathophysiology.

Initially these comparisons were carried out on a preliminary dataset which consisted of the first 92 hybridised microarrays including 3 from each parental strain in the fat tissue. Two normalised data sets were generated using only the 6 parental microarrays: one normalised using RMA with quantile normalisation and the other using MAS5.0 followed by median scaling. This allowed a comparison of the two data processing methods by determining differential expression between the SHR and BN parents. A further data set was derived where all 92 microarrays were normalised using RMA to define some preliminary linkages and eQTL.

Differential expression was determined in two ways, 1) using a t-test and 2) by determining the fold change between the BN and SHR parental strains. A two tailed t-test with correction for unequal variance was carried out on the parental RMA and median scaled data sets to compare the three SHR microarray outputs to the three BN outputs. The results of the t-tests were then compared with each other to determine the extent of overlap using two normalisation methods. The numbers of genes determined to be significantly differentially expressed were similar, approximately 1100 (**Table 3.2**), however only 506 probe sets were differentially expressed in both data sets ($P<0.05$) indicating that they are truly differentially expressed since they were deteced by two different methods. **Figure 3.3** is a comparison of t-test *P*-values

Table 3.2 Differentially expressed genes in fat tissue

P-value[a]	Data transformation method[b]	
	RMA	MAS5.0 & median scaling
0.05	1181	1129
0.01	256	259
0.005	137	121
0.001	27	28
0.0005	13	13
0.0001	3	4
0.00005	1	0
0.00001	0	0

[a] *P*-value from a two tailed t-test with unequal variance between the BN and SHR parental strains
[b] Method used to normalise the data from the six parental microarrays

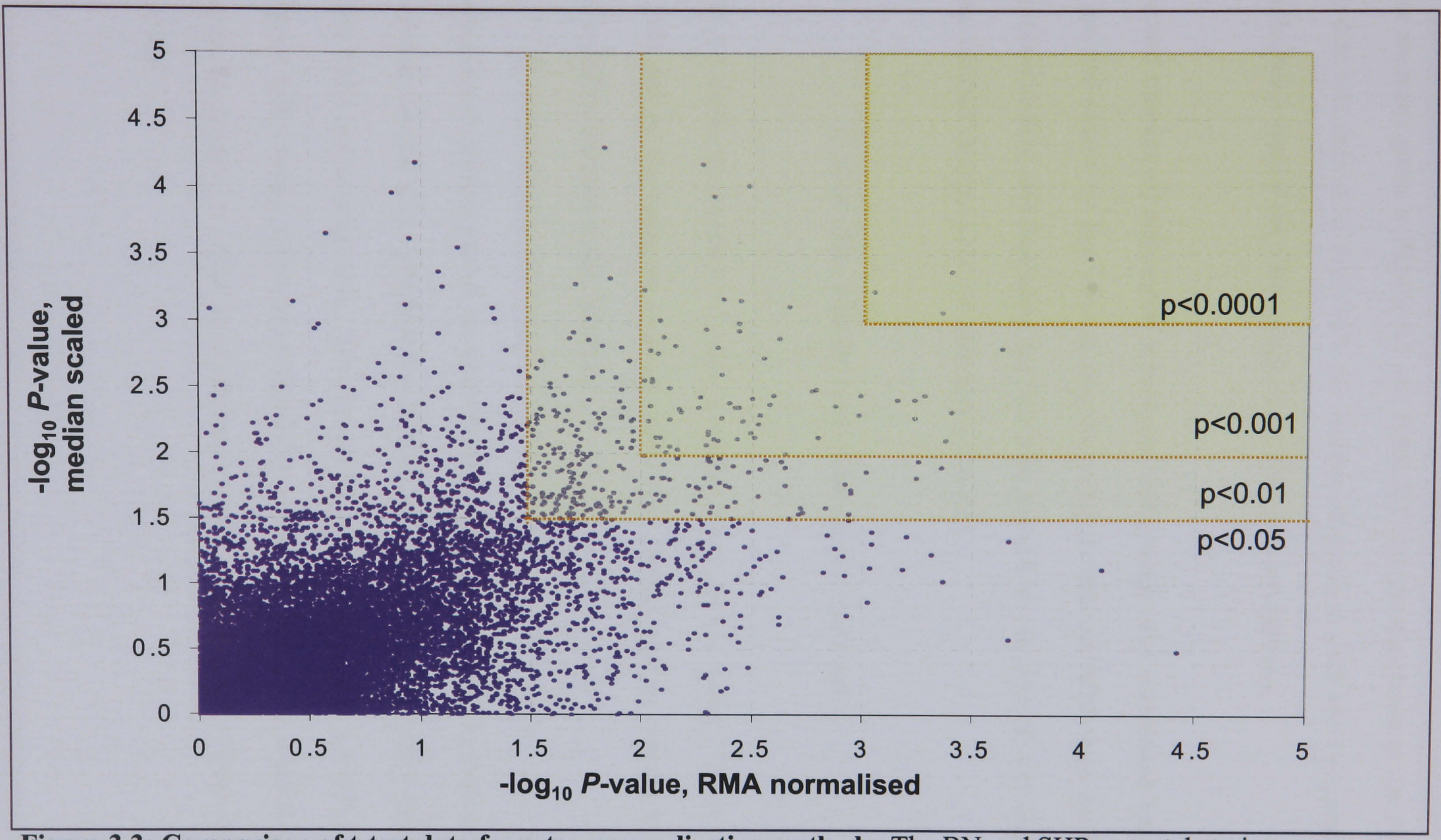


Figure 3.3: Comparison of t-test data from two normalisation methods: The BN and SHR parental strains were normalised using two methods. The above plot is a comparison of the -log of the t-test *P*-value from these methods. The shaded areas highlight the probe sets which have significant *P*-value in both data sets with darker shading indicating higher significance. **Table 3.2** shows the numbers of genes in each data set. There are a significant number of genes which achieve a particular statistical threshold in one but not both data sets.

from both data sets; the yellow shaded areas highlight the probe sets which are significant in both data sets with increased significance depicted by darker shading. There are many genes in **Figure 3.3** and **Table 3.2** that are significant in one but not both data sets indicating that while the number of genes at each level of statistical significance is similar the composition of the gene lists are different.

A second measure of differential expression, fold change, was determined between the parental strains by first determining the mean for BN and SHR from the three expression profiles and then using the following formula to calculate the fold change, so that the fold change is always SHR relative to BN.

IF SHR>BN (SHR/BN) or IF SHR<BN (-1/(SHR/BN))

Table 3.3 shows the numbers of genes at various ranges of fold change. Most genes (14,574) fall in the range of –1.49 to +1.49 fold change and are considered not differentially expressed in this analysis.

Figure 3.4 is a volcano plot of the t-test *P*-value against fold change for both sets of normalised parental data. The blue points represent MAS5.0 median scaled data and the red RMA normalized data. The median scaled data generates a much wider spread of fold change values than the RMA processed data. The shaded areas represent the probe sets which have arbitrary cut-offs of greater than 2 fold change and $p<0.01$.

Table 3.3 Fold Change of fat tissue genes

Fold change	No. of genes
>8	1
7 - 7.99	1
6 - 6.99	3
5 - 5.99	7
4 - 4.99	9
3 - 3.99	25
2 - 2.99	78
1.5 - 1.99	262
-1.49 -1.49	14574
-1.5 - -1.99	793
-2 - -2.99	125
-3 - -3.99	27
-4 - -4.99	10
-5 - -5.99	3
-6 - -6.99	1
-7 - -7.99	1
-13 - -13.99	1
-16 - -16.99	1
-39 - -39.99	1

A comparison of fold change between the parental strains, with 5 arrays per strain was also carried out when the entire data set of 130 microarrays was available and again similar numbers of genes were detected as differentially expressed using both normalisation methods but the composition of the gene lists was different. In collaboration with Dr Enrico Petretto in the Physiological Genomics and Medicine group a comparison of false discovery rate (FDR) and reproducibility both within and between strains using both normalisation methods was performed. At every threshold of significance the expected number of false positives was estimated by the FDR. In this differential expression analysis RMA normalised data showed a consistently lower number of false positives when compared to MAS5.0 normalisation (**Figure 3.5**) and was therefore selected as the method of choice for subsequent analysis.

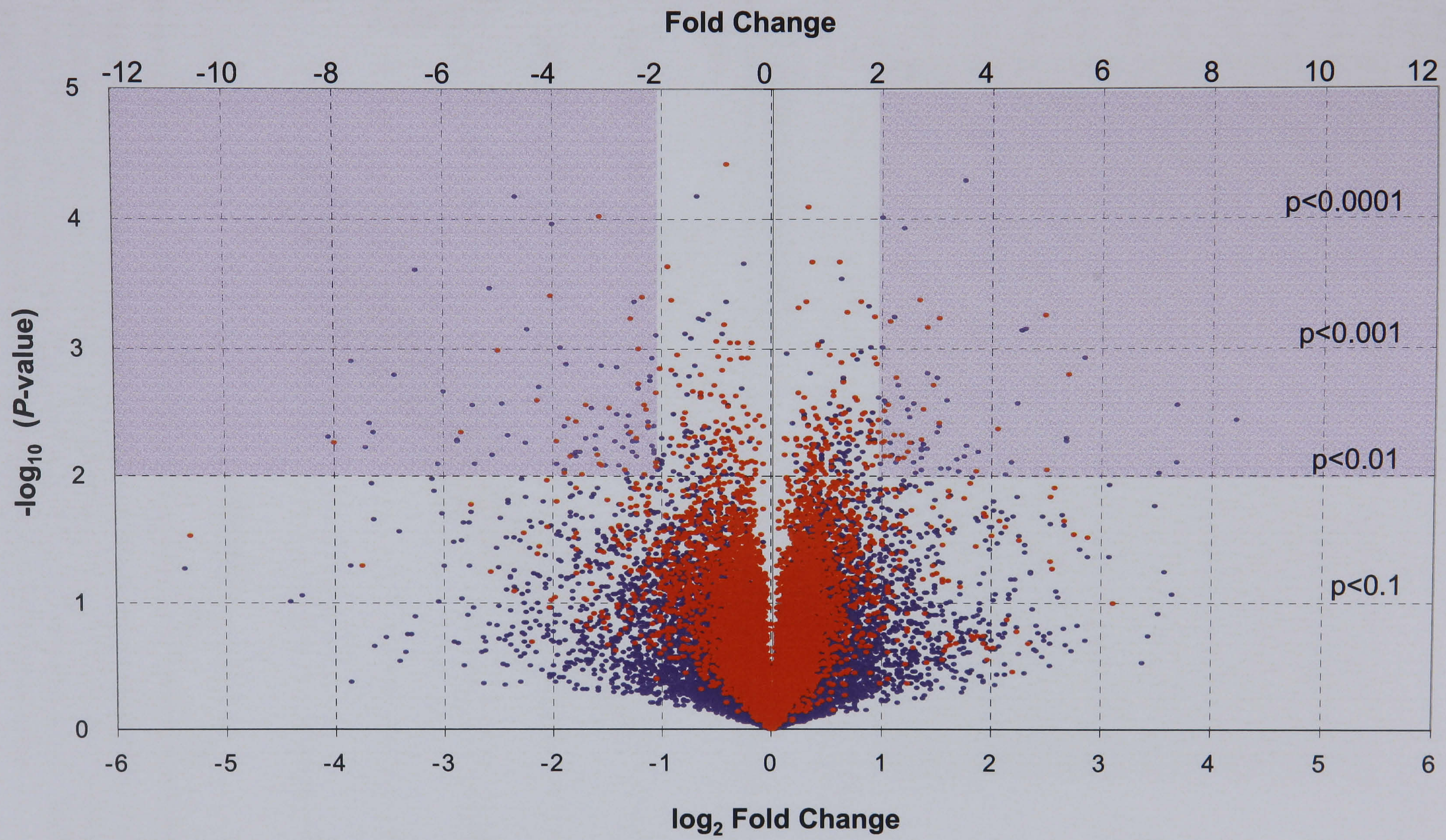
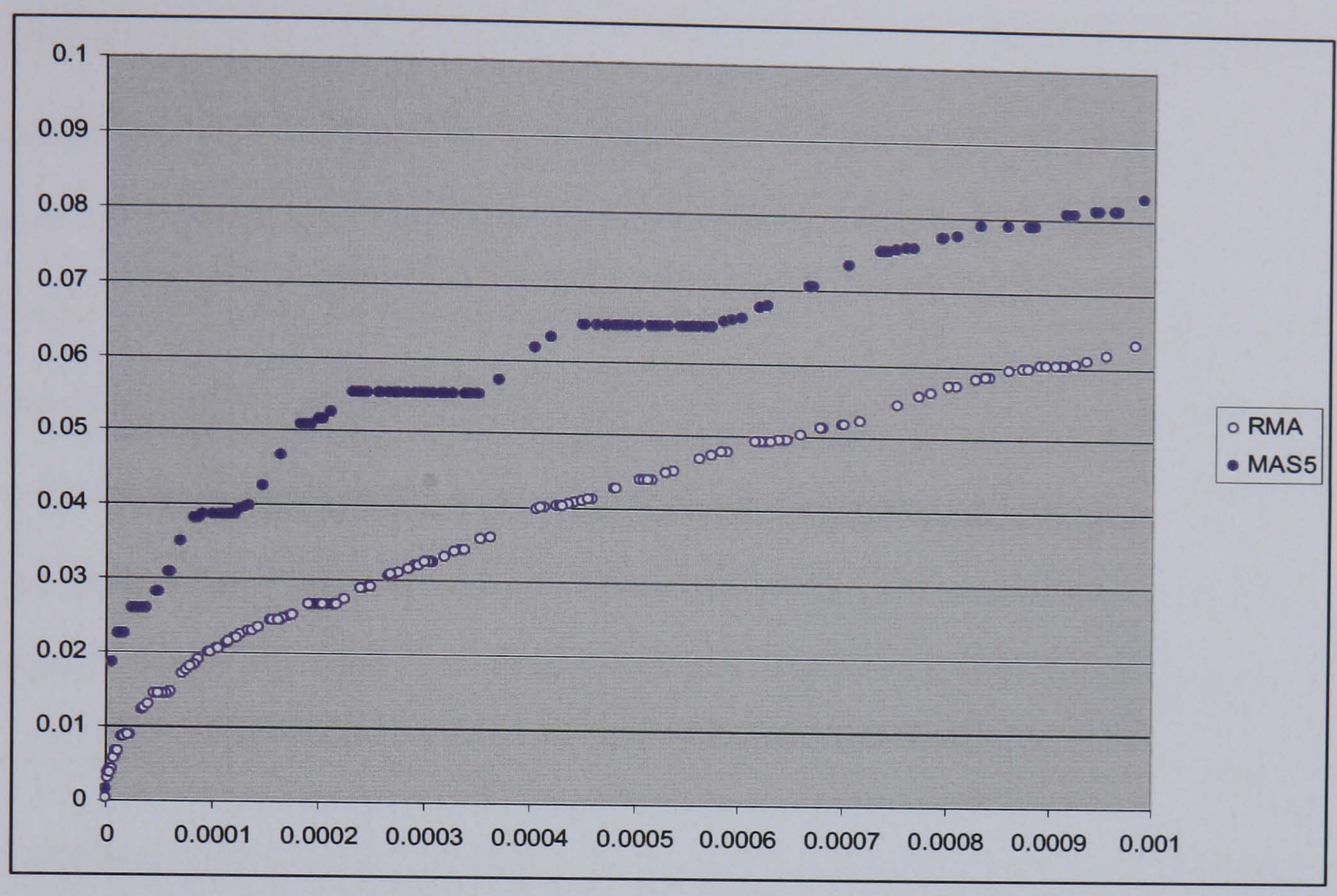


Figure 3.4 *P*-value against fold change in both data sets
In blue is the median scaled and in red the RMA Express data. The shaded areas represent probe sets where the fold change is 2 or more and the $P<0.01$.

a) FDR



b) Reproducibility

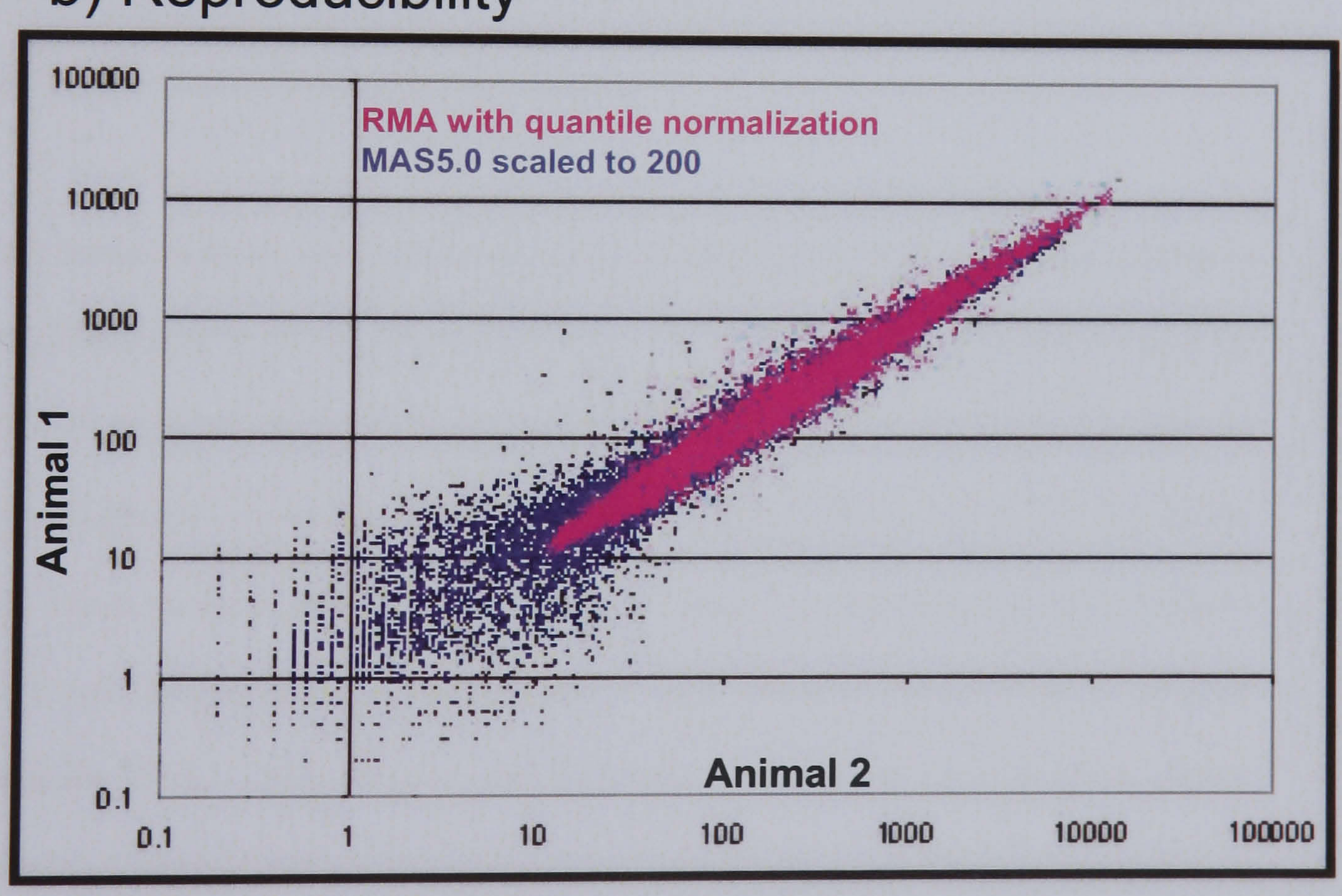


Figure 3.5 Example graphs showing the FDR and reproducibility in the RI strains following different normalisations: a) The FDR at increasing levels of significance in the RMA and MAS5.0 normalised data sets. **b)** This graph shows the reproducibility between animals of the same strain

3.2.3 Linkage analysis

The overall aim of the linkage analysis was to treat the expression profiles generated as quantitative traits and to perform genome-wide linkage analysis to determine the regions of the genome controlling the expression of all 15,923 genes on the microarray. Ultimately QTL Reaper (Wang et al. 2003) was used to achieve this on a genome scale as this program can carry out the linkage analysis for the entire data set in an automated fashion. QTL Reaper is an extended version of Map Manager QTX, the program which was used to detect linkage in the preliminary data set of 92 normalised microarrays. Map Manager performs linkage analysis of one quantitative trait, in this case an expression profile across all the strains, to the genome. Since only one quantitative trait can be tested for linkage at a time genes had to be prioritised before being tested for linkage with Map Manager as it is impractical to manually input 15,923 expression profiles individually.

In the initial analyses, genes were prioritised for linkage analysis across the RI strains in Map Manager by first determining if they were differentially expressed between the SHR/BN parental strains because genes differentially expressed in the parents are likely candidates underlying the heritable phenotypic differences in the RI strains. Forty seven genes (**Table 3.4**) were selected from the preliminary data set discussed above and linkage analysis was carried out using the Map Manager QTX program (**Methods 2.9.1**). This program performs two point linkage analysis using marker regression between all markers in the genetic map and a quantitative phenotype, in this case an expression profile, across all the RI strains within an RI panel. Genes were selected on the basis that they were differentially expressed in the parental

Table 3.4 47 genes initially tested in Map Manager QTX

Probe sets	Probe set Chromosome	SHR mean	BN mean	t-test p-value	Fold change	Best MapManager p-value	Linked marker	Marker Chromosome	Control	Affymetrix Descriptions (December 2003)
Down - regulated										
1369171_at	8	203	896	0.009529	0.226	0.00000	D8Rat202	8	intrachromosomal	gb:NM_024352.1 /DB_XREF=gi:13242238 /GEN=Mst1 /FEA=FLmRNA /CNT=4 /TID=Rn.10613.1 /TIER=FL /STK=2 /UG=Rn.10613 /LL=24566 /DEF=Rattus norvegicus Macrophage stimulating 1 (hepatocyte growth factor-like) (Mst1), mRNA. /PROD=macrophage stimulating 1 (hepatocyte growthfactor-like) /FL=gb:NM_024352.1
1369764_at	13	63	307	0.004374	0.206	0.00002	D13Rat88	13	intrachromosomal	gb:NM_012516.1 /DB_XREF=gi:6978576 /GEN=C4bpa /FEA=FLmRNA /CNT=4 /TID=Rn.10408.1 /TIER=FL /STK=1 /UG=Rn.10408 /LL=24235 /DEF=Rattus norvegicus Complement component 4 binding protein, alpha (C4bpa), mRNA. /PROD=complement component 4 binding protein, alpha /FL=gb:NM_012516.1
1370114_a_at	2	142	653	0.133214	0.218	0.00001	D2Rat94	2	intrachromosomal	gb:D64048.1 /DB_XREF=gi:1246393 /FEA=FLmRNA /CNT=1 /TID=Rn.10599.3 /TIER=FL /STK=1 /UG=Rn.10599 /LL=25513 /UG_GENE=Pik3r1 /DEF=Rattus norvegicus mRNA for phosphatidylinositol 3-kinase p45 subunit, complete cds. /PROD=phosphatidylinositol 3-kinase p45 subunit /FL=gb:D64048.1
1370428_x_at	20	24	384	0.000181	0.062	0.00000	Tnfa	20	intrachromosomal	gb:AJ249701.1 /DB_XREF=gi:5912581 /GEN=RT1-A2q /FEA=mRNA /CNT=4 /TID=Rn.39743.8 /TIER=ConsEnd /STK=0 /UG=Rn.39743 /LL=24737 /DEF=Rattus norvegicus partial mRNA for MHC class Ia A2q antigen (RT1-A2q gene). /PROD=A2q
1370831_at	4	374	1647	0.013533	0.227	0.00275	D11Rat28	11	extrachromosomal	gb:AY081195.1 /DB_XREF=gi:19697885 /GEN=Mgl2 /FEA=FLmRNA /CNT=11 /TID=Rn.22922.1 /TIER=FL /STK=3 /UG=Rn.22922 /DEF=Rattus norvegicus monoglyceride lipase (Mgl2) mRNA, complete cds. /PROD=monoglyceride lipase /FL=gb:AY081195.1
1371033_at	20	35	479	0.013165	0.074	0.00000	Tnfa	20	intrachromosomal	gb:AI715202 /DB_XREF=gi:5032455 /DB_XREF=UI-R-Y0-abn-e-07-0-UI.s1 /CLONE=UI-R-Y0-abn-e-07-0-UI /FEA=mRNA /CNT=8 /TID=Rn.20089.1 /TIER=ConsEnd /STK=0 /UG=Rn.20089 /UG_TITLE=Rat mRNA for MHC class II antigen RT1.B-1 beta-chain
1371960_at	4	106	466	0.001731	0.227	0.00000	D5Rat39	5	unclassified	gb:AI230548 /DB_XREF=gi:3814450 /DB_XREF=EST227258 /CLONE=REMCY16 /FEA=EST /CNT=21 /TID=Rn.6931.1 /TIER=Stack /STK=14 /UG=Rn.6931 /UG_TITLE=ESTs
1376742_at		35	145	0.005451	0.239	0.00100	Ppy	10	unclassified	gb:BF396436 /DB_XREF=gi:11381408 /DB_XREF=UI-R-BS2-bdy-b-04-0-UI.s1 /CLONE=UI-R-BS2-bdy-b-04-0-UI /FEA=EST /CNT=7 /TID=Rn.23219.1 /TIER=ConsEnd /STK=4 /UG=Rn.23219 /UG_TITLE=ESTs
1377452_at	8	53	2129	0.004714	0.025	0.00000	D8Cebr16s5	8	intrachromosomal	gb:AW251280 /DB_XREF=gi:6594871 /DB_XREF=UI-R-BJ0-adg-d-01-0-UI.s1 /CLONE=UI-R-BJ0-adg-d-01-0-UI /FEA=EST /CNT=4 /TID=Rn.23478.1 /TIER=ConsEnd /STK=4 /UG=Rn.23478 /UG_TITLE=ESTs, Highly similar to TETN MOUSE TETRANECTIN PRECURSOR (M.musculus)
1384717_at		27	109	0.000116	0.246	0.00000	D7Rat17	7	unclassified	gb:AA894199 /DB_XREF=gi:3021078 /DB_XREF=EST198002 /CLONE=RSPAS58 /FEA=EST /CNT=2 /TID=Rn.22765.1 /TIER=ConsEnd /STK=2 /UG=Rn.22765 /UG_TITLE=ESTs
1386901_at	4	736	4173	4.44E-05	0.176	0.00000	Cd36	4	intrachromosomal	gb:NM_031561.1 /DB_XREF=gi:13928759 /GEN=Cd36 /FEA=FLmRNA /CNT=38 /TID=Rn.3790.1 /TIER=FL+Stack /STK=23 /UG=Rn.3790 /LL=29184 /DEF=Rattus norvegicus CD36 antigen (collagen type I receptor, thrombospondin receptor) (Cd36), mRNA. /PROD=CD36 antigen /FL=gb:AF072411.1 gb:NM_031561.1 gb:L19658.1 gb:AF113914.1
1387178_a_at	20	146	589	0.041184	0.247	0.00025	D20Cerbp97s7	20	intrachromosomal	gb:NM_012522.1 /DB_XREF=gi:6978610 /GEN=Cbs /FEA=FLmRNA /CNT=9 /TID=Rn.10200.1 /TIER=FL /STK=5 /UG=Rn.10200 /LL=24250 /DEF=Rattus norvegicus Cystathionine beta synthase (Cbs), mRNA. /PROD=cystathionine beta synthase /FL=gb:M88346.1 gb:D01098.1 gb:NM_012522.1 gb:M88344.1
1387341_a_at	18	206	1483	0.000756	0.139	0.01090	Rt2	19	extrachromosomal	gb:NM_017026.1 /DB_XREF=gi:8393758 /GEN=Mbp /FEA=FLmRNA /CNT=44 /TID=Rn.18.1 /TIER=FL /STK=2 /UG=Rn.18 /LL=24547 /DEF=Rattus norvegicus Myelin basic protein (Mbp), mRNA. /PROD=myelin basic protein /FL=gb:NM_017026.1 gb:M25889.1
1389974_at	3	74	310	0.008683	0.239	0.00000	D3Mit3	3	intrachromosomal	gb:BF555171 /DB_XREF=gi:11664901 /DB_XREF=UI-R-E0-ct-a-05-0-UI.r1 /CLONE=UI-R-E0-ct-a-05-0-UI /FEA=EST /CNT=12 /TID=Rn.4231.1 /TIER=Stack /STK=7 /UG=Rn.4231 /UG_TITLE=ESTs, Weakly similar to 0806162N protein URFA6L (M.musculus)
1389986_at	1	270	1384	0.056834	0.195	0.00108	D7Utr1	7	extrachromosomal	gb:AI008409 /DB_XREF=gi:3222241 /DB_XREF=EST202860 /CLONE=REMAW93 /FEA=EST /CNT=9 /TID=Rn.65864.6 /TIER=Stack /STK=7 /UG=Rn.65864 /LL=85383 /UG_GENE=LOC85383 /UG_TITLE=unknown Glu-Pro dipeptide repeat protein
1390048_at	10	362	1489	0.074302	0.243	0.00203	D7Utr1	7	extrachromosomal	gb:BF408990 /DB_XREF=gi:11396965 /DB_XREF=UI-R-BJ2-bpn-g-05-0-UI.s1 /CLONE=UI-R-BJ2-bpn-g-05-0-UI /FEA=EST /CNT=9 /TID=Rn.38798.1 /TIER=Stack /STK=6 /UG=Rn.38798 /UG_TITLE=ESTs, Weakly similar to T31639 hypothetical protein Y57A10A.q - Caenorhabditis elegans (C.elegans)
1390436_at		259	1539	0.032361	0.169	0.00218	D15Rat107	15	unclassified	gb:BI300794 /DB_XREF=gi:14977074 /DB_XREF=UI-R-CV2-cij-g-06-0-UI.s1 /CLONE=UI-R-CV2-cij-g-06-0-UI /FEA=EST /CNT=7 /TID=Rn.22838.2 /TIER=ConsEnd /STK=4 /UG=Rn.22838 /UG_TITLE=ESTs
1398245_at	16	804	5405	0.004048	0.149	0.00025	D16Rat6	16	intrachromosomal	gb:NM_031688.1 /DB_XREF=gi:13928953 /GEN=Sncg /FEA=FLmRNA /CNT=66 /TID=Rn.10421.1 /TIER=FL+Stack /STK=16 /UG=Rn.10421 /LL=64347 /DEF=Rattus norvegicus synuclein, gamma (Sncg), mRNA. /PROD=synuclein, gamma /FL=gb:NM_031688.1
Up-regulated										
1376908_at	1	231	58	0.006254	4.013	0.00032	D15Rat107	15	extrachromosomal	gb:AW531805 /DB_XREF=gi:7174219 /DB_XREF=UI-R-C4-akz-c-07-0-UI.s1 /CLONE=UI-R-C4-akz-c-07-0-UI /FEA=EST /CNT=6 /TID=Rn.35775.1 /TIER=ConsEnd /STK=4 /UG=Rn.35775 /UG_TITLE=ESTs, Moderately similar to S71356 glucocorticoid-attenuated response gene 49 protein - mouse (M.musculus)
1367626_at	1	1213	258	0.131331	4.695	0.00308	D5Rat147	5	extrachromosomal	gb:NM_012530.1 /DB_XREF=gi:6978660 /GEN=Ckm /FEA=FLmRNA /CNT=40 /TID=Rn.10756.1 /TIER=FL+Stack /STK=34 /UG=Rn.10756 /LL=24265 /DEF=Rattus norvegicus Creatine kinase, muscle form (Ckm), mRNA. /PROD=creatine kinase, muscle form /FL=gb:M10140.1 gb:NM_012530.1

Table 3.4 47 genes initially tested in Map Manager QTX										
1368718_at	1	146	26	0.000895	5.651	0.00160	D1Rat296	1	intrachromosomal	gb:M23995.1 /DB_XREF=gi:202845 /FEA=FLmRNA /CNT=4 /TID=Rn.74044.1 /TIER=FL /STK=3 /UG=Rn.74044 /LL=29651 /UG_GENE=Aldh1a4 /UG_TITLE=aldehyde dehydrogenase family 1, subfamily A4 /DEF=Rat aldehyde dehydrogenase mRNA, complete cds. /FL=gb:NM_017272.14 gb:M23995.1
1369110_x_at	20	479	85	0.000396	5.607	0.00000	Tnfa	20	intrachromosomal	gb:NM_012645.1 /DB_XREF=gi:6981493 /GEN=RT1Aw2 /FEA=FLmRNA /CNT=7 /TID=Rn.39743.1 /TIER=FL /STK=2 /UG=Rn.39743 /LL=24737 /DEF=Rattus norvegicus RT1 class Ib gene (RT1Aw2), mRNA. /PROD=RT1 class Ib gene RT1Aw2 /FL=gb:NM_012645.1 gb:M31038.1
1369202_at	11	193	29	0.005739	6.783	0.00026	D11Cebr87s1	11	intrachromosomal	gb:NM_017028.1 /DB_XREF=gi:8393796 /GEN=Mx1 /FEA=FLmRNA /CNT=4 /TID=Rn.10374.1 /TIER=FL /STK=2 /UG=Rn.10374 /LL=24575 /DEF=Rattus norvegicus Myxovirus (influenza) resistance, homolog of murine Mx (also interferon-inducible protein IFI78) (Mx1), mRNA. /PROD=myxovirus (influenza) resistance, homolog ofmurine Mx (also interferon-inducible protein IFI78) /FL=gb:NM_017028.1
1369415_at	4	586	140	0.008724	4.195	0.00216	D19Utr1	19	extrachromosomal	gb:NM_053328.1 /DB_XREF=gi:16758031 /GEN=Bhlhb2 /FEA=FLmRNA /CNT=2 /TID=Rn.81055.1 /TIER=FL /STK=2 /UG=Rn.81055 /LL=79431 /DEF=Rattus norvegicus basic helix-loop-helix domain containing, class B2 (Bhlhb2), mRNA. /PROD=basic helix-loop-helix domain containing, classB2 /FL=gb:NM_053328.1 gb:AF009330.1
1369928_at	19	3421	772	0.120464	4.434	0.00060	D15Rat1	15	extrachromosomal	gb:NM_019212.1 /DB_XREF=gi:9506370 /GEN=Acta1 /FEA=FLmRNA /CNT=54 /TID=Rn.39438.1 /TIER=FL+Stack /STK=46 /UG=Rn.39438 /LL=29437 /DEF=Rattus norvegicus actin, alpha 1, skeletal muscle (Acta1), mRNA. /PROD=actin, alpha 1, skeletal muscle /FL=gb:NM_019212.1
1370900_at		103	22	0.288658	4.609	0.00249	Cacna1s	4	unclassified	gb:BM391169 /DB_XREF=gi:18191222 /DB_XREF=UI-R-DY0-ckq-e-03-0-UI.s1 /CLONE=UI-R-DY0-ckq-e-03-0-UI /FEA=mRNA /CNT=13 /TID=Rn.1072.1 /TIER=Stack /STK=11 /UG=Rn.1072 /UG_TITLE=Rattus norvegicus myosin heavy chain mRNA, 3 end
1371015_at		259	35	0.005608	7.347	0.00173	D15Rat107	15	unclassified	gb:X52711.1 /DB_XREF=gi:56720 /FEA=mRNA /CNT=4 /TID=Rn.10373.1 /TIER=ConsEnd /STK=1 /UG=Rn.10373 /UG_TITLE=Rat mRNA for Mx1 protein /DEF=Rat mRNA for Mx1 protein.
1371213_at		196	47	0.002148	4.192	0.00000	Tnfa	20	unclassified	gb:AJ005023.1 /DB_XREF=gi:3006073 /FEA=mRNA /CNT=1 /TID=Rn.39743.19 /TIER=ConsEnd /STK=0 /UG=Rn.39743 /LL=24737 /UG_GENE=RT1Aw2 /DEF=Rattus norvegicus mRNA for mature MHC class Ib alpha chain, clone cc9. /PROD=MHC class Ib alpha chain
1371241_x_at	8	1926	474	0.198212	4.068	0.00162	D15Rat1	15	extrachromosomal	gb:AF370889.1 /DB_XREF=gi:14134066 /FEA=mRNA /CNT=1 /TID=Rn.1033.8 /TIER=ConsEnd /STK=0 /UG=Rn.1033 /LL=24851 /UG_GENE=Tpm1 /DEF=Rattus norvegicus tropomyosin-alpha mRNA, partial cds; alternatively spliced. /PROD=tropomyosin-alpha
1371776_at		1383	271	0.147368	5.100	0.00000	D2Rat94	2	unclassified	gb:AA819268 /DB_XREF=gi:4227927 /DB_XREF=UI-R-A0-al-d-12-0-UI.s2 /CLONE=UI-R-A0-al-d-12-0-UI /FEA=EST /CNT=20 /TID=Rn.2676.1 /TIER=Stack /STK=16 /UG=Rn.2676 /UG_TITLE=ESTs
1373053_at		469	81	0.001592	5.793	0.00016	Drd2	8	unclassified	gb:AI169620 /DB_XREF=gi:3709660 /DB_XREF=EST215506 /CLONE=RKIBV33 /FEA=EST /CNT=14 /TID=Rn.25001.1 /TIER=Stack /STK=9 /UG=Rn.25001 /UG_TITLE=ESTs
1374527_at		437	69	0.00365	6.311	0.00000	D5Rat79	5	unclassified	gb:AI172274 /DB_XREF=gi:3712314 /DB_XREF=EST218275 /CLONE=RMUBX21 /FEA=EST /CNT=10 /TID=Rn.1408.1 /TIER=Stack /STK=6 /UG=Rn.1408 /UG_TITLE=ESTs, Weakly similar to ENOYL-COA HYDRATASE, MITOCHONDRIAL PRECURSOR (R.norvegicus)
1374932_at	8	1526	347	0.004845	4.398	0.00016	D14Rat52	14	extrachromosomal	gb:BI282731 /DB_XREF=gi:14933790 /DB_XREF=UI-R-CW0s-cce-e-07-0-UI.s1 /CLONE=UI-R-CW0s-cce-e-07-0-UI /FEA=EST /CNT=6 /TID=Rn.66850.1 /TIER=Stack /STK=6 /UG=Rn.66850 /UG_TITLE=ESTs
1376845_at	6	3514	606	0.013721	5.800	0.00058	D15Rat107	15	extrachromosomal	gb:AA819034 /DB_XREF=gi:4227699 /DB_XREF=UI-R-A0-ba-e-12-0-UI.s1 /CLONE=UI-R-A0-ba-e-12-0-UI /FEA=EST /CNT=6 /TID=Rn.16842.1 /TIER=ConsEnd /STK=4 /UG=Rn.16842 /UG_TITLE=ESTs, Weakly similar to AF154572 1 ERG2 protein (R.norvegicus)
1377334_at		499	77	0.000236	6.476	0.00000	Tnfa	20	unclassified	gb:BG378249 /DB_XREF=gi:13302721 /DB_XREF=UI-R-CV1-bvp-h-01-0-UI.s1 /CLONE=UI-R-CV1-bvp-h-01-0-UI /FEA=EST /CNT=4 /TID=Rn.39138.1 /TIER=ConsEnd /STK=4 /UG=Rn.39138 /UG_TITLE=ESTs, Moderately similar to I54411 MHC RT1-B A-alpha chain - rat (R.norvegicus)
1383564_at	1	1290	269	0.006391	4.804	0.00007	D15Rat107	15	extrachromosomal	gb:BF411036 /DB_XREF=gi:11399011 /DB_XREF=UI-R-CN0-bmi-h-12-0-UI.s1 /CLONE=UI-R-CN0-bmi-h-12-0-UI /FEA=EST /CNT=5 /TID=Rn.6246.1 /TIER=ConsEnd /STK=0 /UG=Rn.6246 /UG_TITLE=ESTs, Highly similar to IRF7 MOUSE INTERFERON REGULATORY FACTOR 7 (M.musculus)
1387134_at	10	1002	169	0.016706	5.919	0.00013	D15Rat107	15	extrachromosomal	gb:NM_053687.1 /DB_XREF=gi:16758505 /GEN=Slfn4 /FEA=FLmRNA /CNT=8 /TID=Rn.54422.1 /TIER=FL+Stack /STK=7 /UG=Rn.54422 /LL=114247 /DEF=Rattus norvegicus schlafen 4 (Slfn4), mRNA. /PROD=schlafen 4 /FL=gb:AF168795.1 gb:NM_053687.1
1388694_at	20	1825	306	0.002992	5.969	0.00054	D10Rat6	10	extrachromosomal	gb:AI233121 /DB_XREF=gi:3817001 /DB_XREF=EST229809 /CLONE=RKICL08 /FEA=EST /CNT=18 /TID=Rn.40217.1 /TIER=Stack /STK=12 /UG=Rn.40217 /UG_TITLE=ESTs, Moderately similar to MHC class I (M.musculus)
1391544_at		908	105	0.045628	8.618	0.00050	D8Mit5	8	unclassified	gb:BI285941 /DB_XREF=gi:14940094 /DB_XREF=UI-R-CX0s-ccp-f-02-0-UI.s1 /CLONE=UI-R-CX0s-ccp-f-02-0-UI /FEA=EST /CNT=5 /TID=Rn.74079.1 /TIER=ConsEnd /STK=2 /UG=Rn.74079 /UG_TITLE=ESTs
Top ten p- value										
1386901_at	4	736	4173	4.44E-05	0.176	0.00000	Cd36	4	intrachromosomal	gb:NM_031561.1 /DB_XREF=gi:13928759 /GEN=Cd36 /FEA=FLmRNA /CNT=38 /TID=Rn.3790.1 /TIER=FL+Stack /STK=23 /UG=Rn.3790 /LL=29184 /DEF=Rattus norvegicus CD36 antigen (collagen type I receptor, thrombospondin receptor) (Cd36), mRNA. /PROD=CD36 antigen /FL=gb:AF072411.1 gb:NM_031561.1 gb:L19658.1 gb:AF113914.1
1368334_at	10	77	103	3.59E-05	0.752	0.00028	D16Rat14	16	extrachromosomal	gb:NM_053403.1 /DB_XREF=gi:16758131 /GEN=Grb7 /FEA=FLmRNA /CNT=8 /TID=Rn.28109.1 /TIER=FL+Stack /STK=6 /UG=Rn.28109 /LL=84427 /DEF=Rattus norvegicus growth factor receptor binding protein GRB7 (Grb7), mRNA. /PROD=growth factor receptor binding protein GRB7 /FL=gb:AF190121.1 gb:NM_053403.1
1373253_at	10	390	739	4.73E-05	0.528	0.00001	Slc4a	10	intrachromosomal	gb:BE111646 /DB_XREF=gi:8503751 /DB_XREF=UI-R-BJ1-avv-b-11-0-UI.s1 /CLONE=UI-R-BJ1-avv-b-11-0-UI /FEA=EST /CNT=10 /TID=Rn.18284.1 /TIER=Stack /STK=9 /UG=Rn.18284 /UG_TITLE=ESTs, Weakly similar to ACBP RAT ACYL-COA-BINDING PROTEIN (R.norvegicus)

Table 3.4 47 genes initially tested in Map Manager QTX										
1374668_at	14	133	106	7.99E-05	1.255	0.00270	D4Rat202	4	extrachromosomal	gb:AI230180 /DB_XREF=gi:3814067 /DB_XREF=EST226875 /CLONE=REMCT31 /FEA=EST /CNT=8 /TID=Rn.24287.1 /TIER=Stack /STK=6 /UG=Rn.24287 /UG_TITLE=ESTs, Weakly similar to T31425 C-terminal domain-binding protein rA4, splice form 2 - rat (fragment) (R.norvegicus)
1376635_at	16	1232	427	6.01E-05	2.885	0.00008	D16Mit2	16	intrachromosomal	gb:BF283341 /DB_XREF=gi:11214411 /DB_XREF=EST447932 /CLONE=RGIEA66 /FEA=EST /CNT=8 /TID=Rn.26865.1 /TIER=ConsEnd /STK=4 /UG=Rn.26865 /UG_TITLE=ESTs
1379896_at		33	99	3.45E-05	0.333	0.00000	D10Utr2	10	unclassified	gb:BM390256 /DB_XREF=gi:18190309 /DB_XREF=UI-R-CN1-cjq-m-15-0-UI.s1 /CLONE=UI-R-CN1-cjq-m-15-0-UI /FEA=EST /CNT=4 /TID=Rn.43551.1 /TIER=ConsEnd /STK=2 /UG=Rn.43551 /UG_TITLE=ESTs
1384717_at		27	109	0.000116	0.246	0.00000	D7Rat17	7	unclassified	gb:AA894199 /DB_XREF=gi:3021078 /DB_XREF=EST198002 /CLONE=RSPAS58 /FEA=EST /CNT=2 /TID=Rn.22765.1 /TIER=ConsEnd /STK=2 /UG=Rn.22765 /UG_TITLE=ESTs
1386507_at		754	490	0.00014	1.538	0.00095	D5Cebr10s12	5	unclassified	gb:AA799858 /DB_XREF=gi:2862813 /DB_XREF=EST189355 /CLONE=RHEAG21 /FEA=EST /CNT=1 /TID=Rn.6245.1 /TIER=ConsEnd /STK=1 /UG=Rn.6245 /UG_TITLE=EST, Moderately similar to ODPB RAT PYRUVATE DEHYDROGENASE E1 COMPONENT BETA SUBUNIT, MITOCHONDRIAL PRECURSOR (R.norvegicus)
1388602_at	7	10160	13223	2.42E-05	0.768	0.00066	D14Rat52	14	extrachromosomal	gb:AI237358 /DB_XREF=gi:3830864 /DB_XREF=EST233920 /CLONE=RPLCY05 /FEA=EST /CNT=15 /TID=Rn.16172.1 /TIER=Stack /STK=14 /UG=Rn.16172 /UG_TITLE=ESTs, Highly similar to I55608 complement factor D (EC 3.4.21.46) precursor - rat (R.norvegicus)
1390562_s_at		39	20	8.23E-05	1.924	0.00043	D20Mgh5	20	unclassified	gb:BE102350 /DB_XREF=gi:8494449 /DB_XREF=UI-R-BT1-aql-e-07-0-UI.s1 /CLONE=UI-R-BT1-aql-e-07-0-UI /FEA=EST /CNT=5 /TID=Rn.6745.1 /TIER=ConsEnd /STK=4 /UG=Rn.6745 /UG_TITLE=ESTs

strains either, $p<0.00014$ (n=10) from the t-test or >4 fold differential expression (n=39), two of which were in both data sets. Mean expression values per RI strain from the RMA normalized data set was the input for the analysis. The Map Manager output lists the markers which are significantly linked to each trait. All 47 probe sets had a linkage with a $p<0.05$ and the map locations of the most significantly linked marker per probe set are shown in **Figure 3.6**. During this preliminary analysis linkages were defined as either inter- or intra-chromosomal depending on whether the most significantly linked marker was on the same or a different chromosome to the probe set as a *cis/trans* definition in terms of chromosomal proximity to probe set location had not been finalised. **Figure 3.6** shows the locations of the most significantly linked marker for the 47 genes studied, 15 (shown in red) were intra-chromosomal linkages with the inter-chromosomal linkages shown in blue. Only the chromosomal location of the marker is known for those linkages in yellow, that is the probe sets genomic location is unknown. As a result these linkages cannot be defined as inter or intra chromosomal. All of the most significantly linked ($p<0.00001$) markers where the probe set location is known are intra-chromosomal.

As stated above Map Manager QTX can only perform linkage analysis between the expression profiles from 1 probe set to the genome-wide markers at a time, making it impractical to use to analyse all 15,923 expression profiles. Consequently QTL Reaper was used to carry out the linkage analysis for all probe sets (**Methods 2.9.1**).

The 120 microarrays generated for the RI strains were normalised using RMA, then mean expression values were derived for each strain and these 30 values used in the QTL Reaper analysis to detect linkages across the genome. QTL Reaper uses two

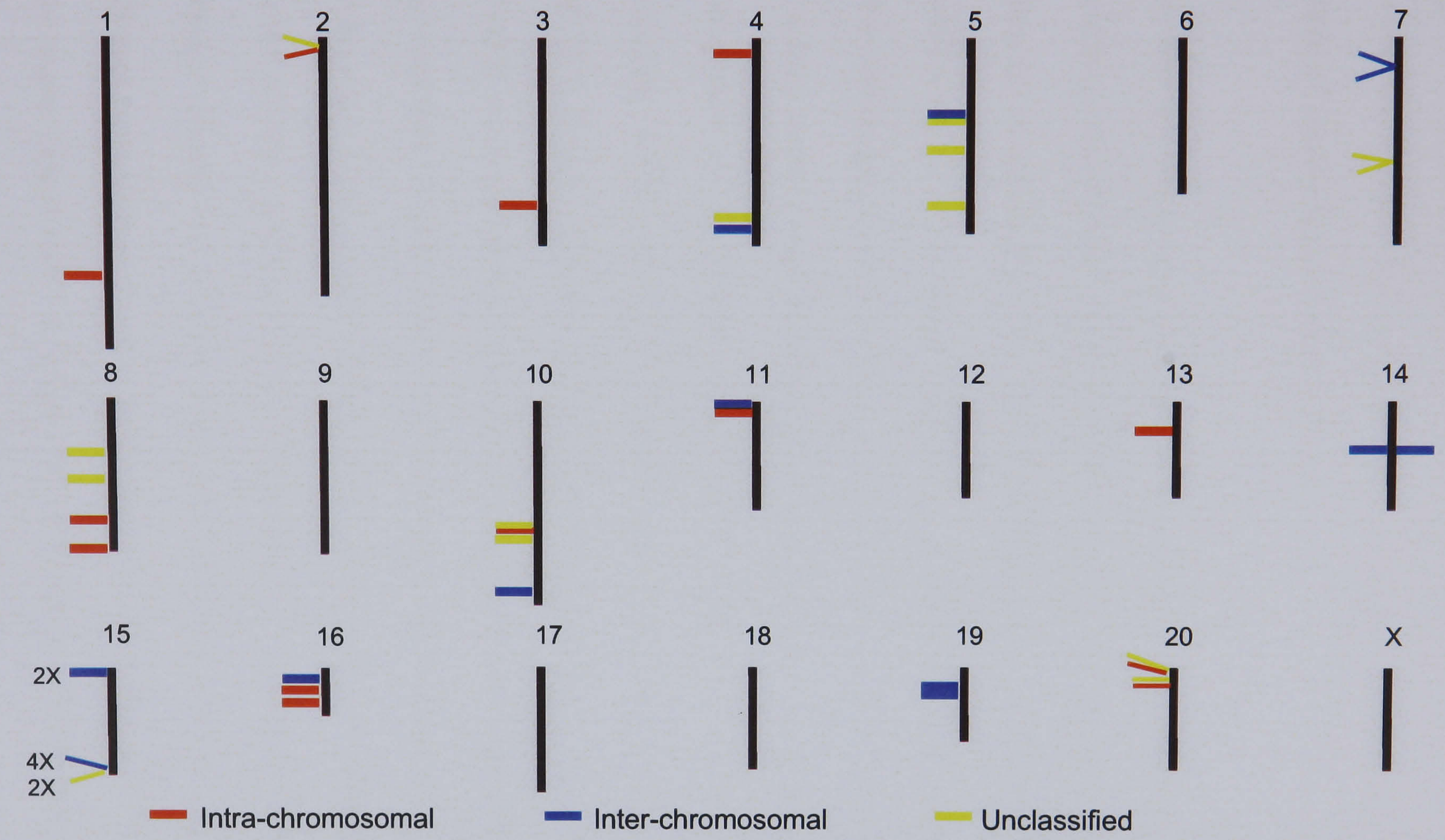


Figure 3.6 Chromosome map of markers significantly linked to differentially expressed genes using Map Manager QTX
The 47 most differentially expressed genes between SHR and BN as determined by either fold change or *P*-value were subjected to linkage analysis using Map Manager QTX. The map location of the most significantly linked marker for each probe set was used to plot the above figure because at the time this analysis was carried out not all probe set locations were mapped. In red are intra-chromosomal linkages (n=15). In blue are the linkages where the most linkage is inter-chromosomal (n=17). In yellow are the linked markers where the probe set location is unknown (n=15) and therefore the linkage cannot be classified.

point linkage analysis to generate a likelihood ratio statistic (LRS) for each marker on the RI strain map and determines genome wide significance by permutation testing. This permutation testing corrects for multiple testing across the markers but not for testing all 15,923 probe sets. Therefore Dr Enrico Petretto carried out FDR analysis to account for multiple testing across 15,923 transcripts and thus to estimate a suitable threshold of significance for this experiment.

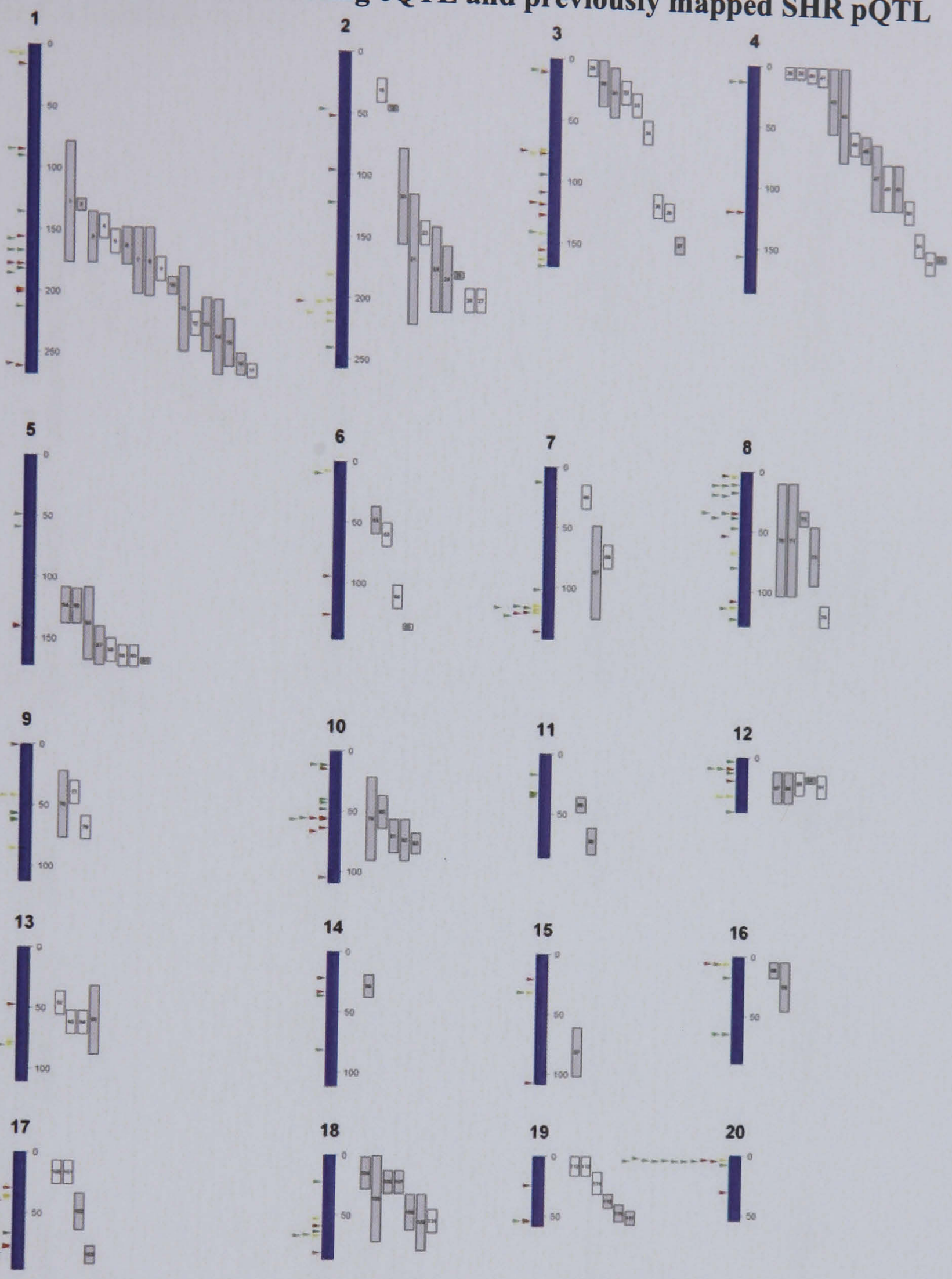
Following QTL Reaper analysis, marker redundancy was removed from the data sets and *cis/trans* classification of the eQTL was carried out. Marker redundancy removal is necessary as often a probe set will be linked to multiple tightly linked markers, that is markers which are in the same genomic location, and this results in an inflated number of linkages for each tissue. It was decided to merge linkages where markers were within 10cM of each other together into one locus and call this locus the eQTL. The algorithm was designed by this student in collaboration with Drs Enrico Petretto and Caroline Wallace and Michael Muller wrote a program to automate this process for all QTL Reaper linkages. This generated non redundant data sets of 2118 and 2490 eQTL in fat and kidney respectively, at a genome wide significance of $P<0.05$. These were then characterised as *cis* or *trans* eQTL by defining a *cis*-acting eQTL as one where the linkage peak location for a given probe set is within 10Mb (up or downstream) of its genomic location (**Methods 2.9.2**). **Table 3.5** shows the numbers of linkages detected in both the kidney and fat tissue data sets. At genome wide significance (permutated *P*-value 0.05) there are 2118 and 2490 eQTL in fat and kidney tissue respectively of which 311 are shared. **Table 3.5** also shows the numbers of eQTL linked in *cis* or *trans* at increasing levels of significance. At higher significance levels *cis* acting eQTL are more prevalent. **Figures 3.7 and 3.8** show

Table 3.5 Number and *cis/trans* classification of non-redundant eQTL detected in fat and kidney datasets at different genome-wide thresholds of significance

Threshold[a]	Fat					Kidney					Shared[c]				
	cis	*trans*	Total (*cis*/*trans*)	Unknown[b]	*Total*	*cis*	*trans*	Total (*cis*/*trans*)	Unknown[b]	*Total*	*cis*	*trans*	Total (*cis*/*trans*)	Unknown[b]	*Total*
$P < 0.05$	622	1211	1833	285	2118	800	1251	2051	439	2490	222	44	266	45	311
$P \leq 0.01$	392	268	660	94	754	514	295	809	140	949	138	18	156	27	183
$P \leq 10^{-3}$	189	46	235	36	271	270	56	326	58	384	75	9	84	13	97
$P \leq 10^{-4}$	97	18	115	22	137	129	10	139	29	168	45	3	48	7	55
$P \leq 10^{-5}$	49	4	53	8	61	62	4	66	6	72	23	0	23	2	25
$P \leq 10^{-6}$	18	0	18	2	20	28	3	31	3	34	5	0	5	3	8

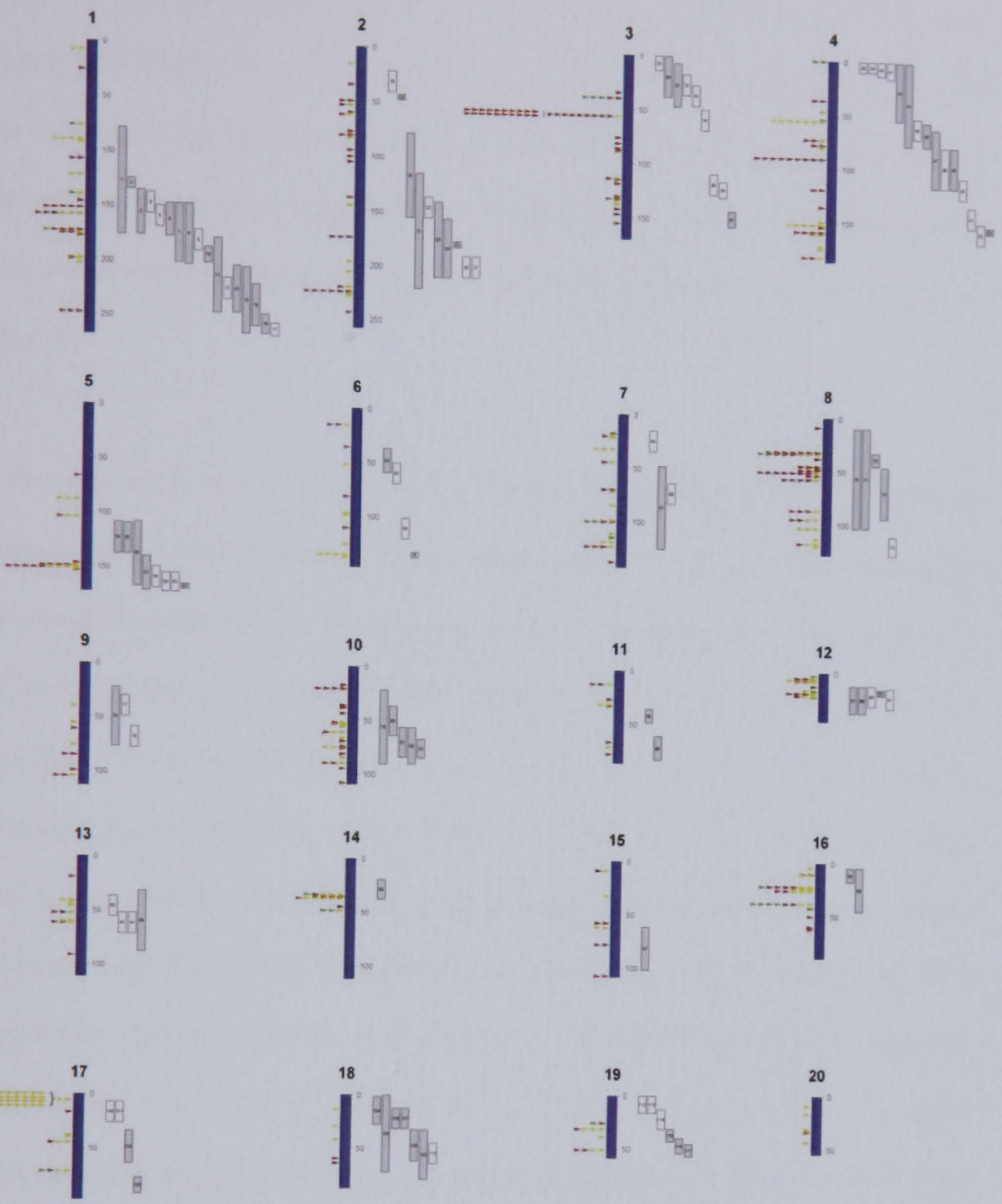
[a]*P*-value generated by permutation testing thus 0.05 is genome wide significance.
[b]These eQTL could not be defined as *cis* or *trans* because the physical location of either the probe set or the peak marker is unknown. [c]For each threshold of significance, overlapping eQTL between fat and kidney datasets were identified according to the following criteria: i) the eQTL are linked to the same probe set, ii) the eQTL have the same set of genetic markers identifying the locus, *or* they have both the same genetic marker and *P*-value threshold at the peak of linkage.

Figure 3.7 Location of *cis*-acting eQTL and previously mapped SHR pQTL



Chromosomes are shown in blue. The arrowheads on the left of each chromosome (yellow, fat; red, kidney; green, shared) represent the location of the probe set for each *cis*-acting eQTL with a $P \leq 10^{-4}$. Previously identified pQTL in SHR and the RI strains are shown on the right of each chromosome. Grey boxes denote pQTL where both flanking markers are mapped. White boxes denote pQTL where incomplete flanking marker information was available in which case the flanking marker(s) are estimated to be at a distance of 10Mbp from the peak of linkage. The numbers on the pQTL bars correspond to the pQTL table in **Appendix II**. pQTL in this Table which do not appear in the above Figure are due to lack of physical mapping information for the genetic markers that define the pQTL. Gene names, fold change in parental strains and public database references for probe sets are given in **Tables 3.7** (fat) and **3.8** (kidney).

Figure 3.8 Location of *trans*-acting eQTL and previously mapped SHR pQTL



pQTL are as described as in **Figure 3.7**. The arrowheads on the left of each chromosome (colours as in **Figure 3.7**) represent the location of the marker at the peak of linkage for each *trans*-regulated eQTL with $P \leq 10^{-2}$. This threshold of significance was chosen because *trans*-acting eQTLs are expected to have smaller effects than *cis*-acting eQTL. Gene names, fold change in parental strains and public database references for probe sets are reported on the **Data CD.**

genome maps illustrating the locations of all *cis* and *trans* eQTL with a genome wide significance of 10^{-4} and 10^{-2} respectively. Details of the *cis*-acting eQTL are given in **Tables 3.7 and 3.8** and this information was used as the basis for candidate gene selection in **Chapter 4**. The physiological QTL presented in **Figures 3.7** and **3.8** were curated by this student and represent all metabolic and cardiovascular QTL in SHR and are presented in **Appendix II**. This work is described in **Chapter 5**. The eQTL which were defined as *trans*-acting are shown in **Appendix IV**, which is on the **Data CD**.

I reclassified the 47 initial linkages (as in **Figure 3.6**) as either *cis* or *trans* using the definition above and the current release of the genome. The location of both the marker and the probe set were determined again and therefore could be designated as *cis/trans* eQTL linkages. **Figure 3.9** and **Table 3.6** shows the preliminary 47 linkages as in **Table 3.4** and **Figure 3.6** reclassified according to the above *cis/trans* definition. Following this reanalysis the inter-chromosomal linkages were redefined as being linked in *trans* and one intra chromosomal linkage is now classed as a *trans* linkage because the most significantly linked marker is more than 10Mb up or downstream of the probe set. Only two unclassified linkages remain where the probe set location is still unknown. Six of the 47 linkages are not shown in **Figure 3.9** as the marker location is not in the current genome assembly. Finally eight of these *cis* eQTL from this preliminary data set remain in the final data set of *cis* linkages generated when the entire data set was available (**Tables 3.7 and 3.8**). Those which did not show significance in QTL Reaper analysis were no longer considered significant and discarded from further analysis.

Table 3.6 Initial MapManager linkages reclassified

Probesets	Probe set Chromosome	Linked marker	Marker Chromosome	Control as in table 3.4	*cis or trans*
Down - Regulated					
1377452_at	8	D8Cebr16s5	8	intrachromosomal	*cis*
1370428_x_at	20	Tnfa	20	intrachromosomal	*cis*
1371033_at	20	Tnfa	20	intrachromosomal	*cis*
1387341_a_at	18	Rt2	19	extrachromosomal	*trans*
1398245_at	16	D16Rat6	16	intrachromosomal	*trans*
1390436_at	4	D15Rat107	15	unclassified	*trans*
1386901_at	4	Cd36	4	intrachromosomal	*cis*
1389986_at	1	D7Utr1	7	extrachromosomal	*trans*
1369764_at	13	D13Rat88	13	intrachromosomal	*cis*
1370114_a_at	2	D2Rat94	2	intrachromosomal	*cis*
1369171_at	8	D8Rat202	8 (not mapped)	intrachromosomal	-
1371960_at	5	D5Rat39	5	unclassified	*cis*
1370831_at	4	D11Rat28	11	extrachromosomal	*trans*
1376742_at	10	Ppy	10	unclassified	*cis*
1389974_at	3	D3Mit3	3	intrachromosomal	*cis*
1390048_at	10	D7Utr1	7	extrachromosomal	*trans*
1384717_at	7	D7Rat17	7	unclassified	*cis*
1387178_a_at	20	D20Cerbp97s7	20 (not mapped)	intrachromosomal	-
Up - regulated					
1376908_at	1	D15Rat107	15	extrachromosomal	*trans*
1371241_x_at	8	D15Rat1	15	extrachromosomal	*trans*
1371213_at	20	Tnfa	20	unclassified	*cis*
1369415_at	4	D19Utr1	19	extrachromosomal	*trans*
1374932_at	8	D14Rat52	14	extrachromosomal	*trans*
1369928_at	19	D15Rat1	15	extrachromosomal	*trans*
1370900_at	10	Cacna1s	4	unclassified	*trans*
1367626_at	1	D5Rat147	5	extrachromosomal	*trans*
1383564_at	1	D15Rat107	15	extrachromosomal	*trans*
1371776_at	2	D2Rat94	2	unclassified	*cis*
1369110_x_at	20	Tnfa	20	intrachromosomal	*cis*
1368718_at	1	D1Rat296	1 (not mapped)	intrachromosomal	-
1373053_at	hits multiple chromosomes	Drd2	8	unclassified	-
1376845_at	6	D15Rat107	15	extrachromosomal	*trans*
1387134_at	10	D15Rat107	15	extrachromosomal	*trans*
1388694_at	20	D10Rat6	10	extrachromosomal	*trans*
1374527_at	5	D5Rat79	5	unclassified	*trans*
1377334_at	not mapped	Tnfa	20	unclassified	-
1369202_at	not mapped	D11Cebr87s1	11 (not mapped)	intrachromosomal	-
1371015_at	11	D15Rat107	15	unclassified	*trans*
1391544_at	7	D8Mit5	8	unclassified	*trans*
Top Ten p - value					
1386901_at	4	Cd36	4	intrachromosomal	*cis*
1388602_at	7	D14Rat52	14	extrachromosomal	*trans*
1379896_at	10	D10Utr2	10	unclassified	*cis*
1368334_at	10	D16Rat14	16	extrachromosomal	*trans*
1373253_at	10	Slc4a	10	intrachromosomal	*cis*
1376635_at	16	D16Mit2	16	intrachromosomal	*cis*
1374668_at	14	D4Rat202	4	extrachromosomal	*trans*
1390562_s_at	20	D20Mgh5	20(not mapped)	unclassified	-
1384717_at	7	D7Rat17	7	unclassified	*cis*
1386507_at	not mapped	D5Cebr10s12	5(not mapped)	unclassified	-

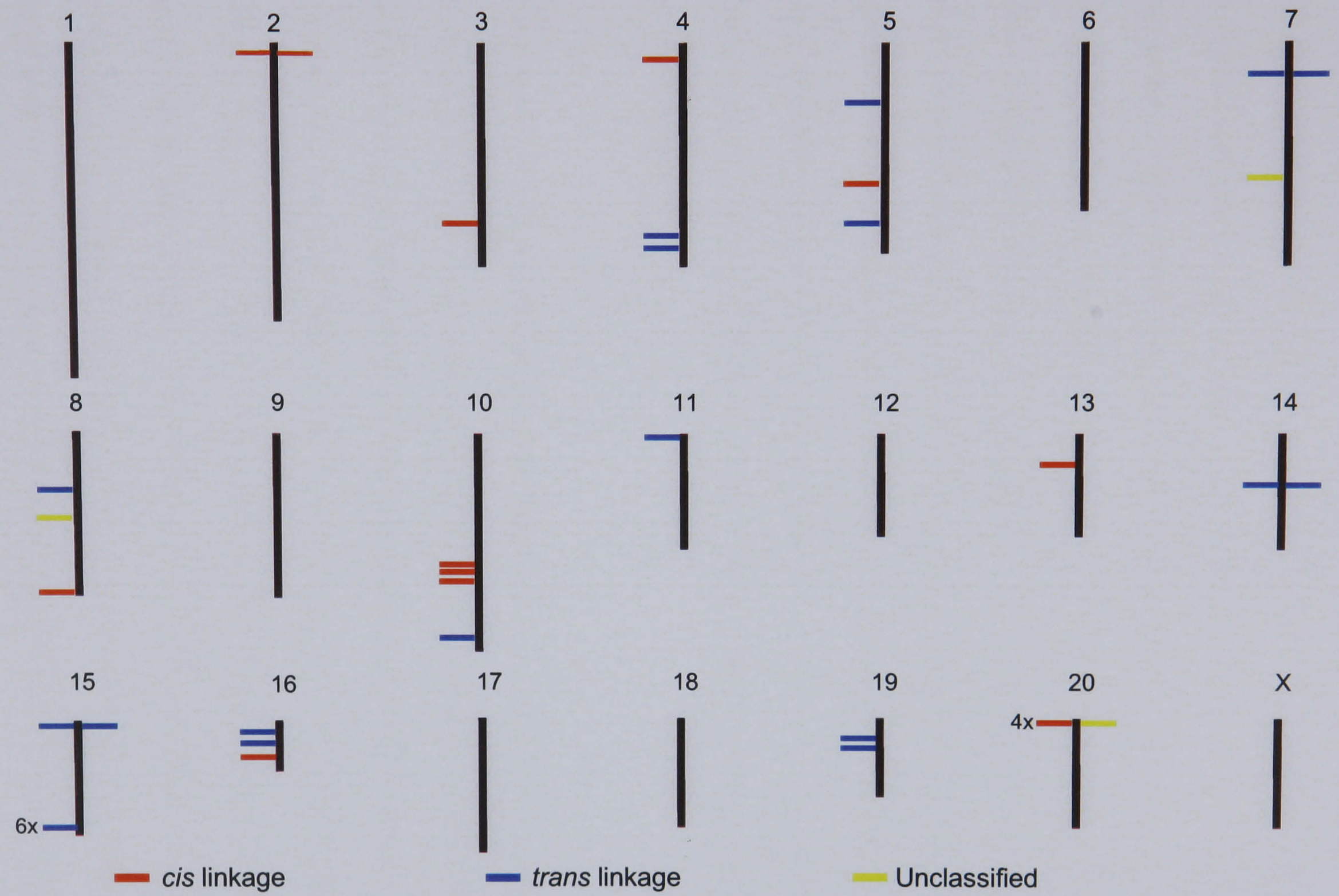


Figure 3.9 Chromosome map of the initial 47 eQTL linkages reclassified according to *cis/trans* definition.
This is **Figure 3.6** reclassified according to the *cis/trans* definition and with updated marker and probe set mapping information from Ensembl. 41 eQTL linkages are mapped above with only 2 now being unclassified as *cis* or *trans*. The remaining six linkages are not present as the marker is not mapped to the current assembly.

3.2.4 Target probe set sequencing

In the preliminary analyses (**Table 3.4**) and the final QTL Reaper analysis (**Table 3.7**) one of the most differentially expressed and significantly linked probe sets is 1386901_at which maps to the *Cd36* gene. This probe set is in a region of the SHR genome which contains a deletion and therefore the probe set contains multiple sequence variants. Thus the differential expression detected by this probe set is due to sequence variation causing differential hybridisation and rather than true differential expression. To determine if this occurs repeatedly and if differential expression could often be due to strain differences in the Affymetrix probe set region, a subset of genes were selected for sequencing from the preliminary data set of 92 microarrays. These probe sets were selected because they were linked in *cis* (with *cis* meaning on the same chromosome in this case) in both the kidney and fat data sets and had significant *P*-values in the initial analyses.

The entire probe set region was sequenced from cDNA in both the BN and SHR parental strains for 20 probe sets. These probe sets had been found to be differentially expressed in a Wilcoxon-Mann-Whitney test carried out by Dr Enrico Petretto. This WMW test has not been discussed in this thesis because it was carried out on the entire data set prior to QTL Reaper analysis before the QTL Reaper program was working and was the first analysis of the entire data set. Thirteen of these were found to have no sequence variants in the target region. Seven had sequence differences with one (1389156_at) having no differences between SHR and BN only with both strains to the reference sequence. Therefore six probe sets have sequence differences between the BN and SHR strains within the target sequence from which the

Table 3.7 *cis*-acting eQTLs in fat tissue mapped at $\geq 10^{-4}$

Gene Name	Gene Symbol	QTL Reaper p_value	Mean SHR Fat[b]	Mean BN Fat[b]	Fold change Fat[c]	t-test p-value in Fat	Mean SHR Kidney[b]	Mean BN Kidney[b]	Fold change Kidney[c]	t-test p-value in Kidney	probeset	peak marker	Chromosome position of probeset	Physical position of probeset	Chomosome Position of peak marker	Physical position of peak marker	Physical distance probeset to peak marker	Representative Public ID	Archival UniGene Cluster	UniGene ID
interleukin enhancer binding factor 3	*Ilf3*	0.0000001	335.58	94.46	3.55	0.00069955	139.31	52.64	2.63	0.00421879	1387366_at	*D8Rat68*	8	20500409	8	19460546	1039863	NM_053412	24511	Rn.24511
Transcribed sequences	---	0.0000001	23.84	109.77	4.60	0.00000336	32.91	56.27	1.71	0.00378399	1384717_at	*D7Rat17*	7	101521387	7	100292585	1228802	AA894199	22765	Rn.22765
Transcribed sequences	---	0.0000001	109.40	174.52	1.60	0.00014305	40.11	48.22	1.20	0.02716092	1388630_at	*D18Rat13*	18	67232914	18	66678736	554178	BI296506	31250	Rn.31250
Transcribed sequences	---	0.0000001	156.45	218.16	1.39	0.00481003	134.59	156.67	1.16	0.06561724	1380404_at	*D2Rat236*	2	213126553	2	204098590	9029963	H31722	14586	Rn.14586
Transcribed sequences	---	0.0000001	483.96	372.49	1.30	0.00000334	153.03	174.85	1.14	0.20526429	1388437_at	*D7Cebr77s1*	7	115173438	7	115036128	137310	AI103153	11888	Rn.11888
Similar to NRH quinone oxidoreductase (LOC291084), mRNA	---	0.0000001	205.60	304.18	1.48	0.01788183	316.92	382.84	1.21	0.14996509	1374959_at	*D17Mit2*	17	37258758	17	38356829	-1098071	AA945624	7882	Rn.7882
Transcribed sequences	---	0.0000001	119.92	88.71	1.35	0.00121598	97.36	78.64	1.24	0.00794355	1374907_at	*D1Rat47*	1	158396431	1	158590248	-193817	BI274338	41004	Rn.41004
acyl-CoA dihydroxyacetonephosphate acyltransferase	*Gnpat*	0.0000001	148.59	260.73	1.78	0.00350172	49.55	59.35	1.20	0.01316417	1368891_at	*D19Cebrp150s1*	19	55051904	19	53569812	1482092	AI014001	1739	Rn.1739
Transcribed sequences	---	0.0000001	58.66	101.36	1.73	0.00024936	88.14	76.75	1.13	0.21593154	1390364_at	*D11Mit2*	11	32782152	11	30888403	1893749	AI044316	20637	Rn.20637
apoptosis inhibitor 2	*Api2*	0.000001	306.09	281.59	1.09	0.11346728	114.39	105.44	1.08	0.42481810	1368291_at	*D8Rat58*	8	4658632	8	3039650	1618982	NM_021752	43867	Rn.43867
Similar to zinc finger protein 95 (LOC304275), mRNA	---	0.000001	118.04	181.02	1.53	0.00000501	80.76	98.50	1.22	0.01879836	1377244_at	*D12Cebrp97s4*	12	9727956	12	5552993	4174963	BM390979	38399	Rn.38399
prochymosin	LOC56825	0.000001	28.63	25.40	1.13	0.04381896	26.92	26.93	1.00	0.99304877	1369866_at	*D2Rat236*	2	202649345	2	204098590	-1449245	NM_020091	48676	Rn.48676
cd36 antigen	*Cd36*[a]	0.000001	656.05	3294.06	5.02	0.00000009	56.16	171.24	3.05	0.00066710	1386901_at	*Cd36*	4	13513148	4	13513439	-291	NM_031561	3790	Rn.3790
cd36 antigen	*Cd36*[a]	0.000001	2078.78	3885.80	1.87	0.00005691	66.81	140.67	2.11	0.00017586	1367889_a_at	*Cd36*	4	13513148	4	13513439	-291	NM_031561	3790	Rn.10241 8
Transcribed sequences	---	0.000001	47.25	35.34	1.34	0.20175081	33.76	30.80	1.10	0.03357709	1377212_at	*D12Rat56*	12	3569265	12	3017040	552225	BF416289	62985	Rn.62985
postsynaptic protein Cript	*Cript*	0.000001	295.31	393.68	1.33	0.00632077	227.53	299.31	1.32	0.00124067	1367222_at	*D6Cebrp124s2*	6	10479027	6	15163788	-4684761	NM_019907	30508	Rn.30508

Similar to DEAD (Asp-Glu-Ala-Asp) box polypeptide 51, Dead box protein 73D-like (LOC304570), mRNA	---	0.000001	231.22	182.80	1.26	0.03073809	151.30	131.13	1.15	0.01837126	1373682_at	*D12Rat22*	12	46020509	12	45907504	113005	BI280764	15921	Rn.15921
achaete-scute complex homolog-like 3 (Drosophila)	*Ascl3*	0.000001	47.85	98.28	2.05	0.00016444	94.96	124.08	1.31	0.00043350	1371085_at	*D1Rat356*	1	167456363	1	173621619	-6165256	AB048449	19594	Rn.19594
solute carrier family 7 (cationic amino acid transporter, y+system), member 7	*Slc7a7*	0.000003	263.47	79.80	3.30	0.00016223	1545.63	1885.92	1.22	0.19882594	1387808_at	*D15Rat6*	15	32467225	15	32084408	-187183	AF200684	18058	Rn.18058
Similar to mitochondrial ribosomal protein L4 isoform a (LOC363023), mRNA	---	0.000003	558.72	1331.66	2.03	0.00196498	306.77	591.35	1.93	0.00044739	1368366_at	*D8Rat58*	8	20021842	8	19460546	571296	BI275903	13113	Rn.13113
Transcribed sequence with strong similarity to protein ref:NP_112577.1 (H.sapiens) hypothetical protein MGC3133 [Homo sapiens]	---	0.000003	402.06	495.14	1.23	0.19172774	250.39	349.38	1.40	0.00656800	1371324_at	*D1Rat327*	1	7747020	1	8210299	-463279	BG673668	2083	Rn.2083
Similar to CHMP1.5 protein (LOC364890), mRNA	---	0.000003	331.04	261.68	1.27	0.01789579	211.73	183.34	1.19	0.05511491	1390128_at	*D19Rat89*	18	63788794	18	63770207	18587	BF557618	39132	Rn.39132
Transcribed sequences	---	0.000003	88.37	288.86	3.27	0.00009587	82.74	105.65	1.66	0.00093107	1375655_at	*D9Rat15*	9	63385214	9	62781727	603487	BE107208	22668	Rn.22668
Transcribed sequences	---	0.000003	620.78	186.96	3.32	0.00222211	155.30	58.20	2.67	0.00002702	1375343_at	*D18Rat13*	18	66437340	18	66678736	-241396	BE116572	7718	Rn.7718
Similar to RIKEN cDNA 2010012D11 (LOC361239), mRNA	---	0.000003	301.29	186.48	1.62	0.00864505	627.82	491.38	1.28	0.05706830	1388617_at	*D17Rat144*	17	37146280	17	37974570	-928290	AW914746	22050	Rn.22050
Transcribed sequences	---	0.000003	30.90	22.44	1.38	0.04060744	24.82	21.02	1.18	0.01107155	1376550_at	*D12Mit5*	12	13196112	12	12800256	395856	BF405084	58248	Rn.58248
amyotrophic lateral sclerosis 2 (juvenile) chromosome region, candidate 3	*Als2cr3*	0.000003	168.70	126.46	1.33	0.38211064	317.84	211.79	1.50	0.04893454	1370176_at	*D9Cebrp60s12*	9	57042696	9	59010552	-1967856	BG378620	22429	Rn.12357.1
integrin alpha 1	*Itga1*	0.000003	178.67	185.44	1.04	0.77205037	185.65	108.81	1.71	0.12973055	1387144_at	*D2Rat201*	2	47085288	2	49619880	-2534592	NM_030994	11491	Rn.91044
LOC361111 (LOC361111), mRNA	---	0.000003	217.47	134.02	1.62	0.01118083	99.89	61.68	1.46	0.00219000	1390435_at	*D16Mit2*	16	6555704	16	4304396	2251308	AA996717	12381	Rn.12381
aldehyde oxidase	*Aox1*	0.000003	2899.87	1046.47	2.77	0.00035238	130.81	93.87	1.39	0.00044379	1387376_at	*D9Rat93*	9	56880889	9	55895160	985729	NM_019363	15681	Rn.15681
peroxisomal biogenesis factor 3	*Pex3*	0.000003	140.26	222.65	1.59	0.00005704	103.33	122.29	1.18	0.00254582	1368526_at	*D1Rat327*	1	8337072	1	8210299	126773	NM_031350	11773	Rn.11773
synaptogyrin 1	*Syngr1*	0.000005	134.02	167.02	1.25	0.03225551	367.50	323.11	1.14	0.12843313	1386511_s_at	*Bsrp*	7	118388943	7	121630597	-3241654	BF403383	15631	Rn.88235
kinesin 1C	*Kif1c*	0.000005	449.26	745.00	1.66	0.00399849	403.44	584.44	1.45	0.00972230	1370930_at	*Abpa*	10	57615722	10	58447585	1168037	BF417285	15980	Rn.89698
flavin-containing monooxygenase 3	*Fmo3*	0.000005	281.32	115.16	2.44	0.00009834	2284.94	1454.85	1.57	0.00391055	1368304_at	*D13Wrl3*	13	78682851	13	78084836	598015	NM_053433	11678	Rn.28357

Transcribed sequences	---	0.000005	193.17	110.12	1.75	0.01054545	136.18	111.32	1.22	0.08544799	1375194_at	*D14Rat52*	14	38611129	14	37229534	618405	AI231460	12107	Rn.12107
four and a half LIM domains 2	*Fhl2**	0.000006	93.72	107.16	1.14	0.14292782	268.86	309.23	1.15	0.28650109	1371651_at	*D9Rat104*	9	42321222	9	42292159	29063	AA800031	3849	Rn.3849
Similar to zinc finger protein 366 (LOC293044), mRNA	---	0.000006	86.22	137.18	1.59	0.00087685	129.83	229.11	1.76	0.00356380	1389650_at	*D1Rat42*	1	135946190	1	135496747	449443	BI295153	43682	Rn.43682
GNAS complex locus	*Gnas*	0.000006	63.93	52.54	1.22	0.44927926	58.89	59.76	1.01	0.86115420	1387906_a_at	*D3Rat132*	3	165120034	3	164970696	149338	AF107845	31	Rn.31
Macrophage stimulating 1 (hepatocyte growth factor-like)	*Mst1*	0.000006	193.07	888.81	4.85	0.00000124	60.01	61.41	1.02	0.79462520	1369171_at	*Apeh*	8	113369944	8	113376046	-6102	NM_024352	10613	Rn.10613
sex hormone binding globulin	*Shbg*	0.000008	56.69	76.17	1.34	0.00030858	71.27	87.18	1.22	0.02822549	1368285_at	*Abpa*	10	56447684	10	56447685	-1	NM_012650	37473	Rn.37473
RT1 class Ib, locus Aw2	*RT1-Aw2**	0.000008	222.58	527.13	2.37	0.00105206	119.56	256.82	2.16	0.00164360	1389734_x_at	*D20Rat71*	20	2837388	20	4741042	-1903654	BI282965	40492	Rn.83667
cytochrome P450 4F5	*Cyp4f5*	0.000009	41.85	78.91	1.89	0.00041188	28.81	31.83	1.11	0.04342743	1392720_at	*D7Mit17*	7	13111312	7	13075668	35644	BG376949	43212	Rn.10171
RT1 class Ib, locus Aw2	*RT1-Aw2**	0.000009	20.15	387.51	19.74	0.00000001	27.55	148.72	5.40	0.00000029	1370428_x_at	*D20Rat41*	20	5023682	20	4741042	282020	AJ249701	39743	Rn.10796 0
Similar to trinucleotide repeat containing 6 EDIE, GW182 autoantigen (LOC308971), mRNA	---	0.000009	278.88	112.49	2.48	0.00146996	113.97	47.82	2.38	0.00002418	1375664_at	*D1Rat287*	1	182191778	1	190264130	-8072352	BI284732	2857	Rn.6839
Transcribed sequences	---	0.000009	201.92	174.81	1.16	0.11256396	146.65	134.78	1.09	0.06726290	1390827_at	*D8Mit12*	8	67834471	8	58049018	9784553	AA997679	12544	Rn.12544
Transcribed sequence with weak similarity to protein pir S52790 (H.sapiens) S52790 kynurenine--oxoglutarate transaminase	---	0.00001	357.90	240.74	1.49	0.00002960	215.13	166.63	1.29	0.03998280	1374005_at	*D2Rat66*	2	240840214	2	241762826	-922612	BI295878	28263	Rn.28263
Transcribed sequences	---	0.00001	30.36	18.93	1.60	0.01846817	46.47	19.12	2.43	0.00145391	1378840_at	*D12Cebrp97s4*	12	9698220	12	5552993	4145227	AI103040	22444	Rn.22444
Similar to tetracycline transporter-like protein (LOC305449), mRNA	---	0.00001	110.29	71.12	1.55	0.02261117	48.39	41.13	1.18	0.00825318	1382667_at	*D14Rat37*	14	81749981	14	81450719	299262	AA866425	3100	Rn.3100
---	---	0.00001	482.50	576.92	1.20	0.01788885	163.81	209.77	1.28	0.06025336	1399157_at	*D1Rat20*	1	90508428	1	90360305	148123	AA849312	---	---
dermatopontin	*Dpt*	0.000012036	4189.61	3492.99	1.20	0.02718834	1009.60	877.08	1.15	0.26273485	1371732_at	*D13Rat131*	13	80605153	13	80769618	-164465	BI285495	20274	Rn.8396
Transcribed sequences	---	0.000012232	131.15	178.35	1.36	0.00114824	76.75	90.68	1.18	0.00842986	1379927_at	*D1Rat212*	1	84967045	1	86588680	-1621635	BE100774	12898	Rn.12898
Transcribed sequences	---	0.000012397	442.54	513.16	1.19	0.60415811	51.13	46.18	1.11	0.29079309	1390443_at	*D11Rat20*	11	17447184	11	16670413	776751	BE105991	12252	Rn.12252

Similar to 2610203E10Rik protein (LOC362188), mRNA	---	0.000013319	58.56	74.42	1.27	0.01903586	45.43	50.31	1.11	0.05308761	1383440_at	*D3Rat35*	3	78140579	3	68184508	9955071	BF418459	19446	Rn.19446
mRNA decapping enzyme	*Dcps**	0.000013995	148.96	297.49	2.00	0.00003877	101.09	160.31	1.59	0.00038906	1390185_at	*D8Rat219*	8	34913410	8	38221201	-3307791	BE109709	21813	Rn.10198.6
Transcribed sequence with strong similarity to protein pir:JC2473 (H.sapiens) JC2473 doc2 protein - human	---	0.000014139	125.15	102.36	1.22	0.05468901	76.95	61.10	1.26	0.11520136	1377614_at	*D1Rat287*	1	186181355	1	190264130	-4082775	BI283147	3179	Rn.94076
SAC2 (supressor of actin mutations 2, homolog)-like (S. cerevisiae)	*Vps52**	0.000014235	186.12	120.75	1.53	0.00950337	113.28	74.42	1.52	0.00285457	1394386_s_at	*D20Rat41*	20	5074917	20	4741042	333875	BE107955	51354	Rn.51354
acyl-Coenzyme A dehydrogenase family, member 9	*Acad9*	0.000014589	419.34	259.33	1.62	0.00173740	224.41	167.56	1.34	0.00292120	1373389_at	*D2Rat136*	2	122529659	2	121330528	1199131	BM391807	13649	Rn.13649
Transcribed sequences	---	0.000014765	173.33	197.54	1.14	0.02705579	173.32	182.32	1.05	0.29087223	1372415_at	*D12Rat36*	12	32985603	12	30695058	2290545	BM385649	9533	Rn.73201
GM2 ganglioside activator protein	*Gm2a*	0.00001757	2197.28	3360.62	1.53	0.00089825	366.56	430.88	1.17	0.02096778	1371803_at	*D10Rat215*	10	40528161	10	38404352	2123809	AI717483	11732	Rn.98783
Transcribed sequences	---	0.000017667	222.85	194.96	1.14	0.14683687	135.17	104.57	1.29	0.01712300	1373537_at	*D3Cebr26s1*	3	9983732	3	9359972	623760	BE113175	20445	Rn.20445
RT1 class I, A3	*RT1-A3*	0.000018154	440.17	69.58	6.33	0.00001879	87.59	33.10	2.65	0.00285349	1369110_x_at	*D20Rat41*	20	5012216	20	4741042	271174	NM_012845	39743	---
v-ets erythroblastosis virus E26 oncogene like (avian)	*Erg*	0.000018644	170.41	208.50	1.21	0.46103334	142.05	114.78	1.24	0.02681287	1387789_at	*D11Cebr87s1*	11	35764038	11	32334036	3430002	NM_133397	50673	Rn.50673
RT1 class II, locus Ba	*RT1-Ba*	0.000019508	397.98	58.17	6.84	0.00051633	741.66	65.52	11.32	0.00000487	1377334_at	*D20Rat41*	20	4702361	20	4741042	-38681	BG378249	39138	Rn.25717
Transcribed sequence with weak similarity to protein sp:Q16600 (H.sapiens) Z239_HUMAN Zinc finger protein 239	---	0.000020333	32.78	29.48	1.11	0.04126769	34.75	30.87	1.13	0.06963991	1377501_at	*D8Rat58*	8	18082519	8	19460546	-1378027	BM383704	57387	Rn.12335.8
solute carrier family 3, member 1	*Slc3a1*	0.000020710	99.27	60.04	1.60	0.07130561	9270.09	8708.17	1.07	0.22922613	1368440_at	*D6Cebrp424s2*	6	8402310	6	15163788	-6761478	NM_017216	11186	Rn.11186
RT1 class II, locus Bb	*RT1-Bb*	0.000022381	40.28	513.25	12.74	0.00001023	32.39	493.42	15.23	0.00000001	1371033_at	*D20Rat41*	20	4733596	20	4741042	-7446	AI715202	20089	Rn.20089
---	---	0.000024048	331.62	273.28	1.21	0.00203991	115.33	97.91	1.18	0.06995183	1377040_a_at	*D3M2*	3	143523707	3	143120242	403465	AI172225	---	---
hypoxia up-regulated 1	*Hyou1*	0.000024439	251.05	301.96	1.56	0.02472409	158.30	280.66	1.77	0.00155671	1371442_at	*D8Mit1*	8	47367303	8	46015473	1548170	BI282904	983	Rn.10542
RT1 class Ib, locus Aw2	*RT1-Aw2**	0.000028205	178.05	40.59	4.38	0.00034166	79.18	49.97	1.58	0.00020344	1371213_at	*D20Rat41*	20	5029602	20	4741042	288560	AJ005023	39743	Rn.11679.2
mRNA decapping enzyme	*Dcps**	0.000028302	289.03	123.90	2.33	0.00006243	196.97	98.88	2.03	0.00053527	1374583_at	*D8Rat219*	8	34908365	8	38221201	-3312836	BI288131	7059	Rn.10198.6

Similar to RIKEN cDNA 1110011E12 (LOC310645), mRNA	---	0.000028986	442.44	304.47	1.45	0.05263669	71.48	73.08	1.02	0.73589094	1373243_at	*D2Rat150*	2	181614414	2	190305036	-8690622	BG378288	22202	Rn.22202
brain-enriched SH3-domain protein Besh3	*Besh3*	0.000029268	130.71	114.61	1.14	0.27807969	88.61	85.65	1.03	0.64686745	1387824_at	*D9Rat4*	9	86479486	9	90901633	-4422147	AF254801	20670	Rn.20670
Transcribed sequence with strong similarity to protein sp P00722 (E coli) BGAL_ECOLI Beta-galactosidase	---	0.000029557	793.68	982.09	1.24	0.02252325	418.64	510.64	1.23	0.01733594	1372728_at	*D2Rat236*	2	203893953	2	204098590	-204637	BE103745	18004	Rn.18004
Similar to RIKEN cDNA 3110038B19 (LOC291444), mRNA	---	0.000034006	196.79	158.18	1.24	0.04984263	127.70	112.53	1.13	0.05825667	1374560_at	*D18Rat55*	18	53212805	18	54928884	-1715079	AI137296	22764	Rn.22764
four and a half LIM domains 2	*Fhl2**	0.000035088	54.30	68.14	1.25	0.11405661	243.46	295.85	1.22	0.03578442	1369313_at	*D9Rat104*	9	42335939	9	42292159	43780	NM_031677	3849	Rn.3849
Transcribed sequence with weak similarity to protein pir RGECDW (E coli) RGECDW transcription activator of D-serine dehydratase -	---	0.000035699	42.04	133.87	3.18	0.00739684	37.96	41.75	1.10	0.15886338	1376398_at	*D2Cebrf58*	2	219739471	2	220393840	-654369	BF417784	7699	Rn.7699
Transcribed sequences	---	0.000037415	17.88	19.53	1.09	0.04633230	21.51	22.18	1.03	0.48110306	1384309_at	*D9Rat158*	9	61941251	9	61488829	452422	BF393535	7284	Rn.7284
guanine nucleotide binding protein beta 5	*Gnb5*	0.000039711	322.43	297.96	1.08	0.45962445	158.42	102.40	1.55	0.00935933	1382105_at	*D8Rat21*	8	80188907	8	87142271	-6953364	AI603217	23042	Rn.23042
RT1 class Ib, locus Aw2	*RT1-Aw2**	0.000048583	34.33	84.47	2.46	0.00145264	24.99	48.23	1.93	0.00014937	1388202_at	*D20Rat41*	20	3441453	20	4741042	-1299589	BI385698	39743	Rn.39743
SAC2 (supressor of actin mutations 2, homolog)-like (S. cerevisiae)	*Vps52**	0.000049587	252.84	176.31	1.43	0.00521172	271.26	179.60	1.51	0.00261584	1369867_at	*D20Rat41*	20	5078591	20	4741042	337549	NM_033087	21933	Rn.51354
CUG triplet repeat, RNA-binding protein 2	*Cugbp2*	0.000063571	781.90	1336.04	1.71	0.01458467	185.41	125.84	1.47	0.02304054	1389867_at	*D17Rat62*	17	83039499	17	83490772	-451273	BI281086	37864	Rn.30004
cytochrome P450, family 2, subfamily d, polypeptide 9	*Cyp2d9*	0.000063571	36.66	28.91	1.27	0.00207722	158.71	102.52	1.55	0.00049799	1370377_at	*D7Rat129*	7	120806126	7	119195750	1610376	M25143	10842	Rn.91133
Transcribed sequence with weak similarity to protein pir T02345 (H.sapiens) T02345 hypothetical protein KIAA0324 - human	---	0.0000563	302.20	251.22	1.20	0.01007597	316.11	266.64	1.19	0.04003848	1374675_at	*D1Rat356*	1	167465427	1	173621619	-6156192	BG374992	8151	Rn.8151
Transcribed sequences	---	0.000065556	83.07	105.27	1.67	0.00621172	175.36	255.98	1.46	0.01527223	1374126_at	*D17Rat50*	17	67022332	17	66987506	34826	BG374261	27862	Rn.27862
Transcribed sequence with moderate similarity to protein pdb 1LBG (E coli) B Chain B, Lactose Operon Repressor Bound To 21-Base Pair	---	0.000056338	299.67	72.22	4.15	0.00109714	382.91	47.28	8.08	0.00003230	1375676_at	*D3Rat166*	3	95234620	3	101256409	-6021789	AA956897	9461	Rn.9461
Similar to mitochondrial ribosomal protein L52 CG1577-PA (LOC361037), mRNA	---	0.000058537	526.49	302.10	1.74	0.00011563	524.14	301.25	1.74	0.00011494	1388654_at	*D15Rat6*	15	32507604	15	32654408	-146804	AI409180	6660	Rn.6660
Transcribed sequences	---	0.000059	156.45	218.10	1.39	0.00481003	134.59	156.67	1.16	0.08561724	1380404_at	*D2Rat54*	2	213128553	2	209974717	3153836	H31722	14586	Rn.14586
Similar to Peci protein (LOC2810751 mRNA	---	0.000060109	894.17	1209.55	1.35	0.01839492	360.07	618.17	1.71	0.00134838	1388908_at	*D17Rat144*	17	36213022	17	37974570	-1761548	BI278258	3194	Rn.10802 9

Similar to M-RdgB2 retinal degeneration protein B subtype 2 (LOC288655), mRNA	---	0.000060914	402.61	520.08	1.29	0.00662733	237.24	222.98	1.06	0.21550420	1372532_at	*D12Rat36*	12	33468361	12	30895058	2771303	BI289477	19419	Rn.75968
Transcribed sequence with weak similarity to protein ref NP_501117.1 (C elegans) C34D4.11.p (Caenorhabditis elegans)	---	0.000072289	60.26	77.65	1.29	0.00667334	40.70	38.03	1.07	0.48251511	1374097_at	*D3Rat35*	3	75357533	3	68184508	7173025	AW251791	41066	Rn.41066
alpha 1,3-fucosyltransferase Fuc-T (similar to mouse Fut4)	*Fut4*	0.000073333	99.77	80.70	1.24	0.01862562	40.10	33.00	1.22	0.07151617	1373838_at	*D8Utr3*	8	11542040	8	9172072	2369968	AI237192	37309	Rn.44467
Transcribed sequences	---	0.000075342	174.04	264.41	1.52	0.00050831	36.51	42.10	0.15	0.00401114	1380293_at	*D18Rat13*	18	67235253	18	66878736	556517	BF412754	6667	Rn.31250
Similar to KIAA1731 protein (LOC363018), mRNA	---	0.000078571	75.50	54.53	1.38	0.00013353	57.29	47.04	1.22	0.00113026	1380286_x_at	*D8Utr3*	8	12117542	8	9172072	2945470	BM386844	36574	Rn.86809
Similar to RIKEN cDNA 2700091N06 (LOC312677), mRNA	---	0.000091603	251.08	191.72	1.31	0.00669408	169.41	114.96	1.47	0.00056373	1390284_at	*D4Rat202*	4	156949551	4	159380532	-2410981	BM389683	19815	Rn.19815
Similar to RIKEN cDNA C730048E16 (LOC362957), mRNA	---	0.000092308	84.57	62.49	1.35	0.00110563	111.15	96.94	1.15	0.04452594	1389264_at	*D7Rat129*	7	117060690	7	119195750	-2135060	BI278594	28758	Rn.28758
phosphatidylinositol glycan, class L	*Pigl*	0.000095588	264.56	213.20	1.24	0.01412200	216.43	194.67	1.11	0.08147403	1396309_at	*D10Rat73*	10	48673210	10	46352552	2320658	D88364	10774	Rn.10774
Similar to RIKEN cDNA 1500015O10 (LOC363225), mRNA	---	0.000096	204.97	97.93	2.09	0.01398184	448.02	485.36	1.08	0.29634922	1372646_at	*D9Rat104*	9	42642305	9	42282159	650147	BF418649	16593	Rn.16593

Footnotes

Each row in the table contains the eQTL description (gene name, gene symbol), P-value of the peak of linkage from eQTL Reaper, mean expression values, fold change and t-test P-values in the parental strains for fat and kidney; probeset and genebc marker at the peak of linkage, chromosome and physical position of the probset and of the marker at the peak of linkage, distance between probeset and marker at the peak of linkage. Public ID relating to the probeset as assigned by Affymetrix, Archival UniGene Cluster number relating to the probeset as assigned by Affymetrix; UniGene ID relating to the probeset as assigned by Affymetrix

[a]Five genes with multiple eQTLs (*Cd36*, *Dcps*, *Vps52*, *FN7* and *R71Aw2*). These are eQTL Reaper linkages to two or more distinct Affymetrix probesets representing the same gene

[b]Mean parental strains expression values are calculated from antilog RMA values

[c]Fold change in the parental strains was calculated such that the overexpressed strain was divided by the underexpressed strain

Table 3.8. *cis*-acting eQTLs in kidney tissue mapped at P≤10[-4]

Gene Name	Gene Symbol	QTL Reaper p_value	Mean BHR Kidney[b]	Mean BN Kidney[b]	Fold change Kidney[c]	t-test p-value in Kidney	Mean BHR Fat[b]	Mean BN Fat[b]	Fold change Fat[c]	t-test p-value in Fat	probeset	peakmarker	Chromosome position of probeset	Physical position of probeset	Chomosome Position of peak marker	Physical position of peak marker	Physical distance probeset to peak marker	Representative Public ID	Archival UniGene Cluster	UniGene ID
aldehyde oxidase	*Aox1*	0.0000001	130.81	93.87	1.39	0.00044379	2899.87	1046.47	2.77	0.00036238	1387376_at	*D9Rat93*	9	56890689	9	55895160	995729	NM_019363	15681	Rn.15681
cd36 antigen	*Cd36*[a]	0.0000001	56.16	171.24	3.05	0.00006710	666.05	3294.06	5.02	0.00000009	1386901_at	*Cd36*	4	13513148	4	13613439	-291	NM_031561	3790	Rn.3790
guanylate cyclase 2g	*Gucy2g*	0.0000001	24.52	34.67	1.41	0.00250054	18.99	18.36	5.03	0.41055148	1370671_at	*D1Rat225*	1	261719476	1	261391328	328148	U33847	10198	Rn.10198
interleukin enhancer binding factor 3	*Ilf3*	0.0000001	138.31	52.64	2.63	0.00421879	335.58	94.46	3.55	0.00069955	1387366_at	*D8Rat68*	8	20500409	8	19460546	1039863	NM_053412	24511	Rn.24511
phosphatidylinositol 3-kinase	*Pik3c3*	0.0000001	114.59	97.60	1.17	0.04679169	199.59	156.46	1.21	0.00905965	1369655_at	*D18Rat103*	18	22564078	18	20164002	2400076	NM_022958	30010	Rn.30010
Similar to 3230401M21Rik protein (LOC287157), mRNA	---	0.0000001	106.11	231.59	2.29	0.00003738	149.64	144.79	1.03	0.59555338	1376136_at	*D10Mit6*	10	15110603	10	14721325	389277	BI295118	12579	Rn.12579
Similar to cDNA sequence BC023151 (LOC309103), mRNA	---	0.0000001	112.18	86.61	1.31	0.00216644	294.94	298.80	1.01	0.84185589	1374831_at	*D1Cebr10s3*	1	201277127	1	198323072	2954055	AI172579	23728	Rn.23728
Similar to FKSG26 protein (LOC292062), mRNA	---	0.0000001	890.50	791.16	1.13	0.05291423	1305.12	1218.07	1.07	0.10120008	1371870_at	*D19Utr7*	19	53737293	19	53156456	580837	AI169493	7782	Rn.107920
Similar to hypothetical protein MGC20446 (LOC361729), mRNA	---	0.0000001	89.02	63.62	1.38	0.00535622	201.34	140.20	1.44	0.00048101	1372646_at	*D1Rat293*	1	213238035	1	213040485	197550	AI179227	13816	Rn.13816
Similar to mitochondrial ribosomal protein L4 isoform a (LOC363023), mRNA	---	0.0000001	306.77	591.35	1.93	0.00044739	558.72	1131.65	2.03	0.00196498	1388395_at	*D8Rat68*	8	20031842	8	19460546	571296	BI275903	13113	Rn.13113
Similar to RIKEN cDNA 1500010J02 (LOC360544), mRNA	---	0.0000001	288.49	196.53	1.37	0.00506498	305.73	224.25	1.09	0.20624756	1373955_at	*D10Rat59*	10	55802253	10	55017963	784290	AI598089	11557	Rn.11557
Transcribed sequences	---	0.0000001	32.91	56.27	1.71	0.00378399	23.84	109.77	4.60	0.00000336	1384717_at	*D7Rat17*	7	101521387	7	100292585	1228802	AA894199	22765	Rn.22765
Transcribed sequences	---	0.0000001	99.10	154.22	1.56	0.00141337	199.97	208.92	1.04	0.72470008	1376259_at	*D17Rat92*	17	78497260	17	83490772	-4993512	AI171093	8227	Rn.8227
achaete-scute complex homolog-like 3 (Drosophila)	*Ascl3*	0.000001	94.96	124.08	1.31	0.00043350	47.85	98.28	2.05	0.00016444	1371085_at	*D1Rat356*	1	167456353	1	173521619	-6165266	AB046449	19594	Rn.19594
cytochrome P450, family 2, subfamily d, polypeptide 26	*Cyp2d26*	0.000001	832.30	1845.05	2.22	0.00006997	44.02	40.28	1.09	0.23328638	1367917_at	*D7Rat129*	7	120832474	7	119195750	1636724	NM_012730	40137	Rn.91355
dopa/tyrosine sulfotransferase	*LOC64305*	0.000001	839.97	182.59	4.60	0.00082517	34.29	27.39	1.25	0.11008161	1387314_at	*D11Utr6*	14	22025564	14	18265386	3759168	NM_022513	1183	Rn.1183
dynein, cytoplasmic, heavy chain 2	*Dnchl2*	0.000001	149.05	178.62	1.20	0.07027394	152.06	144.59	1.05	0.39071713	1388676_at	*D8Rat58*	8	3949870	8	3039650	910220	NM_023024	44806	Rn.44806
kinesin 1C	*Kif1c*	0.000001	403.44	584.44	1.45	0.00872230	449.26	745.00	1.66	0.00360849	1370330_at	*Abpa*	10	59611723	10	59447685	1168017	BF417285	15980	Rn.99698

Similar to estradiol 17beta-dehydrogenase (EC 1.1.1.62), A-specific - mouse (LOC307092), mRNA	---	0.000001	291.83	1192.08	4.08	0.00019233	16.21	15.65	1.04	0.27113226	1393902_at	*D17Rat151*	17	77104669	17	71580725	5523943	AI072107	20058	Rn.102679
Similar to PKE protein kinase (LOC365381), mRNA	---	0.000001	31.63	34.51	1.09	0.00972546	29.43	30.29	1.03	0.60914848	1377534_at	*D1Rat69*	1	199033868	1	199262391	-228523	BE097305	49654	Rn.49654
Similar to Stromal cell-derived factor 2 precursor (SDF-2) (LOC287470), mRNA	---	0.000001	469.12	353.78	1.31	0.00750751	656.90	643.16	1.02	0.72014032	1371645_at	*D10Mgh6*	10	64397520	10	67691615	-3294095	BI284339	3965	Rn.3965
Transcribed sequence with weak similarity to protein pir S52790 (H.sapiens) S52790 kynurenine--oxoglutarate transaminase	---	0.000001	215.13	169.63	1.26	0.03998280	357.90	240.74	1.49	0.00602960	1374006_at	*D2Rat56*	2	240840214	2	241762826	-922612	BI295878	28263	Rn.28263
Transcribed sequences	---	0.000001	267.84	361.48	1.35	0.00789216	193.27	213.17	1.10	0.04058250	1377076_at	*D8Rat219*	8	38897833	8	38221201	676732	AI716131	24258	Rn.24258
Transcribed sequences	---	0.000001	62.74	105.65	1.68	0.00003107	88.37	288.86	3.27	0.00009587	1375655_at	*D9Rat15*	9	63385214	9	62781727	503487	BE107209	22668	Rn.22668
Transcribed sequences	---	0.000001	155.36	58.20	2.67	0.00002702	620.76	186.96	3.32	0.00222211	1375343_at	*D18Rat13*	18	66437340	18	66678736	-241396	BE116572	7718	Rn.7718
Transcribed sequences	---	0.000001	112.88	88.13	1.28	0.00382223	103.20	74.14	1.39	0.00019016	1390690_at	*D8Rat7*	8	123006013	8	120852323	2153690	AA893902	14812	Rn.14812
Transcribed sequences	---	0.000001	97.36	78.64	1.24	0.00794355	119.92	58.71	1.35	0.00121598	1374901_at	*D1Rat47*	1	158396431	1	158590248	193817	BI274338	41004	Rn.41004
Transcribed sequences	---	0.000001	33.76	30.80	1.10	0.03357709	47.25	35.34	1.34	0.20175081	1377212_at	*D12Rat56*	12	3569265	12	3017040	552225	BF416289	62985	Rn.62985
cd36 antigen	Cd36[4]	0.000003	66.81	140.67	2.11	0.00017586	2078.78	3885.80	1.87	0.00000691	1368685_a_at	*Cd36*	4	13513148	4	13513439	-291	NM_031561	3790	Rn.102418
coagulation factor 2	*F2*	0.000003	40.56	64.62	1.59	0.00864018	20.65	19.27	1.07	0.25810962	1368442_at	*D3Rat35*	3	75908106	3	68184508	7723597	NM_022924	54498	Rn.54498
glutathione S-transferase, mu 1	*Gstm1*	0.000003	235.28	659.30	2.38	0.00007438	452.18	397.25	1.14	0.35058368	1386985_at	*D2Rat236*	2	203540915	2	204098590	-557675	M28241	625	Rn.93760
ProSAPiP1 protein	*ProSAPiP 1*	0.000003	468.47	660.03	1.41	0.00002231	146.06	141.48	1.03	0.50290999	1373057_at	*D3Rat159*	3	118168659	3	119342992	-1174033	AI229142	7646	Rn.7646
putative N-acetyltransferase Camello 2	*Cml2*	0.000003	614.39	1150.98	1.79	0.00006528	19.05	19.32	1.01	0.70943991	1368966_at	*D4Rat44*	4	120007763	4	121888870	-1582107	NM_021668	42037	Rn.42037
RT1 class Ib, locus Aw2	*RT1-Aw2*[1]	0.000003	119.56	258.82	2.16	0.00164360	222.58	527.13	2.37	0.00105266	1389734_x_at	*D20Rat41*	20	2837388	20	4741042	-1903654	BI282965	40492	Rn.83667
RT1 class Ib, locus Aw2	*RT1-Aw2*[1]	0.000003	79.18	49.97	1.58	0.00026344	178.08	40.69	4.38	0.00034166	1371213_at	*D20Rat41*	20	5029602	20	4741042	288560	AJ005023	39743	Rn.116792
Similar to hypothetical protein FLJ23263 (LOC290811), mRNA	---	0.000003	26.88	23.42	1.15	0.00217020	43.57	34.59	1.26	0.02407338	1378907_a_at	*D16Rat53*	16	64892156	16	66606445	-1714287	BG380154	12075	Rn.12067
Similar to RIKEN cDNA 2010012D11 (LOC361239), mRNA	---	0.000003	627.82	491.38	1.28	0.05708830	201.20	186.48	1.62	0.00864565	1388617_at	*D17Rat143*	17	37146290	17	37974670	-828290	AW914746	22060	Rn.22050
sulfotransferase family, cytosolic, 1C, member 2	*Sult1c2*	0.000003	5499.93	3833.57	1.43	0.00279093	26.37	27.43	1.04	0.56200081	1370943_at	*D9Rat88*	9	1040728	9	6330236	-5289508	AI168983	22471	Rn.108183

Transcribed sequence with strong similarity to protein sp P00722 (E. coli) BGAL_ECOLI Beta-galactosidase	---	0.000003	39.12	55.23	1.41	0.00515731	55.82	52.10	1.02	0.81986977	1374123_at	D13Cebr5s3	13	47959731	13	45617080	2442651	BM384680	25048	Rn.25048
Transcribed sequences	---	0.000003	604.85	649.69	1.11	0.24860432	397.79	516.67	1.30	0.00346151	1372814_at	D13Rat131	13	80939524	13	90769515	169605	BF283084	8424	Rn.8424
Transcribed sequences	---	0.000003	46.47	19.12	2.43	0.00145381	30.36	18.93	1.60	0.01845817	1376840_at	D12Cebrp97s4	12	9698220	12	5552993	4145227	AI103040	22444	Rn.22444
Transcribed sequences	---	0.000003	18.22	18.64	1.02	0.15352467	14.29	15.18	1.06	0.17482872	1377181_at	D14Rat13	14	33578048	14	34942157	1364109	AI502604	27825	Rn.27825
Transferrin	Tf	0.000003	156.79	122.51	1.27	0.00568514	1099.73	1225.72	1.12	0.51001038	1370229_at	D8Rat123	8	108250583	8	108651868	3301285	AA945178	2514	Rn.91296
acyl-CoA dihydroxyacetonephosphate acyltransferase	Gnpat	0.000005	49.59	59.35	1.20	0.01316417	146.59	200.73	1.78	0.00350172	1368991_at	D19Cebrp150s1	19	59051904	19	53668812	1482092	AI014001	1739	Rn.1739
SAC2 (suppressor of actin mutations 2, homolog)-like (S. cerevisiae)	Vps52	0.000005	271.26	179.60	1.51	0.00261584	252.84	176.31	1.43	0.00521172	1369967_at	D20Rat11	20	5078591	20	4741042	337549	NM_033097	21933	Rn.51354
Similar to DEAD (Asp-Glu-Ala-Asp) box polypeptide 51; Dead box protein 73D-like (LOC304570), mRNA	---	0.000005	151.39	181.13	1.19	0.01837126	221.22	182.80	1.26	0.03073809	1373682_at	D12Rat22	12	46020509	12	45907504	113005	BI290764	15021	Rn.15021
Similar to exostoses (multiple) 2 (LOC311215), mRNA	---	0.000005	160.03	190.96	1.19	0.02476723	339.48	349.42	1.03	0.72541193	1371611_at	D3Rat29	3	78011667	3	80223021	4211354	BM387363	2878	Rn.2878
Similar to zinc finger protein 365 (LOC293044), mRNA	---	0.000005	129.83	229.11	1.76	0.00356380	86.22	137.18	1.59	0.00087685	1369650_at	D1Rat42	1	135946190	1	135496747	449443	BI295153	43682	Rn.43682
Transcribed sequences	---	0.000005	460.26	280.87	1.64	0.00074253	36.17	38.70	1.07	0.44101149	1377155_at	Spi	4	120294031	4	119611925	682106	BI284805	18519	Rn.18519
cleavage and polyadenylation specific factor 4, 30kD subunit	Cpsf4	0.000005	221.85	289.11	1.30	0.00409694	521.21	578.65	1.11	0.28304541	1386878_at	D12Cebrp97s4	12	9781083	12	5552993	4228090	Y17306	22470	Rn.104788
RT1 class II, locus Ba	RT1.Ba	0.000006	741.85	66.52	11.32	0.00000487	397.98	58.17	6.84	0.00051633	1377334_at	D20Rat41	20	4702351	20	4741042	38691	BG378249	39138	Rn.25717
Transcribed sequence with weak similarity to protein pir T02345 (H.sapiens) T02345 hypothetical protein KIAA0324 - human	---	0.000008	316.11	266.84	1.19	0.04003548	302.23	251.22	1.20	0.01007597	1374675_at	D1Rat256	1	167465427	1	173621619	6156192	BG374902	8151	Rn.8151
mRNA decapping enzyme	Dcps*	0.000009	101.09	160.31	1.59	0.00035506	148.96	292.49	2.00	0.00003877	1390185_at	D8Rat213	8	3491340	8	38221201	3307791	BE109709	31813	Rn.101986
postsynaptic protein Cript	Cript	0.000009	227.53	299.31	1.32	0.00124067	226.31	299.58	1.33	0.00032007	1387222_at	D6Cebrp32s2	6	104799027	6	15163789	4984761	NM_019907	30508	Rn.30508
Similar to Cytosolic nonspecific dipeptidase (Glutamate carboxypeptidase-like protein 1) (LOC291394), mRNA	---	0.000009	2314.66	2341.34	1.01	0.88946244	803.37	830.72	1.03	0.58442858	1372132_at	D18Rat5	18	81200287	18	79278807	1764480	BI297721	9814	Rn.98700
Similar to Hypothetical protein MGC38513 (LOC292781), mRNA	---	0.000009	161.89	194.03	1.20	0.00010902	57.68	60.05	1.04	0.47333999	1374249_at	D1Rat212	1	85342309	1	96188940	1246371	AI169962	2550	Rn.2550
Transcribed sequences	---	0.000009	58.14	76.25	1.13	0.21503154	58.66	101.36	1.73	0.00024936	1390364_at	D11Rat11	11	82782152	11	83736689	960537	AI044310	20637	Rn.20637
v-ros UR2 sarcoma virus oncogene homolog 1 (avian)	Ros1	0.000009	39.05	49.03	1.29	0.00160283	48.85	42.74	1.14	0.04587307	1368432_a_at	D20Rat9	20	30846951	20	31944899	1099247	NM_012074	10705	Rn.87416

Similar to hypertension-associated protein SA - rat (LOC309036), mRNA	---	0.00001	440.78	40.03	8.99	0.00000001	15.81	15.56	1.02	0.69730337	1383303_at	Sa	1	178220794	1	178167709	53085	BI282211	28004	Rn.88644
Similar to tetracycline transporter-like protein (LOC305449), mRNA	---	0.00001	48.39	41.13	1.18	0.00825318	110.29	71.12	1.55	0.02261117	1382667_at	*D14Rat37*	14	81749981	14	81450719	299262	AA956425	3100	Rn.3100
Similar to trinucleotide repeat containing 6; EDXE, GW182 autoantigen (LOC308971), mRNA	---	0.00001	113.97	47.82	2.38	0.00002418	278.88	112.49	2.48	0.00146996	1375664_at	*D1Rat287*	1	192181778	1	190264130	-8072352	BI294732	2857	Rn.6839
Transcribed sequence with weak similarity to protein pir T48275 (H.sapiens) T48275 hypothetical protein DKFZp564L0578.1 - human	---	0.00001	178.29	196.54	1.10	0.22455793	179.05	200.94	1.12	0.18222885	1389005_at	*Apeh*	8	113872092	8	113376046	496046	BI282843	9337	Rn.41503
casein kinase II, alpha 1 polypeptide	*Csnk2a1*	1.251E-05	35.10	86.88	2.48	0.00007641	72.30	299.37	4.02	0.00209402	1389974_at	*D3Mit2*	3	142518200	3	143120242	602042	BF565171	4231	Rn.4231
benzodiazepin receptor	*Bzrp*	1.313E-05	582.30	400.88	1.45	0.00025608	4488.14	4087.86	1.10	0.49590019	1370249_at	*D7Rat129*	7	121632443	7	119196750	2436693	AI008680	1920	Rn.1920
zinc finger protein HIT-4	*Hit4*	1.376E-05	82.48	40.60	2.03	0.00023549	67.58	47.78	1.42	0.00241099	1370705_at	*D10Utr3*	10	11918556	10	10129010	1789546	AF277900	64641	Rn.64641
Similar to helicase-like protein NHL isoform 2 (LOC362288), mRNA	---	1.414E-05	189.38	98.13	2.03	0.00084203	241.93	131.63	1.84	0.00064510	1374921_at	*D3Rat1*	3	170402696	3	169922504	480192	BI274548	19882	Rn.98315
legumain	*Lgmn*	1.427E-05	2651.71	1788.52	1.48	0.00585948	387.96	398.96	1.03	0.64678350	1368430_at	*D6Rat111*	6	126585634	6	126025080	60554	AF154349	44616	Rn.44616
Transcribed sequence with moderate similarity to protein sp P50539 (H.sapiens) MXI1_HUMAN MAX INTERACTING PROTEIN 1	---	1.437E-05	756.28	540.52	1.47	0.00031922	492.16	438.17	1.12	0.34913619	1372093_at	*D1Rat225*	1	259801039	1	261391328	1590289	AI409308	1297	Rn.1297
Transcribed sequence with weak similarity to protein sp Q16600 (H.sapiens) Z239_HUMAN Zinc finger protein 239	---	1.441E-05	34.75	30.87	1.13	0.06963991	32.78	29.48	1.11	0.04126769	1377501_at	*D8Rat58*	8	19082519	8	19460646	1378027	BM383764	57387	Rn.123358
Similar to zinc finger protein 95 (LOC304275), mRNA	---	1.465E-05	80.76	98.50	1.22	0.01879836	118.04	181.02	1.53	0.00000501	1377244_at	*D12Cebvp97s4*	12	9727056	12	5552993	4174963	BM390979	38399	Rn.38399
RT1 class Ib, locus *Aw2*	*RT1-Aw2*[1]	1.489E-05	27.66	148.72	5.40	0.00000029	23.15	387.51	16.74	0.00000001	1370428_x_at	*D20Rat41*	20	5023662	20	4741042	282620	AJ249701	35743	Rn.107960
cytochrome P450 4F5	*Cyp4f5*	1.678E-05	28.81	31.83	1.11	0.04342743	41.85	78.91	1.89	0.00043189	1392720_at	*D7Mit17*	7	13111312	7	13075668	35644	BG376949	43212	Rn.10171
Similar to mitochondrial ribosomal protein L52 CG1577-PA (LOC361037), mRNA	---	1.846E-05	524.14	301.25	1.74	0.00011494	526.49	302.10	1.74	0.00011563	1388654_at	*D15Rat6*	15	32507804	15	32664408	-146604	AI409180	6660	Rn.6660
Transcribed sequences	---	1.905E-05	259.56	214.51	1.21	0.05588801	103.26	114.73	1.11	0.07280066	1389744_at	*D3Rat12*	3	128518971	3	127894069	624902	AW527194	43161	Rn.43161
prion protein	*Prnp*	1.964E-05	646.19	325.44	1.68	0.00710709	1301.67	1291.69	1.01	0.90980331	1370156_at	*D3Rat159*	3	119090036	3	119342992	-247044	BI218802	4936	Rn.3936
Similar to RIKEN cDNA 2410002K23 (LOC287069), mRNA	---	1.978E-05	591.08	516.76	1.14	0.00480663	969.00	742.18	1.35	0.01698874	1374467_at	*D10Utr3*	10	11759834	10	10120010	1639824	BM389084	508	Rn.99648
Transcribed sequence with moderate similarity to protein pir T17245 (H.sapiens) T17245 hypothetical protein DKFZp586J0917.1 - human	---	1.993E-05	163.11	117.78	1.39	0.00024584	255.70	216.21	1.23	0.24624120	1374087_at	*D7Rat3*	7	136577153	7	137480820	903007	AI411088	6748	Rn.6748
RT1 class II, locus Bb	*RT1-Bb*	2.048E-05	32.39	493.42	15.23	0.00000001	41.24	511.37	12.74	0.00001023	1370033_at	*D20Rat11*	20	4733506	20	4741042	-7536	AI411203	29380	Rn.29380

dynamin 1	*Dnm1*	2.348E-05	178.40	482.46	2.70	0.00004810	651.02	736.01	1.13	0.19312057	1368292_at	*D3Ceb/26s1*	3	11352965	3	9359072	1993013	NM_080689	10354	Rn.10354
Similar to RIKEN cDNA 6330575F11 (LOC309041), mRNA	---	2.415E-05	682.19	250.26	2.46	0.00008228	725.47	523.97	1.38	0.00263280	1388567_at	*Se*	1	178136754	1	178167709	-30955	BF282947	12345	Rn.12345
Similar to Galactose-1-phosphate uridylyltransferase (Gal-1-P uridylyltransferase) (UDP-glucose--hexose-1-phosphate	---	2.419E-05	484.85	440.33	1.10	0.02088676	481.23	436.32	1.10	0.07779842	1371251_at	*D5Mit1*	5	59187244	5	58223366	963878	L05541	2632	Rn.102781
Transcribed sequences	---	2.455E-05	21.61	20.18	1.03	0.48110306	17.88	19.53	1.09	0.04633230	1384309_at	*D9Rat156*	9	61941251	9	61488829	452422	BF393536	7284	Rn.7284
RT1 class I, A3	*RT1.A3*	2.655E-05	87.59	33.10	2.65	0.00285349	440.17	69.58	6.33	0.00001879	1369110_x_at	*D20Rat41*	20	5012218	20	4741042	271174	NM_012645	39743	---
similar to mannosidase 2, alpha B1	*MGC7256 1*	2.673E-05	906.13	871.01	1.04	0.43012020	730.49	685.54	1.07	0.41695485	1388699_at	*D19Rat14*	19	24716647	19	25697934	-981287	BE114558	1034	Rn.102540
propionyl-coenzyme A carboxylase, alpha polypeptide	*Pcca*	2.697E-05	727.86	603.15	1.20	0.11714935	950.07	797.58	1.12	0.34988305	1371963_at	*D15Rat29*	15	108037783	15	108403192	-365409	BF395042	6033	Rn.6033
SAC2 (suppressor of actin mutations 2 homolog)-like (S. cerevisiae)	*Vps52*	2.817E-05	113.28	74.42	1.52	0.00265457	185.12	120.75	1.53	0.00950337	1394386_s_at	*D20Rat41*	20	5074917	20	4741042	333875	BE107956	51354	Rn.51354
interleukin 18	*Il18*	2.821E-05	75.48	56.63	1.33	0.00104946	48.95	37.89	1.29	0.03016847	1369665_a_at	*D8Mit12*	8	53949356	8	58049918	-4100562	AJ222813	11118	Rn.11118
mRNA decapping enzyme	*Dcps"*	2.916E-05	196.07	96.88	2.03	0.00053527	269.03	129.90	2.33	0.00008243	1374583_at	*D8Rat219*	8	34908365	8	38221201	-3312836	BI288131	7059	Rn.101996
Similar to Tumor necrosis factor ligand superfamily member 12 (TNF-related weak inducer of apoptosis) (TWEAK) (LOC360548), mRNA	---	2.989E-05	150.63	178.05	1.18	0.07919433	332.16	348.89	1.05	0.54508563	1372503_at	*Abpa*	10	56517571	10	56447685	69886	BF283668	3211	Rn.3211
growth hormone receptor	*Ghr*	3.046E-05	178.81	174.04	1.03	0.82814445	718.20	819.05	1.14	0.42485363	1373803_a_at	*D2Rat201*	2	52434108	2	49619880	2814228	AI170771	2178	Rn.2178
Transcribed sequence with moderate similarity to protein ref NP_112197.1 (H.sapiens) TSC-22-like [Homo sapiens]	---	3.39E-05	100.10	283.97	2.94	0.00040676	131.24	416.03	3.17	0.00190658	1375958_at	*D12Mit7*	12	19803204	12	17763275	2039929	BF412297	20818	Rn.20818
carbonyl reductase 1	*Cbr1*	3.408E-05	140.77	161.45	1.15	0.04860230	976.17	1020.91	1.05	0.76828122	1368037_at	*D11Cebr15s1*	11	33863027	11	32794012	1069015	NM_019170	3425	Rn.3425
Similar to Gamma-soluble NSF attachment protein (SNAP-gamma) (N-ethylmaleimide-sensitive factor attachment protein, gamma)	---	3.429E-05	232.80	257.62	1.11	0.43305375	301.99	336.39	1.11	0.02725646	1374075_at	*D18Rat41*	18	59210406	18	59352487	-142081	AI175533	24090	Rn.24090
Similar to KIAA1731 protein (LOC363018), mRNA	---	3.54E-05	52.29	47.04	1.22	0.00113026	75.50	54.53	1.38	0.00013353	1380286_x_at	*D8Rat56*	8	12117542	8	8295034	3822508	BM386844	36574	Rn.86809
diacylglycerol O-acyltransferase homolog 2 (mouse)	*Dgat2*	3.548E-05	411.99	268.04	1.54	0.00404814	7725.07	7117.20	1.09	0.56868415	1371610_at	*D1Rat47*	1	156526202	1	156590248	-2063956	BI279069	9523	Rn.9523
Similar to CHMP1.5 protein (LOC364890), mRNA	---	3.56E-05	217.13	153.34	1.19	0.05511491	331.04	261.68	1.27	0.01789579	1390128_at	*D18Rat89*	18	63788794	18	63770207	18587	BF567618	39132	Rn.39132
Similar to Microfibril-associated glycoprotein 3 precursor (LOC287299), mRNA	---	3.560E-05	99.12	63.56	1.62	0.00267406	131.75	137.28	1.04	0.58254606	1390063_at	*D16Rat165*	16	43132115	16	43937645	-805429	BG380473	43071	Rn.43071
apolipoprotein M	*Apom*	3.591E-05	1489.72	1395.41	1.07	0.34505417	29.35	31.20	1.06	0.18002057	1386980_at	*D20Rat41*	20	4756132	20	4741042	-984910	NM_019373	262	Rn.262

Transcribed sequence with strong similarity to protein ref NP_006338.1 (H.sapiens) peptidyl prolyl isomerase H	---	4.65E-05	87.33	108.96	1.25	0.05487700	197.81	215.72	1.09	0.30583122	1374840_at	*D9Mit14*	5	139893357	5	138300648	1583709	BG377716	41174	Rn.41174
Similar to KIAA1731 protein (LOC363018), mRNA	---	4.66E-05	33.45	27.58	1.21	0.03965532	76.15	45.61	1.67	0.00000316	1398460_at	*D9Rat56*	8	12117388	8	8295034	3822304	AI146147	36574	Rn.86809
proviral integration site 1	*Pmt*	4.83E-05	93.75	71.07	1.32	0.02304820	620.18	176.75	3.51	0.03067544	1374429_at	*D20Mgh5*	20	7822798	20	9302208	-1479410	AI596401	7368	Rn.34868
LOC361111 (LOC361111), mRNA	---	5.19E-05	89.89	61.68	1.46	0.00219000	217.47	134.02	1.62	0.01118083	1390435_at	*D16Rat87*	16	6555704	16	1732626	4823078	AA996717	12381	Rn.12381
Similar to RIKEN cDNA 4832420O05 (LOC292085), mRNA	---	5.33E-05	21.44	24.30	1.13	0.02767916	49.33	56.94	1.15	0.24447446	1375071_at	*D19Cebrp150s1*	19	54096490	19	53569812	526678	BE119719	58778	Rn.91990
Transcribed sequences	---	0.000054	267.84	361.48	1.35	0.00789016	193.27	213.17	1.10	0.04058250	1377076_at	*Kcnj1*	8	38897933	8	32171503	6726430	AI716131	24258	Rn.24258
CUG triplet repeat,RNA-binding protein 2	*Cugbp2*	5.45E-05	185.41	125.84	1.47	0.02304054	781.90	1336.04	1.71	0.01450467	1389867_at	*D17Rat62*	17	83039499	17	83490772	-451273	BI281086	37864	Rn.30004
Transcribed sequence with strong similarity to protein sp Q9UJY5 (H.sapiens) GGA1_HUMAN ADP-ribosylation factor binding protein	---	5.61E-05	66.63	39.24	1.70	0.00005576	64.56	42.31	1.53	0.00563527	1374915_at	*D7Cebr77s1*	7	116891735	7	115036128	1855607	BE113321	45356	Rn.45356
I-branching beta 1,6-acetylglucosaminyltransferase	*Ignt*	5.99E-05	235.73	411.60	1.75	0.00125613	196.30	286.57	1.46	0.02174977	1374903_at	*D17Rat11*	17	29770683	17	30932396	1161653	AI234819	8807	Rn.8807
Transcribed sequence with strong similarity to protein ref NP_062298.1 (M.musculus) glucosamine-phosphate N-acetyltransferase 1, glucosamine-6-	---	5.7E-05	402.01	356.78	1.39	0.09803912	400.27	262.02	1.53	0.00607570	1376530_at	*D15Mit3*	15	21272629	15	19348150	1924479	BG374612	14702	Rn.14702
CDK104 mRNA	---	5.88E-05	719.36	916.44	1.27	0.00445364	308.45	373.04	1.21	0.00135406	1372064_at	*Abpa*	10	5731013	10	6447685	982428	BI295385	7070	Rn.7070
nuclear receptor coactivator 3	*Ncoa3*	6.06E-05	51.21	49.11	1.25	0.06348681	110.53	101.59	1.14	0.09770574	1388178_at	*D3Rat143*	3	157088946	3	152806528	4283418	AF322224	46924	Rn.20681
Transcribed sequences	---	6.09E-05	30.79	29.63	1.04	0.35232493	43.20	48.17	1.12	0.22762239	1379594_at	*D1Rat186*	1	16444480	1	15833536	610974	AW524408	42056	Rn.63081
Transcribed sequence with moderate similarity to protein ref NP_085199.1 (H.sapiens) hypothetical protein dJ122O8.2 [Homo sapiens]	---	6.81E-05	157.32	207.91	1.32	0.09149836	283.00	353.15	1.25	0.04908292	1373440_at	*D5Rat6*	5	49086196	5	52531032	-3435896	AI175525	22183	Rn.22183
protein phosphatase 1A, magnesium dependent, alpha isoform	*Ppm1a*	6.91E-05	151.59	115.91	1.31	0.10827804	467.04	440.41	1.06	0.60345204	1368959_at	*D6Rat165*	6	95197179	6	93822650	1374529	NM_017038	37403	Rn.37403
Transcribed sequences	---	7.01E-05	135.57	104.57	1.29	0.01712300	222.05	194.06	1.14	0.14683687	1373537_at	*D3Cebr26s1*	3	9893732	3	9359012	620760	BE113175	20445	Rn.20445
Similar to 2210021J22Rik protein (LOC362979), mRNA		7.10E-05	99.87	102.52	1.14	0.07250052	240.42	285.31	1.14	0.09238399	1371749_at	*D7Rat129*	7	123848101	7	119195750	4652351	AI169439	3452	Rn.3452
protein kinase inhibitor, alpha	*Pkia*	7.14E-05	54.44	47.30	1.15	0.04125526	242.27	188.73	1.31	0.06177538	1373082_at	*D2Rat24*	2	96467284	2	100673233	-4205949	AA893743	8002	Rn.94808
RT1 class II, locus Ob1	*RT1.Ob1*	7.26E-05	14.97	16.29	1.09	0.31001169	11.15	12.14	1.09	0.01247404	1370382_at	*D20Rat41*	20	4676471	20	4741042	64571	BI276526	33311	Rn.33311
nuclear transcription factor-Y gamma	*Nfyc*	7.31E-05	433.29	612.34	1.41	0.00044836	776.85	783.96	1.01	0.84166255	1398268_at	*D5Rat169*	5	141383735	5	141457807	-74072	BI296061	1457	Rn.1457

Similar to sulfide quinone reductase-like (LOC317630), mRNA	---	7.317E-05	261.31	367.05	1.40	0.00046437	408.50	738.04	1.81	0.00002181	1389142_at	*D3Rat257*	3	109655900	3	115920214	-6273314	AI013361	1416	Rn.1416
Rsec5 protein	*Sec5l1*	7.483E-05	129.70	116.56	1.11	0.19183859	206.05	156.55	1.32	0.00019655	1368225_at	*D17Rat144*	17	40041127	17	37974570	2066557	NM_134414	2869	Rn.2869
myosin, heavy polypeptide 9	*Myh9*	7.801E-05	1186.44	1051.71	1.13	0.11191106	1618.36	1296.20	1.25	0.03033091	1371725_at	*D7Mit15*	7	115714834	7	119064158	-3369324	BM392410	7812	Rn.11365
Transcribed sequence with weak similarity to protein ref:NP_060531.1 (H.sapiens) hypothetical protein FLJ10330 [Homo sapiens]	---	7.657E-05	368.67	284.57	1.30	0.19739696	22.44	21.24	1.06	0.20682616	1394229_at	*D10Mgh6*	10	67274417	10	67691615	-417198	AA893147	13478	Rn.13478
Similar to HN1 (LOC287528), mRNA	---	7.879E-05	313.28	244.42	1.28	0.00484266	453.40	485.85	1.07	0.20505259	1388581_at	*D10Cebr29s2*	10	105594656	10	106325464	-730808	AI317868	1650	Rn.104570
Transcribed sequence with strong similarity to protein pir:JC2473 (H.sapiens) JC2473 doc2 protein - human	---	8.271E-05	76.95	61.10	1.26	0.11520136	125.15	102.36	1.22	0.05468901	1377614_at	*Mll1pa*	1	186181355	1	185840840	340515	BI283147	3179	Rn.94076
Hypothetical LOC315547 (LOC315547), mRNA	---	8.333E-05	86.49	125.04	1.45	0.00202828	220.38	218.08	1.01	0.91570564	1376921_at	*D8Rat219*	8	34993723	8	38221201	-3227478	BG381722	42737	Rn.42737
Similar to RIKEN cDNA C730048E16 (LOC362957), mRNA	---	9.524E-05	111.15	96.94	1.15	0.04452594	84.57	62.49	1.35	0.00110563	1389264_at	*D7Rat129*	7	117060690	7	119195750	-2135060	BI278594	28758	Rn.28758
---	---	9.677E-05	20.36	21.62	1.06	0.15280396	16.67	17.51	1.05	0.44132390	1386362_at	*D16Rat87*	16	5893024	16	1732626	4160398	AA892204	---	---
Similar to hypothetical protein FLJ23263 (LOC290811), mRNA	---	9.756E-05	24.98	22.16	1.13	0.00527685	23.70	20.69	1.15	0.00021806	1376856_at	*D16Rat53*	16	64895811	16	66606445	-1710534	AI111579	12987	Rn.12987
hypoxia up-regulated 1	*Hyou1*	0.0001	158.30	280.68	1.77	0.00155671	251.05	391.96	1.56	0.02472409	1371442_at	*D8Mit4*	8	47367303	8	48815473	-1548170	BI282904	983	---

Footnotes

Each row in the table contains the eQTL description (gene name, gene symbol), *F*-value of the peak of linkage from eQTL Reaper, mean expression values, fold change and t-test *p*-values in the parental strains for fat and kidney probeset and genetic marker at the peak of linkage, chromosome and physical position of the probeset and of the marker at the peak of linkage, distance between probeset and marker at the peak of linkage, Public ID relating to the probeset as assigned by Affymetrix, Archival UniGene Cluster number relating to the probeset as assigned by Affymetrix, UniGene ID relating to the probeset as assigned by Affymetrix)

[a]Three genes with multiple linkages (*Cd36*, *Dcps*, and *RT1Aw2*) these are to two or more distinct Affymetrix probesets representing the same gene

[b]Mean parental strains expression values are calculated from antilog RMA values

[c]Fold change in the parental strains was calculated such that the overexpressed strain was divided by the underexpressed strain

Table 3.9: Significance of probe set sequence variants

Probe set	probe set t-test *P*-value	Probe	Sequence variant detected	Position within Probe (1-25)	Probe t-test *P*-value PM
1374713_at	0.691367355	1	Yes	18	0.0420566
		2			0.5064360
		3	Yes	15	0.7712621
		4			0.8589285
		5			0.3766888
		6			0.4809685
		7			0.2768582
		8			0.2441476
		9			0.5662843
		10			0.5436316
		11			0.8482066
1376091_at	0.932556116	1			0.2682925
		2			0.3523577
		3			0.5391923
		4			0.4327482
		5			0.4491701
		6	Yes	4	0.0051597
		7			0.3432729
		8			0.1690251
		9			0.0070072
		10			0.3591597
		11			0.4008780
1376732_at	0.015961505	1			0.1910972
		2			0.0249391
		3			0.0078297
		4	Yes	25	0.0023640
		5			0.1324638
		6			0.0130989
		7			0.0597369
		8			0.3397299
		9			0.0344075
		10			0.2615884
		11			0.3428103
1388909_at	0.038332063	1	Yes	11	0.0005084
		2			0.4499054
		3			0.7102861
		4			0.6782551
		5			0.8491689
		6			0.6498156
		7			0.4009926
		8			0.9832181
		9	Yes	10	0.0018778
		10			0.3319991
		11			0.5008514
1389575_at	0.61064634	1			0.1766404
		2			0.3697129
		3			0.2864417
		4			0.4590462
		5			0.2975048
		6	Yes	12	0.0003624
		7			0.3943435
		8			0.1946102
		9			0.2600508
		10			0.3310310
		11			0.3930660
1398387_at	0.073906424	1			0.1344671
		2			0.1524257
		3			0.1655052
		4			0.0955107
		5			0.0886343
		6	Yes	3	0.0604937
		7			0.1410897
		8			0.1477562
		9			0.1822567
		10			0.2140814
		11			0.1478108

Affymetrix probes are designed which could account for the differential hybridisation. **Table 3.9** shows a t- test between the 5 parental strains in fat tissue for the probe set and each (PM) probe from the raw *.cel file expression data. This data was extracted from the raw MAS5.0 files for the entire probe set and using DChip (http://biosun1.harvard.edu/complab/dchip/) for each probe pair. A t-test with unequal variance was carried out using excel and the significantly (P <0.05) differentially expressed probe sets and pairs are highlighted in the table. This data is not normalised and therefore it is difficult to compare with the final data sets. Of the eight polymorphisms detected 6 lie in differentially expressed probe pairs (determined using the t-test) and in three of the probe sets these are the only differentially expressed probe pairs (1374713_at, 1388909_at and 1389575_at). In these three (from 20) cases this differentially hybridisation may account for the differentially expression detected. None of the 6 genes in **Table 3.9** are included in the data presented as the highly significant *cis*-acting eQTL genes in **Tables 3.7** and **3.8**

3.3 Discussion

The studies described above combined genome wide linkage analysis and global gene expression profiling in fat and kidney tissue both of which are important in the pathophysiology of the metabolic syndrome. The BXH/HXB RI strain panel was used to map the genetic determinants of gene expression in the 15,923 genes represented on the RAE230A Affymetrix GeneChip. The fat data set was generated by this student in collaboration with Dr Caroline Wallace of the Physiological Genomics and Medicine group while the kidney data set which was used for comparisons and future work was generated by Norbert Hubner's group at the MDC in Berlin.

Prior to carrying out any RNA extraction, cRNA synthesis or microarray hybridisations RNA extraction methods were compared using samples from the age matched animals and subsequently it was decided to use the Trizol® protocol to extract RNA. The Agilent Bioanalyser was tested and compared to the more commonly used RNA formaldehyde gel to determine if it could be used for assessing the quality of both RNA and cRNA because it requires 10 fold less sample for analysis of RNA quality. **Figure 3.1** shows the results of this comparison and since the results from the formaldehyde gel were similar the Bioanalyser was used for assessment of all samples in the fat tissue data set.

Two different normalisation methods were carried out and compared using the fat data set and the results highlight the importance of normalisation procedure when utilising microarray data (**Figures 3.3, 3.4, Table 3.2, 3.3**). RMA and median scaling normalisation followed by a t-test were applied to a preliminary data set of 6 microarrays, 3 from each of the parental strains. The numbers of probe sets which were determined to be significantly differentially expressed were similar when either normalisation method was used (approximately 1,100 with a $p<0.05$), however only 506 of these probe sets were in common in both data sets. This result gives confidence that, even though a significant false positive rate may be expected, the large majority of these shared probe sets will be truly differentially expressed while also highlighting the difference between data transformation methods. This analysis was repeated on the entire data set when completed and similar results were seen, that is while both methods gave rise to similar numbers of differentially expressed genes the composition of the gene lists was different. The normalisation comparison also

allowed the FDR and reproducibility to be tested (**Figure 3.5**) and RMA had a consistently lower FDR then MAS5.0 at each threshold of significance. Considering these factors the RMA normalised data was used in the linkage analysis. The importance of considering data extraction and normalisation and the effect it will have on subsequent data has been reported by other groups (Bolstad et al. 2003; Irizarry et al. 2003; Irizarry et al. 2005).

Map Manager QTX was used to define some preliminary linkages by first selecting genes which were differentially expressed between the parental strains either by fold change or detection by t-test in the preliminaray data set of 6 parental microarrays normalised with RMA. This filtering was necessary because the Map Manager program can only be used to link one expression profile to the genome at a time and therefore cannot be used to define linkage for all 15,923 genes on the RAE230A chip due to the impracticality. The ten probe sets with the most significant *P*-value and all the genes with 4 fold or greater differential expression (n=39) were selected for linkage analysis using Map Manager QTX, with 2 occurring in both lists this gave a total of 47 (**Table 3.4**). All had linkages to at least one marker with $P<0.05$ and the most significant linkages for each probe set were mapped onto the genome. The distribution of the most significant linkage for each probe set (**Figures 3.6, 3.9**) is widespread across the genome however there were two clusters, one on chromosome 20 at the Tumour necrosis factor α (Tnfα) locus, and another on chromosome 15 at marker D15Rat107. Tnfα has been shown known to be causally related to the hypertensive phenotype in SHR and is associated with the pathogenesis of obesity and obesity-associated hypertension in humans (Pausova et al. 2003). The 6 probe sets that linked to D15Rat107 are ESTs and the region of the genome surrounding

D15Rat107 was not well annotated on ensembl at the time but seemed to contain mainly novel genes. These two clusters were not studied further when the complete data set was available since the primary focus of this students' work is in detecting and analysing highly significant *cis*-acting eQTL; also the chromosome 15 clusters did not occur in the final data set (**Figure 3.7**).

Interestingly of these 47 genes, all probe sets which have a linkage with a highly significant *P*-value ($P<0.00001$) and that had been mapped were intra-chromosomal, indicating that the most significant linkages are likely to be to *cis*-acting eQTL. These linkages remained as *cis* when the *cis/trans* definition was refined so that a *cis* eQTL was defined as the most significantly linked marker within 10Mb of the probe set. This supports the findings of Schadt and Brem (Brem et al. 2002; Schadt et al. 2003) who also found *cis* acting eQTL had higher LOD scores than *trans* acting modifiers of gene expression. Many of these initial linkages were not detected in the final data set which reflects the fact that in this preliminary analysis the genes had to be filtered on parental strain differential expression before linkage analysis could be carried out and at that time only 3 parental microarrays were available which meant that no outlier detection could be carried out. Outlier detection was carried out on the final data set (**Methods 2.9.1.2**) as it resulted in reduction in FDR (Hubner et al. 2005).

Since Map Manager QTX can only perform one linkage at a time a more automated approach was used to carry out the linkages of all 15,923 expression profiles to all 1011 markers. In collaboration with Dr Enrico Petretto QTL Reaper analysis was carried out to detect linkages for each probe set on the microarray. This program can be considered as a modified and extended version of Map Manager QTX which is

able to process multiple quantitative traits simultaneously and has an inbuilt permutation test to account for multiple comparisons which are being made to multiple markers.

The QTL Reaper analysis was carried out on both the fat and kidney RMA normalised data sets to identify linkages for each transcript across the genome. These linkages were then collapsed into loci to define eQTL. This redundancy removal step eliminated multiple linkages of probe sets to tightly linked markers and defined 2,118 linkages in fat and 2,490 in kidney at genome wide significance ($P<0.05$) of which 311 were shared between the two tissues (**Table 3.5**). eQTL shared between tissues are likely to reflect common regulatory mechanisms and therefore represent genes which may be involved in the metabolic syndrome at the whole body level. These non redundant eQTL were then classified as being *cis* or *trans* where a gene is operationally defined as controlled in *cis* if the most significantly linked marker maps to 10Mb up or downstream of the probe set. A gene controlled in *trans* was defined as having the most significantly linked marker on another chromosome or more than 10Mb from the probe set. **Table 3.5** shows the numbers of eQTL linked in *cis* and *trans* at increasing levels of significance, this demonstrates that more significant eQTL are *cis*-regulated. This again confirms the finding of Schadt and Brem who found *cis* eQTL to have higher LOD scores than trans eQTL and may be expected since a first order effect is more likely to be under monogenic control and therefore be more easily detected.

The plots in **Figures 3.7 and 3.8** show the *cis* (at $P<10^{-4}$) and *trans* (at $P<10^{-2}$) eQTL mapped onto the genome alongside all metabolic and CVD SHR pQTL. The different

levels of significance were chosen to reflect the fact that *cis* acting eQTL tend to be more significantly linked and act as candidate genes underlying the SHR pQTL and illustrates the distribution of eQTL across the genome. This will be discussed further in **Chapter 4**. *trans* eQTL are generally less significant because they represent the oligo or polygenic effects of other genes, meaning that since multiple genes are influencing expression of the probe set the linkage will not be the same across all strains, while a gene controlled in *cis* is effectively being controlled as a monogenic trait. The *trans* eQTL genes can be used to build up networks and pathways of gene expression however this is outwith the scope of this PhD.

Finally probe set 1386901_at which represents the Cd36 locus was detected as differentially expressed and highly significantly linked in both the preliminary analysis and the final QTL Reaper analysis. The *Cd36* gene was previously identified in the Aitman group (Aitman et al. 1999) as an insulin resistance gene in SHR and the underlying mutation is due to a duplication at the Cd36 locus which, in normal rat strains, (e.g. BN) results in a single transcribed gene and two pseudogenes. In the SHR strain there is a genomic deletion that creates a chimeric gene at the SHR Cd36 locus (Glazier et al. 2002) which has only one copy of the gene and this copy has the 5' end of the functional Cd36 gene and the 3' end of the second pseudogene which results in sequence variation at the 3' end of the gene. The probe set (1386901_at) is within the part of the gene which is deleted in the SHR strain and is therefore not represented in the SHR genome. Therefore the strongly different hybridisation signals are expected to be due to the genomic deletion at the SHR Cd36 locus and thus differential hybridisation signals may not always only be due to differences in gene expression. Therefore sequencing of the Affymetrix probe set region was

carried out in 20 genes to determine what proportion of the genes showing differential expression between the parental strains was due to sequence variants as opposed to true differential expression. These genes were identified in a WMW test and 6 had sequence variants which altered a probe within the probe set. Three of these had polymorphisms which coincide with significant differential expression as detected by a t-test and one of these three had a SNP in another probe pair which was not differentially expressed. This indicates that polymorphisms within the probe region do not account for significant numbers of differentially expressed genes which is in accordance with other studies, including that of the group at the MDC which generated the kidney data set (Hubner et al. 2005).

3.4 Conclusions

The work in this chapter has tested and validated a number of protocols before using these to generate a large and potentially valuable data set in a well characterised animal model of complex human disease. It has also shown that this integrated genome wide linkage analysis and transcriptional profiling approach is feasible for mapping large numbers of eQTL in the rat.

Chapter 4: Generation and characterisation of a data set of differentially expressed *cis*-acting eQTL genes

4.1 Introduction

The overall aim of the work in this chapter was to select and sequence the putative proximal promoter regions of a subset of the most significant *cis*-acting eQTL genes. These genes were selected from the set of *cis*-acting eQTL genes reported in **Chapter 3** and have a fold change greater than 2 in the parental strains and/or a FDR of less than 5%, that is $P<10^{-4}$ in both kidney and fat tissue. These *cis*-acting eQTL represent genes in which the differential expression of the gene is thought to be controlled by sequence variation in the gene itself. It was decided to sequence up to 2.5Kb of genomic sequence immediately upstream of the most 5' exon of these genes to detect polymorphisms; this region represents the putative promoter of the genes. These polymorphisms could account for the differential expression detected on the microarray and therefore underlie the *cis*-acting eQTL. Ultimately the polymorphisms detected can be tested in reporter assays to determine their effect on transcription.

4.1.1 Control of gene expression by *cis* elements

Gene expression can be controlled by the gene itself (*cis*) or by other regions of the genome (*trans*). *Cis*-regulatory elements which can affect transcription include the proximal promoter, enhancer and repressor elements which can be located both up and downstream and within the gene itself (e.g. in introns). There may also be alternative promoters which are available in a tissue or temporally specific manner.

Gene expression can also be affected by polymorphisms in the coding region which may lead to amino acid differences in the protein. In addition, transcript levels can be regulated by nonsense mediated decay whereby a sequence in the 3'UTR is bound by another molecule, such as a micro RNA, and this targets that transcript for destruction. Polymorphisms which cause loss or gain of sites for nonsense mediated decay would lead to differential expression on the microarray. Since the proximal promoter, that is the region immediately upstream of the transcription start site commonly contains the transcription factor binding sites (including that for RNA polymerase) which control expression of the gene, it was decided to sequence this region in the genes selected to detect polymorphisms which may be the cause of the differential expression seen on the microarray between the strains. Not all genes have a well defined promoter therefore up to 2.5Kb upstream of the most 5' exon was sequenced in order to attempt to capture the promoter.

4.2 Results

4.2.1 Selection and description of genes to be sequenced

The *cis*-acting eQTL identified in **Chapter 3** in the kidney and fat tissue data sets represent a set of candidate genes for metabolic and cardiovascular disease and some of the genes may underlie the previously mapped SHR pQTL (**Figure 3.7, Table 5.1**). Therefore it was decided to select a subset of these eQTL genes for sequencing with the aim of detecting polymorphisms which could underlie the gene expression differences detected on the microarray. These were initially selected from the *cis*-acting eQTL (full data set on data CD) and were prioritised for sequencing using the following criteria, the probe set had a FDR of less than 5% (or approximately $P<10^{-4}$)

and/or a fold change in the parental strains of 2 or more. This fold change was selected as it more likely to be reproducible using qRT-PCR or northern blot and results in a manageable number of genes to be tested. A total of 37 genes which matched these criteria were selected from the 129 kidney and 97 fat eQTL in the data set (**Table 4.1**). In some cases the eQTL was shared between the tissues but not always with the same *P*-value or fold change. These cases are highlighted in **Table 4.1** and are mainly *RT1* genes which are involved in immune function and also the *Cd36* probe set 1386901_at, which is due to polymorphisms in the probe set due to a genomic deletion (**Chapter 3**). Two probe sets, 1374583_at and 1390185_at map to the same gene mRNA decapping enzyme (*Dcps*).

The genes to which these eQTL probe sets mapped were studied using the genome databases (Uniform Resource Locators (URLs) Ensembl - http://www.ensembl.org/, NCBI - http://www.ncbi.nlm.nih.gov/Genomes/ and UCSC - http://genome.ucsc.edu/) to determine the 5'sequence of the gene and thus identify the putative promoter before any sequence analysis was undertaken. The gene structures were compared across the databases and also comparatively with the mouse and human orthologues to determine the most likely 5' exon. The sequence upstream of the most 5' exon was defined as the putative proximal promoter and primers were designed to amplify and sequence up to 2.5Kb of this region in fourteen genes. Twenty-two genes were discarded from any further analysis due to the difficulty in determining the 5' end of the gene or the gene not being well characterised and therefore it could not be sequenced. However a further four *cis*-acting eQTL genes were added to this analysis 1368291_at, 1374429_at, 1371137_at and 1371725_at because they fulfilled either the fold change or significance criteria in one of the tissues and were interesting

Table 4.1 All *cis* eQTL at P <10^{-4} and with fold change >2 in both kidney and fat tissue.

Gene Name	Gene Symbol	QTL Reaper p_value (kidney, Fat)	Mean SHR Kidney[b]	Mean BN Kidney[b]	Fold change Kidney[c]	t-test p-value in Kidney	Mean SHR Fat[b]	Mean BN Fat[b]	Fold change Fat[c]	t-test p-value in Fat	probeset
cd36 antigen	*Cd36*[a]	0.000003	66.81	140.67	2.11	0.00017586	2078.78	3885.80	1.87	0.00005691	1367689_a_at
cytochrome P450, family 2, subfamily d, polypeptide 26	*Cyp2d26*	0.000001	832.30	1845.05	2.22	0.00006997	44.02	40.28	1.09	0.23326638	1367917_at
dynamin 1	*Dnm1*	0.000023483	178.40	482.46	2.70	0.00004810	651.02	736.01	1.13	0.19312067	1368292_at
flavin-containing monooxygenase 3	*Fmo3*	0.000005	2284.94	1454.85	1.57	0.00391055	281.32	115.16	2.44	0.00009834	1368304_at
RT1 class I, A3	*RT1-A3*	0.000026549, 0.000018154	87.59	33.10	2.65	0.00285349	440.17	69.58	6.33	0.00001879	1369110_x_at
RT1 class Ib, locus Aw2	*RT1-Aw2*[a]	0.000014888, 0.000009	27.55	148.72	5.40	0.00000029	23.15	387.51	16.74	0.00000001	1370428_x_at
zinc finger protein HIT-4	*Hit4*	0.000013761	82.48	40.60	2.03	0.00023549	67.68	47.78	1.42	0.00241099	1370705_at
RT1 class II, locus Bb	*RT1-Bb*	0.000020484, 0.00002381	32.39	493.42	15.23	0.00000001	40.28	513.25	12.74	0.00001023	1371033_at
achaete-scute complex homolog-like 3 (Drosophila)	*Ascl3*	0.000001	94.96	124.08	1.31	0.00043350	47.85	98.28	2.05	0.00016444	1371085_at
RT1 class Ib, locus Aw2	*RT1-Aw2*[a]	0.000028205	79.18	49.97	1.58	0.00026344	178.08	40.69	4.38	0.00034166	1371213_at
Similar to RIKEN cDNA 1500015O10 (LOC363225), mRNA	---	0.000096	448.02	485.36	1.08	0.29834922	204.97	97.93	2.09	0.01398184	1372646_at

mRNA decapping enzyme	*Dcps*	0.000029101, 0.000028302	196.97	96.88	2.03	0.00053527	289.03	123.90	2.33	0.00008243	1374583_at
Similar to helicase-like protein NHL isoform 2 (LOC362288), mRNA	---	0.000014139	199.38	98.13	2.03	0.00084203	241.93	131.63	1.84	0.00064510	1374921_at
Transcribed sequences	---	0.000001, 0.000003	155.36	58.20	2.67	0.00002702	620.76	186.96	3.32	0.00222211	1375343_at
Transcribed sequences	---	0.000003	62.74	105.65	1.68	0.00003107	88.37	288.86	3.27	0.00009587	1375655_at
Similar to trinucleotide repeat containing 6; EDIE; GW182 autoantigen (LOC308971), mRNA	---	0.00001, 0.000009	113.97	47.82	2.38	0.00002418	278.88	112.49	2.48	0.00146996	1375664_at
similarity to protein pdb:1LBG (E. coli) B Chain B, Lactose Operon Repressor Bound To 21-Base Pair Symmetric Operator Dna, Alpha	---	0.000056338	382.01	47.28	8.08	0.00003230	299.67	72.22	4.15	0.00109714	1375676_at
Transcribed sequence with moderate similarity to protein ref:NP_112197.1 (H.sapiens) TSC-22-like [Homo sapiens]	---	0.000033898	100.10	293.97	2.94	0.00046676	131.24	416.03	3.17	0.00190658	1375958_at
Similar to 3230401M21Rik protein (LOC287157), mRNA	---	0.0000001	106.11	233.59	2.20	0.00003738	149.64	144.79	1.03	0.59655338	1376136_at
similarity to protein pir:RGECDW (E. coli) RGECDW transcription activator of D-serine dehydratase - Escherichia coli	---	0.000035599	37.96	41.75	1.10	0.15888338	42.04	133.87	3.18	0.00739684	1376398_at
Transcribed sequences	---	0.000003	46.47	19.12	2.43	0.00145391	30.36	18.93	1.60	0.01845817	1376840_at
RT1 class II, locus Ba	*RT1-Ba*	0.000006, 0.000019608	741.85	65.52	11.32	0.00000487	397.98	58.17	6.84	0.00051633	1377334_at
Similar to hypertension-associated protein SA - rat (LOC309038), mRNA	---	0.00001	440.78	49.03	8.99	0.00000001	15.81	15.56	1.02	0.69730337	1383303_at
Transcribed sequences	---	0.0000001	32.91	56.27	1.71	0.00378399	23.84	109.77	4.60	0.00000336	1384717_at

cd36 antigen	*Cd36*[a]	0.0000001, 0.000001	56.16	171.24	3.05	0.00066710	656.05	3294.06	5.02	0.00000009	1386901_at
glutathione S-transferase, mu 1	*Gstm1*	0.000003	235.28	559.30	2.38	0.00007438	452.18	397.25	1.14	0.35058399	1386985_at
dopa/tyrosine sulfotransferase	*LOC64305*	0.000001	839.97	182.59	4.60	0.00082617	34.29	27.39	1.25	0.11008161	1387314_at
interleukin enhancer binding factor 3	*Ilf3*	0.0000001, 0.0000001	138.31	52.64	2.63	0.00421879	335.58	94.46	3.55	0.00069955	1387366_at
aldehyde oxidase	*Aox1*	0.000003	130.81	93.87	1.39	0.00044379	2899.87	1046.47	2.77	0.00036238	1387376_at
solute carrier family 7 (cationic amino acid transporter, y+system), member 7	*Slc7a7*	0.000003	1545.63	1885.92	1.22	0.19882594	263.47	79.80	3.30	0.00016223	1387808_at
RT1 class Ib, locus Aw2	*RT1-Aw2*[a]	0.000048583	24.99	48.23	1.93	0.00014937	34.33	84.47	2.46	0.00145264	1388202_at
Similar to mitochondrial ribosomal protein L4 isoform a (LOC363023), mRNA	---	0.000003	306.77	591.35	1.93	0.00044739	558.72	1131.65	2.03	0.00196498	1388366_at
Similar to RIKEN cDNA 6330575P11 (LOC309041), mRNA	---	0.000024145	682.19	289.26	2.36	0.00008228	725.47	523.97	1.38	0.00283280	1388567_at
RT1 class Ib, locus Aw2	*RT1-Aw2*[a]	0.000003, 0.000008	119.56	258.82	2.16	0.00164360	222.58	527.13	2.37	0.00105266	1389734_x_at
casein kinase II, alpha 1 polypeptide	*Csnk2a1*	0.000012513	36.10	86.88	2.41	0.00007641	72.30	290.37	4.02	0.00209402	1389974_at
mRNA decapping enzyme	*Dcps*[a]	0.000013995	101.09	160.31	1.59	0.00038906	148.96	297.49	2.00	0.00003877	1390185_at
Similar to estradiol 17beta-dehydrogenase (EC 1.1.1.62), A-specific - mouse (LOC307092), mRNA	---	0.000001	291.83	1192.08	4.08	0.00019233	16.21	15.65	1.04	0.27113226	1393902_at

Footnotes

Shaded cells represent the probe sets which are significant in both fat and kidney

candidates or had already been partially investigated in the preliminary analysis discussed in **Chapter 3**. This means there were 19 differentially expressed probe sets in the final data set, however, two of these (1374583_at and 1390185_at) mapped to the same gene, an mRNA decapping enzyme, and therefore 18 genes were investigated.

The 18 genes selected for sequencing in this chapter are detailed in **Table 4.2**, this shows the probe set, the location of the eQTL, which tissue the linkage was detected in, the parental QTL Reaper *P*-value and the gene description used to identify the promoter to be sequenced and the current gene discription.

4.2.2 Sequencing polymorphisms

In order to sequence the 2.5Kb of putative promoter, primers were designed to amplify the region, typically in four overlapping fragments to try and ensure full coverage. Fourteen of the 18 genes have polymorphisms within the putative promoter sequence and as **Table 4.3** shows most have more than one SNP. The Sequencher program was used for sequence analysis and all the sequencing related to a probe set is stored together and the variants are numbered in **Table 4.3** according to the position relative to the A nucleotide of the ATG at the translation start site in most cases. Three of the PCR fragments failed and were not replaced with new primers and a further 2 generated only partial sequence. The final column of **Table 4.3** shows which probe sets were sequenced and only 2 contained BN/SHR polymorphisms.

The Sequencher program allows all sequence generated to be complied into contigs for comparison and a reference sequence can be set, which was the public sequence in

Table 4.2 All probesets to be sequenced

Probeset Public ID	QTL Reaper p-value	Affymetrix Gene Name (when sequenced)	Affymetrix Gene Symbol (when sequenced)	Chromosome	strand	Ensembl ID	Ensembl Gene symbol	Affymetrix gene symbol (updated)	Gene actually sequenced (ensID, symbol, strand)
1367917_at	0.000001	cytochrome P450, family 2, subfamily d, polypeptide 26	*Cyp2d26*	7	-1	ENSRNOG00000008988	*Cyp2d26*	*Cyp2d26*	-
1368291_at	0.000001	apoptosis inhibitor 2	*Api2*	8	-1	ENSRNOG00000010602	*RGD 620690*	*Birc2*	-
1368292_at	0.000023483	dynamin 1	*Dnm1*	3	-1	ENSRNOG00000033835	*Dnm1*	*Dnm1*	-
1368304_at	0.000005	flavin-containing monooxygenase 3	*Fmo3*	13	-1	ENSRNOG00000003620	*fmo3*	*fmo3*	-
1371085_at	0.000001	achaete-scute complex homolog-like 3 (Drosophila)	*Ascl3*	1	-1	ENSRNOG00000013702	*ascl3*	*ascl3*	-
1371137_at	0.000916667	Trihydroxycoprostanoyl-CoA Oxidase	*Acox2*	15	1	ENSRNOG00000007378	*Acox2*	*Acox2*	-
1371725_at	0.0008	ESTs	*Myh9*	7	-1	ENSRNOG00000004860	*myh9*	*myh9*	-
1372646_at	0.000096	Similar to RIKEN cDNA 1500015O10 (LOC363225), mRNA	---	9	1	ENSRNOG00000023576	*RGD1305645_predicted*	*RGD1305645_predicted*	-
1374429_at	0.000387097	ESTs, Moderately similar to PIM1 RAT PROTO-ONCOGENE SERINE/THREONINE-PROTEIN KINASE PIM-1 (R.norvegicus)	*Pim1*	20	-1	-	-	none - transcribed locus	ENSRNOG00000000529, *Pim1*, 1
1374921_at	0.000014139	Similar to helicase-like protein NHL isoform 2 (LOC362288), mRNA	---	3	1	ENSRNOG00000027513	*Q5RJZ1_RAT*	*RGD1306721*	-
1375676_at	0.000056338	Transcribed sequence with moderate similarity to protein pdb:1LBG (E. coli) B Chain B, Lactose Operon Repressor Bound To 21-Base Pair Symmetric Operator Dna, Alpha Carbons Only	---	3	-1	-	-	none - transcribed locus	ENSRNOG00000005131, *Lin7c*, 1
1375958_at	0.000033898	Transcribed sequence with moderate similarity to protein ref:NP_112197.1 (H.sapiens) TSC-22-like [Homo sapiens]	---	12	1	ENSRNOESTG00000018049	-	-	ENSRNOG00000024616, *Tsc22*, 1
1387366_at	0.0000001, 0.0000001	interleukin enhancer binding factor 3	*Ilf3*	8	1	ENSRNOG00000022741	*Ilf3*	*Ilf3*	-
1387376_at	0.000003	aldehyde oxidase	*Aox1*	9	1	ENSRNOG00000015354	*RGD620528*	*Aox1*	-
1388366_at	0.000003	Similar to mitochondrial ribosomal protein L4 isoform a (LOC363023), mRNA	---	8	1	ENSRNOG00000020659	*Mrpl4_predicted*	*Mrpl4_predicted*	-
1388567_at	0.000024145	Similar to RIKEN cDNA 6330575P11 (LOC309041), mRNA	---	1	-1	ENSRNOG00000014946	*Thumpd1*	*Thumpd1*	-
1390185_at	0.000013995	mRNA decapping enzyme	*Dcps*	8	-1	ENSRNOG00000009993	*Dcps*	*Dcps*	-
1393902_at	0.000001	Similar to estradiol 17beta-dehydrogenase (EC 1.1.1.62), A-specific - mouse (LOC307092), mRNA	---	17	-1	ENSRNOG00000038331	*NP_001028869.1*	*Akr1c6*	-

Table 4.3 Polymorphisms detected in the *cis* -acting eQTL genes

Gene	Probeset(s) Public ID	Primer Number	Fragment	Primers	Primer sequence	Fragment size (bp)	Position of nucleotide varient in Sequencher	position relative to ATG	SHR	BN	Probeset Sequenced
											yes
dcps	1390185_at BE109709	19/20	F1/R1	MH19	gtagtggcgcactccatcctt	615	-250	-250	G	A	No SNPs
				MH20	ctccatctctgttaggcatt						
		163/164		MH163	gcttcgagttgcagtccgtt	596	-440	-440	T	A	
				MH164	cctcgttggctacaaatatccaa						
		165/166		MH165	ccacctcaggcttggttgt	493	-810	-810	C	T	
				MH166	cctcaccaaagattccagcaa						
											yes
api2	1368291_at	25/26	F1/R1	MH25	cctgagaacacaggagttctt	742	-				No SNPs
				MH26	ccacggcttggatgtgaacttt						
	NM_021752	27/28	F2/R2	MH27	gggaaatcacatgaacactttct	871	-				
				MH28	ggaactgctgtagttggatgt						
		29/30	F3/R3	MH29	gcactcgtgtacatgaacatt	710	-3322	-3322	T	A	
				MH30	gccgtcaggtccaaagta		-3655	-3655	G	T	
							-3664	-3664	3 more repeats		
		31/32	F4/R4	MH31	ggactgaagctagccaaggg	755	-3852	-3852	ins C	-	
				MH32	ctacatggtcaatagggaaa		-4064	-4064	T	A	
											Yes
	1375676_at AA956897	51/52	F1/R1	MH51	gcagggactcggacggtat	813	221	221	C	T	No SNPs
				MH52	gcagcctctgacctggatt						
		49/50	F2/R2	MH49	ccacctttctgttcaaacaga	667	-740	-740	T	C	
				MH50	cgggtctggcatcttcaa		-1017	-1017	A	C	
		53/54	F3/R4	MH53	ccagaaacttccaggaaact	604	-1080	-1080	G	A	
				MH54	ctggcagcagtcctcaa						
											Yes
lmo3	1368304_at	55/56	F1/R1	MH55	ctgctacttccaagaaca	455	109	109	C	T	SNPs and Del
				MH56	cgagtgttccgaactactct		70	70	G	A	
	NM_053433	55/55	F2/R2	MH55a	gttgccacactgaatgtaga	407	-2111	-2111	Del AATG	-	
				MH55b	gcagggaccaaggagtgaa		-2117	-2117	A	T	
							-2192	-2192	T6/A7	T7/A5	
							-2255	-2255	CG	GA	
							-2379	-2379	A	G	
							-2397	-2397	C	A	
		55/55	F3/R3	MH55c	cctgtagcaagacagcaca	580	-2516	-2516	A	G	
				MH55d	tctgggactagcaagccta		-2559	-2559	insA	-	
							-2578	-2578	C	G	
							-2681	-2681	del A	-	
							"-2860 to -2685"	860 to -2685"	deleted	-	
							-2863	-2863	C	G	
		55/55	F4/R4	MH55e	ccgaagacagctacagtgt	757	Replaced				
				MH55f	cacaccggaatggctcttt						
		388/389	F4/R4	MH388	gctacaaccatgtcccctat	380	-2953	-2953	C	A	
				MH389	cctaggctgctatgtcccaga		-2961	-2961	insC	G	
							-2974	-2974	A	A	
							-2988	-2988	C	T	
							-3005	-3005	C	G	
							-3122	-3122	A	-	
							-3254	-3254	C	T	
		55/55	F5/R5	MH55g	ggagaacaaaaccagcactta	682					
				MH55h	gcaggagaacctaagagatga						
											yes
ecrg4	1372646_at	56/57	F1/R1	MH56	caaggcctgagcctgaa	539 Failed					No SNPs
				MH57	ggatcatatgcaatgcgtgaa						
	BF418649	58/59	F2/R2	MH58	gagaaggtggctccctgaat	647	-479	-479	T	A	
				MH59	caacaggttggacagtga		-551	-551	del ACACAGC	-	
							-765	-765	T	C	
		387/61	F3/R3	MH387	ccactgcgctcagccta	450 None					
				MH61	ctgatggagcacagacaa						
		60/386	F4/R4	MH60	ggcagatgtcactgcagca	502	-1242	-1242	G	A	
				MH386	ggataagacatagcaatctgaa		-1470	-1470	A	T	
		62/63	F5/R5	MH62	gctcatggtgtgcatggtt	747	-1553	-1553	G	C	

				MH63	ggctgactctcaggagcta		-1977	-1977 A	G	
										yes
mrt4	1388366_at	64/65	F1/R1	MH64	gcggatdctctgagttcaa	621	-85	-85 G	C	No SNPs
	BI275903			MH65	gctctcggaagttatcaacct		39	39 A	G	
		66/67	F2/R2	MH66	ggttactcttgactacacaga	817 -				
				MH67	gtggacagacctcctgct					
		68/69	F3/R3	MH68	ggtcaagtgtggaacagaga	674 -				
				MH69	ggctaaccttgagttcagaga					
										yes
aox1	1387376_at	70/71	F1/R1	MH70	ggtgcaggactctggcat	566	-303	303 C	T	No SNPs
				MH71	ccaaactaaggtctcagaga					SNPs nearby
	NM_019363	72/73	F2/R2	MH72	gcagaagcagatagaagcta	690	-785	-785 A	G	
				MH73	gcctgtaccagtgcttgtt					
		137/138	F4/R4	MH137	ggagttgtggaagactgtgaa	441	-871	-871 T	A	
				MH138	ctcctgagtatcttaagcacat		-866	-866 T	C	
							-859	-859 del ACACTCACACTACAC		
		135/136	F3/R3	MH135	ggcctctgcttgctccttt	789				
				MH136	ctcagtggttcagagcacttgt					
										No
cyp2d26	1367917_at	74/75	F1/R1	MH74	gtccacctgcagcaggtt	681	178	-178 T	C	
				MH75	gccclgctgctgcatactgata					
	NM_012730	76/77	F2/R2	MH76	cctgagggcatcgtctca	714 -				
				MH77	gagggtgggcctcagatt					
		78/79	F3/R3	MH78	gaggaggtgacctggtca	635				
				MH79	gtgcctgcaggctaccca					
		80/81	F4/R4	MH80	ccaagagcctgagaaccaga	646				
				MH81	ggtgacactcaagagtcaca					
										yes
akr1c6	1393902_at	82/83	F1/R1	MH82	ccccgatactgaatgccatct	716 -				1 SNP in probe 1
				MH83	ggcgtttgccactgagct					1 SNP in exon but not probe sequence
	AI072107	84/85	F2/R2	MH84	ggtgatcaaaccccaggacct	627	-512	-512 A	T	
				MH85	cgtaagcactgcactagtgt					
		86/87	F3/R3	MH86	gcatatgtgtatgtacagaga	710 -				
				MH87	gctcaggatactgagtggga					
		88/89	F4/R4	MH88	ccctggagatcttaacatca	706	-1475	1475 insT		
							-1548	-1548 G	A	
				MH89	cctctgagagcacaagaca		-1725	1725 T	A	
										yes
rte	1374921_at	90/91	F1/R1	MH90	gggctccaagctcctttga	475				No SNPs
				MH91	cagacagaccctgtctgga					
		92/93	F2/R2	MH92	ggtcctcacctgactgatt	805	1199	-1199 G	C	
				MH93	cggttcagaactgggagaga		-1324	1324 A	G	
		94/95	F3/R3	MH94	cgctgtgcatgctgcctca	662	-1791	1791 A	T	
				MH95	gtggttactgacctcaacct		-2067	2067 T	C	
		96/97	F4/R4	MH96	gctgagggacagtgaagt	486	-2609	2609 AG	TT	
				MH97	ggctactagtgaagctgaca		-2605	2605 ins GTGTGCATTAC		
							2440	2440 del 6(TG)		
							-2426	2426 C	T	
							-2415	2415 C	G	
Dnm1	1368292_at	98/99	F1/R1	MH98	ccagcgaggacctgagta	682				no
				MH99	ttctccttctgcggctttcg					
	NM_080689	100/101	F2/R2	MH100	gatccgagcggagcgctt	592 -				
				MH101	ggtcctgtctatcccacca					
		102/103	F3/R3	MH102	ccaggaaccctttcctggta	518 -				
				MH103	ggagtctactaggtcct					
		104/105	F4/R4	MH104	ggatgatagccacattcaga	725				
				MH105	ggagtcctaactaggtcct					
										yes
Thumpd1	1388567_at	106/107	F1/R1	MH106	ccatactcgttgagcaggct	501				No SNPs
				MH107	cctccttggacaggacta					SNPs in target

		108/109	F2/R2	MH108	ggacicc-agcagggtgta	659	649	-649 G	A		
				MH109	gtcctgggattctgagacct		670	-670 microsatelite didn't sequence through			
		110/110	F3/R3	MH110	cagtgtgggatggattgtaga	864	833	-833 C	T		
				MH111	cccgaatagggcacaagaggt		1051	-1051 T	C		
		111/112	F4/R4	MH112	cncccicactgagatggaca	623					
				MH113	ccgtacgattgctagagcaa						
										Yes	
tsc22	1375958_at	113/114	F1/R1	MH113	gggttgctggatgcctac	655				1 SNP	
				MH114	caccgtaagttcacgagaaa						
	BF412297	115/116	F2/R2	MH115	cacgattgggcaaagctgat	741	749	1558 G	A		
				MH116	ctggatgtatctctgggct		902	1405 A	G		
							970	1337 A	T		
		117/118	F3/R3	MH117	cccgagaaggcaggactt	744	1595	712 G	C		
				MH118	ggagcggagggacagaat						
		119/120	F4/R4	MH119	ctcagggtggccctttca	843	2172	135 A	-		
				MH120	ggtgacgrtggtgatctgaa						
										No	
pim1	1374429_at	121/122	F1/R1	MH121	ggctttgtaacaggctcga	723					
				MH122	cgatgcactgtgggctagat						
	AI598401	123/124	F2/R2	MH123	cctcgcaggcctgtgact	806					
				MH124	ccgccgccaggagtaaag						
		125/126	F3/R3	MH125	ggccgaggtggtgtagct	764					
				MH126	gagaggcacgggacactaa						
										Yes	
acox2	1371137_at	223/224	F1/R1	MH223	ggagtcgttgtggccacact	472				No SNPs	
				MH224	cctcttctgccgtatctgaa					1SNP in exon	
		225/226	F2/R2	MH225	gctgttagctgttgagtcact	489					
				MH226	ccagtttgtctctgcagaaaga						
		227/228	F3/R3	MH227	ggtggccccatgatcctagt	859					
				MH228	ggtagtcacaacacagtgaga						
		229/230	F4/R4	MH229	cctcatggacctacagaa	602	279	-279 G	A		
				MH230	ggaccaaatgagagccatct		-322	322 G	G		
							349	349 insG	-		
							-514	-514 C	T		
							-517	-517 G	A		
ilf3	1387366_at	7/8	F1/R1	MH7	cctatgaatgctctgtgca	832				Yes	
AAF220102				MH8	ggctctaatgaacgccatata					No SNPs	
	NM_053412	9/10	F2/R2	MH9	gtctctcaccctctgacatt	919					
				MH10	cctctcacaaaacgctatgat						
		11/12	F3/R3	MH11	cctgtgtgctcctactaaca	824					
				MH12	cagtggtgagtcacttgctt						
				MH235	gcaagcatcactctcaggact	replaces 7/8					
				MH236	gccagagcaccaacagcatgc						
				MH237	ggtggtcctgtttctctcta	replaces 11/12	493	490 G	T		
				MH238	caagcgcaaggcccttggt						
										No	
asct3	1371085_at	39/40	F1/R1	MH39	gggaaatcctgtaccctca	955					
				MH40	gccctgaactcacagctatt						
	AB046449	127/128	F2/R2	MH127	ggcatcgatatcaaagaagacl	828					
				MH128	ggtccactgtgggtgatcat						
		129/130	F3/R3	MH129	ccatatcattgttcctcatca	784					
				MH130	ccaaacctgcatttcctgaa						
		131/132	F4/R4	MH131	ctaggaaacgtgcaaggccct	715					
				MH132	gagatgctatcaagtgcttac						
		133/134	F5/R5	MH133	caagcgctcttatctgctgaa	817					
				MH134	gacaatctgagccttgccgtt						
										No	
myh9	1371725_at	41/42	F1/R1	MH41	ggagccaacgaccteatct	519					
				MH42	ggcatcggaaggctcaagaa						
	BM392410	43/44	F2/R2	MH43	ctcacctgaacctgatct	751					
				MH44	ttcaacctgtcattccgtgtg						
		45/46	F3/R3	MH45	cccttcgtcctgtcctaaca	710					
				MH46	cacccacaaccacaggatt						
		47/48	F4/R4	MH47	ccacatagaggtacttgtct	223					
				MH48	ggacttcctcatgtgtcacca						

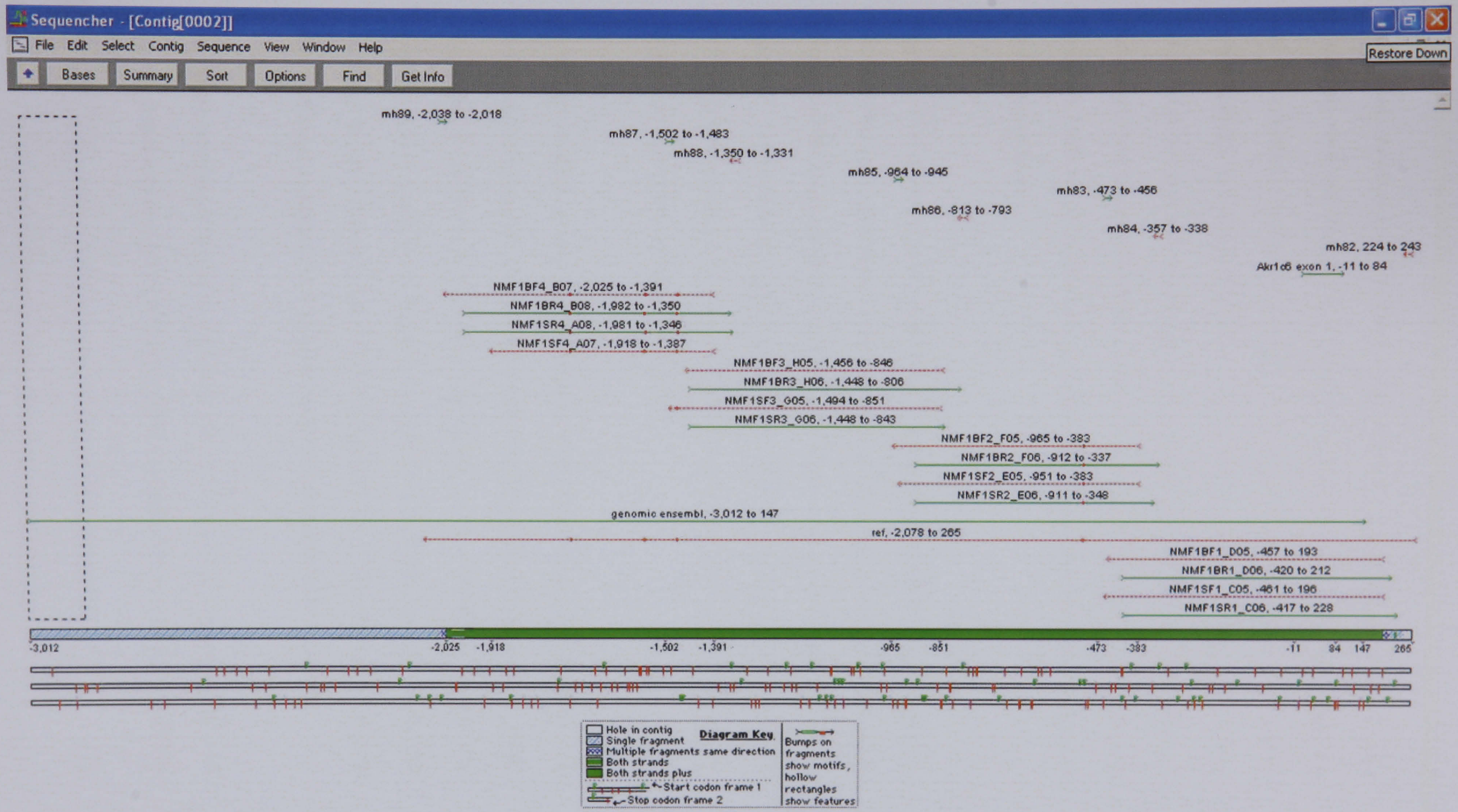


Figure 4.1 Sequencher contig view of the *Akr1c6* gene This is a screenshot from the Sequencher program to demonstrate how it displays the contigs created from sequencing fragments. The primers and the sequence files are presented alongside the reference sequence.

these cases. An example of the Sequencher file format is presented in **Figure 4.1**; this is a contig view and shows all the fragments sequenced in BN and SHR, the primers and the reference sequence for the *Akr1c6* gene (probe set 1393902_at). In total 76 polymorphisms were detected and an example sequence electropherogram generated for a part of the promoter of the above gene in both BN and SHR parental strains in the forward and reverse directions is shown in **Figure 4.2**. This T to A transversion shown on both strands illustrates the quality of sequence generated and coverage of the polymorphisms.

Figure 4.3 is another screen shot from the Sequencher program and displays two SNPs and a repeat difference in the aldehyde oxidase (*Aox1*) gene. As the figure shows the SHR fragments have a deletion of 14bp in a repeat region and 2 SNPs close together. In addition, the genomic reference sequence for this gene has a run of NNNs indicating a gap in the public genome sequence which has now been filled with the sequence generated in fragment F4/R4.

Probe set 1368304_at maps to the Hepatic flavin containing-monooxygenase gene (*Fmo3*) on chromosome 13 and lies within an SHR blood pressure pQTL. When sequenced the promoter of this gene was shown to have 16 SNPs, 5 single base insertions and 5 single base deletions in addition to a 175bp deletion in the SHR relative to the BN strain over the 2.5Kb sequenced. This 175bp deletion can be seen in the gel image in **Figure 4.4** which shows the PCR fragment in BN and SHR.

The probe sets (target region) from 13 of the *cis*-acting eQTL genes selected were sequenced to determine if the differential expression detected could be due to SNPs in

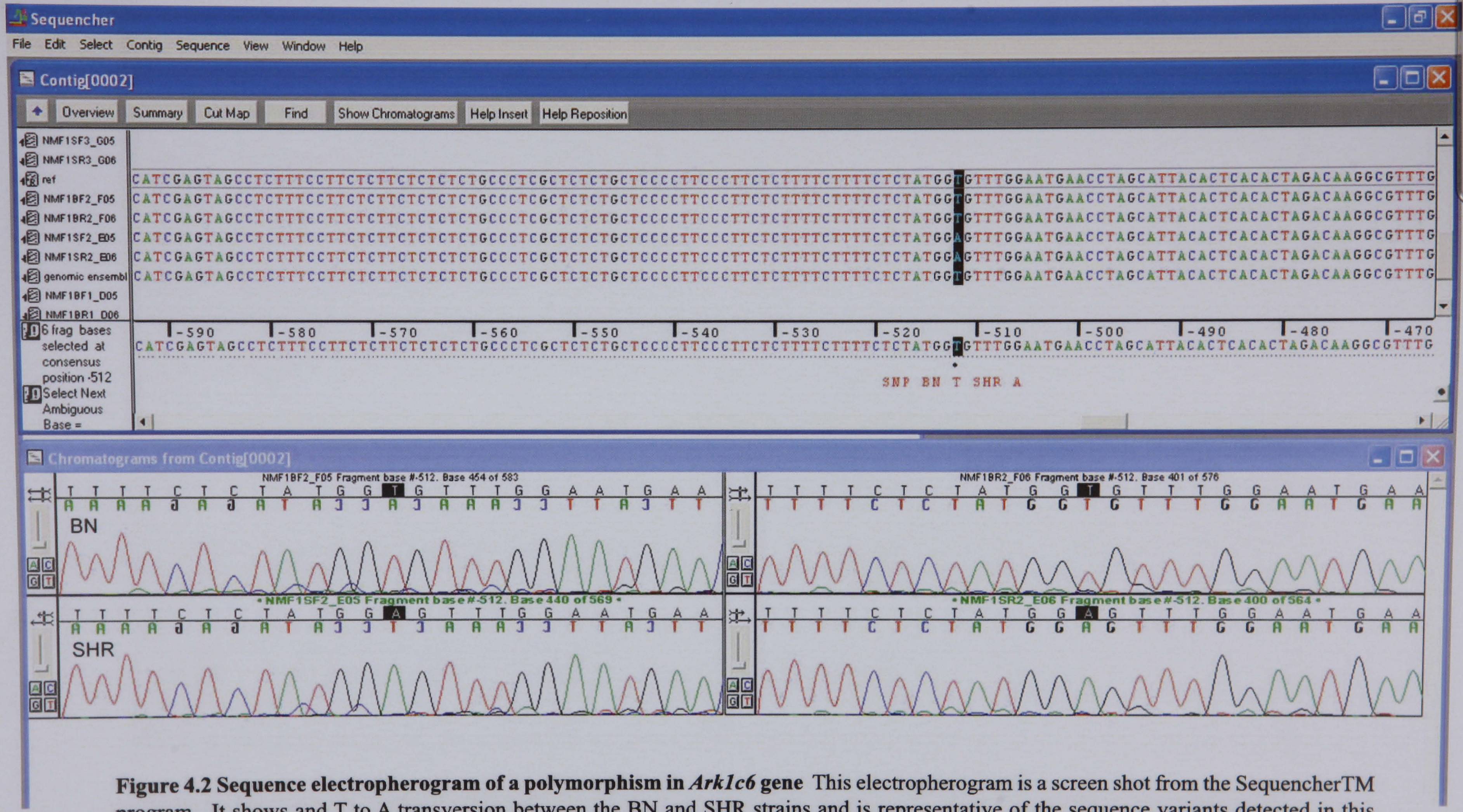


Figure 4.2 Sequence electropherogram of a polymorphism in *Ark1c6* gene This electropherogram is a screen shot from the SequencherTM program. It shows and T to A transversion between the BN and SHR strains and is representative of the sequence variants detected in this study.

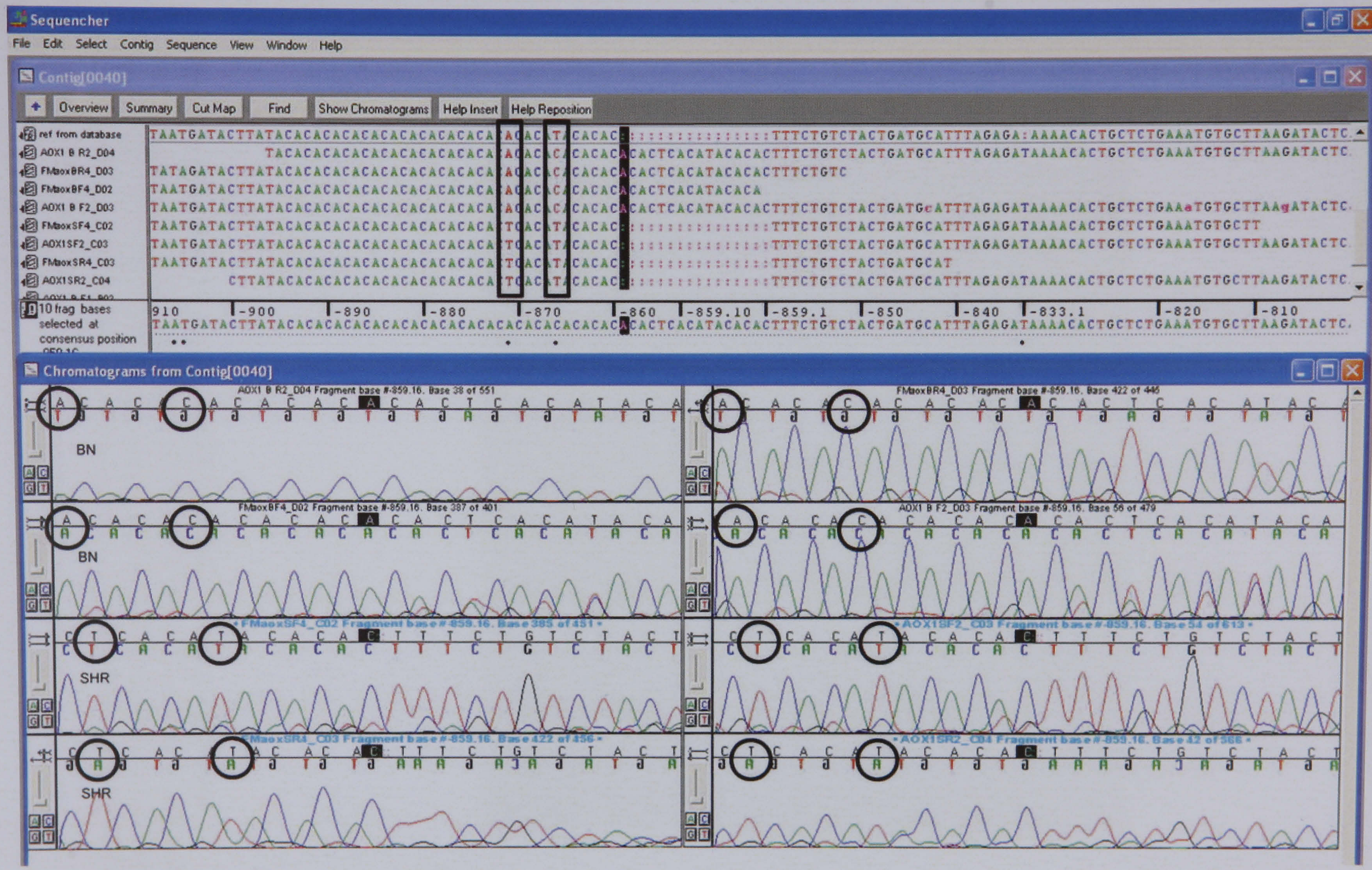


Figure 4.3 Sequence electropherogram of polymorphisms in *Aox1* gene. This electropherogram is another screen shot from the SequencherTM program. This shows 2 SNPs and a 14bp deletion in PCR fragments 2 and 3 of the 5' upstream region which was sequenced in the *Aox1* gene.

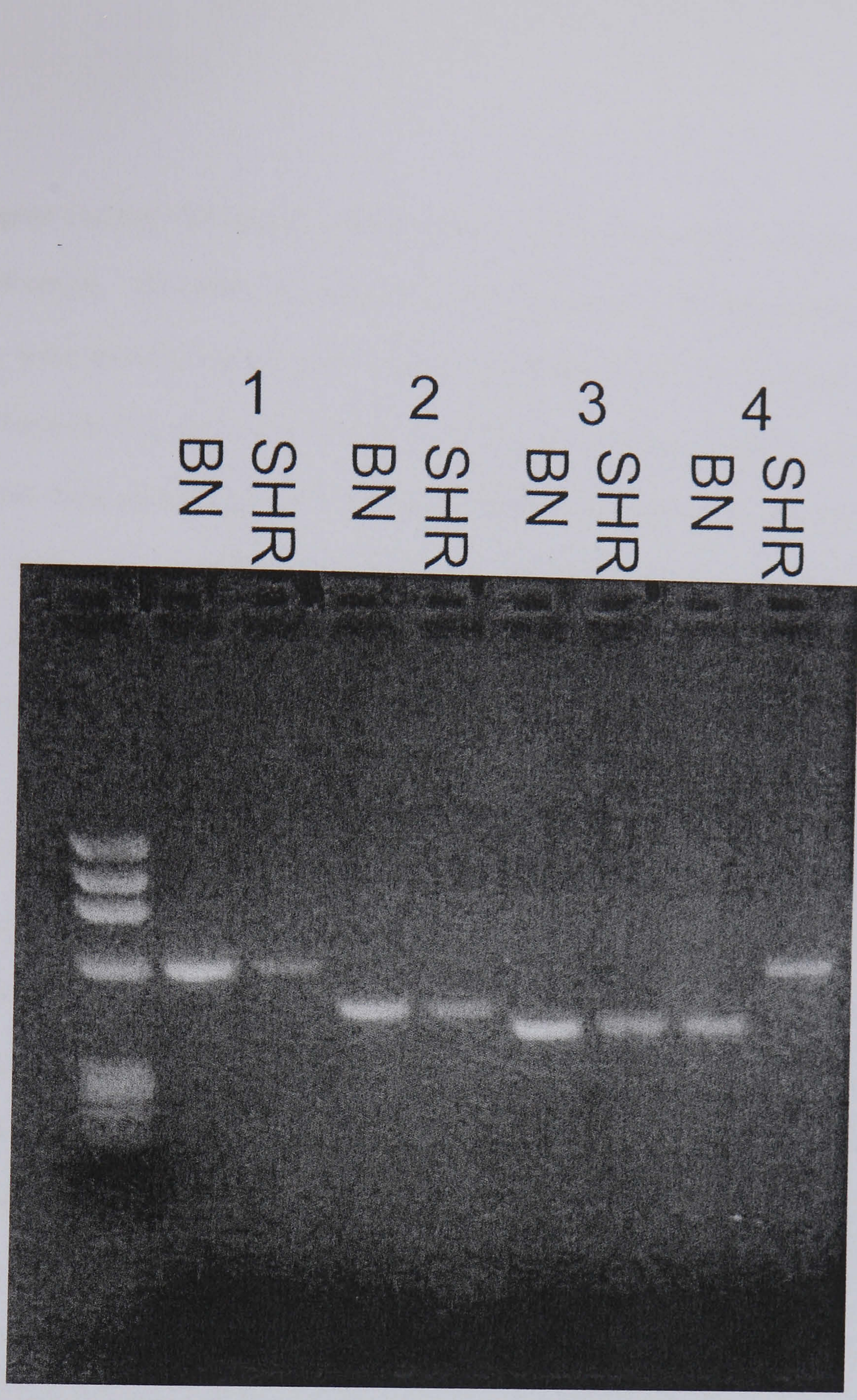


Figure 4.4 Genomic PCR of fragment 3 in SHR and BN in *Fmo3* gene
The expected size of the product of this fragment is 580bp and this is seen in the BN strain PCR. However the SHR strain has a shorter PCR product indicating a deletion in the SHR strain in this region. This deletion is 175bp which was detected during sequencing. (**Table 4.3**)

the probe region causing differential hybridisation and three were found to contain probe set differences. The other five genes were not sequenced in the target region because they were found to contain only one or no polymorphisms in the putative promoter. Therefore this part of the promoter was not the cause of the differential expression thus these genes were discarded from further consideration at this time. Prior to this they were mapped onto the genome and this detected some errors, 1374583_at, 1371725_at and 1375676_at do not map to the genes *Dcps, Pim-1* and *Lin7c* respectively as the orginal annotations suggested.

4.2.3 *Trans* gene sequencing

Overall 76 polymorphisms in approximately 40Kb of sequence from the putative promoters of *cis* eQTL were detected. To determine if this number of polymorphisms is higher than would be expected by chance the 2.5Kb upstream of 8 genes which were controlled in *trans* were selected. These genes are controlled in *trans* and therefore the differential expression is regulated by another gene or element located elsewhere in the genome. Thus these genes were expected to have fewer polymorphisms in their proximal promoters. **Table 4.4** details the genes selected and the 14 polymorphisms detected within in approximately 20Kb of sequence. This means that the there were 69 SNPs in 40Kb (mean 1.725) in the *cis*-acting genes promoters and 13 in 20Kb (mean 0.65) in the *trans*. Therefore this analysis determined that *cis* eQTL genes have a higher frequency of SNPs within promoters than *trans* genes. The sequence of the primers for the 8 *trans*-regulated genes is shown in **Appendix III**.

4.3 Discussion

The genes selected for sequencing in this chapter were *cis*-acting eQTL genes because these represent a group of genes which control their own expression, that is, polymorphisms within these genes result in the eQTL linkage and differential expression seen in the RI and parental strains. Since the proximal promoter is the primary region of a gene involved in gene expression it was decided to sequence this in the 18 genes selected. The proximal promoters of genes are not always well defined therefore 2.5Kb upstream of the most 5'exon in the genes which represented the eQTL probe sets selected was sequenced.

Thirty-seven genes from both the fat and kidney data sets were found to be both 2 fold differentially expressed between the parental strains and have an QTL Reaper $P<10^{-4}$. Many genes had a significant eQTL linkage in both fat and kidney data sets, however, most of these were *RT* type genes which are involved in antigen presentation in the immune system and probably reflect the different immune backgrounds of these strains so they were not studied further here. Hubner *et al* (Hubner et al. 2005) demonstrated that *RT* type genes have many SNPs in the probe region and may therefore be causing differential hybridisation. In addition, *Cd36,* a gene which has been shown to be involved in insulin resistance in SHR appears in **Table 4.1** but since the molecular basis for this differential gene expression is already known it was also not studied. The probe sets which mapped to genes that were poorly characterised or where there was difficulty in defining the most 5'exon were also not sequenced. These probe sets are still good candidates for hypertension and the metabolic syndrome however further investigation is needed to identify the most 5' exon and therefore the upstream sequence and the promoter. The aim here was to detect

Table 4.4 *trans* genes selected to test overall polymorphism rate

Gene	Probeset(s) Public ID	Primer Number	Fragment	PCR Worked	Sequenced	Position of nucleotide varient in Sequencher	SHR	BN	comments
Mst1	1369171_at	33/34	F1/R1	yes	yes	-			complete no sequence differences
		35/36	F2/R2	yes	yes	-			complete no sequence differences
		37/38	F3/R3	yes	yes	-			complete no sequence differences
thyroid peroxidase	1369784_at	FM1/2	F1/R1	yes	yes	187	A	G	Complete
		FM3/4	F2/R2	yes	yes				complete no sequence differences
		FM5/6	F3/R3	yes	yes				complete no sequence differences
		FM7/8	F4/R4	yes	yes				complete no sequence differences
		FM9/10	F5/R5	yes	yes	2025	A	G	complete
						2204	G	T	
						2227	A	G	
similar to peroxisomal biosis factor	1372923_at	FM11/12	F1/R1	yes	yes				complete no sequence differences
		FM13/14	F2/R2	yes	yes	1156	A	C	complete
						1403	G	T	
		FM15/16	F3/R3	yes	yes				complete no sequence differences
		FM17/18	F4/R4	yes	yes				complete no sequence differences
		FM19/20	F5/R5	yes	yes	2213	C	G	complete
glycoprotein m6a	1373773_at	FM21/22	F1/R1	yes	yes				only SHR not different
		FM23/24	F2/R2	yes	yes				complete no sequence differences
		FM25/26	F3/R3	yes	yes				poor seq not enough DNA
		FM27/28	F4/R4	yes	yes				complete no sequence differences
Podocalyxin	1385823_s_a	FM29/30	F1/R1	yes	yes	microsat			complete
		FM31/32	F2/R2	yes	yes				complete no sequence differences
		FM33/34	F3/R3	yes	yes	1606	G	A	complete
Similar to hypothetica protein MGC17337 (predicted)	1390346_at	FM39/40	F1/R1	No					
		FM41/42	F2/R2	yes	yes				complete no sequence differences
		FM43/44	F3/R3	yes	yes	1970	T	C	complete
		FM45/46	F4/R4	yes	yes	2130			complete
		FM47/46	F4a/R4	yes	yes	2493	C	T	complete
Atrial natriuretic peptide clearance receptor	1369273_a_a	NPR3F1/R1	F1/R1	no					
		NPR3F2/R2	F2/R2	yes	yes				complete
		NPR3F3/R3	F3/R3	yes	yes				complete
		NPR3F4/R4	F4/R4	yes	yes	2490	A	-	
		NPRF5/R5	F5/R5	yes	yes	2780	C	T	
		NPR3F6/R6	F6/R6	yes	yes				complete
	1370806_at	RMT7F1/R1	F1/R1	no					
		RMT7F2/R2	F2/R2	yes	yes				complete no sequence differences
		RMT7F3/R3	F3/R3	yes	yes				complete no sequence differences

polymorphisms in the promoter region which may underlie differential expression therefore genes were selected which had a distinct 5' exon.

A further four *cis*-acting eQTL genes were sequenced despite the fact that they only met either the fold change or significance criteria in at least one tissue, 1368291_at and 1371725_at had already been partially investigated in the preliminary analysis described in **Chapter 3**, while 1371137_at maps to the Acyl-coenzyme A oxidase 2, (*Acox2*) which is involved in lipid metabolism making it an interesting candidate. The final probe set 1374429_at was significant in kidney (P =0.00064) and had 2 fold differential expression in fat and was thought to map to the Proto-oncogene serine/threonine-protein kinase (*Pim-1*) gene.

Sequencing was carried out in the putative proximal promoter of 18 *cis*-acting eQTL genes and 76 polymorphisms of which 69 were SNPs were detected in approximately 40Kb. To ensure that this is not the number of SNPs which would be detected by chance in a promoter region promoters from eight of the genes which are classified as *trans*-regulated, that is their expression is controlled by another region of the genome and not the gene itself were sequenced too. These genes are shown in **Table 4.4** and were found to contain 13 SNPs in the 20Kb sequenced. This higher frequency of SNPs in *cis*-acting genes is expected since gene expression is controlled by the gene itself and therefore would be expected to be mutated in a control region such as the promoter, while *trans*-regulated genes are controlled by a factor which maps somewhere else in the genome. A recent study by Doss *et al.* (Doss et al. 2005) shows that *cis*-acting eQTL reside in regions of the genome which are not inherited

IBD from the parental strains. This could also explain the increased SNP frequency seen in the *cis* genes sequenced here.

Figure 4.2 and 4.3 show sequence electropherograms demonstrating examples of a few of the sequence variants detected in the *Akr1c6* and *Aox1* genes. Within the 2.5Kb sequence upstream of the *Aox1* gene there is a run of NNNs in the Ensembl database and the primers to amplify fragment F4/R4 were designed to span this region and fill in the gap. This gap was filled and the sequence generated was used in genome alignments using BLAST against the rat genome at Ensembl and found to hit all chromosomes excluding 12 indicating it is a repeat sequence of some kind.

Figure 4.4 depicts PCR amplification of the upstream region of the *Fmo3* gene. It shows a 175bp deletion in this fragment in the SHR strain which was also demonstrated on sequencing. There were also 16 SNPs, 5 single base insertions and 5 single base deletions in this gene between the SHR and BN alleles.

The target regions that the probe sets are designed from were sequenced in 13 of the 18 eQTL genes, those genes in which SNPs were detected. This was carried out to determine if the differential expression could actually be due to differential hybridisation of the BN and SHR transcripts caused by these polymorphisms. Only 3 of the 13 genes had a SNP in the probe set and the differential expression should be verified in these genes first in case it is in fact differential hybridisation. This analysis could be started by determining if the differential hybridisation detected is due to the specific probe of the 11 on the microarray which contains the SNP as in **Chapter 3**.

Prior to sequencing the target regions the mapping of all the probe sets was determined because it is known that there are errors in probe set mapping (personal communication, NJD), therefore all probe sets were studied in Ensembl and mapped independently using BLAST to get the best location possible. There were two probe sets in **Table 4.1** that originally mapped to the *Dcps* gene the putative promoter of which was sequenced. However it was found that 1374583_at does not map to *Dcps* and therefore was not sequenced as part of the target region sequencing. The *Pim-1* promoter was sequenced and no polymorphisms were found. Rechecking the mapping of this Affymetrix probe set (1374429_at) on a more recent genome assembly found that it is on the opposite strand to the *Pim-1* gene and there is no known annotated gene in the region. The probe set which maps to the *Pim-1* gene (1369862_at) is not differentially expressed and does not have a *cis*-acting eQTL in either tissue. This is further proof that *cis*-acting eQTL genes have higher frequency of SNPs in the promoter regions. The final probe set which was mapped to the 'wrong' gene was 1375676_at which mapped to *Lin7c*, in fact this probe set is on the other strand from *Lin7c* and has no gene associated with it. This was one of the 2 probe sets which had a SNP in the target region. The sequencing of the *Lin7c* putative promoter detected 4 polymorphisms and probe set 1369610_at maps to the *Lin7c* gene with an eQTL reaper *P*-value of 0.0012 in the kidney data set therefore narrowly missing the $P<10^{-4}$ cut-off for significance. Therefore this gene may still merit further investigation.

Not routinely sequencing the target region and the inconsistency in probe set mapping highlights some of the limitations of the strategy used. In addition, the coding region of the gene was not sequenced and no validation of the differential expression or the

effect the sequence variants detected have on gene expression has been carried out. However, the polymorphisms detected are being investigated by others in the Physiological Genomics and Medicine Group by generating luciferase reporter assay constructs of the BN and SHR alleles which can then be tested for expression. The luciferase assay will allow the polymorphisms to be investigated to identify which are crucial in altering gene expression.

4.4 Conclusion

In conclusion the work in this chapter has investigated sequence variants in the proximal promoters of *cis*-acting eQTL genes to identify polymorphisms which may cause the altered expression between the BN and SHR parental strains and provide the eQTL linkage in the RI strains. The variants detected are being tested by others in the Physiological Genomics and Medicine group in luciferase reporter assays to determine their effect on gene expression.

Chapter 5: *In silico* studies: pQTL determination and comparative mapping

5.1 Introduction

The overall aim of the work in this chapter was to investigate the applicability of the data set of *cis*-acting eQTL generated in the rat to human disease gene identification. Comparative mapping of SHR pQTL to human QTL for hypertension, dyslipidemia, insulin resistance, CVD and type 2 diabetes to determine the syntenic overlap was undertaken. As discussed in the **Introduction 1.2** these phenotypes are involved in the highly prevalent metabolic syndrome and cardiovascular diseases which are a major health burden in Westernised society and the SHR is a well established animal model of these disorders. Rat *cis*-acting eQTL genes with genome wide significance ($p<0.05$) which had human orthologues within syntenic regions were identified to delineate a data set of potential candidate genes for human hypertension, insulin resistance, CVD and metabolic syndrome. This was achieved in a stepwise process where first human QTL regions were mapped syntenically to the rat genome and then *cis*-acting eQTL genes were identified within these rat regions. Following this, the human ortholog of the rat gene was identified and checked to ensure it was situated within the human QTL. In addition the rat pQTL (if any) which overlapped the *cis*-acting eQTL was identified. The genes identified during this process represent potential candidates for the metabolic syndrome and CVD and some of these candidate genes have an already known biological association in human and/or rat with the phenotypes mapped.

In order to carry out this analysis relevant pQTL were collated from the literature in both SHR and human populations. These data were then used to determine which of the human QTL were syntenic and had *cis*-acting rat eQTL within them.

5.1.1 QTL analysis

QTL analysis (**Introduction 1.1.1**) has been utilised to identify regions of the genome influencing complex traits such as hypertension and other metabolic phenotypes, in both humans and animal models of human disease. This means there are already many regions of the genome implicated in susceptibility to these highly prevalent human disorders. Over 300 studies have been carried out in human populations for traits involved in the metabolic syndrome and CVD; however due to the difficulties discussed in **Introduction 1.1.1** relatively few genes underlying these QTL have been identified. There has been more success in identifying the genes underlying animal models of human diseases since they are inbred strains and can be more easily studied for example in experimental crosses.

5.1.2 Comparative mapping

Comparative mapping is a powerful tool in genomics because if a DNA sequence is conserved across species it is more likely to be involved in gene function and or control. In the examples discussed here the organisms being compared are also in similar disease states therefore the DNA sequences being investigated may be more likely to be involved in the pathogenesis of that disease. One example of this approach, published by Wang *et al* (Wang et al. 2005) used mouse/human comparisons to identify novel candidate genes for atherosclerosis. This thesis extends the use of comparative mapping by suggesting candidate human genes from the rat

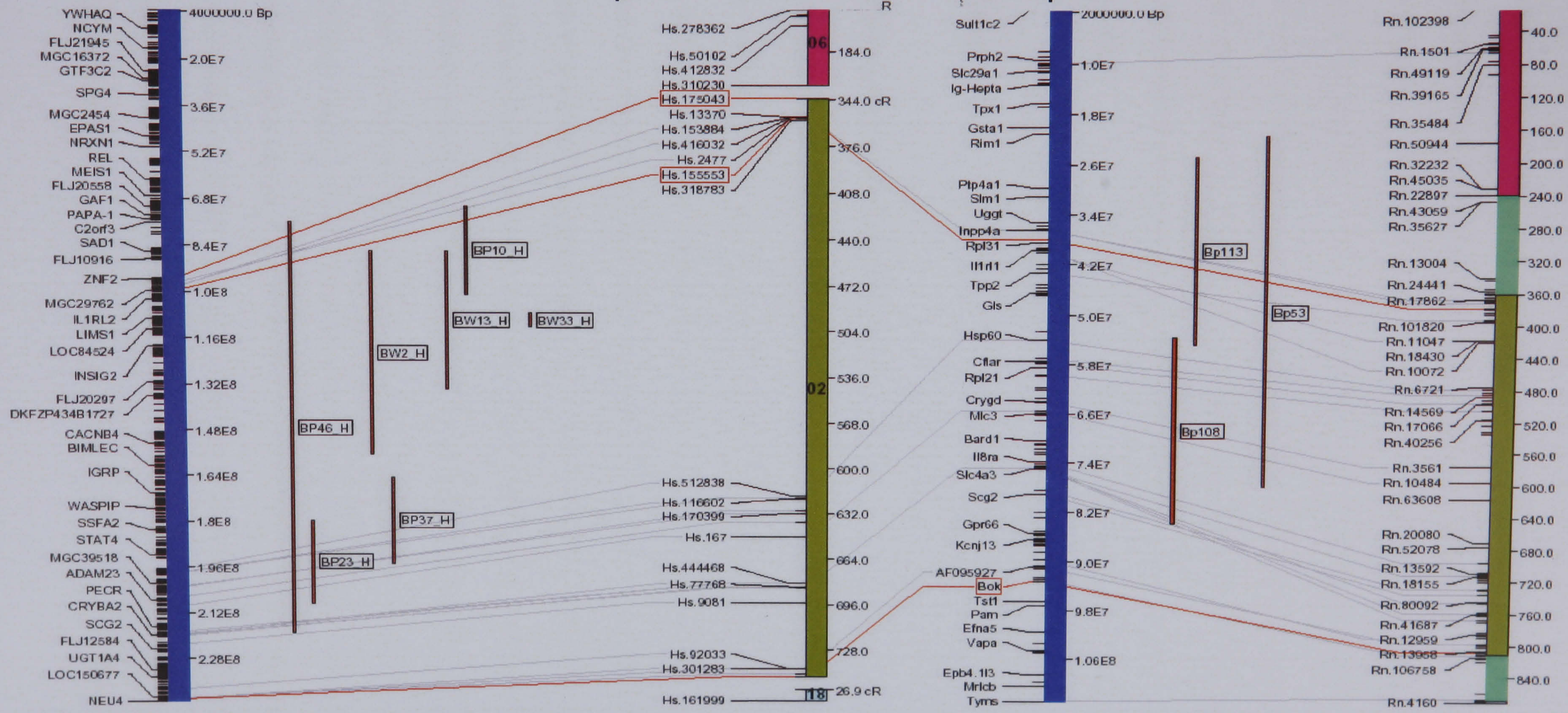


Figure 5.1 Synteny map derived from Rat Genome Database VC map

This shows an example of the regions of synteny between rat and human. The example show here shows Rat chromosome 9 which is syntenic to human chromosomes 2 (Olive) and 6 (Pink). The blue bars represent the pQTL maps in human and rat which correspond to the syntenic regions and the hypertension (BP) QTL are marked (red bars).

eQTL identified in **Chapter 3** using the method outlined in **Section 5.1**. **Figure 5.1** shows an example of one rat chromosome (9) with the human syntenic regions plotted. Alongside are the QTL maps in both rat and human with only the hypertension QTL selected. This plot was derived using VCmap which is a visualisation tool available at the Rat Genome Database (RGD) (http://rgd.mcw.edu/). In this chapter the genomic locations of the syntenic blocks and thus the syntenic QTL were used to identify rat eQTL that reside within the syntenic regions and which may underlie these QTL.

5.2 Results

5.2.1 Collation of QTL in SHR and humans

As previously mentioned SHR is a widely studied rat model of insulin resistance and hypertension and therefore many genome wide studies have been conducted identifying SHR pQTL for hypertension, insulin resistance and dyslipidemia. Many of these published studies have been reported in RGD therefore I used this as the starting point in generating a comprehensive list of all SHR pQTL. The data from RGD was checked against the original reference to ensure the reported peak of linkage, including markers, and LOD score or *P*-value were consistent with the published paper. At the time this analysis was carried out RGD had 98 QTL in SHR strains which were relevant to hypertension, metabolic and CVD phenotypes in the database. An extensive literature search was also carried out to identify any further SHR pQTL not already in RGD. Also pQTL mapped in the RI strains themselves were included, some of which were unpublished (M. Pravenec, unpublished). This generated a list of 120 pQTL in SHR for CVD and metabolic phenotypes (**Table 5.1**)

Table 5.1 Physiological SHR QTLs mapped in previous genome screens

pQTL ID[a]	pQTL Symbol[b]	RGD ID[c]	Chromosome	Peak Marker	Flanking Markers	Crossed With	Cross Type	LOD	*P*-Value
1	BpQTLcluster1	70188	1	-	D1Mgh5 - D1Mgh9	WKY, DRY	F2	4.82	-
2	Bp29	61344	1	-	D1Wox29 - D1Mit2	WKY	reciprocal congenics	7.50	-
3	Lvm2/Cm23	631199	1	D1Mit 3	D1Mit2 - D1Mgh9	DRY	F2	4.60	-
4	Str1	61442	1	D1Mit3	- - -	SHRSP	F2	7.40	-
5	Kw (Hamet et al ,1998)	-	1	RT6	- - -	BN.Lx	RI strains	-	0.05000
6	Bp44	61379	1	-	D1Mit3 - Sah	BB/OK	F2	19.20	-
7	Weight2/Bw18	70160	1	-	D1Mit3 - Sct	BB/OK	backcross	5.70	-
8	Bp26	61341	1	-	D1Mit3 - Igf2	BN/Cr	congenic	-	-
9	Bp (Huang et al , 1995)	-	1	Scnn1b	- - -	Dahl	RI strains	2.70	-
10	Bp43	61378	1	-	Spn - Sct	BB/OK	F2	22.60	-
11	Weight3/Bw20	70163	1	-	Sah - D1Mgh12	BB/OK	backcross	5.10	-
12	Cm7	631658	1	D1Rat73	- - -	SS/JrRkb	F2	5.32	-
13	Bp42	61376	1	-	Igf2 - D1Mgh12	BB/OK	F2	23.40	-
14	Map1/Bp196	631205	1	D1Mit 3	D1Mgh10 - D1Mgh1	DRY	F2	4.00	-
15	Km1	724538	1	D1Rat86	D1U1A5 - D1Rat86	Dahl	backcross	3.40	-
16	Iddm10	631670	1	D1Mgh13	D1Mgh12 - D1Mit14	BB/OK	backcross	2.60	-
17	Scl5/Stl3	631690	1	D1Mit14	- - -	BB/OK	backcross	2.10	-
18	Arr3	738012	2	D2Rat94	- - -	BN.Lx	RI strains	3.60	-
19	Bp18	61471	2	-	D2Mgh14 - Mt1pb	BN	F2	3.00	-
20	Lvm1/Cm22	631198	2	D2Mgh15	D2Mit9 - D2Mgh7	DRY	F2	4.30	-
21	Bp105	631507	2	-	D2Rat171 (D2Rat33) - D2Arb24	BN.Lx	-	-	0.00100
22	Arr7	738007	2	D2Rat221	- - -	BN.Lx	RI strains	4.40	-
23	BpQTLcluster3	70175	2	-	D2Mgh24 - D2Mgh12	WKY, DRY	F2	2.76	-
24	Bp101	631501	2	D2Mit10	D2Mit10 - D2Mgh12	wild	backcross	2.40	-
25	Bp19	61473	2	-	D2Mgh8 (D2Got87) - Npr1	BN	F2	6.30	-
26	Bp16	61469	2	D2Wox7	- - -	WKY	F2	5.64	-
27	Bp63	70162	2	D2 Wox7	- - -	WKY	F2	5.64	-
28	Bp17/DBP	61470	2	-	- - -	BN.Lx	RI strains	-	0.00060
29	Bp112/Cm10	631679	3	D3Mgh9	- - -	SS/JrRkb	F2	7.34	-
30	Bp92	631568	3	D3Rat53	D3Rat53 - D3Rat45	SS/Jr	F2	2.20	-
31	BpQTLcluster4	70191	3	Scn2a1	D3Mgh16 - D3Mit8	WKY, DRY	F2	2.58	-
32	Kw (Hamet et al ,1998)	-	3	D3Mit9	- - -	BN.Lx	RI strains	-	0.01000
33	Bp110/Cm8	631676	3	D3Mgh7	- - -	SS/JrRkb	F2	7.03	-
34	Arr10	61335	3	D3Rat180	- - -	BN.Lx	RI strains	3.90	-
35	Cm17	724532	3	D3Rat160	- - -	Dahl	backcross	2.00	-
36	Ir1 (Pravenec et al , 2002)	-	3	D3mit4	D3Mit14 - -	BN.Lx	RI strains	3.80	-
37	Bp20	32767	3	-	D3Mit2 - D3Mgh2	BN	F2	3.00	-
38	Bp33	61351	4	Il6	- - -	BN	RI strains	-	0.01800
39	Scl23/Spl1	61408	4	Il6	- - -	BN.Lx	RI strains	-	0.00050
40	Irg1	61333	4	D4Arb13	- - -	WKY	F2	4.30	-
41	Irg1	61333	4	Ae2 (Slc4a2)	- - -	WKY	backcross	3.10	-
42	Irg1	61333	4	Rt6	- - -	BN.Lx	RI strains	-	0.00004
43	Stl2	631642	4	-	Il6 - D4Mit9	BB/OK	backcross	3.30	-
44	Bp79	619616	4	-	Il6 - D4Mit24	BB/OK	F2	-	0.02920
45	Str3	61445	4	D4Mgh16	- - -	SHRSP	F2	3.00	-
46	Bp21	61336	4	-	D4Mit2 - Npy	BN	F2	4.60	-

pQTL ID[a]	pQTL Symbol[b]	RGD ID[c]	Chromosome	Peak Marker	Flanking Markers	Crossed With	Cross Type	LOD	*P*-Value
47	BpQTLcluster5	70192	4	Slc4a2	D4Mgh16 - D4Mit23	WKY, DRY	F2	2.60	-
48	Bw22	70167	4	D4Mit24	D4Mit6 - Spr	BB/OK	backcross	3.10	-
49	Bp135	631556	4	Npy	Npy - Spr	WKY	-	-	0.05000
50	Scl4	631689	4	Spr	- - -	BB/OK	backcross	1.90	-
51	Crn18	724535	4	D4Rat195	- - -	Dahl	backcross	2.60	-
52	Hematocrit (Pravenec et al., 1997)	-	4	Eno2	- - -	BN.Lx	RI strains	-	0.00010
53	Crg2	61446	4	-	Pparg - Eno2 (D4Arb40)	BN.Lx	RI strains	3.50	-
54	Bp139	634349	5	-	D5Wox20 - D5Wox12	BN.Lx	congenic	-	0.00100
55	BpQTLcluster6	70194	5	-	D5Mit5 - D5Mgh13 (D5Mit13)	WKY, DRY	F2	2.60	-
56	Bp155	129809	5	-	D5Mgh22 (D15Got98) - D5Mgh15	WKY	-	3.80	-
57	Bp103	631505	5	D5Mgh8	Slc2a1 - D5Mgh8	wild	backcross	3.20	-
58	Arr4	738018	5	D5Rat45	- - -	BN.Lx	RI strains	5.50	-
59	Str2	61444	5	Nppa	- - -	SHRSP	F2	4.70	-
60	Bp131	631553	5	Nppa	- - -	SS	congenic	4.20	-
61	Bp100	631498	5	Nppb	Nppb - Nppa	WKY	F2	4.90	-
62	Crg1	70199	6	D6Mit9	D6Rat37 - D6Mit4	BN.Lx	RI strains	11.80	-
63	Bp141	634307	6	D6Mit4	- - -	Lew	F2	4.00	-
64	Arr5	738033	6	D6Rat117	- - -	BN.Lx	RI strains	5.90	-
65	Iddm6	71111	6	D6Mgh2	D6Mgh2 - -	BB/OK	congenic	1.90	-
66	Bp102	631503	7	Igf1	- - -	wild	backcross	5.90	-
67	Lvm7/Cm27	631504	7	D7Got60	D7Rat107 - D7Rat9	Lew	F2	3.45	-
68	Arr8	738030	7	D7Rat110	- - -	BN.Lx	RI strains	4.10	-
69	Basal Glucose uptake (Artman et al., 1997)	-	7	-	- - -	BN.Lx	RI strains	-	0.00010
70	Bp. HDL phospholipid, serum cholesterol (Bottger et al., 1996)	-	8	-	D8Mit9 - Rbp2	BN.Lx	congenic	-	-
71	Bp35	61353	8	-	D8Mit5 - Rbp2	BN.Lx	congenic	-	-
72	Bp22	61337	8	-	D8Mit5 - D8Mit3	BN	F2	5.10	-
73	Bp62	70161	8	-	D8Mit3 - D8Rat133	SS/Jr	F2	2.90	-
74	Inf1	631842	8	-	- - -	BN.Lx	RI strains	4.10	-
75	Arr9	738011	8	D8Rat7	- - -	BN.Lx	RI strains	6.10	-
76	Bp53/Cm28	70218	9	D9Rat26	D9Rat70 - D9Uia6	SS/Jr	F2	7.30	-
77	Bp113/Cm11	631680	9	D9Rat125	- - -	SS/JrRkb	F2	3.10	-
78	Bp108	631656	9	D9Rat10	- - -	SS/JrRkb	-	5.97	-
79	BpQTLcluster9	70195	10	-	D10Mit5 - Slc4a1	WKY, DRY	F2	2.94	-
80	Trs1	61441	10	Myh3	D10Mit4 - D10Mit2	BN.Lx	RI strains	4.00	-
81	Bw10	631515	10	-	Abp1 - D10Wox23	wild	backcross	3.50	-
82	Bp45	61381	10	-	Abp1 - Ppy	BB/OK	F2	23.20	-
83	Bp71	70363	10	-	D10Mgh6 - D10Rat106	SS	congenic	-	0.04000
84	Liver Cholesterol (Bottger et al., 1998)	-	10	-	- - -	BN.Lx	RI strains	3.30	-
85	BpQTLcluster10	70180	11	-	D11Mit1 - D11Mit5	WKY, DRY	F2	3.19	-
86	Bp104	631506	11	Sst	D11Mgh4 - D11Wox6	wild	backcross	2.80	-
87	Irg2	61334	12	-	D12Mgh1 - D12Mit8 (D12Mgh10)	WKY	F2	6.30	-
88	Irg2	-	12	-	D12Mgh1 - D12Mit8 (D12Mgh10)	WKY	backcross	1.90	-
89	Heart weight (Hamet et al., 1996)	-	12	Hspb1	- - -	BN.Lx	RI strains	-	0.01120
90	Trs2	61443	12	-	D12Mit1 - Hspb1	BN.Lx	RI strains	3.30	-
91	Bp142	634310	12	D12Mit3	- - -	Lew	F2	4.20	-
92	Bp (Pravenec et al., 1991)	-	13	Renin	- - -	BN.Lx	RI strains	-	-
93	Bp55	70220	13	D13Mit2	- - -	WKY	F2	5.75	-

pQTL ID[a]	pQTL Symbol[b]	RGD ID[c]	Chromosome	Peak Marker	Flanking Markers	Crossed With	Cross Type	LOD	*P*-Value
94	Bp31	61349	13	D13Mit2	- - -	WKY	F2	5.75	-
95	Bp80	619615	13	D13Uwm1	D3Mit1 - Atp1a2	BB/OK	-	0.04	-
96	Scl11	631528	14	Gc	Afp - D14Wox4	wild	backcross	4.90	-
97	Bg1	631516	15	Ednrb	D15Mgh3 - D15Mgh5	wild	backcross	7.00	-
98	Bp23	61338	16	-	D16Mit2 - D16Mit5	BN	F2	4.30	-
99	BpQTLcluster13	70183	16	Glud1	D16Mit2 - D16Mgh1	WKY, DRY	F2	3.33	-
100	Kw (Hamet et al ,1998)	-	17	D1a	- - -	BN Lx	RI strains	-	0.01000
101	Heart weight (Pravenec et al , 1995)	-	17	Drd1a	- - -	BN Lx	RI strains	-	0.00038
102	Stl1	631499	17	-	D17Wox3 - Aldoc	wild	backcross	3.60	-
103	Lvm6/Cm26	631502	17	D17Rat52	D17Rat39 - D17Rat52	Lew	F2	3.71	-
104	Bw11	631518	18	D18Rat36	Spink1 - D18Wox18 (D18Rat36)	wild	backcross	2.60	-
105	Bp & Ss (Pravenec et al , 2001)	-	18	-	D18Rat32 - D18Rat12	BN.Lx	congenic	-	-
106	Scl22	631264	18	Ttr	Ttr - Nr3c1	BB/OK	F2	6.20	-
107	Bp46	61382	18	-	Ttr - Nr3c1	BB/OK	F2	18.80	-
108	Bp47	61383	18	-	Nr3c1 - Gja1 (D18Wox13)	BB/OK	F2	17.30	-
109	Spl3	631274	18	D18Mit17	Nr3c1 - D18Mit9	BB/OK	F2	5.30	-
110	Arr11	738005	18	D18Rat55	- - -	BN Lx	RI strains	3.40	-
111	Bp48	61384	19	-	Ofl - D18Mit9	BB/OK	F2	19.40	-
112	Cm9/Bp111	631678	19	D19Rat16	- - -	SS/JrRkb	F2	4.27	-
113	Cm12/Bp114	631681	19	D19Rat18	- - -	SS/JrRkb	F2	3.33	-
114	Scl12	61407	19	D19Mit2	Es2 - Ednra	BN Lx	RI strains	-	0.00008
115	Km3	724546	19	D19Uia1	D19Rat29 - D19Uia1	Dahl	backcross	3.10	-
116	Bp32	61350	19	D19Mit7	Tat - Agt	BN Lx	RI strains	-	0.01200
117	Ir2	-	19	D19Mit7	D19Mit7 - D19Rat57	BN Lx	RI strains	4.10	-
118	Bp144	724519	20	-	D20Utr1 - Tnf	BN Lx	RI strains	-	0.00100
119	Scl13	61409	20	-	- - -	BN.Lx	RI strains	-	0.00035
120	Heart weight (Kunes et al , 1990)	-	20	-	- - -	BN Lx	-	-	-

Footnotes

[a]pQTL identifier as reported in **Figures 3.7 and 3.8**

[b]pQTL names from Rat Genome Database (www.rgd.mcw.edu). Where pQTL names have not been assigned the publication reference is given

[c]pQTL identifier from Rat Genome Database (www.rgd.mcw.edu)

Key to QTLs

arr	anxiety related response
bg	blood glucose
bp	blood pressure
bw	body weight
cm	cardiac mass
crg	compensatory renal growth
iddm	insulin dependant diabetes mellitus
lef	infertility severity
ir	insulin resistance
irg	insulin resistance gene
km	kidney mass
kw	kidney weight
lvm	left ventricular mass
map	mean arterial pressure
scl	serum cholestrol level
ss	salt sensitivity
stl	serum triglyceride
str	sensitivity to stroke
tc	total cholestrol
trs	thermal response to stress

References

Ashman T. et al. 1997. Nat Genet 16:197-201
Bottger A. et al. 1996. J Clin Invest 98:856-862
Bottger A. et al. 1998. Biochem Biophys Res Commun 246:272-275
Hamet P. et al. 1996. Hypertension 28:1112-1117
Hamet P. et al. 1998. Kidney Int 53:1488-1492
Huang H. et al. 1995. J Hypertens 13:1247-51
Kunes J. et al. 1990. J Hypertens 1990:1091-1095
Pravenec M. et al. 1991. Genomics 9:466-472
Pravenec M. et al. 1995. J Clin Invest 96:1973-1978
Pravenec M. et al. 1997. Mamm Genome 8:367-369
Pravenec M. et al. 2001. Am J Hypertens 14:82A
Pravenec M. et al. 2002. Mamm Genome 13:253-256

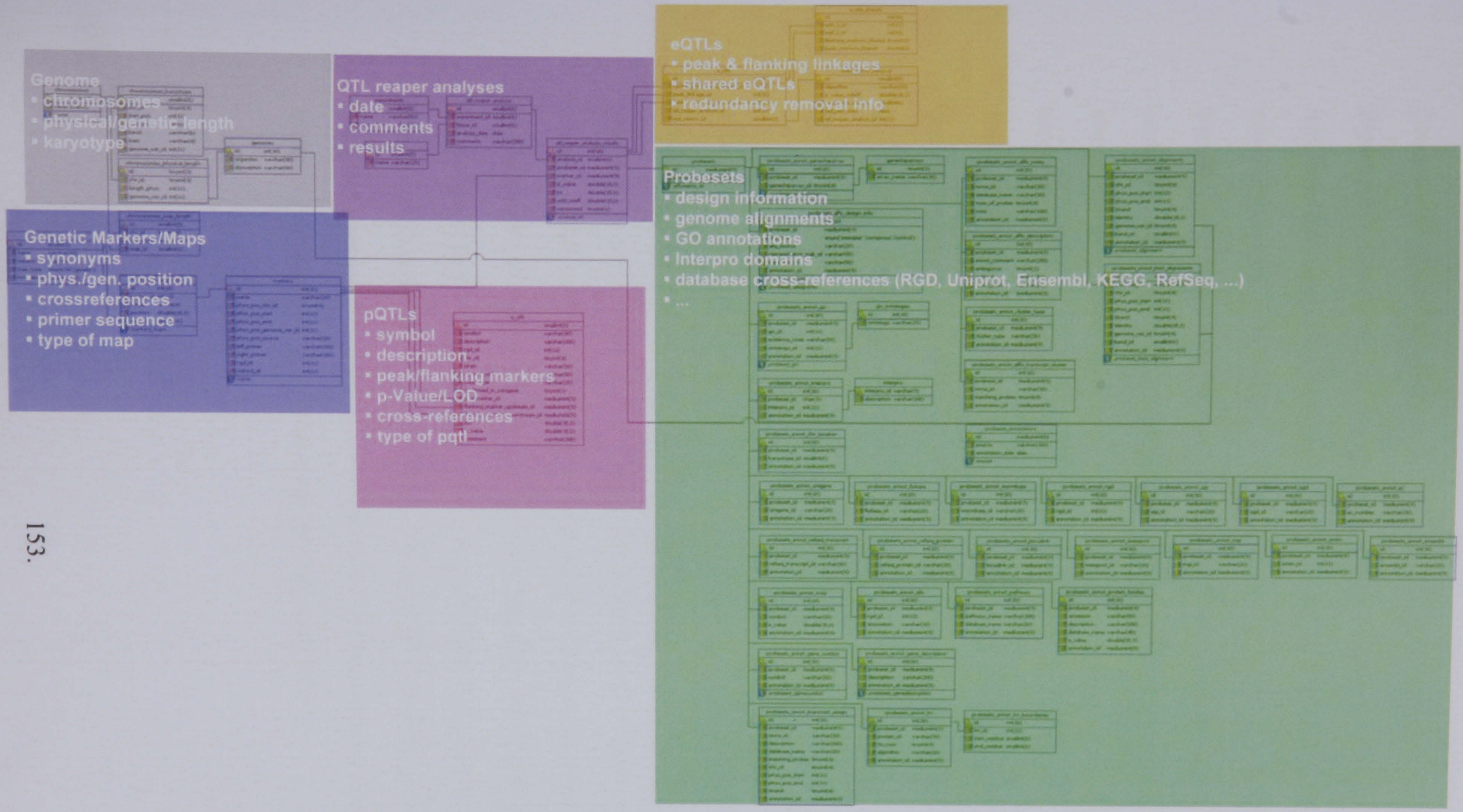


Figure 5.2 Original Database Schema for eQTL explorer This schema shows how the data generated in this thesis has been utilised to develop a visualisation tool which is now extensively used by members of the Physiological Genomics and Medicine group to mine the eQTL data. There are 6 main types of data in the database. I generated all of the pQTL data (pink) and in collaboration with Dr Enrico Petretto the eQTL data (orange) and QTL Reaper (purple) analysis data. I also checked and cross referenced any of the data when necessary.

which was used in the eQTL explorer program developed in the Physiological Genomics and Medicine group to visualise the eQTL data (Mueller et al. 2006). **Figure 5.2** is the original database schema for eQTL explorer and demonstrates how data generated for this thesis, both in this chapter and in **Chapter 3** has been used by others in the group. Identification of eQTL within these already known pQTL makes that eQTL more interesting as a potential candidate gene underlying the pQTL. An example is shown in **Figure 5.3** where the probe sets (1367689_a_at and 1386901_at) shown map to the *Cd36* gene and are controlled in *cis* in both fat and kidney tissue. Mutations within this gene are known to underlie the insulin resistance and altered fatty acid metabolism seen in the SHR strain, which is shown by the pQTL (pQTL Irg1) on the right side of the chromosome. Therefore other *cis*-acting eQTL which lie within these already mapped pQTL make attractive candidates genes for those particular phenotypes.

This collation was repeated for QTL in human populations by again looking at RGD initially and then carrying out an extensive literature search for genome wide scans for metabolic syndrome, hypertension, obesity, dyslipidemia, insulin resistance and type 2 diabetes. The results of this are shown in **Table 5.2** which reports a list of 286 QTL in multiple human populations. This list includes some QTL which do not reach the Lander and Kruglyak thresholds of 3.3 and 1.9 for significant and suggestive linkage respectively (Lander and Kruglyak 1995). The non significant QTL were included when they overlapped with other QTL and therefore provide supporting evidence that this region of the genome is important in the phenotype studied.

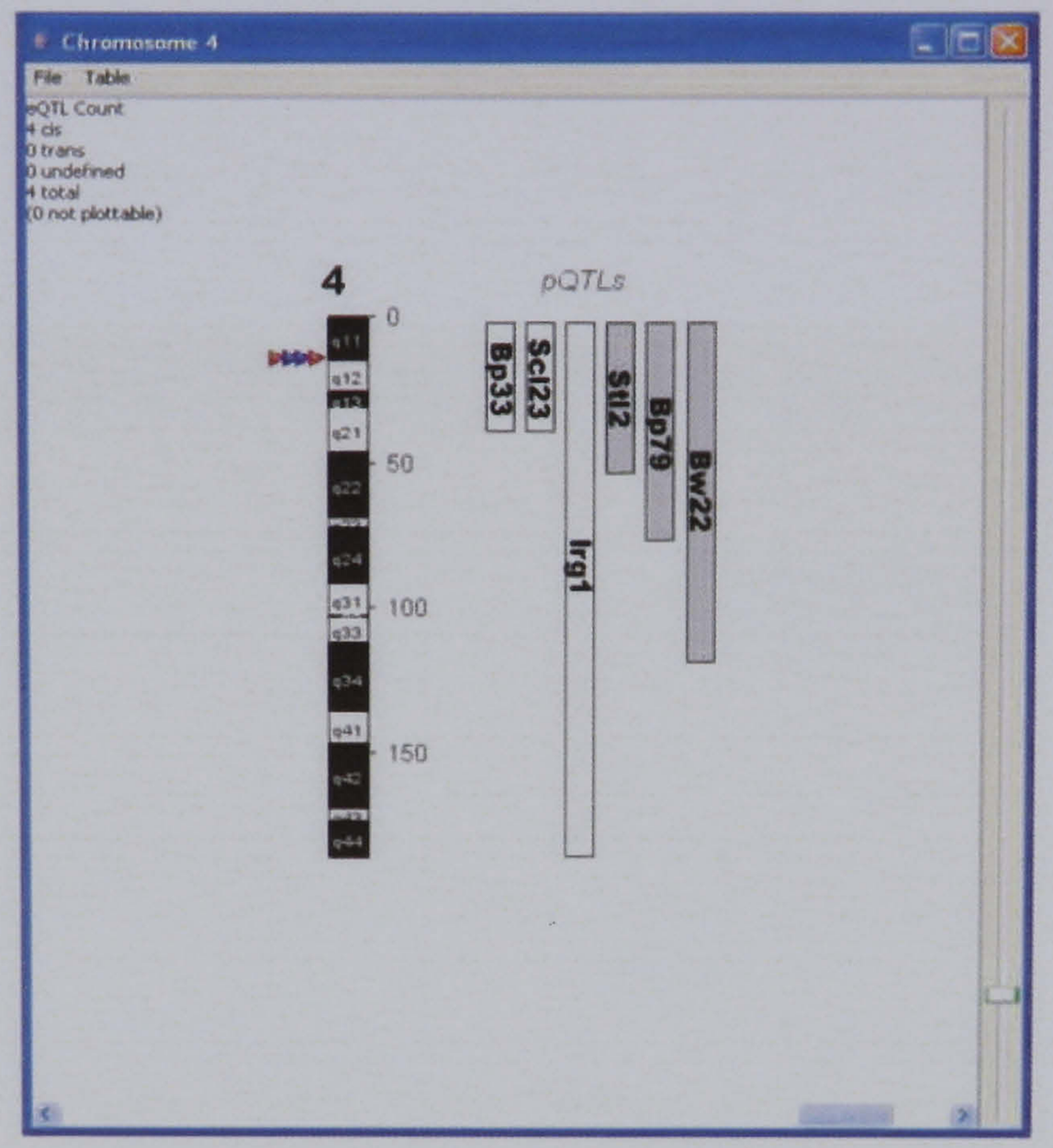


Figure 5.3 Screenshot from eQTL explorer
This screenshot shows the 4 *cis* acting eQTL which map to the *Cd36* gene in fat and kidney tissue. There are 2 eQTL in each tissue which represent the 2 probe sets for the *Cd36* gene. Also shown are the 5 pQTL that overlap this region.

Table 5.2 Human pQTL list

Human QTL number	RGD_QTL_ SYMBOL	CHROMOSOME	LOD	P_VALUE	QTL_classification
QTL1	BP17_H	1		0.0162	Blood Pressure
QTL2	BP18_H	1		0.0263	Blood Pressure
QTL3	BP25_H	1	3.5		Blood Pressure
QTL4	BP45_H	1	1.77	0.00218	Blood Pressure
QTL5	BP5_H	1	2.96		Blood Pressure
QTL6	BP9_H	1	3.9		Blood Pressure
QTL7	BW12_H	1	2.32		Body Weight
QTL8	BW19_H	1	3.7		Body Weight
QTL9	BW3_H	1	2.09	0.00095	Body Weight
QTL10	BW57_H	1	1	0.0001	Body Weight
QTL11	BW60_H	1	2.2	0.0006	Body Weight
QTL12	BW71_H	1		0.0099	Body Weight
QTL13		1	2.6		Type 2 Diabetes
QTL14		1	1.9		Type 2 Diabetes
QTL15		1	2.5	0.000078	Lipid Levels
QTL16		1	3.81		Lipid Levels
QTL17		1	4.295		Type 2 Diabetes
QTL18		1	3.7		Lipid Levels
QTL19		1	2.33		Type 2 Diabetes
QTL20		1	4.5	<.00001	Metabolic Syndrome
QTL21		1	3.71	0.00002	Metabolic Syndrome
QTL22		1	3.04	0.00018	Type 2 Diabetes
QTL23	BP10_H	2	3.92		Blood Pressure
QTL24	BP23_H	2	3.36		Blood Pressure
QTL25	BP37_H	2	2.96		Blood Pressure
QTL26	BP46_H	2	2.26	0.00062	Blood Pressure
QTL27	BW13_H	2	3.42		Body Weight
QTL28	BW2_H	2	4.04		Body Weight
QTL29	BW33_H	2	1.5	0.0075	Body Weight
QTL30	BW53_H	2	1.16		Body Weight
QTL31	SLEP2_H	2	4.95	0.0000009	Hormone level
QTL32		2	NULL	0.0089	Blood Pressure
QTL33		2	NULL	0.0193	Blood Pressure
QTL34		2	NULL	0.0353	Blood Pressure
QTL35		2	NULL	0.0007	Blood Pressure
QTL36		2	2.84	NULL	Blood Pressure
QTL37		2	2		Body Weight
QTL38		2	2.6		Lipid Levels
QTL39		2	2.3		Type 2 Diabetes
QTL40		2	3.5		Body Weight
QTL41		2	2.32	0.000009	Lipid Levels
QTL42		2	3.9		Type 2 Diabetes
QTL43		2	2.56		Obesity in T2 diabetics
QTL44		2	3.3		Obesity in T2 diabetics
QTL45		2	2.3		Lipid Levels
QTL46		2	2.68		Body Weight
QTL47		2	3.59		Blood Pressure
QTL48		2	2.22	0.0007	Metabolic Syndrome
QTL49		2	2.07	0.001	Metabolic Syndrome
QTL50		2	3.34	0.0004	Metabolic Syndrome
QTL51		2		0.0047	Blood Pressure
QTL52		2	2.7	0.0006	Type 2 Diabetes
QTL53		2	3.4		Lipid Levels
QTL54	BP24_H	3	2.9		Blood Pressure
QTL55	BP30_H	3	2.03		Blood Pressure
QTL56	BP36_H	3	4.04		Blood Pressure
QTL57	BP51_H	3	1.27	0.00787	Blood Pressure
QTL58	BFD1_H	3	3.94		Body fluid distribution
QTL59	BW10_H	3	1.8	0.002	Body Weight
QTL60	BW11_H	3	4.3	0.000004	Body Weight
QTL61	BW23_H	3	3.45		Body Weight
QTL62	BW24_H	3	2.03		Body Weight
QTL63	BW88_H	3	3.54	0.009	Body Weight
QTL64		3	3.91		Type 2 Diabetes
QTL65		3	2.64	0.0004	Lipid Levels
QTL66		3	2.6		Body Weight
QTL67		3	4.11		Lipid Levels
QTL68		3	4.67	0.000004	Type 2 Diabetes
QTL69		3	3.13	0.0001	Type 2 Diabetes
QTL70		3	2.47	0.0004	Lipid Levels
QTL71	BP1_H	4	3.9		Blood Pressure
QTL72	BP26_H	4	3.2		Blood Pressure
QTL73	BW1_H	4	9.2		Body Weight
QTL74	BW15_H	4	4.5	0.000003	Body Weight
QTL75	BW4_H	4	2.09	0.00096	Body Weight
QTL76	HRTRT1_H	4	3.9	0.00435	Heart rate

Human QTL number	RGD_QTL_ SYMBOL	CHROMOSOME	LOD	P_VALUE	QTL_classification
QTL77		4	2.55	0.266	Body Weight
QTL78		4	2.39		Obesity in T2 diabetics
QTL79		4	2.3	0.004	Type 2 Diabetes
QTL80		4	4.11		Lipid Levels
QTL81		4	2.73		Type 2 Diabetes
QTL82		4	2.44		Blood Pressure
QTL83	BP21_H	5		0.0785	Blood Pressure
QTL84	BP47_H	5	2.8	0.00016	Blood Pressure
QTL85	BFD2_H	5	4.49		Body fluid distribution
QTL86	BW61_H	5	2.4	0.0004	Body Weight
QTL87	BW72_H	5		0.0039	Body Weight
QTL88	BW8_H	5	1.9	0.0017	Body Weight
QTL89		5	2.28	0.397	Body Weight
QTL90	NULL	5	NULL	0.0076	Blood Pressure
QTL91	NULL	5	NULL	0.0309	Blood Pressure
QTL92		5	2.56		Obesity in T2 diabetics
QTL93		5	4.1	0.0000077	Hormone level
QTL94		5	2.8		Type 2 Diabetes
QTL95		5	2.36	0.0004	Type 2 Diabetes
QTL96		5	2.4	0.00044	Type 2 Diabetes
QTL97		5	3.64	0.0002	Lipid Levels
QTL98		5	3.64	0.0000212	Type 2 Diabetes
QTL99	BP4_H	6	2.5		Blood Pressure
QTL100	BFD3_H	6	2.79		Body fluid distribution
QTL101	BFD4_H	6	2.52		Body fluid distribution
QTL102	BW14_H	6	1.24		Body Weight
QTL103	BW5_H	6	1.77	0.0021	Body Weight
QTL104	BW54_H	6	1.53		Body Weight
QTL105	BW9_H	6	2.7	0.0002	Body Weight
QTL106	NULL	6	NULL	<0.0001	Blood Pressure
QTL107		6	2.49	0.297	Body Weight
QTL108		6	2.2		Body Weight
QTL109		6	2.9		Blood Pressure
QTL110		6	4.9		Type 2 Diabetes
QTL111		6	4.2		Type 2 Diabetes
QTL112		6	5.5		Body Weight
QTL113		6	26.99		Lipid Levels
QTL114		6	2.96		Type 2 Diabetes
QTL115		6	3.21	0.042	Blood Pressure
QTL116		6	3.3		Body Weight
QTL117		6	2.99	0.0001	Metabolic Syndrome
QTL118		6	2.58	0.0003	Metabolic Syndrome
QTL119		6	2.9		Lipid Levels
QTL120		6	2.66	0.00023	Type 2 Diabetes
QTL121		6	3.7		Lipid Levels
QTL122		6	4.1	0.0001	Type 2 Diabetes
QTL123		6	3.1		Lipid Levels
QTL124	BP48_H	7	2.26	0.00063	Blood Pressure
QTL125	BFD5_H	7	3.16		Body fluid distribution
QTL126	BW18_H	7	2.26	0.0006	Body Weight
QTL127	BW25_H	7	2.66		Body Weight
QTL128	BW26_H	7	2.36		Body Weight
QTL129	BW27_H	7	1.29		Body Weight
QTL130	BW35_H	7	3.83		Body Weight
QTL131	BW68_H	7		0.0011	Body Weight
QTL132		7	3.7		Body Weight
QTL133		7	4.73		Blood Pressure
QTL134		7	2.6		Blood Pressure
QTL135		7	3.2		Type 2 Diabetes
QTL136		7	1.9		Type 2 Diabetes
QTL137		7	2.72		Body Weight
QTL138		7	4.9	0.00001	Body Weight
QTL139		7	2.4		Lipid Levels
QTL140		7	2.2		Body Weight
QTL141		7	2.3		Body Weight
QTL142		7	2.05		Lipid Levels
QTL143	BP11_H	8	1.92		Blood Pressure
QTL144	BP19_H	8	0.0284		Blood Pressure
QTL145	BP2_H	8	1.8		Blood Pressure
QTL146	BP49_H	8	2.24	0.00066	Blood Pressure
QTL147	BW20_H	8	3.43		Body Weight
QTL148	BW52_H	8	3.21		Body Weight
QTL149	BW59_H	8	2	0.001	Body Weight
QTL150	BW69_H	8		0.0065	Body Weight
QTL151	BW7_H	8	1.95	0.014	Body Weight
QTL152	SLEP1_H	8	3.4	0.00004	Hormone level
QTL153		8	4.87	0.000001	Lipid Levels

Human QTL number	RGD_QTL_ SYMBOL	CHROMOSOME	LOD	P_VALUE	QTL_classification
QTL154		8	3.1	0.0003	Type 2 Diabetes
QTL155		8	3.09		Type 2 Diabetes
QTL156		8	3.6		Lipid Levels
QTL157		8	3.3		Lipid Levels
QTL158		8	2.55	0.00036	Type 2 Diabetes
QTL159	BW62_H	9	2.3		Body Weight
QTL160	NULL	9	NULL	0.0171	Blood Pressure
QTL161		9	2.8		Type 2 Diabetes
QTL162		9	2		Type 2 Diabetes
QTL163		9	3.9	0.0002	Type 2 Diabetes
QTL164		9	3.4	0.00004	Lipid Levels
QTL165	BP52_H	10	1.41	0.00547	Blood Pressure
QTL166	BP6_H	10	2.5		Blood Pressure
QTL167	BW16_H	10	1.46		Body Weight
QTL168	BW22_H	10	2.32		Body Weight
QTL169	BW63_H	10	2.7		Body Weight
QTL170		10	2.9	0.153	Body Weight
QTL171		10	3.75		Type 2 Diabetes
QTL172		10	2.39		Type 2 Diabetes
QTL173		10	4.85		Body Weight
QTL174		10	2.07		Type 2 Diabetes
QTL175		10	3.11	0.0001	Type 2 Diabetes
QTL176		10	3.2		Lipid Levels
QTL177	BP22_H	11		0.004	Blood Pressure
QTL178	BP31_H	11	2.07		Blood Pressure
QTL179	BW36_H	11	3.35		Body Weight
QTL180	BW64_H	11	2.2		Body Weight
QTL181	BW70_H	11		0.0079	Body Weight
QTL182		11	2.6		Lipid Levels
QTL183		11	3.22	0.001213	Lipid Levels
QTL184		11	2.11	0.000009	Lipid Levels
QTL185		11	2.5	0.0007	Blood Pressure
QTL186		11	3.6		Lipid Levels
QTL187		11	3.08	0.00017	Type 2 Diabetes
QTL188		11	3.21		Type 2 Diabetes
QTL189		11	2.21		Type 2 Diabetes
QTL190		11	2.98	0.00011	Type 2 Diabetes
QTL191	BP53_H	12	1.33	0.00673	Blood Pressure
QTL192	BP7_H	12	3.44		Blood Pressure
QTL193	BW37_H	12	4.08	0.00001	Body Weight
QTL194	BW58_H	12	2.3	0.0001	Body Weight
QTL195	BW65_H	12	2.2		Body Weight
QTL196		12	2.6		Body Weight
QTL197		12	2.2		Type 2 Diabetes
QTL198		12	4.06	0.000746	Lipid Levels
QTL199		12	2.81		Type 2 Diabetes
QTL200		12	2.05	0.0011	Metabolic Syndrome
QTL201		12	2.56		Type 2 Diabetes
QTL202		13	2.98		Type 2 Diabetes
QTL203		13	3.2	0.00006	Body Weight
QTL204		13	2.36	0.004	Lipid Levels
QTL205		13	3.28		Type 2 Diabetes
QTL206		13	2.86		Type 2 Diabetes
QTL207		13	3.2		Blood Pressure
QTL208	BP54_H	14	1.43	0.00514	Blood Pressure
QTL209	BW28_H	14	2.15		Body Weight
QTL210	BW66_H	14	2.4		Body Weight
QTL211		14	3.9	0.00015	Myocardial infarction
QTL212		14	3.2	0.0000062	Hormone level
QTL213		14	2.74	0.00019	Type 2 Diabetes
QTL214	BP32_H	15	2.69		Blood Pressure
QTL215	BP8_H	15		0.000016	Blood Pressure
QTL216	NULL	15	NULL	0.0033	Blood Pressure
QTL217	NULL	15	2.94	NULL	Blood Pressure
QTL218		15	2.3		Body Weight
QTL219		15	3.29	0.000054	Lipid Levels
QTL220		15	2		Type 2 Diabetes
QTL221		15	3.11	0.000288	Lipid Levels
QTL222		15	2.19	0.00138	Type 2 Diabetes
QTL223		15	2.28		Type 2 Diabetes
QTL224		15	3.88		Lipid Levels
QTL225	BP27_H	16	3.2		Blood Pressure
QTL226	BP33_H	16	2.74		Blood Pressure
QTL227	BW29_H	16	2.55		Body Weight
QTL228	NULL	16	NULL	0.0314	Blood Pressure
QTL229		16	2.1		Body Weight
QTL230		16	3.21		Body Weight

Human QTL number	RGD_QTL_ SYMBOL	CHROMOSOME	LOD	P_VALUE	QTL_classification
QTL231		16	4.33	0.0000008	Lipid Levels
QTL232		16	3.6	0.00002	Lipid Levels
QTL233	BP14_H	17	1.9	0.0029	Blood Pressure
QTL234	BP15_H	17		0.025	Blood Pressure
QTL235	BP16_H	17		0.0004	Blood Pressure
QTL236	BP20_H	17		0.0256	Blood Pressure
QTL237	BP34_H	17	2.16		Blood Pressure
QTL238	BP35_H	17		0.0001	Blood Pressure
QTL239	BW30_H	17	2.47		Body Weight
QTL240	BW31_H	17	1.87		Body Weight
QTL241	BW55_H	17	1.34		Body Weight
QTL242	BW56_H	17	2.33		Body Weight
QTL243	SLEP3_H	17	4.97	0.0003	Hormone level
QTL244		17	3	0.0119	Body Weight
QTL245		17	3	0.003	Type 2 Diabetes
QTL246		17	2.29		Lipid Levels
QTL247		17	2.48		Lipid Levels
QTL248		17	2.4		Body Weight
QTL249		17	3.3	0.0001	Type 2 Diabetes
QTL250		17	3.2		Type 2 Diabetes
QTL251	BP13_H	18	2.09		Blood Pressure
QTL252	BW32_H	18	2.3	0.00097	Body Weight
QTL253	BW67_H	18	1.9	0.001	Body Weight
QTL254	NULL	18	NULL	0.016	Blood Pressure
QTL255	NULL	18	4.6	NULL	Blood Pressure
QTL256	NULL	18	2.6	NULL	Blood Pressure
QTL257	NULL	18	3.21	NULL	Blood Pressure
QTL258		18	3.53		Body Weight
QTL259		18	4.6	0.0000021	Blood Pressure
QTL260		18	3.79		Type 2 Diabetes
QTL261		18	2.6		Blood Pressure
QTL262		18	2.05		Lipid Levels
QTL263	BP3_H	19	2.4		Blood Pressure
QTL264	BP50_H	19	2.1	0.00094	Blood Pressure
QTL265	BW34_H	19	1.5	0.0063	Body Weight
QTL266		19	3.59	0.042	Body Weight
QTL267		19	2.6		Type 2 Diabetes
QTL268		19	3.59	0.000005	Lipid Levels
QTL269		19	2.28	0.0004	Type 2 Diabetes
QTL270		19	3.89	0.00002	Lipid Levels
QTL271		19	2.23	0.0007	Lipid Levels
QTL272	BP29_H	20			Blood Pressure
QTL273	BW17_H	20	1.5		Body Weight
QTL274	BW6_H	20	1.64	0.003	Body Weight
QTL275		20	3.06	0.00009	Body Weight
QTL276	NULL	20	NULL	0.024	Blood Pressure
QTL277		20	2.3		Type 2 Diabetes
QTL278	BP12_H	21	2.82		Blood Pressure
QTL279	NULL	21	NULL	0.0234	Blood Pressure
QTL280		21	3.9		Lipid Levels
QTL281	BP38_H	22	2.07		Blood Pressure
QTL282	BP55_H	22	1.54	0.00383	Blood Pressure
QTL283		22	3.4	0.0001	Type 2 Diabetes
QTL284		22	2.49	0.0004	Metabolic Syndrome
QTL285	BP28_H	X	3.3		Blood Pressure
QTL286	BP39_H	X	2.41		Blood Pressure

Footnotes
Shading is to separate chromosomes
An extended version can be found on data CD containg more description and references

5.2.2 Identification of *cis*-acting human candidate genes

cis-acting human candidate genes were identified by first determining the region of the rat genome which were syntenic to the 286 human QTL in **Table 5.2** and identifying any *cis*-acting eQTL within these regions. This analysis was carried out in collaboration with Nicholas J Dickens of the Physiological Genomics and Medicine group. I defined the strategy and thus algorithm for identifying syntenic regions and then mapping eQTL within these regions and looking for human orthologs and Nicholas Dickens automated the process. The 286 human QTL identified as described above were mapped onto the rat genome and 253 of these have a defined syntenic rat region. *cis*-acting eQTL with a genome wide significant *P*-value which were located within these syntenic regions were then identified. This detected 711 eQTL which had complete marker and probe set mapping information within 163 of the human QTL. A further 93 eQTL were identified; however these were discarded due to insufficient mapping information, such as that the markers or probe set position could not be determined. Therefore the syntenic regions and eQTL information could not be derived. The rat gene corresponding to the eQTL was determined using the Ensembl annotations of the Affymetrix probe sets. The human orthologs of these genes were identified and checked to ensure they were located within the corresponding human QTL. This data was collapsed for presentation in the thesis (complete table on data CD) and **Table 5.3** contains 142 rat eQTL (with the human and rat orthologs and their GO classifications) that are contained within human QTL and are therefore candidate genes for human disease. The rat pQTL are also shown in this table and 34 of these 142 genes do not map within a rat QTL for metabolic syndrome traits. The table was collapsed because many of the rat eQTL mapped to multiple human QTL. Also 23 rat eQTL occur in both kidney and fat tissue and are

classified as shared in **Table 5.3** while 53 and 71 rat eQTL were unique to fat or kidney respectively. During this analysis a further 8 probe sets were removed for various reasons, which are shown in **Table 5.4**. Of the 142 eQTL linkages in **Table 5.3** there are 135 unique probe sets and of these 98 and 93 have Ensembl gene

Table 5.3 Gene List generated by comparative mapping

Human pQTL Name	Human pQTL Type	Human Chromosome	Overlapping eQTL - Tissue	Overlapping eQTL - probe set	Overlapping eQTL - Chromosome	Overlapping eQTL - Linked marker	eQTL P-Value (fat, kidney)	Rat pQTL	Rat Genes in this region	Homologous Human Genes
QTL125_BFD5_H; QTL141; QTL142;	Body fluid distribution; Body Weight; Lipid Levels	7 shared	1367589_a_at	4 D4S3	1e-06 3e-06	Bp79, Bw22, Sd2	ENSRNOG00000005908|CD36_RAT|Platelet glycoprotein IV (GPIV) (CD36 antigen) (PAS IV) (PAS- 4 protein) (Fatty acid transport protein) (Fatty acid translocase) (Adipocyte membrane protein) [Source:UniprotSWISSPROT Acc:Q07969|]	ENSG00000135218|CD36|Platelet glycoprotein IV (GPIV) (Glycoprotein IIIb) (GPIIIB) (Leukocyte differentiation antigen CD36) (CD36 antigen) (PAS IV) (PAS- 4 protein) (Platelet collagen receptor) (Fatty acid translocase) (FAT) (Thrombospondin receptor). [Source:UniprotSWISSPROT Acc:P16671|]
QTL282_BP55_H; QTL284;	Blood Pressure; Metabolic Syndrome	22, kidney	1367917_at	7 D7Rat129		0.000001	Lvm7 (Cm27)	ENSRNOG00000005988|CP2D2_RAT|Cytochrome P450 2D2 (EC 1.14.14.1) (CYPIID2) (P450-DB2) (P450-CMF2) (Debrisoquine 4-hydroxylase) [Source:UniprotSWISSPROT Acc:P10834|] ENSRNOG00000032165|DHCA_RAT|Carbonyl reductase [NADPH] 1 (EC 1.1.1.184) (NADPH-dependent carbonyl reductase 1). [Source:UniprotSWISSPROT Acc:P47727|],ENSRNOG00000003283|Q9QZI4_RAT|Carbonyl reductase isoform I (Fragment) [Source:UniprotSPTREMBL Acc:Q9QZI6|]	ENSG00000100197|CYP2D6|Cytochrome P450 2D6 (EC 1.14.14.1) (CYPIID6) (P450-DB1) (Debrisoquine 4-hydroxylase) [Source:UniprotSWISSPROT Acc:P10635|] ENSG00000159228|CBR1|Carbonyl reductase [NADPH] 1 (EC 1.1.1.184) (NADPH-dependent carbonyl reductase 1) (Prostaglandin-E(2) 9-reductase) (EC 1.1.1.189) (Prostaglandin 9-ketoreductase) (15-hydroxyprostaglandin dehydrogenase [NADP+]) (EC 1.1.1.197). [Source:UniprotSWISSPROT Acc:P16152|]
QTL278_BP12_H; QTL280;	Blood Pressure; Lipid Levels;	21 kidney	1368037_at	11 D11Gabr156		0.000034056	NONE			
QTL234_BP15_H; QTL237_BP34_H; QTL239_BW30_H; QTL241_BW55_H; QTL243_SLEP3_H; QTL250;	Blood Pressure; Blood Pressure; Body Weight; Body Weight; Hormone level; Type 2 Diabetes;	17 fat	1368785_at	10 D10Wox13		0.000006	Tbs1, BpQTLcluster9	ENSRNOG00000011357|SHBG_RAT|Sex hormone-binding globulin (SHBG) (Sex steroid-binding protein) (SBP) (Testis-specific androgen-binding protein) (ABP) [Source:UniprotSWISSPROT Acc:P06686|]	ENSG00000129214|SHBG|Sex hormone-binding globulin precursor (SHBG) (Sex steroid-binding protein) (SBP) (Testis-specific androgen-binding protein) (ABP) (Testosterone-estrogen binding globulin) (Testosterone-retradiol binding globulin) (TeBG) [Source:UniprotSWISSPROT Acc:P04278|]
QTL178_BP31_H; QTL179_BW36_H; QTL180_BW64_H; QTL181_BW70_H;	Blood Pressure; Body Weight; Body Weight; Body Weight;	11 fat	1368291_at	8 D8Rat58		0.000009	Inf1	ENSRNOG00000010602|NP_068520.1|apoptosis inhibitor 2 [Source:RefSeq_peptide;Acc:NP_068520|]	ENSG00000110330|BIRC2|Baculoviral IAP repeat-containing protein 2 (Inhibitor of apoptosis protein 2) (IAP2) (HIAP-2) (C-IAP1) (TNFR2-TRAF signaling complex protein 2) (IAP homolog B). [Source:UniprotSWISSPROT Acc:Q13490|] ENSG00000007933|FMO3|Dimethylaniline monooxygenase [N-oxide-forming] 3 (EC 1.14.13.8) (Hepatic flavin-containing monooxygenase 3) (FMO 3) (Dimethylaniline oxidase 3) (FMO form 2) (FMO II). [Source:UniprotSWISSPROT Acc:P31513|]
QTL1, BW19_H; QTL10_SW57_H; QTL11_SW60_H; QTL14; QTL20; QTL21 QTL22	Body Weight; Body Weight; Body Weight; Type 2 Diabetes; Metabolic Syndrome; Metabolic Syndrome; Type 2 Diabetes;	1 fat	1368304_at	19 D19M43		0.000005	Bp80	ENSRNOG00000009620|XP_345449.1|ENSRNOG00000007086|LGMN_RAT|Legumain (EC 3.4.22.34) (Asparaginyl endopeptidase) (Protease, cystatin 1) [Source:UniprotSWISSPROT Acc:O88QJ8|]	ENSG00000100600|LGMN|Legumain precursor (EC 3.4.22.34) (Asparaginyl endopeptidase) (Protease, cystatin 1) [Source:UniprotSWISSPROT Acc:Q99538|]
QTL206_BW28_H; QTL211; QTL213;	Body Weight; Myocardial infarction; Type 2 Diabetes;	14 kidney	1368430_at	6 D6Rat111		0.0000142767	NONE	ENSRNOG00000000408|NP_037005.1|v-ros UR2 sarcoma virus oncogene homolog 1, Rat heart - derived c - ros - 1 proto - onco [Source:RefSeq_peptide;Acc:NP_037005|]	ENSG00000047936|ROS1|Proto-oncogene tyrosine-protein kinase ROS precursor (EC 2.7.1.112) (c-ros-1) [Source:UniprotSWISSPROT Acc:P08922|]
QTL101_BFD4_H; QTL121;	Body fluid distribution; Lipid Levels;	6 kidney	1369432_a_at	20 D20Rat9		0.000009	NONE	ENSRNOG00000007008|SLC31_RAT|Neutral and basic amino acid transport protein rBAT (B(0,+)-type amino acid transport protein) (NAA-TR) (D2) [Source:UniprotSWISSPROT Acc:O54319|]	ENSG00000138079|SLC3A1|Neutral and basic amino acid transport protein rBAT (B(0,+)-type amino acid transport protein) (NBAT) (D2H) [Source:UniprotSWISSPROT Acc:Q07837|]
QTL30_BW53_H; QTL31_SLEP2_H; QTL32; QTL43; QTL44; QTL46;	Body Weight; Hormone level; Blood Pressure; Obesity in T2 diabetes; Obesity in T2 diabetes; Metabolic Syndrome;	2 fat	1369440_at	6 D6Cebrp424s2		0.0000720716	NONE	ENSRNOG00000018325|THRB_RAT|Prothrombin (EC 3.4.21.5) [Source:UniprotSWISSPROT Acc:P18292|]	ENSG00000180210|F2|Prothrombin precursor (EC 3.4.21.5) (Coagulation factor II) [Source:UniprotSWISSPROT Acc:P00734|]
QTL178_BP31_H; QTL182; QTL178_BP31_H; QTL178_BW36_H; QTL180_BW64_H; QTL181_BW70_H; QTL31_SLEP2_H;	Blood Pressure; Lipid Levels;	11 kidney	1368442_at	3 D3Rat35		0.000003	Anr10	ENSRNOG00000032070|Q8JJ79_RAT|Cytoplasmic dynein heavy chain [Source:UniprotSPTREMBL Acc:Q8JJ79|]	
	Blood Pressure; Body Weight; Body Weight; Body Weight; Hormone level;	11 kidney	1368676_at	8 D8Rat56		0.000001	Inf1		
QTL209_BW28_H; QTL211; QTL213;	Body Weight; Myocardial infarction; Type 2 Diabetes;	14 kidney	1368859_at	6 D6Cebr36a1		0.0000656364	NONE	ENSRNOG00000005918|PP2CA_RAT|Protein phosphatase 2C alpha isoform (EC 3.1.3.16) (PP2C-alpha) (IA) (Protein phosphatase 1A) [Source:UniprotSWISSPROT Acc:P20650|]	ENSG00000100814|PPM1A|Protein phosphatase 2C alpha isoform (EC 3.1.3.16) (PP2C-alpha) (IA) (Protein phosphatase 1A) [Source:UniprotSWISSPROT Acc:P35813|]
QTL231; QTL232;	Lipid Levels; Lipid Levels;	18 shared	1368891_at	19 D19Cebrp150s1	1e-07 5e-06	WZ, BpJ2	ENSRNOG00000019208|GNPAT_RAT|Dihydroxyacetone phosphate acyltransferase (EC 2.3.1.42) (DHAP-AT) (DAP-AT) (Glycerone-phosphate O-acyltransferase) (Acyl-CoA dihydroxyacetonephosphateacyltransferase) [Source:UniprotSWISSPROT Acc:Q9ES71|]	ENSG00000116906|GNPAT|Dihydroxyacetone phosphate acyltransferase (EC 2.3.1.42) (DHAP-AT) (DAP-AT) (Glycerone-phosphate O-acyltransferase) (Acyl-CoA dihydroxyacetonephosphateacyltransferase). [Source:UniprotSWISSPROT Acc:O15228|]
QTL178_BP31_H; QTL179_BW36_H; QTL180_BW64_H; QTL181_BW70_H;	Blood Pressure; Body Weight; Body Weight; Body Weight;	11 fat	1369171_at	8 D8Rat102		0.000005	Inf1	ENSRNOG00000019560|NP_077328.1|macrophage stimulating 1 (hepatocyte growth factor-like) [Source:RefSeq_peptide;Acc:NP_077328|]	ENSG00000173531|MST1|Hepatocyte growth factor-like protein precursor (Macrophage stimulatory protein) (MSP) (Macrophage stimulating protein) [Source:UniprotSWISSPROT Acc:P26927|]
								ENSRNOG00000010886|SLIM3_RAT|Skeletal muscle LIM-protein 3 (SLIM 3) (LIM-domain protein DRAL) (Four and a half LIM domains protein 2) (FHL-2) [Source:UniprotSWISSPROT Acc:O35115|]	ENSG00000115641|FHL2|Skeletal muscle LIM-protein 3 (SLIM 3) (LIM-domain protein DRAL) (Four and a half LIM domains protein 2) (FHL-2) [Source:UniprotSWISSPROT Acc:Q14192|]
QTL28_BP46_H	Blood Pressure	2 fat	1369313_at	9 D9Rat104	0.0000035088	Bp46, Bw11, Sd22, Chr18 congenic (Bp), Bp113 (Cm11), Bp53 (Cm25)	ENSRNOG00000017840|NP_075247.1|phosphatidylinositol 3-kinase [Source:RefSeq_peptide;Acc:NP_075247|]	ENSG00000078142|PIK3C3|phosphoinositide-3-kinase, class 3 [Source:RefSeq_peptide;Acc:NP_002638|]	
QTL253_BW87_H; QTL257; QTL258	Body Weight; Blood Pressure; Body Weight;	18 kidney	1369655_at	18 D18Rat103	0.0000001	Bp46			

Human pQTL Name	Human pQTL Type	Human Chromosome	Overlapping eQTL - Tissue	Overlapping eQTL - probe set	Overlapping eQTL - Chromosome	Overlapping eQTL - Linked marker	eQTL P-Value (fat, kidney)	Rat pQTL	Rat Genes in this region	Homologous Human Genes
QTL214_BP32_H	Blood Pressure	15	kidney	1369665_a_at	8	D8Mit12	0.000028205	Chr8 congenic (Bp, cardiac mass, HDL phospholipid, serum cholesterol), Cardiac002, Bp35, Bp82	ENSRNOG00000009848|IL18_RAT|Interleukin-18 (IL-18) (Interferon-gamma inducing factor) (IFN-gamma-inducing factor) (Interleukin-1 gamma) (IL-1 gamma). [Source:Uniprot/SWISSPROT;Acc:P97636|]	ENSG00000150782|IL18|Interleukin-18 precursor (IL-18) (Interferon-gamma inducing factor) (IFN-gamma-inducing factor) (Interleukin-1 gamma) (IL-1 gamma). [Source:Uniprot/SWISSPROT;Acc:Q14116|]
QTL10_BW57_H	Body Weight	1	fat	1369866_at	2	D2Rat236	0.000001	Bp16, Bp101, BpQTLcluster3, Bp63	ENSRNOG00000018236|NP_064476.1|prochymosin [Source:RefSeq_peptide;Acc:NP_064476|]	ENSG00000168607||
QTL24_BP23_H; QTL26_BP46_H; QTL42; QTL50;	Blood Pressure; Blood Pressure; Type 2 Diabetes; Metabolic Syndrome;	2	fat	1370176_at	9	D9Cebrp60s12	0.000003	Bp53 (Cm28), Bp106	ENSRNOG00000010881|AL253_RAT|Amyotrophic lateral sclerosis 2 chromosomal region candidate (GABA-A receptor interacting factor-1) (GRIF-1) (O GlcNAc transferase-interacting protein of 98 kDa). [Source:Uniprot/SWISSPROT;Acc:Q8R2H7|]	ENSG00000115993|ALS2CR3|Amyotrophic lateral sclerosis 2 chromosomal region candidate gene protein 3. [Source:Uniprot/SWISSPROT;Acc:O60296|] ENSG00000144867|SRPRB|Signal recognition particle receptor beta subunit (SR-beta) (Protein APMCF1). [Source:Uniprot/SWISSPROT;Acc:Q9Y5M8|],ENSG00000091513|TF|Serotransferrin precursor (Transferrin) (Siderophilin) (Beta-1-metal binding globulin) (PRO1400). [Source:Uniprot/SWISSPROT;Acc:P02787|]
QTL81_BW23_H	Body Weight	3	kidney	1370226_at	6	D6Rat123	0.000003	Cm19 (Hw4)	ENSRNOG00000030825|TRFE_RAT|Serotransferrin (Transferrin) (Siderophilin) (Beta-1-metal binding globulin). [Source:Uniprot/SWISSPROT;Acc:P12346|]	
QTL282_BP55_H; QTL284;	Blood Pressure; Metabolic Syndrome;	22	kidney	1370249_at	7	D7Rat129	0.000013129	Lvm7 (Cm27)	ENSRNOG00000010549|BZRP_RAT|Peripheral-type benzodiazepine receptor (PBR) (PKBS) (Mitochondrial benzodiazepine receptor). [Source:Uniprot/SWISSPROT;Acc:P16257|]	ENSG00000100300|BZRP|Peripheral-type benzodiazepine receptor (PBR) (PKBS) (Mitochondrial benzodiazepine receptor). [Source:Uniprot/SWISSPROT;Acc:P30536|]
QTL282_BP55_H; QTL284;	Blood Pressure; Metabolic Syndrome;	22	fat	1370377_at	7	D7Rat129	0.000053571	Lvm7 (Cm27)	ENSRNOG00000030182|CP2D1_RAT|Cytochrome P450 2D1 (EC 1.14.14.1) (CYPIID1) (P450-DB1) (P450-CMF1A) (P450-UT-7) (Debrisoquine 4-hydroxylase). [Source:Uniprot/SWISSPROT;Acc:P10833|],ENSRNOG00000029128|CP2D5_RAT|Cytochrome P450 2D5 (EC 1.14.14.1) (CYPIID5) (P450-DB5) (P450-CMF1B) (Debrisoquine 4-hydroxylase). [Source:Uniprot/SWISSPROT;Acc:P12939|]	ENSG00000100197|CYP2D6|Cytochrome P450 2D6 (EC 1.14.14.1) (CYPIID6) (P450-DB1) (Debrisoquine 4-hydroxylase). [Source:Uniprot/SWISSPROT;Acc:P10635|]
QTL171	Type 2 Diabetes	10	kidney	1370671_at	1	D1Rat225	0.0000001	Map1 (Bp196), Iddm10, Scl5/Sll3	ENSRNOG00000015724|GCYK_RAT|Kinase-like domain containing soluble guanylyl cyclase (EC 4.6.1.2) (48 kDa chain) (KSGC). [Source:Uniprot/SWISSPROT;Acc:P55205|]	ENSG00000080218||
QTL234_BP15_H; QTL237_BP34_H; QTL239_BW30_H; QTL241_BW55_H; QTL243_SLEP3_H; QTL250;	Blood Pressure; Blood Pressure; Body Weight; Body Weight; Hormone level; Type 2 Diabetes;	17	shared	1370930_at	10	D10Wox13	5e-06;1e-06	Trs1, BpQTLcluster9	ENSRNOG00000031364|KIF1D_RAT|Kinesin-like protein KIF1D. [Source:Uniprot/SWISSPROT;Acc:O35787|] ENSRNOG00000026808|21 kDa protein [Source:IPI;Acc:IPI00478777|],ENSRNOG00000032506|21 kDa protein [Source:IPI;Acc:IPI00188771|]	ENSG00000129250|KIF1C|Kinesin-like protein KIF1C. [Source:Uniprot/SWISSPROT;Acc:O43896|] ENSG00000198203|SULT1C1|Sulfotransferase 1C1 (EC 2.8.2.-) (SULT1C#1) (ST1C2) (humSULTC2). [Source:Uniprot/SWISSPROT;Acc:O00338|]
QTL99_BP4_H	Blood Pressure	6	kidney	1370943_at	9	D9Rat88	0.000003	Cm4 (SCOR50, 88, 118)		
QTL163	Type 2 Diabetes	9	kidney	1371251_at	5	D5Mit1	0.000024194	NONE	ENSRNOG00000014788|GALT_RAT|Galactose-1-phosphate uridylyltransferase (EC 2.7.7.12) (Gal-1-P uridylyltransferase) (UDP-glucose--hexose-1-phosphate uridylyltransferase). [Source:Uniprot/SWISSPROT;Acc:P43424|]	ENSG00000137104|GALT|Galactose-1-phosphate uridylyltransferase (EC 2.7.7.12) (Gal-1-P uridylyltransferase) (UDP-glucose--hexose-1-phosphate uridylyltransferase). [Source:Uniprot/SWISSPROT;Acc:P07902|]
QTL177_BP22_H; QTL178_BP31_H; QTL179_BW36_H; QTL180_BW64_H; QTL186;	Blood Pressure; Blood Pressure; Body Weight; Body Weight; Lipid Levels;	11	fat	1371442_at	8	D8Cebrp1s4	0.000003486	Chr8 congenic (Bp, cardiac mass, HDL phospholipid, serum cholesterol), Cardiac002, Bp22, Bp35	ENSRNOG00000010844|OXRP_RAT|150 kDa oxygen-regulated protein (Orp150) (Hypoxia up-regulated 1). [Source:Uniprot/SWISSPROT;Acc:Q63617|]	ENSG00000149428|HYOU1|150 kDa oxygen-regulated protein precursor (Orp150) (Hypoxia up-regulated 1). [Source:Uniprot/SWISSPROT;Acc:Q9Y4L1|]
QTL177_BP22_H; QTL178_BP31_H; QTL179_BW36_H; QTL180_BW64_H; QTL186;	Blood Pressure; Blood Pressure; Body Weight; Body Weight; Lipid Levels;	11	shared	1371442_at	8	D8Mit1	2.43e-06;0.0001	Chr8 congenic (Bp, cardiac mass, HDL phospholipid, serum cholesterol), Cardiac002, Bp35, Bp82	ENSRNOG00000010844|OXRP_RAT|150 kDa oxygen-regulated protein (Orp150) (Hypoxia up-regulated 1). [Source:Uniprot/SWISSPROT;Acc:Q63617|]	ENSG00000149428|HYOU1|150 kDa oxygen-regulated protein precursor (Orp150) (Hypoxia up-regulated 1). [Source:Uniprot/SWISSPROT;Acc:Q9Y4L1|]
QTL178_BP31_H; QTL184; QTL185; QTL187;	Blood Pressure; Lipid Levels; Blood Pressure; Type 2 Diabetes;	11	kidney	1371611_at	3	D3Rat29	0.000005	NONE	ENSRNOG00000006944|XP_230303.2|similar to exostoses (multiple) 2 [Source:RefSeq_peptide;Acc:XP_230303|]	ENSG00000151348|EXT2|Exostosin-2 (EC 2.4.1.224) (EC 2.4.1.225) (Glucuronosyl-N-acetylglucosaminyl-proteoglycan/N-acetylglucosaminyl-proteoglycan 4- alpha-N-acetylglucosaminyltransferase) (Putative tumor suppressor protein EXT2) (Multiple exostoses protein 2). [Source:Uniprot/SWISSPROT;Acc:Q93063|]
QTL178_BP31_H; QTL179_BW36_H; QTL180_BW64_H; QTL187; QTL180;	Blood Pressure; Body Weight; Body Weight; Type 2 Diabetes; Type 2 Diabetes;	11	kidney	1371615_at	1	D1Rat47	0.000035484	Weight2 (Bw18), Kw, Lvm2 (Cm23), Bp26, BpQTLcluster1, Bp44	ENSRNOG00000016573|XP_341888.1|similar to diacylglycerol O-acyltransferase 2; diacylglycerol acyltransferase 2 [Source:RefSeq_peptide;Acc:XP_341888|]	ENSG00000062282|DGAT2|diacylglycerol O-acyltransferase homolog 2 [Source:RefSeq_peptide;Acc:NP_115953|]
QTL237_BP34_H; QTL245; QTL249;	Blood Pressure; Type 2 Diabetes; Type 2 Diabetes;	17	kidney	1371645_at	10	D10Mit2	0.000001	Trs1, BpQTLcluster9	ENSRNOG00000012121|XP_213377.1|similar to Stromal cell-derived factor 2 (SDF-2) [Source:RefSeq_peptide;Acc:XP_213377|]	ENSG00000132581|SDF2|Stromal cell-derived factor 2 precursor (SDF-2). [Source:Uniprot/SWISSPROT;Acc:Q99470|]
QTL8_BW19_H; QTL10_BW57_H; QTL11_BW60_H; QTL14; QTL20; QTL21; QTL22;	Body Weight; Body Weight; Body Weight; Type 2 Diabetes; Metabolic Syndrome; Metabolic Syndrome; Type 2 Diabetes;	1	fat	1371732_at	13	D13Rat131	0.000012036	Bp80	ENSRNOG00000002947|XP_213925.1|similar to early quiescence protein-1 [Source:RefSeq_peptide;Acc:XP_213925|]	ENSG00000143196|DPT|Dermatopontin precursor (Tyrosine-rich acidic matrix protein) (TRAMP). [Source:Uniprot/SWISSPROT;Acc:Q07507|]

Human pQTL Name	Human pQTL Type	Human Chromosome	Overlapping eQTL - Tissue	Overlapping eQTL - probe set	Overlapping eQTL - Chromosome	Overlapping eQTL - Linked marker	eQTL P-Value (fat, kidney)	Rat pQTL	Rat Genes in this region	Homologous Human Genes
QTL282_BP55_H; QTL284;	Blood Pressure; Metabolic Syndrome;	22 kidney	1371749_at		7 D7Rat129	0.000071006	Lvm7 (Cm27)	ENSRNOG00000002846[XP_343306.1[similar to 2210021J22Rik protein [Source:RefSeq_peptide;Acc:XP_343306]]	
QTL86_BW81_H	Body Weight	5 fat	1371803_at		10 D10Rat215	0.00001757	Cmr1 (Liver Cholesterol, Trs1, BpQTLcluster9	ENSRNOG00000011864[NP_758638.1[GM2 ganglioside activator protein [Source:RefSeq_peptide;Acc:NP_758638]]	ENSG00000019743[GM2A[Ganglioside GM2 activator precursor (GM2-AP) (Cerebroside sulfate activator protein) (Sphingolipid activator protein 3) (SAP-3). [Source:UniprotSWISSPROT;Acc:P17900]]
QTL231; QTL232;	Lipid Levels; Lipid Levels;	16 kidney	1371870_at		19 D19Utr7	0.0000001	Il2; Bp32		
QTL28_BP46_H	Blood Pressure	2 fat	1371951_at		9 D9Rat104	0.000066	Bp113 (Cm11), Bp53 (Cm26)	ENSRNOG00000016866[LIM3_RAT]Skeletal muscle LIM-protein 3 (SLIM 3) (LIM-domain protein DRAL) (Four and a half LIM domains protein 2) (FHL-2). [Source:Uniprot/SWISSPROT;Acc:O35115]]	ENSG00000115641[FHL2[Skeletal muscle LIM-protein 3 (SLIM 3) (LIM-domain protein DRAL) (Four and a half LIM domains protein 2) (FHL-2). [Source:Uniprot/SWISSPROT;Acc:Q14192]]
QTL203; QTL205;	Body Weight; Type 2 Diabetes;	13 kidney	1371953_at		15 D15Rat29	0.000026996	NONE	ENSRNOG00000014146[5 kDa protein [Source:IPI;Acc:IPI00400653]]	ENSG00000175198[PCCA[Propionyl-CoA carboxylase alpha chain, mitochondrial precursor (EC 6.4.1.3) (PCCase alpha subunit) (Propionyl-CoA carbon dioxide ligase alpha subunit) [Source:Uniprot/SWISSPROT;Acc:P05165]]
QTL234_BP15_H; QTL237_BP34_H; QTL239_BW30_H; QTL241_BW55_H; QTL243_SLEP3_H; QTL250;	Blood Pressure; Blood Pressure; Body Weight; Body Weight; Hormone level; Type 2 Diabetes;	17 kidney	1372084_at		10 D10Wox13	0.000058824	Trs1, BpQTLcluster9	ENSRNOG00000026847[Q6AXU5_RAT] ENSRNOG00000031846[MXI1_RAT]MAX interacting protein 1 (MXI1 protein) [Source:UniprotSWISSPROT;Acc:O08015]] ENSRNOG00000034078[MXI1_RAT]MAX interacting protein 1 (MXI1 protein) [Source:UniprotSWISSPROT;Acc:O08015]] ENSRNOG00000028634[P70542_RAT]Mxi1 protein (Fragment). [Source:UniprotSPTREMBL;Acc:P70542]]	ENSG00000019167[NFE2L1[Nuclear factor erythroid 2-related factor 1 (NF-E2-related factor 1) (Transcription factor LCR-F1) (TCF-11) (Locus control region factor 1) (LCR-F1) [Source:UniprotSWISSPROT;Acc:Q14494]] ENSG00000163349[MXI1[MAX interacting protein 1 (MXI1 protein) [Source:UniprotSWISSPROT;Acc:Q99P4]
QTL171	Type 2 Diabetes	10 kidney	1372093_at		1 D1Rat225	0.000014371	Map1 (Bp196), Iddm10, Scl5/SR3		
QTL253_BP13_H; QTL234_BP15_H; QTL237_BP34_H; QTL239_BW30_H; QTL241_BW55_H; QTL243_SLEP3_H; QTL250;	Blood Pressure; Blood Pressure; Body Weight; Body Weight; Hormone level; Type 2 Diabetes;	17 kidney	1372563_at		10 D10Wox13	0.000029891	Trs1, BpQTLcluster9	ENSRNOG00000023576[XP_343563.1[]	ENSG00000119147[NP_115787.1[esophageal cancer related gene 4 protein [Source:RefSeq_peptide;Acc:NP_115787]]
QTL26_BP46_H	Blood Pressure	2 fat	1372648_at		9 D9Rat104	0.000066	Bp113 (Cm11), Bp53 (Cm26)		
QTL10_BW57_H	Body Weight	1 fat	1372726_at		2 D2Rat230	0.000029557	Bp18, Bp101, BpQTLcluster3, Bp53		
QTL214_BP32_H; QTL218; QTL219; QTL8_BW19_H; QTL10_BW57_H; QTL11_BW90_H; QTL14; QTL20; QTL21; QTL22;	Blood Pressure; Body Weight; Lipid Levels; Body Weight; Body Weight; Body Weight; Type 2 Diabetes; Metabolic Syndrome; Metabolic Syndrome; Type 2 Diabetes;	15 kidney	1372794_at		8 D8Utr5	0.000016979	Chr5 congenic (Bp cardiac mass, HDL phospholipid, serum cholesterol), Bp35, Bp82	ENSRNOG00000017332[XP_217177.1[similar to death-associated kinase 2 [Source:RefSeq_peptide;Acc:XP_217177]]	ENSG00000035664[DAPK2_HUMAN[Death-associated protein kinase 2 (EC 2.7.1.37) (DAP kinase 2) (DAP- kinase related protein 1) (DRP-1) [Source:UniprotSWISSPROT;Acc:Q9UIK4]]
QTL178_BP31_H; QTL182;	Blood Pressure; Lipid Levels;	13 kidney	1372848_at		1 D1Rat293	0.0000001	Map1 (Bp196), Bp42, Weight3 (Bw20)	ENSRNOG00000020702[XP_342023.1[] ENSRNOG00000021216[XP_227468.1[similar to peroxisomal biosis factor	ENSG00000182144[NP_705839.2[cytochrome b, ascorbate dependent 3 [Source:RefSeq_peptide;Acc:NP_705839]] ENSG00000131779[PEX11B[peroxisomal biogenesis factor 11B
QTL10_BW57_H; QTL20;	Body Weight; Metabolic Syndrome;	1 fat	1372923_at		2 Atp1a1	0.0001	Bp80	[Source:RefSeq_peptide;Acc:XP_227468]]	[Source:RefSeq_peptide;Acc:NP_003837]]
QTL10_BW57_H; QTL13; QTL20;	Body Weight; Type 2 Diabetes; Metabolic Syndrome;	1 fat	1373243_at		2 Nor1	0.000026698	Bp101, Bp19, BpQTLcluster3	ENSRNOG00000020696[XP_227421.1[]	ENSG00000183044[PMVK[Phosphomevalonate kinase (EC 2.7.4.2) (PMKase) [Source:UniprotSWISSPROT;Acc:Q15126]]
QTL63_BW86_H	Body Weight	3 fat	1373589_at		2 D2Rat136	0.000014589	Lvm1 (Cm22)	ENSRNOG00000014178[NP_861433.1[Hypgf2 protein [Source:RefSeq_peptide;Acc:NP_861433]]	ENSG00000177848[ACAD9[Acyl-CoA dehydrogenase family member 9, mitochondrial precursor (EC 1.3.99.-) (ACAD-9). [Source:UniprotSWISSPROT;Acc:Q9H845]]
QTL163	Type 2 Diabetes	9 kidney	1373440_at		5 D5Rat6	0.000058182	NONE	ENSRNOG00000007099[10 kDa protein [Source:IPI;Acc:IPI00189612]]	ENSG00000063099[NP_065199.1[]
QTL234_BP15_H; QTL237_BP34_H; QTL239_BW30_H; QTL241_BW55_H; QTL243_SLEP3_H; QTL250;	Blood Pressure; Blood Pressure; Body Weight; Body Weight; Hormone level; Type 2 Diabetes;	17 kidney	1373865_at		10 D10Rat56	0.0000001	Trs1, BpQTLcluster9	ENSRNOG00000005357[XP_340922.1[]	ENSG00000178921[NM_025099.2[]
QTL193_BW37_H; QTL194_BW58_H;	Body Weight; Body Weight;	12 shared	1373882_at		12 D12Rat53	1e-06.5e-06	Cm5 (SCOR120)		
QTL84_BP47_H; QTL85_BFD2_H; QTL86_BW81_H; QTL88_BWR_H; QTL93; QTL95; QTL97	Blood Pressure; Body fluid distribution; Body Weight; Body Weight; Hormone level; Type 2 Diabetes; Lipid Levels	5 kidney	1373903_s_at		2 D2Rat201	0.0000030457	NONE	ENSRNOG00000015954[GHR_RAT]Growth hormone receptor (GH receptor) (Somatotropin receptor) (Contains: Growth hormone binding protein (GH binding protein) (GHBP) (Serum binding protein)] [Source:UniprotSWISSPROT;Acc:P16310]]	ENSG00000112964[GHR[Growth hormone receptor precursor (GH receptor) (Somatotropin receptor) [Contains: Growth hormone binding protein (GH binding protein) (GHBP) (Serum binding protein)] [Source:UniprotSWISSPROT;Acc:P10912]]

Human pQTL Name	Human pQTL Type	Human Chromosome	Overlapping eQTL - Tissue	Overlapping eQTL - probe set	Overlapping eQTL - Chromosome	Overlapping eQTL - Linked marker	eQTL P-Value (fat, kidney)	Rat pQTL	Rat Genes in this region	Homologous Human Genes
QTL178_BP31_H; QTL179_BW36_H; QTL180_BW64_H; QTL181_BW70_H; QTL195_BW65_H; QTL198; QTL178_BP31_H; QTL182; QTL8_BW19_H; QTL10_BW57_H; QTL11_BW60_H; QTL14; QTL20; QTL21; QTL22; QTL265_BW34_H; QTL271;	Blood Pressure; Body Weight; Body Weight; Body Weight; Body Weight; Body Weight; Blood Pressure; Lipid Levels; Body Weight; Body Weight; Body Weight; Type 2 Diabetes; Metabolic Syndrome; Metabolic Syndrome; Type 2 Diabetes; Body Weight; Lipid Levels;	11 fat 12 kidney 11 fat 1 kidney 19 kidney	1373838_at 1374087_at 1374089_at 1374123_at 1374249_at	8 D8Utr3 7 D7Rat3 3 D3Rat36 13 D13Cebr5s3 1 D1Rat212	0.000073333 0.000019928 0.000072289 0.000003 0.000009	Inf1 NONE Arr10 Bp80, Bp (Pravenec et al) BpQTLcluster1	ENSRNOG00000009157|FUT4_RAT|Alpha-(1,3)-fucosyltransferase (EC 2.4.1.-) (Galactoside 3-L-fucosyltransferase) (Fucosyltransferase 4) (FUCT-IV) (Fuc-TIV). [Source:Uniprot/SWISSPROT;Acc:Q62994]] ENSRNOG00000008192|XP_217033.1[] ENSRNOG00000020830|XP_238097.2[]	ENSG00000196371|FUT4|Alpha-(1,3)-fucosyltransferase (EC 2.4.1.-) (Galactoside 3-L-fucosyltransferase) (Fucosyltransferase 4) (FUCT-IV) (Fuc-TIV) (ELAM-1 ligand fucosyltransferase). [Source:Uniprot/SWISSPROT;Acc:P22083]] ENSG00000181392|Q7Z7L3_HUMAN[]
QTL15 QTL278; QTL280;	Lipid Levels Blood Pressure; Lipid Levels;	1 kidney 21 kidney	1374379_at 1374429_at	13 D13Rat152 20 D20Mgh5	0.000003 0.000048387	NONE Scl13	ENSRNOG00000022143|XP_341157.1[]	ENSG00000158716|DUSP23|dual specificity phosphatase 23 [Source:RefSeq_peptide;Acc:NP_060293]]
QTL159_BW62_H	Body Weight	9 kidney	1374508_at	3 D3Cebr83s1	0.000094488	Bp112 (Cm10), Cardiac001, Bp92	ENSRNOG00000013040|[71 kDa protein [Source:IPI;Acc:IPI00388472]]	ENSG00000125485|DDX31|DEAD (Asp-Glu-Ala-Asp) box polypeptide 31 isoform 2 [Source:RefSeq_peptide;Acc:NP_619526]]
QTL86_BW61_H	Body Weight	5 fat	1374560_at	18 D18Rat55	0.000034056	Bp46, Arr11, Spl3, Chr18 congenic (Bp), Bp47	ENSRNOG00000016433|XP_214541.2[]	ENSG00000164244|Q8NEZ7_HUMAN[]
QTL177_BP22_H; QTL178_BP31_H; QTL179_BW36_H; QTL180_BW64_H; QTL186;	Blood Pressure; Blood Pressure; Body Weight; Body Weight; Lipid Levels;	11 shared	1374583_at	8 D8Rat219	2.6302e-05,2.9101e	Chr8 congenic (Bp, cardiac mass, HDL phospholipid, serum cholesterol), Cardiac002, Bp22, Bp35		
QTL159_BW62_H	Body Weight	9 fat	1374653_at	3 D3Cebr204s4	0.000008	Bp112 (Cm10), Cardiac001, Bp92	ENSRNOG00000017513|XP_231128.2|similar to CG12125-PA [Source:RefSeq_peptide;Acc:XP_231128]]	ENSG00000148343|C9orf54[]
QTL170; QTL171; QTL1_BP17_H; QTL3_BP25_H; QTL7_BW12_H; QTL9_BW3_H; QTL12_BW71_H;	Body Weight; Type 2 Diabetes;	10 kidney	1374831_at	1 D1Cebr10s3	0.0000001	Weight2 (Bw18), Bp26, Bp43, Weight3 (Bw20)	ENSRNOG00000014783|XP_219475.2|similar to cDNA sequence BC023151 [Source:RefSeq_peptide;Acc:XP_219475]]	ENSG00000142102|Q8TEG9_HUMAN[]
QTL177_BP22_H; QTL178_BP31_H; QTL179_BW36_H; QTL180_BW64_H; QTL186;	Blood Pressure; Blood Pressure; Body Weight; Body Weight; Body Weight; Blood Pressure; Blood Pressure; Body Weight; Body Weight; Type 2 Diabetes; Type 2 Diabetes;	1 kidney 11 fat	1374840_at 1374907_at	5 Slc2a1 1 D1Rat47	0.000046512 0.0000001	Bp155, Bp103 Weight2 (Bw18), Kw, Lvm2 (Cm23), Bp26, BpQTLcluster1, Bp44		
QTL272_BP29_H; QTL273_BW17_H; QTL274_BW6_H; QTL276; QTL277;	Blood Pressure; Body Weight; Body Weight; Blood Pressure; Type 2 Diabetes;	20 kidney	1374921_at	3 D3Rat1	0.000014139	NONE	ENSRNOG00000027513|XP_215972.2|similar to helicase-like protein NHL isoform 2 [Source:RefSeq_peptide;Acc:XP_215972]]	ENSG00000028036|TNFRSF6B|Regulator of telomere elongation helicase 1 (Helicase-like protein NHL). [Source:Uniprot/SWISSPROT;Acc:Q9NZ71]]
QTL231; QTL232; QTL71_BP1_H; QTL73_BW1_H; QTL74_BW15_H; QTL75_BW4_H;	Lipid Levels; Lipid Levels; Blood Pressure; Body Weight; Body Weight; Body Weight;	16 kidney 4 fat	1375071_at 1375194_at	19 D19Cebrp150s1 14 D14Rat52	0.000053333 0.000005	Ir2, Bp32 Scl11	ENSRNOG00000017919|XP_214721.2[]	ENSG00000069248|NUP133|Nuclear pore complex protein Nup133 (Nucleoporin Nup133) (133 kDa nucleoporin). [Source:Uniprot/SWISSPROT;Acc:Q8WUM0]]
QTL256; QTL261; QTL24_BP23_H; QTL26_BP48_H; QTL42; QTL50;	Blood Pressure; Blood Pressure; Blood Pressure; Blood Pressure; Type 2 Diabetes; Metabolic Syndrome;	18 fat 2 shared	1375343_at 1375665_at	18 D18Rat13 9 D9Rat15	0.000003 3e-06,1e-06	Bp46, Spl3, Chr18 congenic (Bp) Bp53 (Cm28), Bp106	ENSRNOG00000029250|[],ENSRNOG00000031069|[]	
QTL170; QTL171; QTL214_BP32_H; QTL218; QTL219; QTL220; QTL222; QTL124_BP48_H; QTL126_BW18_H; QTL126_BW26_H; QTL129_BW27_H; QTL131_BW68_H; QTL134; QTL136; QTL138; QTL142;	Body Weight; Type 2 Diabetes; Blood Pressure; Body Weight; Lipid Levels; Type 2 Diabetes; Type 2 Diabetes;	10 shared 15 fat	1375684_at 1375676_at	1 D1Rat287 3 D3Rat166	9e-06,1e-05 0.000056338	Weight2 (Bw18), Bp26, Bp43, Bp (huang et al), Weight3 (Bw20) NONE		
QTL265_BW34_H; QTL271; QTL192_BP7_H; QTL196;	Blood Pressure; Body Weight; Body Weight; Body Weight; Body Weight; Blood Pressure; Type 2 Diabetes; Body Weight; Lipid Levels; Body Weight; Lipid Levels; Blood Pressure; Body Weight;	7 shared 19 fat 12 fat	1375724_at 1375927_at 1376200_at	4 D4Utr3 1 D1Rat212 4 Pthln	7.2368e-05,9.4488e 0.000012232 0.00004896	Bp79, Str3, Bw22, S82 BpQTLcluster1 NONE	ENSRNOG00000008524|[Protein [Source:IPI;Acc:IPI00388788]]	ENSG00000106415|GLCCI1|glucocorticoid induced transcript 1 [Source:RefSeq_peptide;Acc:NP_612435]]
QTL166_BP8_H; QTL168_BW22_H; QTL169_BW63_H; QTL173; QTL175; QTL72_BP26_HQTL76_HRTRT1_H; QTL77; QTL78; QTL80; QTL82; QTL148_BW52_H	Blood Pressure; Body Weight; Body Weight; Body Weight; Type 2 Diabetes; Blood Pressure; Body Weight; Obesity in T2 diabetics; Lipid Levels; Blood Pressure; Body Weight	10 kidney 4 fat 8 kidney	1376259_at 1376396_at 1376607_a_at	17 D17Rat62 2 D2Cebr4s8 16 D16Rat53	0.0000001 0.000035599 0.000003	Lvm6 (Cm26) Bp17 (DBP) NONE	ENSRNOG00000019057|Q9WTQ0_RAT|PKC theta protein (Fragment). [Source:Uniprot/SPTREMBL;Acc:Q9WTQ0]] ENSRNOG00000023494|Q86H58_RAT[]	ENSG00000065675|PRKCQ|Protein kinase C, theta type (EC 2.7.1.-) (nPKC-theta). [Source:Uniprot/SWISSPROT;Acc:Q04759]] ENSG00000129996|NP_079381.1[]
QTL202; QTL204; QTL207; QTL146_BW52_H	Type 2 Diabetes; Lipid Levels; Blood Pressure; Body Weight	13 fat 8 kidney	1376640_at 1376656_at	12 D12Cebrp97s9 16 D16Rat53	0.00001 0.000097561	NONE NONE		
QTL177_BP22_H; QTL178_BP31_H; QTL179_BW36_H; QTL180_BW64_H; QTL186;	Blood Pressure; Blood Pressure; Body Weight; Body Weight; Lipid Levels;	11 kidney	1376821_at	8 D8Rat219	0.000083333	Chr8 congenic (Bp, cardiac mass, HDL phospholipid, serum cholesterol), Cardiac002, Bp22, Bp35	ENSRNOG00000010372|XP_235868.1[]	ENSG00000110074|NP_060017.1[]
QTL177_BP22_H; QTL178_BP31_H; QTL179_BW36_H; QTL180_BW64_H; QTL186;	Blood Pressure; Blood Pressure; Body Weight; Body Weight; Lipid Levels;	11 kidney	1377079_at	8 Kcnj1	0.000054	Chr8 congenic (Bp, cardiac mass, HDL phospholipid, serum cholesterol), Bp35	ENSRNOG00000031268|[]	ENSG00000110013|O9HAU7_HUMAN|cytosolic sialic acid 9-O-acetylesterase forming [Source:RefSeq_peptide;Acc:NP_733748]]

Human pQTL Name	Human pQTL Type	Human Chromosome	Overlapping eQTL - Tissue	Overlapping eQTL - probe set	Overlapping eQTL - Chromosome	Overlapping eQTL - Linked marker	eQTL P-Value (fat, kidney)	Rat pQTL	Rat Genes in this region	Homologous Human Genes
QTL177_BP22_H; QTL178_BP31_H; QTL179_BW35_H; QTL180_BW64_H; QTL186;	Blood Pressure; Blood Pressure; Body Weight; Body Weight; Lipid Levels;	11 kidney	1377078_at	8	D8Rat219	0.000001	Chr8 congenic (Bp, cardiac mass, HDL phospholipid, serum cholesterol), Cardiac002, Bp22, Bp35	ENSRNOG00000031266|[]	ENSG00000110013|O94AU7_HUMAN|Cytosolic sialic acid 9-O-acetylesterase homolog [Source:RefSeq_peptide;Acc:XP_733746|]
QTL75_BW4_H; QTL79;	Body Weight; Type 2 Diabetes;	4 kidney	1377181_at	14	D14Rat8	0.000003	Scl11	ENSRNOG00000002139|XP_214022.1|	ENSG00000108265|XP_044178.4|PREDICTED: KIAA1211 protein [Source:RefSeq_peptide;Acc:XP_044178|]
QTL269; QTL270;	Type 2 Diabetes; Lipid Levels;	19 shared	1377212_at	12	D12Rat56	1e-06,1e-06	NONE		
QTL202; QTL204; QTL207;	Type 2 Diabetes; Lipid Levels; Blood Pressure;	13 shared	1377244_at	12	D12Cebrp97s9	1e-06,1.4647e-05	NONE	ENSRNOG00000031121|XP_221910.2|similar to zinc finger protein 95 [Source:RefSeq_peptide;Acc:XP_221910|] ENSRNOG00000006982|XP_344971.1|similar to PKE protein kinase [Source:RefSeq_peptide;Acc:XP_344971|]	ENSG00000196652|ZFP95|Zinc finger protein 95 homolog (Zfp-95). [Source:Uniprot/SWISSPROT;Acc:Q9Y2L8|] ENSG00000165752|STK32C|serine/threonine kinase 32C [Source:RefSeq_peptide;Acc:NP_775845|]
QTL170; QTL171;	Body Weight; Type 2 Diabetes;	10 kidney	1377534_at	1	D1Rat69	0.000001	Weight2 (Bw18), Bp26, Bp43, Weight3 (Bw20)		
QTL170; QTL171;	Body Weight; Type 2 Diabetes;	10 fat	1377614_at	1	D1Rat267	0.000014139	Weight2 (Bw18), Bp26, Bp43, Bp (huang et al), Weight3 (Bw20)	ENSRNOG00000015993|O9AYH2_RAT|	ENSG00000188522|NP_775889.1|
QTL225_BP27_H; QTL228_BP33_H; QTL229;	Blood Pressure; Blood Pressure; Body Weight;	16 kidney	1377614_at	1	Myo2	0.000062707	Weight2 (Bw18), Bp26, Bp43, Bp (huang et al), Weight3 (Bw20)	ENSRNOG00000015993|O9AYH2_RAT|	ENSG00000188522|NP_775889.1|
QTL143_BP11_H; QTL146_BP49_H; QTL147_BW20_H; QTL149_BW59_H; QTL150_BW69_H; QTL152_SLEP1_H; QTL153; QTL156; QTL157;	Blood Pressure; Blood Pressure; Body Weight; Body Weight; Body Weight; Hormone level; Lipid Levels; Lipid Levels; Lipid Levels;	8 kidney	1377823_at	7	D7Rat17	0.000008	Lvm7 (Cm27)		
QTL178_BP31_H; QTL179_BW35_H; QTL180_BW64_H; QTL181_BW70_H;	Blood Pressure; Body Weight; Body Weight; Body Weight;	11 kidney	1380286_x_at	8	D8Rat56	0.000035398	Inf1	ENSRNOG00000010999|(103 kDa protein [Source:IPI;Acc:IPI00389510|]	ENSG00000188004|Q8N819_HUMAN|PREDICTED: KIAA1731 protein [Source:RefSeq_peptide;Acc:XP_374822|]
QTL178_BP31_H; QTL179_BW35_H; QTL180_BW64_H; QTL181_BW70_H;	Blood Pressure; Body Weight; Body Weight; Body Weight;	11 fat	1380286_x_at	8	D8Jth3	0.000078371	Inf1	ENSRNOG00000010999|(103 kDa protein [Source:IPI;Acc:IPI00389510|]	ENSG00000188004|Q8N819_HUMAN|PREDICTED: KIAA1731 protein [Source:RefSeq_peptide;Acc:XP_374822|]
QTL256; QTL261;	Blood Pressure; Blood Pressure;	18 fat	1380293_at	18	D18Rat13	0.000075342	Bp48, Spl3, Chr18 congenic (Bp)		ENSG00000166845|C18orf54|]
QTL10_BW57_H; QTL10_BW57_H;	Body Weight; Body Weight;	1 fat; 1 fat;	1380404_at; 1380404_at;	2; 2	D2Rat236; D2Rat54;	0.0000001; 0.000059	Bp16, Bp101, BpQTLcluster3, Bp63; Bp101, BpQTLcluster3		
QTL99_BP4_H; QTL100_BFD3_H;	Blood Pressure; Body fluid distribution;	6 fat	1382105_at	6	D6Rat21	0.000039711	Chr6 congenic (Bp, cardiac mass, HDL phospholipid, serum cholesterol), Bp35, Bp52, Weight2 (Bw18), Bp26, Bp (huang et al), Weight3 (Bw20)	ENSRNOG00000008926|[32 kDa protein [Source:IPI;Acc:IPI00392887|]	ENSG00000069996|GNB5|Guanine nucleotide-binding protein beta subunit 5 (Transducin beta chain 5) (Gbeta5). [Source:Uniprot/SWISSPROT;Acc:O14775|]
QTL225_BP27_H; QTL228_BP33_H; QTL229;	Blood Pressure; Blood Pressure; Body Weight;	16 kidney	1383303_at	1	Sa	0.00001		ENSRNOG00000032246|[] ENSRNOG00000009199|XP_342469.1|similar to 2810203E10Rik protein [Source:RefSeq_peptide;Acc:XP_342469|]	ENSG00000187|SAH|SA hypertension-associated homolog isoform 2 [Source:RefSeq_peptide;Acc:NP_973729|] ENSG00000110455|NP_115981.1|1-aminocyclopropane-1-carboxylate synthase [Source:RefSeq_peptide;Acc:NP_115981|]
QTL176_BP31_H; QTL182;	Blood Pressure; Lipid Levels;	11 fat	1383440_at	3	D3Rat35	0.000013319	Arr10		
QTL24_BP23_H; QTL26_BP46_H; QTL42; QTL50;	Blood Pressure; Blood Pressure; Type 2 Diabetes; Metabolic Syndrome;	2 shared	1384309_at	9	D9Rat156	3.7415e-05,2.4948e	Bp53 (Cm28), Bp106		
QTL143_BP11_H; QTL146_BP49_H; QTL147_BW20_H; QTL149_BW59_H; QTL150_BW69_H; QTL152_SLEP1_H; QTL153; QTL156; QTL157;	Blood Pressure; Blood Pressure; Body Weight; Body Weight; Body Weight; Hormone level; Lipid Levels; Lipid Levels; Lipid Levels;	8 shared	1384717_at	7	D7Rat17	1e-07,1e-07	Lvm7 (Cm27)		
QTL2_BP18_H; QTL15; QTL58_BFD1_H	Blood Pressure; Lipid Levels; Body fluid distribution	1 kidney; 3 kidney	1386212_at; 1386352_at	13; 16	D13Rat152; D16Rat87	0.0000001; 0.000096774	NONE; NONE	ENSRNOG00000003537|Q6XDA1_RAT|Erythroid spectrin alpha. [Source:Uniprot/SPTREMBL;Acc:Q6XDA1|]	ENSG00000163554|SPTA1|Spectrin alpha chain, erythrocyte (Erythroid alpha-spectrin). [Source:Uniprot/SWISSPROT;Acc:P02549|]
QTL125_BFD5_H; QTL141; QTL142;	Body fluid distribution; Body Weight; Lipid Levels;	7 shared	1386901_at	4	Cd36	1e-06,1e-07	Bp79, Bw22, Sfl2	ENSRNOG00000005808|CD36_RAT|Platelet glycoprotein IV (GPIV) (GPIIIB) (CD36 antigen) (PAS IV) (PAS- 4 protein) (Fatty acid transport protein) (Fatty acid translocase) (Adipocyte membrane protein). [Source:Uniprot/SWISSPROT;Acc:Q07969|]	ENSG00000135218|CD36|Platelet glycoprotein IV (GPIV) (Glycoprotein IIIb) (GPIIIB) (Leukocyte differentiation antigen CD36) (CD36 antigen) (PAS IV) (PAS- 4 protein) (Platelet collagen receptor) (Fatty acid translocase) (FAT) (Thrombospondin receptor). [Source:Uniprot/SWISSPROT;Acc:P16671|]
QTL10_BW57_H	Body Weight	1 kidney	1386985_at	2	D2Rat236	0.000003	Bp16, Bp101, BpQTLcluster3, Bp63	ENSRNOG00000029728|GSTM1_RAT|Glutathione S-transferase Yb-1 (EC 2.5.1.18) (Chain 3) (GST Yb1) (GST M1-1) (GST class-mu 1). [Source:Uniprot/SWISSPROT;Acc:P04905|]	ENSG00000134184|GSTM1|Glutathione S-transferase Mu 1 (EC 2.5.1.18) (GSTM1-1) (GST class-mu 1) (GSTM1a-1a) (GSTM1b-1b) (HB subunit 4) (GTH4). [Source:Uniprot/SWISSPROT;Acc:P09488|]
QTL84_BP47_H; QTL85_BFD2_H; QTL86_BW61_H; QTL88_BW8_H; QTL91; QTL93; QTL95; QTL97;	Blood Pressure; Body fluid distribution; Body Weight; Body Weight; Blood Pressure; Hormone level; Type 2 Diabetes; Lipid Levels;	5 fat	1387144_at	2	D2Rat201	0.000003	NONE	ENSRNOG00000012080|ITA1_RAT|Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a). [Source:Uniprot/SWISSPROT;Acc:P18614|]	ENSG00000152684|PELO|Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a). [Source:Uniprot/SWISSPROT;Acc:P56199|]
QTL30_BW53_H; QTL31_SLEP2_H; QTL32; QTL43; QTL44; QTL49;	Body Weight; Hormone level; Blood Pressure; Obesity in T2 diabetics; Obesity in T2 diabetics; Metabolic Syndrome;	2 shared	1387222_at	6	D6Cebrp424s2	1e-06,9e-06	NONE	ENSRNOG00000015215|NM_019907.1|Rattus norvegicus postsynaptic protein Cript (Cript), mRNA. [Source:RefSeq_dna;Acc:NM_019907|]	ENSG00000119678|NP_054890.1|postsynaptic protein CRIPT [Source:RefSeq_peptide;Acc:NP_054890|]
QTL75_BW4_H; QTL79;	Body Weight; Type 2 Diabetes;	4 kidney	1387314_at	14	D14Utr6	0.000001	NONE	ENSRNOG00000001987|ST1B1_RAT|Sulfotransferase family cytosolic 1B member 1 (EC 2.8.2.-) (Sulfotransferase 1B1) (DOPA/tyrosine sulfotransferase). [Source:Uniprot/SWISSPROT;Acc:P52847|]	ENSG00000173597|ST1B1_HUMAN|Sulfotransferase family cytosolic 1B member 1 (EC 2.8.2.-) (Sulfotransferase 1B2) (Thyroid hormone sulfotransferase). [Source:Uniprot/SWISSPROT;Acc:O43704|]
QTL24_BP23_H; QTL26_BP46_H; QTL42; QTL50;	Blood Pressure; Blood Pressure; Type 2 Diabetes; Metabolic Syndrome;	2 shared	1387376_at	9	D9Rat93	3e-06,1e-07	Bp53 (Cm28)	ENSRNOG00000015354|ADO_RAT|Aldehyde oxidase (EC 1.2.3.1). [Source:Uniprot/SWISSPROT;Acc:Q9ZOU5|]	ENSG00000138356|AOX1|Aldehyde oxidase (EC 1.2.3.1). [Source:Uniprot/SWISSPROT;Acc:Q06278|]

Human pQTL Name	Human pQTL Type	Human Chromosome	Overlapping eQTL - Tissue	Overlapping eQTL - probe set	Overlapping eQTL - Chromosome	Overlapping eQTL - Linked marker	eQTL P-Value (fat, kidney)	Rat pQTL	Rat Genes in this region	Homologous Human Genes
QTL278_BP12_H; QTL280;	Blood Pressure; Lipid Levels;	21	fat	1387789_at	11	D11Cebr87s2	0.000018644	NONE	ENSRNOG00000001652|NP_596888.1|v-ets erythroblastosis virus E26 onco like [Source:RefSeq_peptide;Acc:NP_596888]]	ENSG00000157554|ERG|Transcriptional regulator ERG (Transforming protein ERG). [Source:Uniprot/SWISSPROT;Acc:P11308]]
QTL208_BP54_H; QTL212;	Blood Pressure; Hormone level;	14	fat	1387808_at	15	D15Rat6	0.000003	NONE	ENSRNOG00000010296|NP_112631.1|solute carrier family 7 (cationic amino acid transporter, y+syst [Source:RefSeq_peptide;Acc:NP_112631]]	ENSG00000155465|SLC7A7|Y+L amino acid transporter 1 (y(+)L-type amino acid transporter 1) (y+LAT-1) (Y+LAT1) (Monocyte amino acid permease 2) (MOP-2). [Source:Uniprot/SWISSPROT;Acc:Q9UM01]]
QTL272_BP29_H; QTL273_BW17_H;	Blood Pressure; Body Weight; Body Weight;									
QTL274_BW6_H; QTL276; QTL277;	Blood Pressure; Type 2 Diabetes;	20	fat	1387906_a_at	3	D3Rat132	0.000006	Bp81		
QTL24_BP23_H; QTL26_BP48_H; QTL42; QTL50;	Blood Pressure; Blood Pressure; Type 2 Diabetes; Metabolic Syndrome;	2	fat	1387924_at	9	D9Rat4	0.000029266	Cm21 (Hw6)	ENSRNOG00000016653|XP_346855.1|ENSRNOG00000005618|NCOA3_RAT|Nuclear receptor coactivator 3 (EC 2.3.1.48) (NCoA-3) (Amplified in breast cancer-1 protein homolog) (AIB1) (Fragment).	ENSG00000066248|NGEF|neuronal guanine nucleotide exchange factor [Source:RefSeq_peptide;Acc:NP_062824]]
QTL272_BP29_H; QTL273_BW17_H;	Blood Pressure; Body Weight; Body Weight;								[Source:Uniprot/SWISSPROT;Acc:Q9EPU2]]	ENSG00000124151|NP_006525.2|nuclear recepter coactivator 3 isoform b [Source:RefSeq_peptide;Acc:NP_006525]]
QTL274_BW6_H; QTL276; QTL277;	Blood Pressure; Type 2 Diabetes;	20	kidney	1388178_at	3	D3Rat143	0.000060606	Bp81, Bp20, Ir1 Weight2 (Bw18), Bp26, Bp (huang et al), Weight3 (Bw20)		ENSG00000066654|THUMPD1|THUMP domain containing protein 1. [Source:Uniprot/SWISSPROT;Acc:Q9NXG2]]
QTL225_BP27_H; QTL228_BP33_H; QTL229;	Blood Pressure; Blood Pressure; Body Weight;	16	kidney	1388567_at	1	Sa	0.000024145		ENSRNOG00000014946|XP_219396.2|	ENSG00000169159|HN1|hematological and neurological expressed 1 isoform 2 [Source:RefSeq_peptide;Acc:NP_001002032]]
QTL236_BP20_H; QTL237_BP34_H; QTL248;	Blood Pressure; Blood Pressure; Body Weight;	17	kidney	1388581_at	10	D10Cebr39s2	0.000078788	Bw10 Bp46, Spl3, Chr18 congenic (Bp)	ENSRNOG00000003661|XP_213527.1|similar to HN1 [Source:RefSeq_peptide;Acc:XP_213527]]	
QTL256; QTL261;	Blood Pressure; Blood Pressure;	18	fat	1388830_at	18	D18Rat13	0.0000001			
QTL208_BP54_H; QTL212;	Blood Pressure; Hormone level;	14	shared	1388654_at	15	D15Rat6	5.8537e-05,1.8482e	NONE		
QTL159_BW62_H	Body Weight	9	shared	1388912_at	3	D3Cebr204s4	9e-06,2.6258e-05	Bp112 (Cm10), Cardiac001, Bp92	ENSRNOG00000027867|XP_231123.2|similar to XPMC2 prevents mitotic catastrophe 2 homolog; Xenopus prevents mitotic catastrophe 2 homolog [Source:RefSeq_peptide;Acc:XP_231123]] ENSRNOG00000020361|Q5XIU6_RAT|Inositol hexakisphosphate kinase 2 (EC 2.7.4.21) (InsP6 kinase 2) (Inositol hexakisphosphate kinase 2) (PKI) [Source:Uniprot/SWISSPROT;Acc:Q9R0U1]]	ENSG00000146300|XPMC2H|XPMC2 prevents mitotic catastrophe 2 homolog [Source:RefSeq_peptide;Acc:NP_065118]] ENSG00000068745|IP6K2_HUMAN|Inositol hexakisphosphate kinase 2 (EC 2.7.4.21) (InsP6 kinase 2) (Inositol hexakisphosphate kinase 2) (PKI) (IP6K2). [Source:Uniprot/SWISSPROT;Acc:Q9UHH9]]
QTL178_BP31_H; QTL179_BW36_H;	Blood Pressure; Body Weight; Body Weight;									
QTL180_BW64_H; QTL181_BW70_H;	Body Weight;	11	kidney	1389005_at	8	D8Rat202	0.00001	Inf1		
QTL26_BP46_H	Blood Pressure	2	kidney	1389142_at	3	D3Rat257	0.000073171	Cm17, Ir1	ENSRNOG00000000172|XP_229231.2|similar to sulfide quinone reductase-like [Source:RefSeq_peptide;Acc:XP_229231]]	ENSG00000104184|PLD3|Pallidin (Pallid protein homolog) (Syntaxin 13-interacting protein). [Source:Uniprot/SWISSPROT;Acc:Q9UL45]] ENSG00000137767|SQRDL|Sulfide quinone oxidoreductase, mitochondrial precursor (EC 1.-.-.-) (CGI-44). [Source:Uniprot/SWISSPROT;Acc:Q9Y6N5]]
QTL214_BP32_H; QTL218; QTL219; QTL220; QTL222;	Blood Pressure; Body Weight; Lipid Levels; Type 2 Diabetes; Type 2 Diabetes;	15	kidney	1389142_at	3	D3Uh9	0.000009537	Cm17	ENSRNOG00000000172|XP_229231.2|similar to sulfide quinone reductase-like [Source:RefSeq_peptide;Acc:XP_229231]]	ENSG00000104184|PLD3|Pallidin (Pallid protein homolog) (Syntaxin 13-interacting protein). [Source:Uniprot/SWISSPROT;Acc:Q9UL45]] ENSG00000137767|SQRDL|Sulfide quinone oxidoreductase, mitochondrial precursor (EC 1.-.-.-) (CGI-44). [Source:Uniprot/SWISSPROT;Acc:Q9Y6N5]]
QTL282_BP55_H; QTL284;	Blood Pressure; Metabolic Syndrome;	22	shared	1389264_at	7	D7Rat129	9.2308e-05,9.5238e	Lvm7 (Cm27)	ENSRNOG00000010821|[]	ENSG00000100124|NP_620152.1[]
QTL282_BP55_H; QTL284;	Blood Pressure; Metabolic Syndrome;	22	fat	1389511_s_at	7	Bzrp	0.000005	Lvm7 (Cm27)		
QTL159_BW62_H	Body Weight	9	shared	1389816_at	3	D3Cebr204s4	3.6328e-05,2.8916e	Bp112 (Cm10), Cardiac001, Bp92	ENSRNOG00000016033|XP_342405.1|similar to endonuclease G [Source:RefSeq_peptide;Acc:XP_342405]]	ENSG00000167136|ENDOG|Endonuclease G, mitochondrial precursor (EC 3.1.30.-) (Endo G). [Source:Uniprot/SWISSPROT;Acc:Q14249]]
QTL166_BP6_H; QTL168_BW22_H;	Blood Pressure; Body Weight; Body Weight;									
QTL169_BW63_H; QTL173; QTL175;	Body Weight; Type 2 Diabetes;	10	shared	1389867_at	17	D17Rat62	5.3571e-05,5.4545e	Lvm6 (Cm26)		
QTL177_BP22_H; QTL178_BP31_H; QTL179_BW36_H; QTL180_BW64_H;	Blood Pressure; Blood Pressure; Body Weight;							Chr8 congenic (Bp, cardiac mass, HDL phospholipid, serum cholesterol), Cardiac002,	ENSRNOG00000006993|DCPS_RAT|Scavenger mRNA decapping enzyme DcpS (EC 3.-.-.-) (DCS-1) (Hint-related 7meGMP-directed hydrolase) (Histidine triad protein member 5) (HINT-5).	ENSG00000110063|DCPS|Scavenger mRNA decapping enzyme DcpS (EC 3.-.-.-) (DCS-1) (Hint-related 7meGMP-directed hydrolase) (Histidine triad protein member 5) (HINT-5) (HSPC015).
QTL186;	Body Weight; Lipid Levels;	11	shared	1390185_at	8	D8Rat219	1.3995e-05,9e-06	Bp22, Bp35	[Source:Uniprot/SWISSPROT;Acc:Q8K4F7]]	[Source:Uniprot/SWISSPROT;Acc:Q98C86]]
QTL191_BP53_H	Blood Pressure	12	fat	1390284_at	4	D4Rat202	0.000091603	Hematocrit		
QTL278_BP12_H; QTL280;	Blood Pressure; Lipid Levels;	21	kidney	1390364_at	11	D11Rat16	0.000009	NONE		
QTL278_BP12_H; QTL280;	Blood Pressure; Lipid Levels;	21	fat	1390364_at	11	D11Mit2	0.0000001	NONE		
QTL58_BFD1_H	Body fluid distribution	3	shared	1390435_at	16	D16Rat87	0.000577,5.1948e-0	NONE	ENSRNOG00000028178|[Rattus norvegicus LOC361111 (LOC361111), mRNA. [Source:RefSeq_dna;Acc:XM_341397]]	ENSG00000168273|XP_496844.1|PREDICTED: similar to CG32736-PA [Source:RefSeq_peptide;Acc:XP_496844]]
QTL278_BP12_H; QTL280;	Blood Pressure; Lipid Levels;	21	fat	1390443_at	11	D11Rat20	0.000012397	NONE	ENSRNOG00000001554|[Rattus norvegicus similar to EURL protein homolog (LOC360892), mRNA. [Source:RefSeq_dna;Acc:XM_340963]]	ENSG00000154642|C21orf91|EURL protein homolog. [Source:Uniprot/SWISSPROT;Acc:Q9NYK6]]
QTL214_BP32_H	Blood Pressure	15	fat	1390827_at	8	D8Mit12	0.000009	Chr8 congenic (Bp, cardiac mass, HDL phospholipid, serum cholesterol), Cardiac002, Bp35, Bp82	ENSRNOG00000013628|[],ENSRNOG00000021778|[11 kDa protein [Source:IPI;Acc:IPI00389763]],ENSRNOG00000008424|NP_599225.1|p34 protein [Source:RefSeq_peptide;Acc:NP_599225]]	ENSG00000103591|NP_078942.2|[]

Human pQTL Name	Human pQTL Type	Human Chromosome	Overlapping eQTL - Tissue	Overlapping eQTL - probe set	Overlapping eQTL - Chromosome	Overlapping eQTL - Linked marker	eQTL P-Value (fat, kidney)	Rat pQTL	Rat Genes in this region	Homologous Human Genes
QTL58_BFD1_H; QTL84; QTL89; QTL166_BP6_H; QTL168_BW22_H	Body fluid distribution, Type 2 Diabetes, Type 2 Diabetes; Blood Pressure, Body Weight, Body Weight,	3	kidney	1390860_at	6	D6Rat7	0.000001	Am9, Cm19 (Hw4)	ENSRNOG00000010096|(Rattus norvegicus similar to 5-azacytidine induced 2 (LOC316081) mRNA (Source:RefSeq_dna;Acc:XM_239986)]	ENSG00000162512|AZI2[5-azacytidine induced 2 isoform b (Source:RefSeq_peptide;Acc:NP_078871)]
QTL168_BW83_H; QTL173; QTL175	Body Weight, Type 2 Diabetes	10	kidney	1393902_at	17	D17Rat151	0.000001	NONE		
QTL237_BP34_H; QTL245; QTL249;	Blood Pressure; Type 2 Diabetes; Type 2 Diabetes	17	kidney	1394228_at	10	D10Mit2	0.000078571	Trs1, BpQTLcluster9	ENSRNOG00000003865|XP_343945 1|]	ENSG00000182271|NP_996863.1|AWKS9372 (Source:RefSeq_peptide;Acc:NP_996863)]
QTL1_BP17_H; QTL3_BP28_H; QTL7_BW12_H; QTL9_BW3_H; QTL12_BW71_H;	Blood Pressure, Blood Pressure, Body Weight; Body Weight, Body Weight,	1	kidney	1398288_at	5	D5Rat189	0.000073171	Bp155, Bp103	ENSRNOG00000010735|CBFC_RAT|Nuclear transcription factor Y subunit gamma (NF-Y protein chain C) (Nuclear factor YC) (NF-YC) (CCAAT-binding transcription factor subunit C) (CBF-C) (Source:Uniprot/SWISSPROT:Acc:O82725)] ENSRNOG00000003070|PIGL_RAT|N-acetylglucosaminyl-phosphatidylinositol de-N-acetylase (EC 3.5.1.89) (Phosphatidylinositol-glycan biosynthesis, class L protein) (PIG-L) [Source:Uniprot/SWISSPROT:Acc:O35790)]	ENSG00000139641|NFYC[Nuclear transcription factor Y subunit gamma (NF-Y protein chain C) (Nuclear factor YC) (NF-YC) (CCAAT-binding transcription factor subunit C) (CBF-C) (Transactivator HSM-1/2). (Source:Uniprot/SWISSPROT:Acc:Q13952)] ENSG00000108474|PIGL[N-acetylglucosaminyl-phosphatidylinositol de-N-acetylase (EC 3.5.1.89) (Phosphatidylinositol-glycan biosynthesis, class L protein) (PIG-L). (Source:Uniprot/SWISSPROT:Acc:Q9Y2B2)]
QTL237_BP34_H; QTL240_BW31_H;	Blood Pressure; Body Weight;	17	fat	1398309_at	10	D10Rat173	0.000095588	Cm11, Trs1, BpQTLcluster9		ENSG00000166004|Q8N919_HUMAN[PRED:CTED: KIAA1731 protein (Source:RefSeq_peptide;Acc:XP_374922)]
QTL178_BP31_H; QTL179_BW36_H; QTL180_BW84_H; QTL181_BW70_H;	Blood Pressure, Body Weight, Body Weight; Body Weight;	11	kidney	1398460_at	8	D8Rat56	0.000046693	Inf1	ENSRNOG00000010999|(103 kDa protein (Source:IPI;Acc:IPI00389510)]	
QTL202; QTL204; QTL207;	Type 2 Diabetes, Lipid Levels, Blood Pressure,	13	kidney	1398878_at	12	D12Cebrp97s9	0.000006	NONE	ENSRNOG00000000985|(24 kDa protein (Source:IPI;Acc:IPI00388886)],ENSRNOG00000025217|XP_344078.1|similar to CPSF4 protein (Source:RefSeq_peptide;Acc:XP_344078)]	ENSG00000190917|CPSF4[Cleavage and polyadenylation specificity factor, 30 kDa subunit (CPSF 30 kDa subunit) (FIP1 effector domain-binding protein 1) (Neb-1) (No arches homolog) (Source:Uniprot/SWISSPROT:Acc:Q95639)]
QTL265_BW34_H; QTL271;	Body Weight, Lipid Levels,	19	fat	1399157_at	1	D1Rat27	0.000001	Cm16 (Hw11), BpQTLcluster1	ENSRNOG00000014463|XP_218532 2|similar to NIX3 (Source:RefSeq_peptide;Acc:XP_218532)]	ENSG00000105176|C1orf2[RNA polymerase II subunit 5-mediating protein (RPB5-mediating protein) (Source:Uniprot/SWISSPROT:Acc:O94763)]

Table 5.4 eQTLs deleted from Table 5.3

Overlapping eQTLS	Overlapping eQTLS - Tissue	Overlapping eQTLS - Chromosome	Overlapping eQTLS - Linked marker	Number of occurences	Reason for deletion
1372923_at	fat	2	Atp1a1	5	This marker maps to rat chromosome 2 in the current Ensembl assembly However in eQTL explorer it maps to rat chromsome 13 which would make the linkage a *trans* . In this case the human syntenic region is to Rat chromosome 13 therefore this is not of interest
1374996_at	fat	10	D1Rat47	5	This probe set and market are on different chromosome therefore is a *trans* linkage
1373833_at	fat	2	D1Cebr16s1	1	This marker is on chromosome 1 and the probe set on chromosome 2 in eQTL explorer and the probeset hits chromosome 1, 2 and 6 in ensembl
1385823_s_at	kidney	9	D3Rat109	6	This marker and probe set are on different chromosomes and are therefore *trans* linkage.
1393759_at	kidney	3	D14M64	4	This marker and probe set are on different chromosomes and are therefore *trans* linkage
1383734_at	fat	9	D12Cebr21s8	4	In total this probe set has 12 linkages to markers on 8 different rat chromosomes
		9	D13Cebrp312s2	3	
		9	D1Cebr21s2	3	
		9	D3Cebrp113s9	5	
		9	D4Cebrp145s3	9	
		9	D5Rat130	3	
		9	D6Cebr23s11	6	
		9	D7Cebr140_16s3	9	
		9	D9Cebrp10s13	3	
1376655_at	shared	9	D9Rat15	8	This probe set has 2 genomic locations one on chromosome 9 and one on 18 therefore is not definately cis
		18	D9Rat15	8	

The number of occurences column refers to the number of times the eQTL occurred in the analysis.

annotations in the rat and human respectively. Thirty-three are known genes in both rat and human and 37 have no information attached. For the eQTL genes with some information an extensive literature search was carried out to determine possible functions and any known involvement in type 2 diabetes, hypertension and other metabolic syndrome traits. This search included OMIM, Pubmed, Genecards and Google in order to gather as much information as possible.

5.3 Discussion

In this chapter comparative mapping of human QTL for metabolic and CVD phenotypes to both the rat and mouse was carried out to identify genes which may play a causal role in human disease. In the rat, eQTL in the fat and kidney tissue data sets were identified from within the regions of the rat genome which were syntenic to the human QTL. **Table 5.3** details the 142 eQTL, their rat genes and the human orthologs which delineates a set of genes which merit testing in human populations. This analysis also identified a gene (Cd36) which has already been shown to underlie an insulin resistance QTL in both humans and rat.

The first step in the analysis was to gather all the QTL for the relevant traits in both rat and human and this data is presented in **Tables 5.1** and **5.2** respectively. The next stage was to identify syntenic regions of the rat genome to all the human QTL and determine which of these rat regions had *cis*-acting eQTL within them. A complete data set of 711 eQTL was identified which was processed to give the 142 eQTL presented in **Table 5.3**. Ninety three of the 711 eQTL were discarded as good mapping information was not available either for the probe set and/or the marker. Ensembl map locations were used to identify the regions of the genome to which the

probe sets and markers mapped as these have been shown to be more accurate than the Affymetrix locations (NJ Dickens, personal communication). The remaining 618 eQTL were studied and collapsed firstly by removing any problem probe sets, then since many eQTL (and their corresponding genes) mapped to multiple human QTL these were combined to make one entry in column one (Human pQTL) of **Table 5.3**. Also any eQTL which were identified in both fat and kidney tissue were called shared and made into one entry in the table.

The seven probe sets in **Table 5.4** were removed from the gene list (**Table 5.3**) for a number of reasons. In 5 of these the probe set and linked marker are located on different rat chromosomes and therefore the linkage is controlled in *trans*. This should not have occurred as all the eQTL selected were *cis*-acting. However closer inspection showed that they are classified in eQTL explorer as *cis* which is probably due to errors in the original mapping data and *cis/trans* definition algorithm which used an older genome assembly in which the marker and probe set did map to the same chromosome. These would have been classified as *cis*, however on reassembly of the genome the marker or probe set are now mapped to a different chromosome and since the mapping unclear they have been removed. Probe set 1383734_at linked to 42 markers on 8 different chromosomes however all were classified as *cis*-acting eQTL. This could be due to the SDP of the markers, or missing genotypes within the marker set which would create false linkages. All eQTL to this probe set were removed as it is not clear which are real and these markers were not well characterised. The final probe set in this table, 1375655_at mapped to chromosomes 9 and 18 with both of these linking to the same marker and therefore the chromosome 18 linkages were removed as these are a *trans* linkage.

The final gene list in **Table 5.3** consists of 142 eQTL linkages representing 135 probe sets, which means that 7 probe sets occur twice. Three of these have two eQTL linkages in the same tissue but to two different markers with a different *P*-value while two probe sets have a different eQTL in fat and kidney tissue. Probe set 1377614_at has one linkage in fat and one in kidney which are syntenic to human chromosome 10 and 16 respectively. Since the orthologous human gene resides on chromosome 16 the eQTL linkage in kidney which is also syntenic to human chromosome 16 is more likely to be the true synteny in this case. Probe set 1389142_at occurs twice as two eQTL with linkage to different markers with different *P*-values which syntenically link to human chromosomes 2 and 15. The gene (*PLDN*) represented by this probe set is located on human chromosome 15. The duplication in the table seen in the previous two examples occurs because the fat and kidney linkages have different peak markers due to the *cis/trans* definition and can therefore link to syntenically different human chromosomes. Finally, one other gene (*FHL2* in human/*Slim3* in rat) has 2 probe sets associated with it (1369313_at and 1371951_at) and both are linked in *cis* to the same marker with a similar *P*-value.

The annotations of the genes in **Table 5.3** were manually studied, using OMIM, Pubmed and Google, to determine if any know relationship with type 2 diabetes, hypertension or other metabolic syndrome traits existed. One gene which was identified, *Cd36*, has been shown to underlie insulin resistance and altered long chain fatty acid metabolism in both humans and SHR (**Introduction**). This gene was first identified as underlying SHR insulin resistance and defective fatty acid metabolism and these phenotypes were rescued by transgenic expression of *Cd36* in this strain.

When tested in human populations Miyaoka *et al* found Cd36 deficient patients have increased concentrations of plasma lipids and insulin resistance. It has been suggested that Cd36 deficiency is a "monogenic form" of metabolic syndrome which is in accordance with it having a *cis*-acting eQTL with very high *P*-value. There are other genes which were detected which have a known role in metabolic syndrome traits. One of these, glutathione *S*-transferase mu-type 1, (*Gstm1*), is involved in the cellular response to oxidative stress. This gene has a strong *cis*-linkage ($P<10^{-5}$) in kidney tissue, is approximately 2.4 fold differentially expressed between the parental strains, and lies in a cluster of blood pressure QTL on rat chromosome 2. *Gstm1* has been implicated previously in hypertension in genetic studies in SHRSP (McBride et al. 2003; McBride et al. 2005) and humans (Rice et al. 2000). In the SHRSP strain, *Gstm1* shows a reduced expression relative to WKY and BN and thus SHRSP (McBride et al. 2005) has higher levels of superoxide, a reactive oxygen species (ROS); these are implicated in diabetes, obesity and hyperlipidaemias. Another gene Sex Hormone-Binding Precursor (*SHBG*) binds testosterone and thus affects the amount of free testosterone in the blood. Testosterone modulates blood vessel resistance and arterial blood flow and decreased levels of both testosterone and *SHBG* have been associated with increased SBP, DBP and hypertension in both men (Svartberg et al. 2004) and in postmenopausal women (Lapidus et al. 1986; Rexrode et al. 2003). An association was found between decreased levels of *SHBG* and increased CVD events in both studies. Two common polymorphisms in the prothrombin gene, which is also present in **Table 5.4** is associated with elevated plasma prothrombin and increased risk of venous thrombosis and ischemic stroke (Catena et al. 2005; Wakim-Ghorayeb et al. 2005). One final example is that of the growth hormone receptor gene (*GHR*). A SNP within this gene results in a $leu^{526} \rightarrow ile$

substitution which influences HDL levels in family members who also had a mutation in the LDL receptor which causes familial hypercholesterolemia (Takada et al. 2003).

The above examples demonstrate the feasibility of this approach in identifying potential candidate genes using comparative mapping. Many other genes which have no know association exhibit biological characteristics that make them plausible, positional candidates for human metabolic and cardiovascular disease and merit testing in human populations. For example, like *Gstm1*, *Aox1* is also involved in the cellular response to oxidative stress and resides in a blood pressure pQTL in both rat and human. *Gnpat* is a key enzyme in the biosynthesis of plasmalogens which are important in the protection against ROS and again occurs in a relevant pQTL, blood pressure in rat and lipid levels in human. *Dgat2* resides in pQTL for body weight and blood pressure in both human and rat. DGAT enzymes are involved in triglyceride synthesis and catalyse a reaction where diacylglycerol is joined to long chain fatty acids (LCFA). Disruption in LCFA synthesis occurs in CD36 deficient patients therefore mutations in this gene could have a similar effect. In addition, PMVK (phosphomevalonate kinase) and BZRP are key enzyme cholesterol biosynthesis and flow into the mitochondria respectively, mutations in which could have an effect on lipid levels. This means that a group of genes which should be given priority for genetic and functional testing in human populations have been identified and increases the likelihood that genes underlie a pQTL in both human and rat.

Genome databases are dynamic entities and since the above analyses were carried out I have looked at RGD again to determine the number of SHR pQTL represented. As

of 3rd May 2006 there are approximately 1100 rat QTL of which 236 are in SHR strain combinations with 174 occurring in cardiovascular and metabolic traits. Therefore this analysis could be repeated with new QTL added to identify more candidates for pQTL that merit genetic and functional testing.

The data from the human QTL was also used in a comparative mapping study in a mouse model of type 2 diabetes, the *Irs*2 knockout mouse. This mouse model lacks the insulin receptor substrate 2 (Irs2) which is a major intracellular target of the activated insulin receptor. These mice have peripheral insulin resistance and develop diabetes in a few weeks as Irs2 signalling is also required for beta cell adaption to insulin resistance (Withers et al. 1998). Gene expression and metabonomic data was derived in three tissues, liver, white adipose tissue and skeletal muscle, from Irs2 knockout mice as part of the BAIR consortium. The genes identified as differentially expressed when the metabonomic and microarray data was integrated were investigated to determine if they lay within QTL for human and mouse metabolic syndrome phenotypes. The human QTL displayed in **Table 5.2** was filtered for Lander and Kruglyak significance and only for human QTL which were related to insulin resistance. This identified 84 genes from the mouse data set which had human orthologues and are potential candidates underlying human insulin resistance and type 2 diabetes (KM, JM, TJA unpublished data).

5.4 Conclusion

In conclusion the work in this chapter provides valuable insight into the uses of comparative mapping between rat and human for the selection of potential candidate

genes for rat and human physiological and pathophysiological phenotypes. Since this method detected genes with a known role in insulin resistance and hypertension, key features of the metabolic syndrome, this provides validation of this technique for generation of a list of potential candidates for human disease using rat eQTL data.

Chapter 6: General Discussion

The overall aim of this study was to use transcriptional profiling in a segregating population to integrate the substantial body of genetic and phenotypic data available for the BXH/HXB RI strains by generating and mapping gene expression profiles across the strains (and SHR and BN parentals) using DNA microarray technology. Following generation of gene expression data, linkage analysis was used to identify the genetic determinants of gene expression and then *cis*-acting eQTL selected as potential candidate genes for human cardiovascular and metabolic phenotypes. The specific aims were set as: generate microarray data for a subset of the BXH/HXB recombinant inbred strains and SHR and BN parental strains, use linkage analysis of the entire microarray data set to identify differentially expressed genes that are controlled in *cis* and in *trans*, to prioritise genes as candidates for sequencing in the promoter region to identify variants which may underlie the differential expression and thus biological phenotype in *cis*-regulated eQTL and finally to carry out comparative mapping of *cis*-acting eQTL to identify candidate genes which potentially underlie human disease.

In **Chapter 3** the generation, validation and normalisation of the microarray data sets is presented and the linkage analysis and *cis*/*trans* determination was carried out.

In **Chapter 4** a subset of the robustly mapped (FDR < 5%, approximately $P<10^{-4}$) *cis*-acting eQTL genes were examined to identify sequence polymorphisms in the putative promoters that may explain the differential expression.

In **Chapter 5** a comparative mapping study was carried out to investigate the applicability of the data set to human disease gene identification. Human physiological QTL for metabolic syndrome traits were mapped to the rat genome and *cis*-acting eQTL in these syntenic regions identified. The human orthologs were identified and thus merit testing in human populations as potential candidates underlying the pQTL.

The results presented in **Chapter 3** for the microarray and linkage analysis identified 2,118 and 2,490 non redundant eQTL linkages in fat and kidney tissue respectively at genome wide significance ($P<0.05$). In addition, eQTL were classified as *cis* and *trans*-acting and, as demonstrated in previous studies, (Brem et al. 2002; Schadt et al. 2003; Morely et al. 2004) *cis*-acting eQTL are more significant. Initial investigations carried out on 42 differentially expressed genes between the SHR and BN parental strains demonstrated that gene expression is heritable and can be mapped in accordance with previous studies (Brem et al. 2002; Schadt et al. 2003). These data provided the basis to carry out the whole genome linkage analysis using QTL Reaper (Wang et al. 2003). Sequence variation in the probe region was investigated in these initial genes and it was discovered that 7 had polymorphisms in the target region with 6 having (a) polymorphism(s) in an actual probe sequence. In three cases these SNPs could account for the differential expression because the single probe which contained the SNP was determined to be significantly differentially expressed when analysed from unnormalised data. The analysis in **Chapter 3** provided a basis for the work in **Chapters 4** and **5**.

In **Chapter 4** the sequencing strategy was to identify promoter sequence variants in a number of *cis*-acting eQTL genes as these may account for the differential expression detected. The putative promoter was identified by finding the most 5' exon of the gene to which the Affymetrix probe set mapped and sequencing up to 2.5 Kb upstream of this. A selection of *trans* regulated genes were also sequenced in the putative promoter region to determine if the *cis*-acting genes had an higher overall SNP frequency and it was concluded that they did. The sequence variants detected remain to be tested in reporter assays to assess their effect on gene expression.

In **Chapter 5** the comparative mapping detected 142 *cis*-acting eQTL which represent candidate genes underlying both SHR and human metabolic syndrome traits. Some of the genes identified in this analysis have a known association with metabolic syndrome phenotypes such as hypertension and obesity and thus the other genes presented are good potential candidates and merit testing in human populations.

Overall the results show the potential utility of integrated linkage analysis and gene expression, the "genetical genomics" approach, to the study of SHR and human metabolic and cardiovascular disease. However there are some points for discussion and questions which are raised by the work presented here.

Major challenges for geneticists in the post-genome era are to map the genetic polymorphisms responsible for variation in disease susceptibility and to determine their allele frequencies, mode of inheritance and effects in different environments. While it is relatively easy to localise QTL using genome scans, identifying the genetic locus, or gene(s), that underlies each QTL requires high-resolution recombination or

linkage disequilibrium mapping to identify candidate genes, followed by genetic and/or functional complementation and gene expression analyses (Mackay 2001). In the search for genes underlying complex traits, an approach utilising a combined genetic linkage and microarray strategy proved successful in identifying the gene underlying the chromosome 4 insulin resistance QTL in SHR (Aitman et al. 1999). This approach has also been used to identify other disease genes such as a susceptibility gene in a model of experimental allergic asthma (Karp et al. 2000). The success of this integrated approach in previous studies of complex traits led to the use of this strategy in the studies presented here to identify candidate genes for cardiovascular and metabolic phenotypes in SHR on a genome wide scale. The approach used in this study combined microarray expression profiling and genome wide linkage analysis following the suggestion of a 'gentical genomic' approach by Jansen and Nap (Jansen and Nap 2001) and the first example using yeast (Brem et al. 2002).

The gentical genomic approach was utilised in this study in fat and kidney tissue, two tissues which are important in the pathophysiology of the metabolic syndrome. The success of this strategy in identifying genes underlying complex traits is dependant on the robustness of the genetic map and the linkage analysis, the quality of the gene expression data and the ability to prioritise and identify candidate genes. While no one specific gene has yet been definitively identified as the causative gene in any specific phenotype many potential positional candidates underlying previously mapped SHR pQTL were investigated, and some human candidate genes have also been suggested by these analyses. At least two genes with a known involvement in insulin resistance and hypertension in SHR, *Cd36* and *Gstm1* respectively have been identified by this

analysis. This demonstrates the utility of this approach and that some of the genes identified will be involved in the metabolic syndrome and have molecular variants which underlie the SHR pQTL.

The advantage of gene expression studies is that the entire (known) transcriptome can be assayed in one reaction. However microarray experiments are expensive and the large data sets require much data handling and statistics, therefore the experimental design, for example tissue type, age of animals and numbers of replicates must be carefully considered and optimised. There are many tissues involved in the metabolic syndrome and fat tissue plays an important role since it is metabolically active and obesity is a major cause of heart disease and diabetes while the kidney is critical in blood pressure homeostasis. The tissues used in this experiment were extracted from same parental and RI strain animals and several other tissues relevant to the pathogenesis of metabolic syndrome were extracted at the same time, and microarray data sets are now available for left ventricle and adrenal tissue. The animals were all the same age at the time of tissue harvesting and this means the SHR will be both hypertensive and insulin resistant. The number of replicates was also considered in this study and 4 was chosen as this is lowest number which allows outlier detection to be carried out.

Microarray normalisation methods attempt to remove experimental noise which allows easier analysis of the data. Several methods of normalisation exist and many comparisons have been carried out and the differences between methods well documented (Quackenbush 2002; Bolstad et al. 2003; Irizarry et al. 2005). In this thesis RMA and MAS5.0 were compared and RMA selected as it has been shown to

outperform MAS5.0 (Irizarry et al. 2003) and resulted in a consistently lower FDR in this data set.

The previous studies which used genetical genomics (Brem et al. 2002; Schadt et al. 2003) found that differential expression was highly heritable and that approximately one third of genes had at least one eQTL. *cis*-acting eQTL are normally found to be more significant which reflects the fact that these represent a first order effect unlike *trans* eQTL which are second order effects and therefore smaller. The work presented in this thesis detected similar results.

In **Chapter 4** a comparison of SNP frequency between the *cis* and *trans*-regulated eQTL genes was undertaken and the higher SNP frequency seen in *cis*-acting eQTL genes may reflect the fact that these genes have a causative variant in their sequence or regulatory regions. However it may also be due to the fact that *cis*-acting eQTL genes lie in SNP enriched regions as they tend to be in the regions of the genome not inherited IBS between the parental strains. Doss *et al* (Doss et al. 2005) determined that more than 96% of *cis* eQTL in their data set lay within regions not inherited IBD and this could be tested in the BN/SHR strain combination as and when haplotype maps become available for the rat genome. Studying the regions IBD may also be helpful when prioritising candidates as non IBD regions (betwent he parents) are more likely to be involved in disease. The *cis/trans* definition is based purely on the relative position of the eQTL and the probe set and therefore it is possible that some of the *cis*-acting eQTL may actually be under *trans* control. A test was carried out to define the size of the *cis* eQTL window and 10Mb was determined as optimum.

One of the weaknesses of this study is that no differential expression validation has been carried out to determine if the fold change can be replicated. Some validation of differential gene expression in the kidney was carried out by our collaborators at the MDC in Berlin (Hubner et al. 2005) and using qRT-PCR they conformed differential expression in 14 of the 15 genes tested. The final gene was an immune gene in the RT1 locus, which are known to be highly polymorphic and therefore the differential expression could be due to these polymorphisms causing differential hybridisation of the probe.

Probe set sequencing was carried out following the promoter sequencing in **Chapter 4**. SNPs were detected in 3 of the genes. These SNPs may account for the differential expression seen on the microarray and it may be beneficial to verify the expression before further analysis. However as it has been previously demonstrated that SNPs in the probe region account for very small amounts of differential expression (Hubner et al. 2005) it was not decided to routinely sequence the target regions.

While the aim of the sequencing strategy employed here was to detect promoter variants polymorphisms in other parts of the gene could be equally important. Variation detected in the 3'UTR could have an effect on post transcription modification and nonsense mediated decay while coding SNPs could lead to changes in amino acids and intronic SNPs at intron/exon boundaries may lead to alternative splicing. Any or all of these could lead to altered transcription however the promoter was focused on in **Chapter 4** of this thesis as this is the most likely region for transcription factor binding sites and alterations in these regions would lead to differential expression.

Two recent publications presented strategies which may have been useful when selecting genes to sequence. Yaguchi *et al* (Yaguchi et al. 2005) combined eQTL with SNP analysis in the coding region (cSNPs) to accentuate discovery of candidate genes in a mouse model of type 2 diabetes. This could be used to identify likely SNPs in region under study; however this approach may not be as useful in rat studies as there is much less SNP information available. Xie *et al* (Xie et al. 2005) used a systematic comparative genomic method to identify regulatory elements in both the promoter and the 3'UTR. They compared human, mouse, rat and dog genomes and detected 174 promoter motifs, including some known transcription factor binding sites and 106 3'UTR motifs. This type of analysis might be useful when deciding which genes to select for sequencing and which nucleotides may be important in regulation.

6.1 Concluding remarks

Complex traits, such as hypertension and metabolic syndrome are a substantial burden on public health. The ability to identify genes which underlie these traits will lead to improved prevention and treatment of these disorders and their associated cardiovascular risk.

The results described in this thesis demonstrate the use of a genome wide combined genetic linkage analysis and gene expression using microarrays to identify candidate genes which may underlie complex genetic traits in a rat disease model.

References

(1998). "Influence of pravastatin and plasma lipids on clinical events in the West of Scotland Coronary Prevention Study (WOSCOPS)." Circulation **97**(15): 1440-5.

(2002). "Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report." Circulation **106**(25): 3143-421.

Abbasi, F., B. W. Brown, Jr., C. Lamendola, T. McLaughlin and G. M. Reaven (2002). "Relationship between obesity, insulin resistance, and coronary heart disease risk." J Am Coll Cardiol **40**(5): 937-43.

Affymetrix (2002). Statistical algorithms document.

Aitman, T. J., A. M. Glazier, C. A. Wallace, L. D. Cooper, P. J. Norsworthy, F. N. Wahid, et al. (1999). "Identification of *Cd36* (*Fat*) as an insulin resistance gene causing defective fatty acid and glucose metabolism in hypertensive rats." Nat Genet **21**: 76-83.

Aitman, T. J., I. F. Godsland, B. Farren, D. Crook, H. J. Wong and J. Scott (1997a). "Defects of insulin action on fatty acid and carbohydrate metabolism in familial combined hyperlipidemia." Arterioscler Thromb Vasc Biol **17**: 748-754.

Aitman, T. J., T. Gotoda, A. L. Evans, H. Imrie, K. Heath, P. M. Trembling, et al. (1997b). "Quantitative trait loci for cellular defects in glucose and fatty acid metabolism in hypertensive rats." Nat Genet **16**: 197-201.

Aitman, T. J., M. Pravenec, T. W. Kurtz and J. Scott (1998). "The spontaneously hypertensive rat: a model to dissect cellular defects in glucose and fatty acid metabolism." Pathol Biol (Paris) **46**: 693-694.

Alberti, K. G. and P. Z. Zimmet (1998). "Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation." Diabet Med **15**(7): 539-53.

Altshuler, D., J. N. Hirschhorn, M. Klannemark, C. M. Lindgren, M.-C. Vohl, J. Nemesh, et al. (2000). "The common PPARγ Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes." Nature Genet **26**: 76-79.

Arner, P. (1997). "Is familial combined hyperlipidaemia a genetic disorder of adipose tissue." Curr Opin Lipidol **8**: 89-94.

Balkau, B., M. Vernay, L. Mhamdi, M. Novak, D. Arondel, S. Vol, et al. (2003). "The incidence and persistence of the NCEP (National Cholesterol Education Program) metabolic syndrome. The French D.E.S.I.R. study." Diabetes Metab **29**(5): 526-32.

Baum, M. (1987). "Insulin stimulates volume absorption in the rabbit proximal convoluted tubule." J Clin Invest **79**(4): 1104-9.

Beevers, D. and G. A. MacGregor (1999). Hypertension in Pratice. London, Martin Dunitz.

Bell, C. G., A. J. Walley and P. Froguel (2005). "The genetics of human obesity." Nat Rev Genet **6**(3): 221-34.

Bolstad, B. M., R. A. Irizarry, M. Astrand and T. P. Speed (2003). "A Comparison of Normalization Methods for High Density Oligonucleotide Array Data Based on Bias and Variance." Bioinfomatics **19**: 185-193.

Bottger, A., H. A. van Lith, V. Kren, D. Krenova, V. Bila, J. Vorlicek, et al. (1996). "Quantitative trait loci influencing cholesterol and phospholipid phenotypes map to chromosomes that contain genes regulating blood pressure in the spontaneously hypertensive rat." J Clin Invest **98**: 856-862.

Bougneres, P., C. Le Stunff, C. Pecqueur, E. Pinglier, P. Adnot and D. Ricquier (1997). "In vivo resistance of lipolysis to epinephrine. A new feature of childhood onset obesity." J Clin Invest **99**: 2568-2573.

Brem, R. B., G. Yvert, R. Clinton and L. Kruglyak (2002). "Genetic dissection of transcriptional regulation in budding yeast." Science **296**: 752-755.

Buchanan, T. A., J. H. Youn, V. M. Campese and G. F. Sipos (1992). "Enhanced glucose tolerance in spontaneously hypertensive rats: pancreatic beta cell function with normal insulin sensitivity." Diabetes **41**: 872-878.

Campbell, D. J., A. M. Duncan, A. Kladis and S. B. Harrap (1995). "Angiotensin peptides in spontaneously hypertensive and normotensive Donryu rats." Hypertension **25**(5): 928-34.

Carey, D. G., A. B. Jenkins, L. V. Campbell, J. Freund and D. J. Chisholm (1996). "Abdominal fat and insulin resistance in normal and overweight women: Direct measurements reveal a strong relationship in subjects at both low and high risk of NIDDM." Diabetes **45**(5): 633-8.

Catena, C., M. Novello, R. Lapenna, S. Baroselli, G. Colussi, E. Nadalini, et al. (2005). "New risk factors for atherosclerosis in hypertension: focus on the prothrombotic state and lipoprotein(a)." J Hypertens **23**(9): 1617-31.

Chen, Y.-D. I., A. Goolay, A. L. M. Swiwlocki and G. M. Reaven (1987). "Resistance to insulin suppression of plasma free fatty acid concentrations and insulin stimulation of glucose uptake in noninsulin-dependent diabetes mellitus." J Clin Endocrinol Metab **64**: 17-21.

Chiappe de Cingolani, G. E. (1988). "Adipocyte responsiveness to norepinephrine in spontaneously hypertensive rats." Metabolism **37**: 318-322.

Churchill, G. A. and R. W. Doerge (1994). "Empirical threshold values for quantitative trait mapping." Genetics **138**: 963-971.

Clausen, J. O., T. Hansen, C. Bjorbaek, S. M. Echwald, S. A. Urhammer, S. Rasmussen, et al. (1995). "Insulin resistance: interactions between obesity and a common variant of insulin receptor substrate-1." The Lancet **346**: 397-402.

Collison, M., A. M. Glazier, D. Graham, J. J. Morton, M. H. Dominiczak, T. J. Aitman, et al. (2000). "*Cd36* and molecular mechanisms of insulin resistance in the stroke prone spontaneously hypertensive rat." Diabetes **49**: 2222-2226.

Dahl, L. K. and M. Heine (1975). "Primary role of renal homografts in setting chronic blood pressure levels in rats." Circ Res **36**(6): 692-6.

Dahl, L. K., M. Heine and L. Tassinari (1962). "Role of genetic factors in susceptibility to experimental hypertension due to chronic excess salt ingestion." Nature **194**: 480-482.

Darvasi, A. (2005). "Dissecting complex traits: the geneticists' "Around the world in 80 days"." Trends Genet **21**(7): 373-6.

DeFronzo, R. A., C. R. Cooke, R. Andres, G. R. Faloona and P. J. Davis (1975). "The effect of insulin on renal handling of sodium, potassium, calcium, and phosphate in man." J Clin Invest **55**(4): 845-55.

Desbois-Mouthon, C., C. Danan, S. Amselem, M. J. Blivet-Van Eggelpoel, C. Sert-Langeron, M. Goossens, et al. (1996). "Severe resistance to insulin and insulin-like growth factor-I in cells from a patient with leprechaunism as a result of two mutations in the tyrosine kinase domain of the insulin receptor." Metabolism **45**: 1493-1500.

Dickson, M. E. and C. D. Sigmund (2006). "Genetic basis of hypertension: revisiting angiotensinogen." Hypertension **48**(1): 14-20.

Doss, S., E. E. Schadt, T. A. Drake and A. J. Lusis (2005). "Cis-acting expression quantitative trait loci in mice." Genome Res **15**(5): 681-91.

Elliott, P., J. Stamler, R. Nichols, A. R. Dyer, R. Stamler, H. Kesteloot, et al. (1996). "Intersalt revisited: further analyses of 24 hour sodium excretion and blood pressure within and across populations. Intersalt Cooperative Research Group." Bmj **312**(7041): 1249-53.

Falconer, D. S. and T. F. C. Mackay (1996). Introduction to quantitative genetics. Harlow, Addison Wesley Longman.

Farrall, M. (2004). "Quantitative genetic variation: a post-modern view." Hum Mol Genet **13 Spec No 1**: R1-7.

Ferrannini, E., G. Buzzigoli, R. Bonadonna, M. A. Giorico, M. Oleggini, L. Graziadei, et al. (1987). "Insulin resistance in essential hypertension." N Engl J Med **317**: 350-357.

Garrett, M. R., H. Dene, R. Walder, Q.-Y. Zhang, G. T. Cicila, S. Assadnia, et al. (1998). "Genome scan and congenic strains for blood pressure QTL using Dahl salt-sensitive rats." Genome Res **8**: 711-723.

Gibbs, R. A., G. M. Weinstock, M. L. Metzker, D. M. Muzny, E. J. Sodergren, S. Scherer, et al. (2004). "Genome sequence of the Brown Norway rat yields insights into mammalian evolution." Nature **428**(6982): 493-521.

Gillum, R. F. (1996). "Epidemiology of hypertension in African American women." Am Heart J **131**(2): 385-95.

Glazier, A. M., J. Scott and T. J. Aitman (2002). "Molecular basis of the Cd36 chromosomal deletion underlying SHR defects in insulin action and fatty acid metabolism." Mamm Genome **13**: 108-113.

Goldstein, J. H., H. G. Schrott, W. R. Hazzard, E. L. Bierman and A. G. Motulsky (1973). "Hyperlipidemia in coronary heart disease II. Genetic analysis of lipid levels in 176 families and delineation of a new inherited disorder, combined hyperlipidemia." J Clin Invest **52**: 1544-1568.

Goldstein, J. L., W. R. Hazzard, H. G. Schrott, E. L. Bierman and A. G. Motulsky (1973). "Hyperlipidemia in coronary heart disease. I. Lipid levels in 500 survivors of myocardial infarction." J Clin Invest **52**(7): 1533-43.

Golub, T. R., D. K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J. P. Mesirov, et al. (1999). "Molecular classification of cancer: class discovery and class prediction by gene expression monitoring." Science **286**: 531-537.

Guyton, A. C., T. G. Coleman, A. V. Cowley, Jr., K. W. Scheel, R. D. Manning, Jr. and R. A. Norman, Jr. (1972). "Arterial pressure regulation. Overriding dominance of the kidneys in long-term regulation and in hypertension." Am J Med **52**(5): 584-94.

Hara, K., T. Okada, K. Tobe, K. Yasuda, Y. Mori, H. Kadowaki, et al. (2000). "The Pro12Ala polymorphism in PPAR gamma2 may confer resistance to type 2 diabetes." Biochem Biophys Res Commun **27**: 212-216.

Harrap, S. B. (1994). Genetic Approaches to Hypertension. Textbook of Hypertension. J. Swales. Oxford, Blackwell Scientific.

Harrington, C. A., C. Rosenow and J. Retief (2000). "Monitoring gene expression using DNA microarrays." Curr Opin Microbiol **3**(3): 285-91.

Hingorini, A. (2000). Genes, Blood Pressure and Pressor systems. Handbook of hypertension. C. Bulpitt. New York, Elsevier Science Publishers. **20**.

Hirschhorn, J. N. and D. Altshuler (2002). "Once and again-issues surrounding replication in genetic association studies." J Clin Endocrinol Metab **87**(10): 4438-41.

Hirschhorn, J. N. and M. J. Daly (2005). "Genome-wide association studies for common diseases and complex traits." Nat Rev Genet **6**(2): 95-108.

Horie, R., Y. Yamori, Y. Nara, M. Sawamura and S. Mizushima (1991). "Aggravating effects of isolated caging on the development of hypertension and its complications in stroke-prone spontaneously hypertensive rats (SHRSP) and Wistar-Kyoto rats (WKY)." Clin Exp Hypertens A **13**(5): 859-64.

Horikawa, Y., N. Oda, N. J. Cox, X. Li, M. Orho-Melander, M. Hara, et al. (2000). "Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus." Nature Genet **26**: 163-175.

Hubner, N., C. A. Wallace, H. Zimdahl, E. Petretto, H. Schulz, F. Maciver, et al. (2005). "Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease." Nat Genet **37**(3): 243-53.

Hulman, S., B. Falkner and Y. Chen (1991). "Insulin resistance in the spontaneously hypertensive rat." Metabolism **40**: 359-61.

Hulman, S., B. Falkner and N. Freyvogel (1993). "Insulin resistance in the conscious spontaneously hypertensive rat: euglycemic hyperinsulinemic clamp study." Metabolism **42**: 14-18.

Hwang, I. S., H. Ho, B. B. Hoffman and G. M. Reaven (1987). "Fructose-induced insulin resistance and hypertension in rats." Hypertension **10**: 512-516.

Iritani, N., E. Fukuda, Y. Nara and Y. Yamori (1977). "Lipid metabolism in spontaneously hypertensive rats (SHR)." Atherosclerosis **28**: 217-222.

Irizarry, R. A., B. M. Bolstad, F. Collin, L. M. Cope, B. Hobbs and T. P. Speed (2003). "Summaries of Affymetrix GeneChip probe level data." Nucleic Acids Res **31**(4): e15.

Irizarry, R. A., B. Hobbs, F. Collin, Y. D. Beazer-Barclay, K. J. Antonellis, U. Scherf, et al. (2003). "Exploration, normalization, and summaries of high density oligonucleotide array probe level data." Biostatistics **4**: 249-264.

Irizarry, R. A., D. Warren, F. Spencer, I. F. Kim, S. Biswal, B. C. Frank, et al. (2005). "Multiple-laboratory comparison of microarray platforms." Nat Methods **2**(5): 345-50.

Isomaa, B., P. Almgren, T. Tuomi, B. Forsen, K. Lahti, M. Nissen, et al. (2001). "Cardiovascular morbidity and mortality associated with the metabolic syndrome." Diabetes Care **24**(4): 683-9.

Jacob, H. J., K. Lindpaintner, S. E. Lincoln, K. Kusumi, R. K. Bunker, Y. P. Mao, et al. (1991). "Genetic mapping of a gene causing hypertension in the stroke-prone spontaneously hypertensive rat." Cell **67**(1): 213-24.

Jansen, R. C. and J.-P. Nap (2001). "Genetical Genomics: the added value from segregation." Trends Genet **17**: 388-391.

Jeffs, B., J. S. Clark, N. H. Anderson, J. Gratton, M. J. Brosnan, D. Gauguier, et al. (1997). "Sensitivity to cerebral ishchaemic insult in rat model of stroke is determined by a single genetic locus." Nat Genet **16**: 364-367.

Jospe, N., P. B. Kaplowitz and R. W. Furlanetto (1996). "Homozygous nonsense mutation in the insulin receptor gene of a patient with severe congenital insulin resistance: leprechaunism and the role of the insulin-like growth factor receptor." Clin Endocrinol (Oxf) **45**(2): 229-35.

Kannel, W. B. (2000). "Fifty years of Framingham Study contributions to understanding hypertension." J Hum Hypertens **14**(2): 83-90.

Karp, C. L., A. Grupe, E. Schadt, S. L. Ewart, M. Keane-Moore, P. J. Cuomo, et al. (2000). "Identification of complement factor 5 as a susceptibility locus for experimental allergic asthma." Nature Immunol **1**: 221-226.

Kishida, K., H. Nagaretani, H. Kondo, H. Kobayashi, S. Tanaka, N. Maeda, et al. (2003). "Disturbed secretion of mutant adiponectin associated with the metabolic syndrome." Biochem Biophys Res Commun **306**(1): 286-92.

Kondo, H., I. Shimomura, Y. Matsukawa, M. Kumada, M. Takahashi, M. Matsuda, et al. (2002). "Association of adiponectin mutation with type 2 diabetes: a candidate gene for the insulin resistance syndrome." Diabetes **51**(7): 2325-8.

Kooner, J. S., R. R. Baliga, J. Wilding, J. Hulse, D. Crook, C. J. Packard, et al. (1998). "Abdominal obesity, impaired nonesterified fatty acid suppression, and insulin-mediated glucose disposal are early metabolic abnormalities in families with premature myocardial infarction." Arterioscler Thromb Vasc Biol **18**: 1021-1026.

Kvetnansky, R., R. McCarty, N. B. Thoa, C. R. Lake and I. J. Kopin (1979). "Sympatho-adrenal responses of spontaneously hypertensive rats to immobilization stress." Am J Physiol **236**(3): H457-62.

Laird, P. W., A. Zijderveld, K. Linders, M. A. Rudnicki, R. Jaenisch and A. Berns (1991). "Simplified mammalian DNA isolation procedure." Nucleic Acids Research **19**(15): 4293.

Lakka, H. M., D. E. Laaksonen, T. A. Lakka, L. K. Niskanen, E. Kumpusalo, J. Tuomilehto, et al. (2002). "The metabolic syndrome and total and cardiovascular disease mortality in middle-aged men." Jama **288**(21): 2709-16.

Lander, E. and L. Kruglyak (1995). "Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results." Nature Genet **11**: 241-247.

Lander, E. S. and N. J. Schork (1994). "Genetic dissection of complex traits." Science **265**: 2037-2048.

Lapidus, L., G. Lindstedt, P. A. Lundberg, C. Bengtsson and T. Gredmark (1986). "Concentrations of sex-hormone binding globulin and corticosteroid binding globulin in serum in relation to cardiovascular risk factors and to 12-year incidence of cardiovascular disease and overall mortality in postmenopausal women." Clin Chem **32**(1 Pt 1): 146-52.

Lehmann, J. M., L. B. Moore, T. A. Smith-Oliver, W. O. Wilkison, T. M. Willson and S. A. Kliewer (1995). "An antidiabetic thiazolidinedione is a high affinity ligand for peroxisome proliferator-activated receptor gamma (PPAR gamma)." J Biol Chem **270**: 12953-12956.

Lifton, R. P., A. G. Gharavi and D. S. Geller (2001). "Molecular mechanisms of human hypertension." Cell **104**: 545-556.

Lindpaintner, K., M. Lee, M. G. Larson, V. S. Rao, M. A. Pfeffer, J. M. Ordovas, et al. (1996). "Absence of association or genetic linkage between the angiotensin-converting-enzyme gene and left ventricular mass." N Engl J Med **334**(16): 1023-8.

Lipshutz, R. J., S. P. Fodor, T. R. Gingeras and D. J. Lockhart (1999). "High density synthetic oligonucleotide arrays." Nat Genet **21**(1 Suppl): 20-4.

MacGregor, G. A. (1985). "Sodium is more important than calcium in essential hypertension." Hypertension **7**(4): 628-40.

Mackay, T. F. (2001). "The genetic architecture of quantitative traits." Annu Rev Genet **35**: 303-39.

Manly, K. F., R. H. J. Cudmore and J. M. Meer (2001). "Map Manager QTX, cross-platform software for genetic mapping." Mamm Genome **12**: 930-2.

McBride, M. W., M. J. Brosnan, J. Mathers, L. I. McLellan, W. H. Miller, D. Graham, et al. (2005). "Reduction of Gstm1 expression in the stroke-prone spontaneously hypertension rat contributes to increased oxidative stress." Hypertension **45**(4): 786-92.

McBride, M. W., F. J. Carr, D. Graham, N. H. Anderson, J. S. Clark, W. K. Lee, et al. (2003). "Microarray analysis of rat chromosome 2 congenic strains." Hypertension **41**(3 Pt 2): 847-53.

McCarthy, M. I. (2004). "Progress in defining the molecular basis of type 2 diabetes mellitus through susceptibility-gene identification." Hum Mol Genet **13 Spec No 1**: R33-41.

McLaughlin, T., F. Abbasi, K. Cheal, J. Chu, C. Lamendola and G. Reaven (2003). "Use of metabolic markers to identify overweight individuals who are insulin resistant." Ann Intern Med **139**(10): 802-9.

Meirhaeghe, A., L. Fajas, N. Helbecque, D. Cottel, J. Auwerx, S. S. Deeb, et al. (2000). "Impact of the Peroxisome Proliferator Activated Receptor gamma2 Pro12Ala polymorphism on adiposity, lipids and non-insulin-dependent diabetes mellitus." Int J Obes Relat Metab Disord **24**: 195-199.

Merikangas, K. R. and N. Risch (2003). "Genomic priorities and public health." Science **302**: 599-601.

Mondon, C. E. and G. M. Reaven (1988). "Evidence of abnormalities of insulin metabolism in rats with spontaneous hypertension." Metabolism **37**: 303-305.

Montgomery, H. E., P. Clarkson, C. M. Dollery, K. Prasad, M. A. Losi, H. Hemingway, et al. (1997). "Association of angiotensin-converting enzyme gene I/D polymorphism with change in left ventricular mass in response to physical training." Circulation **96**(3): 741-7.

Morely, M., C. M. Molony, T. M. Weber, D. J.L., K. G. Ewans, R. S. Speilman, et al. (2004). "Genetic analysis of genome-wide variation in human gene expression." Nature **430**: 743-747.

Morton, N. E. (2005). "Linkage disequilibrium maps and association mapping." J Clin Invest **115**(6): 1425-30.

Mueller, M., A. Goel, M. Thimma, N. J. Dickens, T. J. Aitman and J. Mangion (2006). "eQTL Explorer: integrated mining of combined genetic linkage and expression experiments." Bioinformatics **22**(4): 509-11.

Munroe, P. B. and M. J. Caulfield (2000). "Genetics of hypertension." Curr Opin Genet Dev **10**: 325-329.

Nelson, K. M., R. E. Shepherd and J. A. Spitzer (1987). "Lipolysis and beta-adrenergic receptor binding on adipocytes of spontaneously hypertensive rats." Biochem Med Metabol Biol **37**: 51-60.

Ohashi, K., N. Ouchi, S. Kihara, T. Funahashi, T. Nakamura, S. Sumitsuji, et al. (2004). "Adiponectin I164T mutation is associated with the metabolic syndrome and coronary artery disease." J Am Coll Cardiol **43**(7): 1195-200.

Okamoto, K., Y. Yamori, A. Goshima, C. Park, H. Haebara, M. Matsumoto, et al. (1972). Establishment of the inbred strain of the spontaneously hypertensive rat and genetic factors involved in hypertension. Tokyo, Igaku Shoin Ltd.

Oliver, W. J., E. L. Cohen and J. V. Neel (1975). "Blood pressure, sodium intake, and sodium related hormones in the Yanomamo Indians, a "no-salt" culture." Circulation **52**(1): 146-51.

Page, L. B., A. Damon and R. C. Moellering, Jr. (1974). "Antecedents of cardiovascular disease in six Solomon Islands societies." Circulation **49**(6): 1132-46.

Paternostro, G., K. Clarke, J. Heath, A. M. Seymour and G. K. Radda (1995). "Decreased GLUT-4 mRNA content and insulin-sensitive deoxyglucose uptake show insulin resistance in the hypertensive rat heart." Cardiovasc Res **30**: 205-211.

Pausova, Z., L. Sedova, J. Berube, P. Hamet, J. Tremblay, M. Dumont, et al. (2003). "Segment of rat chromosome 20 regulates diet-induced augmentations in adiposity, glucose intolerance, and blood pressure." Hypertension **41**: 1047-55.

Peltonen, L., M. Perola, J. Naukkarinen and A. Palotie (2006). "Lessons from studying monogenic disease for common disease." Hum Mol Genet **15 Spec No 1**: R67-74.

Pickering, G. (1961). The nature of Essential Hypertension. London, J & A Churchill.

Pischon, T., C. J. Girman, G. S. Hotamisligil, N. Rifai, F. B. Hu and E. B. Rimm (2004). "Plasma adiponectin levels and risk of myocardial infarction in men." Jama **291**(14): 1730-7.

Pravenec, M., D. Gauguier, J.-J. Schott, J. Buard, V. Kren, V. Bila, et al. (1995). "Mapping of quantitative trait loci for blood pressure and cardiac mass in the rat by genome scanning of recombinant inbred strains." J Clin Invest **96**: 1973-1978.

Pravenec, M., D. Gauguier, J.-J. Schott, J. Buard, V. Kren, V. Bila, et al. (1996). "A genetic linkage map of the rat derived from recombinant inbred strains." Mammalian Genome **7**: 117-127.

Pravenec, M., P. Klir, V. Kren, J. Zicha and J. Kunes (1989). "An analysis of spontaneous hypertension in spontaneously hypertensive rats by means of new recombinant inbred strains." J Hypertens **7**(3): 217-21.

Printz, M. P., M. Jirout, R. Jaworski, A. Alemayehu and V. Kren (2003). "Genetic Models in Applied Physiology. HXB/BXH rat recombinant inbred strain platform: a newly enhanced tool for cardiovascular, behavioral, and developmental genetics and genomics." J Appl Physiol **94**(6): 2510-22.

Qu, X. and R. Donnelly (1997). "Is insulin resistance in the spontaneously hypertensive rat related to changes in protein kinase C in skeletal muscle?" Am J Hypertens **10**: 1053-1057.

Quackenbush, J. (2002). "Microarray data normalization and transformation." Nat Genet **32 Suppl**: 496-501.

Rao, R. H. (1993). "Insulin resistance in spontaneously hypertensive rats: difference in interpretation based on insulin infusion rate or on plasma insulin in glucose clamp studies." Diabetes **42**: 1364-1371.

Rapp, J. P. (2000). "Genetic analysis of inherited hypertension in the rat." Physiol Rev **80**(1): 135-72.

Rapp, J. P. and H. Dene (1985). "Development and characteristics of inbred strains of Dahl salt-sensitive and salt-resistant rats." Hypertension **7**: 340-349.

Rapp, J. P., R. P. McPartland and C. L. Batten (1984). "Isoelectric focusing patterns of urinary kallikrein in Dahl salt-hypertension susceptible and resistant rats." Hypertension **6**(4): 519-25.

Reaven, G. M. (1988a). "Banting Lecture 1988. Role of insulin resistance in human disease." Diabetes **37**: 1595-1607.

Reaven, G. M. (1993). "Role of insulin resistance in human disease (syndrome X): an expanded definition." Annu Rev Med **44**: 121-131.

Reaven, G. M. (1995). "The fourth musketeer - from Alexandre Dumas to Claude Bernard." Diabetologia **38**: 3-13.

Reaven, G. M., H. Chang, B. B. Hoffman and S. Azhar (1989b). "Resistance to insulin-stimulated glucose uptake in adipocytes from spontaneously hypertensive rats." Diabetes **38**: 1155-1160.

Reaven, G. M., H. Lithell and L. Landsberg (1996). "Hypertension and associated metabolic abnormalities - the role of insulin resistance and the sympathoadrenal system." N Engl J Med **334**: 374-381.

Rekhter, M. (2002). "Vulnerable atherosclerotic plaque: emerging challenge for animal models." Curr Opin Cardiol **17**(6): 626-32.

Rekhter, M. D. (2002). "How to evaluate plaque vulnerability in animal models of atherosclerosis?" Cardiovasc Res **54**(1): 36-41.

Rexrode, K. M., J. E. Manson, I. M. Lee, P. M. Ridker, P. M. Sluss, N. R. Cook, et al. (2003). "Sex hormone levels and risk of cardiovascular events in postmenopausal women." Circulation **108**(14): 1688-93.

Reynisdottir, S., K. Ellerfeldt, H. Wahrenberg, H. Lithell and P. Arner (1994). "Multiple lipolysis defects in the insulin resistance (metabolic) syndrome." J Clin Invest **93**: 2590-2599.

Reynisdottir, S., M. Eriksson, B. Angelin and P. Arner (1995). "Impaired activation of adipocyte lipolysis in familial combined hyperlipidemia." J Clin Invest **95**: 2161-2169.

Reynisdottir, S., H. Wahrenberg, K. Carlstrom, S. Rossner and P. Arner (1994). "Catecholamine resistance in fat cells of women with upper-body obesity due to decreased expression of beta2-adrenoceptors." Diabetologia **37**: 428-435.

Rice, T., T. Rankinen, M. A. Province, Y. C. Chagnon, L. Perusse, I. B. Borecki, et al. (2000). "Genome-wide linkage analysis of systolic and diastolic blood pressure: the Quebec Family Study." Circulation 2000 **102**: 1956-63.

Rimm, E. B., M. J. Stampfer, E. Giovannucci, A. Ascherio, D. Spiegelman, G. A. Colditz, et al. (1995). "Body size and fat distribution as predictors of coronary heart disease among middle-aged and older US men." Am J Epidemiol **141**(12): 1117-27.

Risch, N. J. (2000). "Searching for genetic determinants in the new millennium." Nature **405**(6788): 847-56.

Rose, G. (1980). Epidemiology. The hypertensive Patient. A. a. B. Marshall, DW. Kent, Pitman Medical: 1-21.

Rubattu, S., M. Volpe, R. Kreutz, U. Ganten, D. Ganten and K. Lindpaintner (1996). "Chromosomal mapping of quantitative trait loci contributing to stroke in a rat model of complex human disease." Nat Genet **13**: 429-434.

Sattar, N., A. Gaw, O. Scherbakova, I. Ford, D. S. O'Reilly, S. M. Haffner, et al. (2003). "Metabolic syndrome with and without C-reactive protein as a predictor of coronary heart disease and diabetes in the West of Scotland Coronary Prevention Study." Circulation **108**(4): 414-9.

Sawicki, P. T., L. Heinemann, A. Starke and M. Berger (1992). "Hyperinsulinaemia is not linked with blood pressure elevation in patients with insulinoma." Diabetologia **35**(7): 649-52.

Schadt, E. E., S. A. Monks, T. A. Drake, A. J. Lusis, N. Che, V. Colinayo, et al. (2003). "Genetics of gene expression surveyed in maize, mouse and man." Nature **422**: 297-302.

Schulze, M. B., I. Shai, E. B. Rimm, T. Li, N. Rifai and F. B. Hu (2005). "Adiponectin and future coronary heart disease events among men with type 2 diabetes." Diabetes **54**(2): 534-9.

Sealey, J. and J. Laragh (1990). The renin-angiotensin-aldosterone system for normal regulation of blood pressure and sodium and potasssium homeostatis. Hypertension pathopysiology, diagonsis and management. J. Laragh and B. Brenner. New York, Raven Press.

Sechi, L. A. (1999). "Mechanisms of insulin resistance in rat models of hypertension and their relationships with salt sensitivity." Hypertension **17**: 1229-1237.

Sever, P. S., D. Gordon, W. S. Peart and P. Beighton (1980). "Blood-pressure and its correlates in urban and tribal Africa." Lancet **2**(8185): 60-4.

Sharma, P., J. Fatibene, F. Ferraro, H. Jia, S. Monteith, C. Brown, et al. (2000). "A genome-wide search for susceptibility loci to human essential hypertension." Hypertension **35**(6): 1291-6.

Shoulders, C. C., E. L. Jones and R. P. Naoumova (2004). "Genetics of familial combined hyperlipidemia and risk of coronary heart disease." Hum Mol Genet **13 Spec No 1**: R149-60.

Silver, L. M. (1995). Mouse genetics: concepts and applications. New York, Oxford University Press.

Spiegelman, B. M. and J. S. Flier (1996). "Adipogenesis and obesity: rounding out the big picture." Cell **87**: 377-389.

St Lezin, E., L. Simonet, M. Pravenec and T. W. Kurtz (1992). "Hypertensive strains and normotensive 'control' strains. How closely are they related?" Hypertension **19**(5): 419-24.

Starling, E. (1909). The Fluids of the Body. London, Constable.

Storey, J. D. (2002). "A direct approach to false discovery rates." J Royal Stat Soc (B) **64**: 479-498.

Stumvoll, M., O. Tschritter, A. Fritsche, H. Staiger, W. Renn, M. Weisser, et al. (2002). "Association of the T-G polymorphism in adiponectin (exon 2) with obesity and insulin sensitivity: interaction with family history of type 2 diabetes." Diabetes **51**(1): 37-41.

Svartberg, J., D. von Muhlen, H. Schirmer, E. Barrett-Connor, J. Sundfjord and R. Jorde (2004). "Association of endogenous testosterone with blood pressure and left ventricular mass in men. The Tromso Study." Eur J Endocrinol **150**(1): 65-71.

Takada, D., Y. Ezura, S. Ono, Y. Iino, Y. Katayama, Y. Xin, et al. (2003). "Growth hormone receptor variant (L526I) modifies plasma HDL cholesterol phenotype in familial hypercholesterolemia: intra-familial association study in an eight-generation hyperlipidemic kindred." Am J Med Genet A **121**(2): 136-40.

Takahashi, N. and O. Smithies (1999). "Gene targeting approaches to analyzing hypertension." J Am Soc Nephrol **10**(7): 1598-605.

Tanase, H., Y. Yamori, C. T. Hansen and W. Lovenberg (1982). "Heart size in inbred strains of rats. Part 1. Genetic determination of the development of cardiovascular enlargement in rats." Hypertension **4**: 864-872.

Taylor, S. I. and D. E. Moller (1993). Mutations of the insulin receptor gene. Chicester, NY, Wiley.

Thompson, G. R. and P. W. Wilson (1992). Coronary Risk Factors and their Assesment. London, Science Press Ltd.

Timio, M., G. Lippi, S. Venanzi, S. Gentili, G. Quintaliani, C. Verdura, et al. (1997). "Blood pressure trend and cardiovascular events in nuns in a secluded order: a 30-year follow-up study." Blood Press **6**(2): 81-7.

Todd, J. A. (2001). "Tackling common disease." Nature **411**: 537-539.

Tontonoz, P., E. Hu and B. M. Spiegelman (1994). "Stimulation of adipogenesis in fibroblasts by PPAR gamma 2, a lipid-activated transcription factor." Cell **79**: 1147-1156.

Truswell, A. S., B. M. Kennelly, J. D. Hansen and R. B. Lee (1972). "Blood pressures of Kung bushmen in Northern Botswana." Am Heart J **84**(1): 5-12.

Vincent, M., E. H. Boussairi, R. Cartier, M. Lo, A. Sassolas, C. Cerutti, et al. (1993). "High blood pressure and metabolic disorders are associated in the Lyon hypertensive rat." J Hypertens **11**: 1179-1185.

Vincent, M., J. Saquet and J. Sassard (1984). The Lyon strains of hypertensive, normotensive and low-blood pressure rats. Handbook of Hypertension, Vol 4. Experimental and genetic models of hypertension. W. de Jong. New York, Elsevier Science Publishers: 314-327.

Wakim-Ghorayeb, S. F., S. H. Keleshian, G. Timson, R. R. Finan, P. Najm, N. Irani-Hakime, et al. (2005). "Factor V G1691A (Leiden) and prothrombin G20210A single-nucleotide polymorphisms in type 2 diabetes mellitus." Am J Hematol **80**(1): 84-6.

Walker, W. G., P. K. Whelton, H. Saito, R. P. Russell and J. Hermann (1979). "Relation between blood pressure and renin, renin substrate, angiotensin II, aldosterone and urinary sodium and potassium in 574 ambulatory subjects." Hypertension **1**(3): 287-91.

Wang, J., R. W. Williams and K. F. Manly (2003). "WebQTL: web-based complex trait analysis." Neuroinformatics **1**(4): 299-308.

Wang, X., N. Ishimori, R. Korstanje, J. Rollins and B. Paigen (2005). "Identifying novel genes for atherosclerosis through mouse-human comparative genetics." Am J Hum Genet 77(1): 1-15.

Wertheimer, E., S. P. Lu, P. F. Backeljauw, M. L. Davenport and S. I. Taylor (1993). "Homozygous deletion of the human insulin receptor gene results in leprechaunism." Nat Genet **5**(1): 71-3.

Withers, D. J., J. S. Gutierrez, H. Towery, D. J. Burks, J. M. Ren, S. Previs, et al. (1998). "Disruption of IRS-2 causes type 2 diabetes in mice." Nature **391**: 900-904.

Wu, R. and M. Lin (2006). "Functional mapping - how to map and study the genetic architecture of dynamic complex traits." Nat Rev Genet 7(3): 229-37.

Xie, X., J. Lu, E. J. Kulbokas, T. R. Golub, V. Mootha, K. Lindblad-Toh, et al. (2005). "Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals." Nature **434**(7031): 338-45.

Yaguchi, H., K. Togawa, M. Moritani and M. Itakura (2005). "Identification of candidate genes in the type 2 diabetes modifier locus using expression QTL." Genomics **85**(5): 591-9.

Yamori, Y. (1984). Development of the spontaneously hypertensive rat (SHR) and of various spontaneous rat models, and their implications. Handbook of Hypertension, Vol 4. Experimental and genetic models of hypertension. W. de Jong. New York, Elsevier Science Publishers: 224-239.

Yvert, G., R. B. Brem, J. Whittle, J. M. Akey, E. Foss, E. N. Smith, et al. (2003). "Trans-acting regulatory variation in Saccharomyces cerevisiae and the role of transcription factors." Nat Genet **35**(1): 57-64.

Zhang, Y., N. Wat, I. M. Stratton, M. G. Warren-Perry, M. Orho, L. Groop, et al. (1996). "UKPDS 19: heterogeneity in NIDDM: separate contributions of IRS-1 and beta 3-adrenergic-receptor mutations to insulin resistance and obesity respectively with no evidence for glycogen synthase gene mutations. UK Prospective Diabetes Study." Diabetologia **39**(12): 1505-11.

Zoccali, C., F. Mallamaci, G. Tripepi, F. A. Benedetto, S. Cutrupi, S. Parlongo, et al. (2002). "Adiponectin, metabolic risk factors, and cardiovascular events among patients with end-stage renal disease." J Am Soc Nephrol **13**(1): 134-41.

Appendix I Microarray sample preparation record

RI Strain Number	Animal Number	Investigator	RI Strain Number	Animal Number	Investigator
SHR/Ola	1	CW	RI 11C	1	CW
	2	CW		2	CW
	3	FM		3	FM
	4	FM		4	FM
	5	FM	RI 12C	1	CW
BN/Lx	1	CW		2	CW
	2	CW		3	FM
	3	FM		4	FM
	4	FM	RI 13C	1	CW
	5	FM		2	CW
RI 1	1	CW		3	FM
	2	CW		4	FM
	3	FM	RI 15	1	CW
	4	FM		2	CW
RI 2	1	CW		3	FM
	2	CW		4	FM
	3	FM	RI 17	1	CW
	4	FM		2	CW
RI 2C	1	CW		3	FM
	2	CW		4	FM
	3	FM	RI 18	1	CW
	4	FM		2	CW
RI 3	1	CW		3	FM
	2	CW		4	FM
	3	FM	RI 20	1	CW
	4	FM		2	CW
RI 3C	1	CW		3	FM
	2	CW		4	FM
	3	FM	RI 21	1	CW
	4	FM		2	CW
RI 4	1	CW		3	FM
	2	CW		4	FM
	3	FM	RI 22	1	CW
	4	FM		2	CW
RI 5	1	CW		3	FM
	2	CW		4	FM
	3	FM	RI 23	1	CW
	4	FM		2	CW
RI 5C	1	CW		3	FM
	2	CW		4	FM
	3	FM	RI 24	1	CW
	4	FM		2	CW
RI 6C	1	CW		3	FM
	2	CW		4	FM
	3	FM	RI 25	1	CW
	4	FM		2	CW
RI 7	1	CW		3	FM
	2	CW		4	FM

	3	FM	RI 26	1	CW
	4	FM		2	CW
RI 8C	1	CW		3	FM
	2	CW		4	FM
	3	FM	RI 27	1	CW
	4	FM		2	CW
RI 9C	1	CW		3	FM
	2	CW		4	FM
	3	FM	RI 29	1	CW
	4	FM		2	CW
RI 10	1	CW		3	FM
	2	CW		4	FM
	3	FM	RI 31	1	CW
	4	FM		2	CW
RI 10C	1	CW		3	FM
	2	CW		4	FM
	3	FM			
	4	FM			

Footnotes
CW indicates Dr Caroline Wallace prepared this sample
FM indicates this student Fiona Maciver prepared the sample

Appendix II Physiological SHR QTLs mapped in previous genome screens

pQTL ID[a]	pQTL Symbol[b]	RGD ID[c]	Chromosome	Peak Marker	Flanking Markers	Crossed With	Cross Type	LOD	*P*-Value	*cis* acting eQTLs overlapping with pQTL (peak of linkage>Affymetrix probe set ID)[d]	*trans* acting eQTLs overlapping with pQTL (peak of linkage>Affymetrix probe set ID)[d]
1	BpQTLcluster1	70188	1	-	D1Mgh5 - D1Mgh9	WKY, DRY	F2	4.82	-	D1Rat42>1389650_at, D1Rat212>1374249_at, D1Rat212>1375927_at, D1Rat356>1374675_at, D1Rat356>1371085_at, D1Rat27>1399157_at, D1Rat47>1371615_at, D1Rat47>1374907_at	D1Rat266>1383169_at, D1Cebr103s1>1373663_at, D1Cebr21s2>1377002_at, D1Cebr21s2>1390032_at, D1Cebr21s2>1383734_at, D1Arb11>1373301_at, D1Arb11>1374599_at, D1Rat356>1377244_at, D1Rat356>1373541_at, D1Rat356>1376710_at, D1Rat356>1368470_at, D1Rat356>1374907_at, D1Utr9>1388978_at, D1Utr9>1374613_at, D1Utr9>1371820_at, D1Utr9>1377407_at, D1Utr9>1383303_at, D1Rat31>1372117_at, D1Rat31>1377357_at, D1Rat27>1398447_at, D1Ral27>1388996_at, D1Rat27>1372205_at, D1Rat27>1377340_at, D1Cebr16s1>1389871_at, D1Cebr16s1>1373833_at, D1Rat47>1371588_at, D1Rat47>1383303_at, D1Rat47>1388567_at, D1Rat47>1371945_at, D1Rat47>1376036_at, D1Rat47>1374996_at
2	Bp29	61344	1	-	D1Wox29 - D1Mit2	WKY	reciprocal congenics	7.50	-		
3	Lvm2/Cm23	631199	1	D1Mit 3	D1Mit2 - D1Mgh9	DRY	F2	4.60	-	D1Rat42>1389650_at, D1Rat356>1374675_at, D1Rat356>1371085_at, D1Rat47>1371615_at, D1Rat47>1374907_at	D1Cebr103s1>1373663_at, D1Cebr21s2>1377002_at, D1Cebr21s2>1390032_at, D1Cebr21s2>1383734_at, D1Rat356>1377244_at, D1Rat356>1373541_at, D1Rat356>1376710_at, D1Rat356>1368470_at, D1Rat356>1374907_at, D1Utr9>1388978_at, D1Utr9>1374613_at, D1Utr9>1371820_at, D1Utr9>1377407_at, D1Utr9>1383303_at, D1Cebr16s1>1389871_at, D1Cebr16s1>1373833_at, D1Rat47>1371588_at, D1Rat47>1383303_at, D1Rat47>1388567_at, D1Rat47>1371945_at, D1Rat47>1376036_at, D1Rat47>1374996_at
4	Str1	61442	1	D1Mit3	- - -	SHRSP	F2	7.40	-	D1Rat47>1371615_at	D1Cebr103s1>1373663_at, D1Utr9>1388978_at, D1Utr9>1374613_at, D1Utr9>1371820_at, D1Utr9>1377407_at, D1Utr9>1383303_at, D1Cebr16s1>1389871_at, D1Cebr16s1>1373833_at
5	Kw (Hamet et al ,1998)	-	1	RT6	- - -	BN.Lx	RI strains	-	0.05000	D1Rat356>1374675_at, D1Rat356>1371085_at, D1Rat47>1371615_at, D1Rat47>1374907_at	D1Utr9>1388978_at, D1Utr9>1374613_at, D1Utr9>1371820_at, D1Utr9>1377407_at, D1Utr9>1383303_at, D1Rat47>1371588_at, D1Rat47>1383303_at, D1Rat47>1388567_at, D1Rat47>1371945_at, D1Rat47>1376036_at, D1Rat47>1374996_at
6	Bp44	61379	1	-	D1Mit3 - Sah	BB/OK	F2	19.20	-	D1Rat356>1374675_at, D1Rat356>1371085_at, D1Rat47>1371615_at, D1Rat47>1374907_at	D1Cebr21s2>1377002_at, D1Cebr21s2>1390032_at, D1Cebr21s2>1383734_at, D1Rat356>1377244_at, D1Rat356>1373541_at, D1Rat356>1376710_at, D1Rat356>1368470_at, D1Rat356>1374907_at, D1Utr9>1388978_at, D1Utr9>1374613_at, D1Utr9>1371820_at, D1Utr9>1377407_at, D1Utr9>1383303_at, D1Rat47>1371588_at, D1Rat47>1383303_at, D1Rat47>1388567_at, D1Rat47>1371945_at, D1Rat47>1376036_at, D1Rat47>1374996_at
7	Weight2/Bw18	70160	1	-	D1Mit3 - Scl	BB/OK	backcross	5.70	-	D1Cebr10s3>1374831_at, Sa>1388567_at, Sa>1383303_at, D1Rat69>1377534_at, D1Rat356>1374675_at, D1Rat356>1371085_at, D1Rat287>1375664_at, D1Rat287>1377614_at, D1Rat47>1371615_at, D1Rat47>1374907_at	D1Cebr21s2>1377002_at, D1Cebr21s2>1390032_at, D1Cebr21s2>1383734_at, Sa>1368977_a_at, D1Arb22>1371006_at, D1Arb22>1373212_at, D1Rat69>1372876_at, D1Rat69>1376710_at, D1Rat356>1377244_at, D1Rat356>1373541_at, D1Rat356>1376710_at, D1Rat356>1368470_at, D1Rat356>1374907_at, D1Utr9>1388978_at, D1Utr9>1374613_at, D1Utr9>1371820_at, D1Utr9>1377407_at, D1Utr9>1383303_at, D1Rat47>1371588_at, D1Rat47>1383303_at, D1Rat47>1388567_at, D1Rat47>1371945_at, D1Rat47>1376036_at, D1Rat47>1374996_at
8	Bp26	61341	1	-	D1Mit3 - Igf2	BN/Cr	congenic	-	-	D1Cebr10s3>1374831_at, Sa>1388567_at, Sa>1383303_at, D1Rat69>1377534_at, D1Rat356>1374675_at, D1Rat356>1371085_at, D1Rat287>1375664_at, D1Rat287>1377614_at, D1Rat47>1371615_at, D1Rat47>1374907_at	D1Cebr21s2>1377002_at, D1Cebr21s2>1390032_at, D1Cebr21s2>1383734_at, Sa>1368977_a_at, D1Arb22>1371006_at, D1Arb22>1373212_at, D1Rat69>1372876_at, D1Rat69>1376710_at, D1Rat356>1377244_at, D1Rat356>1373541_at, D1Rat356>1376710_at, D1Rat356>1368470_at, D1Rat356>1374907_at, D1Utr9>1388978_at, D1Utr9>1374613_at, D1Utr9>1371820_at, D1Utr9>1377407_at, D1Utr9>1383303_at, D1Rat47>1371588_at, D1Rat47>1383303_at, D1Rat47>1388567_at, D1Rat47>1371945_at, D1Rat47>1376036_at, D1Rat47>1374996_at
9	Bp (Huang et al., 1995)	-	1	Scnn1b	- - -	Dahl	RI strains	2 70	-	Sa>1388567_at, Sa>1383303_at, D1Rat287>1375664_at, D1Rat287>1377614_at	D1Cebr21s2>1377002_at, D1Cebr21s2>1390032_at, D1Cebr21s2>1383734_at, Sa>1368977_a_at, D1Rat356>1377244_at, D1Rat356>1373541_at, D1Rat356>1376710_at, D1Rat356>1368470_at, D1Rat356>1374907_at
10	Bp43	61378	1	-	Spn - Scl	BB/OK	F2	22 60	-	D1Cebr10s3>1374831_at, D1Rat69>1377534_at	D1Arb22>1371006_at, D1Arb22>1373212_at, D1Rat69>1372876_at, D1Rat69>1376710_at
11	Weight3/Bw20	70163	1	-	Sah - D1Mgh12	BB/OK	backcross	5 10	-	D1Cebr10s3>1374831_at, Sa>1388567_at, Sa>1383303_at, D1Rat69>1377534_at, D1Rat287>1375664_at, D1Rat287>1377614_at	Sa>1368977_a_at, D1Arb22>1371006_at, D1Arb22>1373212_at, D1Rat69>1372876_at, D1Rat69>1376710_at, D1Mit27>1383349_at
12	Cm7	631658	1	D1Rat73	- - -	SS/JrRkb	F2	5 32	-		
13	Bp42	61376	1	-	Igf2 - D1Mgh12	BB/OK	F2	23 40	-		D1Mit27>1383349_at
14	Map1/Bp196	631205	1	D1Mit 3	D1Mgh10 - D1Mgh1	DRY	F2	4 00	-	D1Rat225>1372093_at, D1Rat225>1370671_at	D1Rat235>1372650_at, D1Cebrp2[illegible]>1379911_at, D1Cebrp137s2>1377002_at, D1Cebrp1[illegible]>1390032_at, D1Cebrp137s2>1383734_at

15	Km1	724538	1	D1Ra186	D1UA5 - D1Ra186	DaN	backcross	3.40	-		D1Ra1235>13726S0_at, D1Ceb<p2956>137S931_at, D1Ceb<p137s2>137700Z_at, D1Ceb<p137s2>1390032_at, D1Ceb<p137s2>1383734_at
16	Iddm10	631670	1	D1Mgh13	D1Mgh12 - D1Mi1l4	BB/OK	backcross	2.80	-	D1Ra1225>13720Q3_at, D1Ra1225>1370671_at	D1Ra1235>13726S0_at, D1Ceb<p2956>137S931_at, D1Ceb<p137s2>137700Z_at, D1Ceb<p137s2>1390032_at, D1Ceb<p137s2>1383734_at
17	Sds/Sli3	631690	1	D1Mi1l4	---	BB/OK	backcross	2.10	-	D1Ra1225>13720Q3_at, D1Ra1225>1370671_at	
18	Arr3	738012	2	D2Ra194	---	BN.Lx	RI strains	3.80	-		D2Ra111>136897S_at
19	Bp18	61471	2	-	D2Mgh14 - Mt1pb	BN	F2	3.00	-		
20	Lvm1/Cm22	631198	2	D2Mgh15	D2Mii6 - D2Mgh7	DRY	F2	4.30	-	D2Ra124>1373082_at, D2Ra1136>1373389_at	D2Ceb<r1104>138958Q_at, D2Mii6>1371044_at, D2Ra1115>1374883_at, D2Ra124>1370281_at, D2Uh8>1370281_at, D2Uh8>1389580_at, D2Ra195>137S92Z_at, Cpb>136899Z_a_at, D2Ra1222>1372239_at
21	Bp105	631507	2	-	D2Ra1171 (D2Ra133) - D2Ath24	BN.Lx	-	-	0.00100	D2Ra1236>1386985_at, D2Ra1236>1369866_at, D2Ra1236>1380404_at, D2Ra1236>1372728_at, D2Ra1150>1373243_at, D2Ra154>1380404_at, D2Ra1136>1373389_at	Fgg>136889Q_at, Fgg>1374640_at, Fgg>1368178_at, D2Ra1115>1374883_at, D2Ceb<r458>138796Q_at, D2Ceb<r458>1392720_at, D2Ra1157>1371298_at, D2Mn14>1373243_at, D2Ra1222>1372239_at
22	Arr7	738007	2	D2Ra1221	---	BN.Lx	RI strains	4.40	-		D2Ra1115>1374883_at, D2Ra1222>1372239_at
23	BpQTLdus1e3	70175	2	-	D2Mgh24 - D2Mgh12	WKY, DRY	F2	2.76	-	D2Ra1236>1386985_at, D2Ra1236>1369866_at, D2Ra1236>1372728_at, D2Ra1150>1373243_at	Fgg>136889Q_at, Fgg>1374640_at, Fgg>1368178_at, D2Ra1157>1371298_at, D2Mn14>1373243_at, D2Ra1222>1372239_at
24	Bp101	631501	2	D2Mi110	D2Mi110 - D2Mgh12	wild	backcross	2.40	-	D2Ra1236>1386985_at, D2Ra1236>1369866_at, D2Ra1236>1372728_at, D2Ra1150>1373243_at	Fgg>136889Q_at, Fgg>1374640_at, Fgg>1368178_at, D2Ra1157>1371298_at, D2Mn14>1373243_at
25	Bp19	61473	2	-	D2Mgh8 (D2Go187) - Npr1	BN	F2	6.30	-	D2Ra1150>1373243_at	
26	Bp16	61469	2	D2Wox7	---	WKY	F2	5.64	-	D2Ra1236>1386985_at, D2Ra1236>1369866_at, D2Ra1236>1372728_at	D2Ra1157>1371298_at, D2Mi114>1373243_at
27	Bp63	70162	2	D2 Wox7	---	WKY	F2	5.64	-	D2Ra1236>1386985_at, D2Ra1236>1369866_at, D2Ra1236>1372728_at	D2Ra1157>1371298_at, D2Mi114>1373243_at
28	Bp17/DBP	61470	2	-	---	BN.Lx	RI strains	-	0.00080		
29	Bp112/Cm10	631679	3	D3Mgh9	---	SS/JrRkb	F2	7.34	-	D2Ceb<r2651>1368292_at, D2Ceb<r2651>1371537_at	
30	Bp92	631568	3	D3Ra153	D3Ra153 - D3Ra145	SS/Jr	F2	2.20	-	D2Ceb<r2651>1368292_at, D2Ceb<r2651>1371537_at	D.Ra1100>1376955_at
31	BpQTLdus1e#4	70191	3	Scn24t	D3Mgh16 - D3Mi18	WKY, DRY	F2	2.58	-	D2Ceb<r2651>1368292_at, D2Ceb<r2651>1373537_at	D3Ra1Z0>1373785_at, D.Ra1Z0>1390714_at, D3Ra1Z0>1389142_at, D.Ra1100>1376955_at
32	Kw (Hamel et al.,1998)	-	3	D3Mi19	---	BN.Lx	RI strains	-	0 01000		D.Ra1100>1376955_at
33	Bp110/Cm8	631676	3	D3Mgh7	---	SS/JrRkb	F2	7.03	-		D3Ra120>1378655_at, D.Ra1100>1376955_at, D3Ra120>1390717_at, D3Ra120>1389142_at, D.Ra1100>1376955_at
34	Arr10	61335	3	D3Ra1180	---	BN.Lx	RI strains	3.90	-		D3Ra1109>1377279_at, D3Ra1109>1368478_at, D3Ra1109>1390637_at, D3Ra1109>1387623_at, D3Ra1109>1387285_at, D3Ra1109>1376101_at, D3Ra1109>1389851_at, D3Ra1109>1389627_at, D3Ra1109>1390283_at, D3Ra1109>1387021_at, D3Ra1109>1384817_a_at, D3Ra1109>1377293_at, D3Ra1109>1369699_at, D3Ra1109>1375429_at, D3Ra1109>1375813_at, D3Ra1109>1375259_at, D3Ra1109>1368499_at, D3Ra1109>1374443_at, D3Ra1109>1390398_at, D3Ra1109>1368688_at, D3Ra1109>1374471_at, D3Ra1109>1382714_at, D3Ra1109>1375680_at, D3Ra1109>1382454_at, D3Ra1109>1376833_at, D3Ra1109>1373067_at
35	Cm17	724532	3	D3Ra1160	---	DaN	backcross	2.00	-	D3Ra1257>1389142_at, D3Ra1159>1370156_at, D3Ra1159>1373057_at, D3Ra1127>1389744_at	Edn1-1190492_at, Edn1-1178411_at, D1Ra1159>1390717_at, D1Mn11-1188401_at
36	Ir1 (Pravenec et al., 2002)	-	3	D3Mi14	D3Mi14 - -	BN.Lx	RI strains	3.80	-	D3Ra1159>1370156_at, D3Ra1159>1373057_at, D3Ra1127>1389744_at	D1Mn4-1170299_at, D3Ra1159>1390717_at

37	Bp20	32767	3	-	D3Mit2 - D3Mgh2	BN	F2	3.00	-	D3Rat143>1388178_at, D3Mit2>1377040_a_at	D3Rat6>1372422_at, D3Cebr80s1>1394772_at
38	Bp33	61351	4	Il6	---	BN	RI strains	-	0.01800	.	Il6>1386901_at
39	Scl23/Spl1	61408	4	Il6	---	BN.Lx	RI strains	-	0.00050	.	Il6>1386901_at
40	Irg1	61333	4	D4Arb13	---	WKY	F2	4.30	-	.	Il6>1386901_at
41	Irg1	61333	4	Ae2 (Slc4a2)	---	WKY	backcross	3.10	-	Cd36>1386901_at, Cd36>1367689_a_at	Il6>1386901_at
42	Irg1	61333	4	Rt8	---	BN.Lx	RI strains	-	0.00004	.	.
43	Sll2	631642	4	-	Il6 - D4Mit9	BB/OK	backcross	3.30	-	Cd36>1386901_at, Cd36>1367689_a_at	D4Utr3>1375724_at, D4Mgh2>1388968_at, D4Mgh2>1387010_s_at, D4Rat156>1387369_at, Il6>1386901_at
44	Bp79	619616	4	-	Il6 - D4Mit24	BB/OK	F2	-	0.02920	Cd36>1386901_at, Cd36>1367689_a_at	Lep>1376434_at, Lep>1376960_at, Lep>1390649_at, Lep>1388547_at, Lep>1384970_at, Lep>1370507_at, Lep>1368532_at, D4Utr3>1375724_at, D4Cebrp145s3>1377002_at, D4Cebrp145s3>1390032_at, D4Cebrp145s3>1383734_at, D4Mgh2>1388968_at, D4Mgh2>1387010_s_at, D4Rat156>1387369_at, D4Cebr28s5>1367466_at, Il6>1386901_at, Npy>1369937_at, Npy>1392604_at
45	Str3	61445	4	D4Mgh16	---	SHRSP	F2	3.00	-	.	Lep>1376434_at, Lep>1376960_at, Lep>1390649_at, Lep>1388547_at, Lep>1384970_at, Lep>1370507_at, Lep>1368532_at, D4Utr3>1375724_at, D4Cebr28s5>1367466_at
46	Bp21	61336	4	-	D4Mit2 - Npy	BN	F2	4.60	-	.	Lep>1376434_at, Lep>1376960_at, Lep>1390649_at, Lep>1388547_at, Lep>1384970_at, Lep>1370507_at, Lep>1368532_at, D4Cebrp145s3>1377002_at, D4Cebrp145s3>1390032_at, D4Cebrp145s3>1383734_at, D4Cebr28s5>1367466_at, Npy>1369937_at, Npy>1392604_at
47	BpQTLcluster5	70192	4	Slc4a2	D4Mgh16 - D4Mit23	WKY, DRY	F2	2.60	-	.	D4Cebrp145s3>1377002_at, D4Cebrp145s3>1390032_at, D4Cebrp145s3>1383734_at, D4Rat35>1369006_at, D4Rat35>1375272_at, D4Rat35>1389560_at, D4Rat35>1368610_at, D4Rat35>1368648_at, D4Rat35>1398840_at, D4Rat35>1389965_at, D4Rat35>1370806_at, D4Rat35>1389376_at, Hoxa11>1372945_at, Npy>1369937_at, Npy>1392604_at
48	Bw22	70167	4	D4Mit24	D4Mit6 - Spr	BB/OK	backcross	3.10	-	.	D4Rat35>1369006_at, D4Rat35>1375272_at, D4Rat35>1389560_at, D4Rat35>1368610_at, D4Rat35>1368648_at, D4Rat35>1398840_at, D4Rat35>1389965_at, D4Rat35>1370806_at, D4Rat35>1389376_at, Hoxa11>1372945_at
49	Bp135	631556	4	Npy	Npy - Spr	WKY	-	-	0.05000	.	D4Rat35>1369006_at, D4Rat35>1375272_at, D4Rat35>1389560_at, D4Rat35>1368610_at, D4Rat35>1368648_at, D4Rat35>1398840_at, D4Rat35>1389965_at, D4Rat35>1370806_at, D4Rat35>1389376_at, Hoxa11>1372945_at
50	Scl4	631689	4	Spr	---	BB/OK	backcross	1.90	-	Spr>1377155_at, D4Rat44>1368366_at	Spr>1370747_at, Spr>1390216_a_at
51	Cm18	724535	4	D4Rat195	---	Dahl	backcross	2.60	-	.	D4Rat58>1389936_at, Cacna1s>1389980_at, Cacna1s>1372410_at, Cacna1s>1367508_at, Cacna1s>1389293_at, Cacna1s>1376212_at, D4Mit19>1377287_at
52	Hematocrit (Pravenec et al., 1997)	-	4	Eno2	---	BN.Lx	RI strains	-	0.00010	D4Rat202>1390284_at	Cacna1s>1389980_at, Cacna1s>1372410_at, Cacna1s>1367508_at, Cacna1s>1389293_at, Cacna1s>1376212_at, D4Rat202>1371257_at, D4Rat202>1371172_at, D4Rat202>1371332_at, D4Rat202>1377112_at, D4Rat66>1373881_at, D4Rat66>1388493_at, D4Mit19>1377287_at, D4Rat153>1389475_at
53	Crg2	61446	4	-	Pparg - Eno2 (D4Arb40)	BN.Lx	RI strains	3.50	-	D4Rat202>1390284_at	Cacna1s>1389980_at, Cacna1s>1372410_at, Cacna1s>1367508_at, Cacna1s>1389293_at, Cacna1s>1376212_at, D4Rat153>1389475_at
54	Bp139	634349	5	-	D5Wox20 - D5Wox12	BN.Lx	congenic	-	0.00100	.	.
55	BpQTLcluster6	70194	5	-	D5Mit5 - D5Mgh13 (D5Mit13)	WKY, DRY	F2	2.60	-	.	.
56	Bp155	129809	5	-	D5Mgh22 (D15Got98) - D5Mgh15	WKY	-	3.80	-	D5Mit14>1374640_at, D5Rat169>1398268_at	D5Rat93>1372642_at, D5Rat63>1375851_at, D5Rat38>1388223_at, D5Rat38>1373318_at, D5Rat38>1376417_at, D5Rat38>1376172_at, D5Rat38>1367880_at, D5Rat174>1373382_at, D5Rat174>1367857_at, D5Rat174>1367952_at, D5Rat174>1389391_at, D5Rat174>1390374_at, D5Rat174>1370174_at, D5Rat174>1372198_at, D5Rat174>1373446_at, D5Rat174>1371412_a_at, D5Rat93>1371735_at
57	Bp103	631505	5	D5Mgh8	Slc2a1 - D5Mgh9	wild	backcross	3.20	-	D5Mit14>1374640_at, D5Rat169>1398268_at	D5Rat93>1372642_at, D5Rat63>1375851_at, D5Rat38>1388223_at, D5Rat38>1373318_at, D5Rat38>1376417_at, D5Rat38>1376172_at, D5Rat38>1367880_at, D5Rat174>1373382_at, D5Rat174>1367857_at, D5Rat174>1367952_at, D5Rat174>1389391_at, D5Rat174>1390374_at, D5Rat174>1370174_at, D5Rat174>1372198_at, D5Rat174>1373446_at, D5Rat174>1371412_a_at, D5Rat93>1371735_at

70	Bp_HDL phospholipid, serum cholesterol (Bottger et al., 1996)	-	8	-	D8Mi6 - Rbp2	BN.Lx	congenic	-	-	D8Ra|2|9>13769218_at, D8Ra|2|9>13907858_at, D8Ra|2|9>13745838_at, D8Ra|68>13821058_at, D8Ra|68>13873668_at, D8Ra|68>13775018_at, D8Ra|68>13883668_at, D8Mu|2>13696658_a_at, D8Mu|12>13908278_at, D8Ju|2>13738388_at, D8Ju|2>13802868_x_at, D8Mcw1>13712448_at	Mll>13883858_at, Gnk4>13710068_at, Gnk4>13732128_at, D8Ra|150>13682178_at, D8Ra|150>13915448_at, D8Ch6r8|81>13903108_x_at, D8Ch6r8|81>13933038_at, D8Ch6r8|81>13774588_at, D8Ch6r8|81>13730558_at, D8Ch6r8|81>13742208_at, Qrc|6br8|81>13714428_at, D8Na228>13733868_at, D8Ra|219>13751848_at, D8Ra|21>13779378_a_at, D8Ra|21>13746048_at, D8Ra|21>13742078_at, D8Ra|21>13704268_a_at, D8Ra|21>13779378_a_at, D8Mi3>13884048_at, D8Mi3>13873778_a_at, D8Mu|3>13902988_at, D8Mu|12>13796618_at, D8Mu|12>13680438_at, D8Mu|12>13746938_at, D8Mu|12>13727948_at, D8Mu|12>13705398_at, Rbp2>13889488_at, Rbp2>13699388_at, Rbp2>13735568_at, D8Mcw1>13760008_at, D8Mcw1>13722688_at, D8Mcw1>13712848_at, D8Ra|_42>13719118_at, D8Ra|_42>13742658_at, D8Ra|_42>13719078_at, D8Ra|_42>13775008_at, D8Ra|_42>13894558_at, D8Ra|_42>13872638_at, D8Ra|_42>13712848_at, D8Ra|_42>13736918_at, D8Ra|131>13711468_at, D8Ra|131>13727468_at, D8Ra|131>13731218_at, D8Ra|131>13702188_at, D8Ra|131>13727468_at, D8Ra|131>13731218_at, D8Ra|49>13692598_at, D8Ra|P9>13712138_at, D8Ch6r8|54>13772578_at, D8Ch6r8|54>13907108_x_at, D8Ch6r8|54>13730558_at, D8Ch6r8|54>13861488_at, D8Ch6r8|54>13861488_at
69	Basal Glucose uptake (Altman et al., 1997)	-	7	-	---	BN.Lx	RI strains	-	0.00010		
68	Arr8	73803O	7	D7Ra|110	---	BN.Lx	RI strains	4.10	-	D7Ch6r14C1652_at, D7Ch6r14C1652_at	D7Ch6r14C1653>13837334_at, D7Ch6r14C1652>13770028_at, D7Ch6r14C1652>13900328_at
67	Lvm7/Cm27	631504	7	D7Go|60	D7Ra|107 - D7Ra|9	Lew	F2	3.45	-	D7Ra|129>13703778_at, D7Ra|129>13892648_at, D7Ch6r7St>13884378_at, D7Ra|129>13703778_at, D7Ra|129>13892648_at, D7Ch6r7St>13884378_at, Bzrp>13895118_s_at	D7Ra|17>13844718_at, D7Ra|17>13676288_at, D7Ra|129>13700448_at, D7Ra|129>13915888_at, D7Ch6r7St>13847178_at, D7Ra|131>13766708_at, D7Ch6r14C1652>13770028_at, D7Ch6r14C1652>13900328_at, D7Mi|11>13695128_at, Mvc>13720318_at, D7Mi|11>13760448_at, D7Mi|11>13695128_at, D7Mi|11>13690958_at, D7Mi|11>13837348_at, D7Mi|5>13691678_at
66	Bp102	631503	7	Igf1	---	wild	backcross	5.90	-	.	D7Ra|32>13903478_at, D7Ra|35>13769768_at, D7Ra|35>13751838_at, D7Ju|2>13900648_at, D7Ju|>13775308_at, D7Ju|>13906298_at, D7Ju|>13775938_at
65	Iddm8	71111	6	D6Mgh2	D6Mgh2 - -	BB/OK	congenic	1.90	-	.	.
64	Arr5	738033	6	D6Ra|117	---	BN.Lx	RI strains	5.90	-	.	D6Ra|117>13710488_at
63	Bp141	634307	6	D6Mi|4	---	Lew	F2	4.00	-	.	D6Ra|132>13706668_at
62	Crg1	70199	6	D6Mi|9	D6Ra|37 - D6Mi|4	BN.Lx	RI strains	11.80	-	.	D6Ra|132>13706668_at
61	Bp100	631498	5	Nppb	Nppb - Nppa	WKY	F2	4.90	-	.	.
60	Bp131	631553	5	Nppa	---	SS	congenic	4.20	-	.	.
59	Slr2	61444	5	Nppa	---	SHRSP	F2	4.70	-	.	.
58	Arr4	738018	5	D5Ra|45	---	BN.Lx	RI strains	5.50	-		D5Ra|93>13726428_at, D5Ra|38>13882238_at, D5Ra|38>13733188_at, D5Ra|38>13764178_at, D5Ra|38>13767128_at, D5Ra|38>13678808_at, D5Ra|174>13723828_at, D5Ra|174>13738578_at, D5Ra|174>13678958_at, D5Ra|174>13903748_at, D5Ra|174>13701748_at, D5Ra|174>13772198_at, D5Ra|174>13734468_at, D5Ra|174>13714128_a_at, D5Ra|39>13717358_at

71	Bp35	61353	8	-	D8Mit6 - Rbp2	BN.Lx	congenic	-	-	D8Rat219>1376921_at, D8Rat219>1390185_at, D8Rat219>1374583_at, D8Rat21>1382105_at, D8Rat68>1387366_at, D8Rat68>1377501_at, D8Rat68>1388366_at, D8Mit12>1369665_a_at, D8Mit12>1390827_at, D8Utr3>1373838_at, D8Utr3>1380286_x_at, D8Mcw1>1371442_at	Mll>1388385_at, Grik4>1371006_at, Grik4>1373212_at, D8Rat150>1368217_at, D8Rat150>1391544_at, D8Cebr81s1>1390710_x_at, D8Cebr81s1>1377457_a_at, D8Cebr81s1>1393933_at, D8Cebr81s1>1377458_at, D8Cebr81s1>1373055_at, D8Cebr81s1>1374720_at, D8Cebr81s1>1371442_at, D8N228>1373386_at, D8Rat219>1375184_at, D8Rat21>1374604_at, D8Rat21>1374207_at, D8Rat21>1370426_a_at, D8Rat21>1377937_at, D8Mit3>1388404_at, D8Mit3>1387377_a_at, D8Mit3>1390298_at, D8Mit12>1379661_at, D8Mit12>1368043_at, D8Mit12>1374693_at, D8Mit12>1372794_at, D8Mit12>1370539_at, Rbp2>1388948_at, Rbp2>1369938_at, Rbp2>1373556_at, D8Mcw1>1376000_at, D8Mcw1>1372268_at, D8Mcw1>1371284_at, D8Rat_42>1371911_at, D8Rat_42>1374365_at, D8Rat_42>1371707_at, D8Rat_42>1377500_at, D8Rat_42>1389255_at, D8Rat_42>1387263_at, D8Rat_42>1371284_at, D8Rat_42>1373691_at, D8Rat131>1387789_at, D8Rat131>1370183_at, D8Rat131>1377462_at, D8Rat131>1373121_at, D8Rat49>1369259_at, D8Mgh9>1371713_at, D8Cebr81s4>1377457_a_at, D8Cebr81s4>1390710_x_at, D8Cebr81s4>1377458_at, D8Cebr81s4>1373055_at, D8Cebr81s4>1386148_at, D8Cebr81
72	Bp22	61337	8	-	D8Mit5 - D8Mit3	BN	F2	5.10	-	D8Rat219>1376921_at, D8Rat219>1390185_at, D8Rat219>1374583_at	D8N228>1373386_at, D8Rat219>1375184_at, D8Mit3>1388404_at, D8Mit3>1387377_a_at, D8Mit3>1390298_at, D8Mgh9>1371713_at, D8Cebr81s4>1377457_a_at, D8Cebr81s4>1390710_x_at, D8Cebr81s4>1377458_at, D8Cebr81s4>1373055_at, D8Cebr81s4>1386148_at, D8Cebr81s4>1371909_at, D8Cebr81s4>1371844_at, D8Cebr81s4>1371442_at
73	Bp62	70161	8	-	D8Mit3 - D8Rat133	SS/Jr	F2	2.90	-	D8Rat21>1382105_at, D8Mit12>1369665_a_at, D8Mit12>1390827_at, D8Mcw1>1371442_at	Mll>1388385_at, Grik4>1371006_at, Grik4>1373212_at, D8Rat150>1368217_at, D8Rat150>1391544_at, D8Rat21>1374604_at, D8Rat21>1374207_at, D8Rat21>1370426_a_at, D8Rat21>1377937_at, D8Mit12>1379661_at, D8Mit12>1368043_at, D8Mit12>1374693_at, D8Mit12>1372794_at, D8Mit12>1370539_at, D8Mcw1>1376000_at, D8Mcw1>1372268_at, D8Mcw1>1371284_at, D8Rat_42>1371911_at, D8Rat_42>1374365_at, D8Rat_42>1371707_at, D8Rat_42>1377500_at, D8Rat_42>1389255_at, D8Rat_42>1387263_at, D8Rat_42>1371284_at, D8Rat_42>1373691_at
74	Inf1	631842	8	-	- - -	BN.Lx	RI strains	4.10	-		
75	Arr9	738011	8	D8Rat7	- - -	BN.Lx	RI strains	6.10	-	Apeh>1369171_at	Apeh>1368067_at, D8Rat7>1388601_at, D8Cebr46s6>1370252_at, D8Cebr46s6>1387344_at, D8Cebr46s6>1374221_at
76	Bp53/Cm28	70218	9	D9Rat26	D9Rat70 - D9Uia6	SS/Jr	F2	7.30	-	D9Rat93>1387376_at, D9Rat156>1384309_at, D9Cebrp60s12>1370176_at, D9Rat15>1375655_at, D9Rat104>1371951_at, D9Rat104>1369313_at, D9Rat104>1372646_at	D9Rat93>1370193_at, D9Rat156>1375868_at, D9Utr3>1370221_at, D9Rat124>1371623_at
77	Bp113/Cm11	631680	9	D9Rat125	- - -	SS/JrRkb	F2	3.10	-	D9Rat104>1371951_at, D9Rat104>1369313_at, D9Rat104>1372646_at	D9Rat124>1371623_at
78	Bp108	631656	9	D9Rat10	- - -	SS/JrRkb	-	5.97	-	D9Rat156>1384309_at, D9Rat15>1375655_at	D9Rat156>1375868_at, D9Utr3>1370221_at
79	BpQTLcluster9	70198	10	-	D10Mit5 - Slc4a1	WKY, DRY	F2	2.94	-	D10Rat59>1373665_at, D10Rat215>1371803_at, D10Rat73>1398309_at, D10Mgh6>1371645_at, D10Mgh6>1394228_at, Abpa>1372064_at, Abpa>1372503_at, Abpa>1368285_at, Abpa>1370930_at	D10Rat59>1388907_at, D10Rat215>1382200_at, D10Rat166>1370046_at, D10Cebr44s3>1370901_at, D10Cebr44s3>1373490_at, D10Cebr44s3>1388784_at, D10Rat161>1374891_at, D10Rat26>1375994_at, D10Mit7>1390034_at, D10Mit7>1370930_at, D10Mit7>1384517_at, D10Cebrp1016s5>1371645_at, D10Rat31>1387550_a_at, D10Rat31>1372186_a_at, Abpa>1370015_at, D10Cebr23s8>1385808_at, D10Cebr23s8>1372355_at, D10Cebr23s8>1371676_at, D10Cebr4s9>1368749_at, D10Mgh8>1372602_at, D10Mgh8>1374645_at, D10Rat93>1370930_at, D10Mit4>1382200_at, D10Rat145>1374323_at, D10Rat28>1370930_at, D10Rat102>1389085_at, D10Rat102>1389110_at
80	Trs1	61441	10	Myh3	D10Mit4 - D10Mit2	BN.Lx	RI strains	4.00	-	D10Rat59>1373665_at, D10Rat215>1371803_at, D10Rat73>1398309_at, D10Mgh6>1371645_at, Abpa>1372064_at, Abpa>1372503_at, Abpa>1368285_at, Abpa>1370930_at	D10Rat59>1388907_at, D10Rat215>1382200_at, D10Rat166>1370046_at, D10Cebr44s3>1370901_at, D10Cebr44s3>1373490_at, D10Cebr44s3>1388784_at, D10Rat31>1387550_a_at, D10Rat31>1372186_a_at, Abpa>1370015_at, D10Cebr4s9>1368749_at, D10Mgh8>1372602_at, D10Mgh8>1374645_at, D10Mit4>1382200_at, D10Rat102>1389085_at, D10Rat102>1389110_at
81	Bw10	631515	10	-	Abp1 - D10Wox23	wild	backcross	3.50	-	D10Mgh6>1371645_at, D10Mgh6>1394228_at, Abpa>1372064_at, Abpa>1372503_at, Abpa>1370930_at	D10Cebr44s3>1370901_at, D10Cebr44s3>1373490_at, D10Cebr44s3>1388784_at, D10Rat161>1374891_at, D10Rat26>1375994_at, D10Cebrp1016s5>1371645_at, D10Cebr23s8>1385808_at, D10Cebr23s8>1372355_at, D10Cebr23s8>1371676_at, D10Rat93>1370930_at, D10Rat28>1370930_at
82	Bp45	61381	10	-	Abp1 - Ppy	BB/OK	F2	23.20	-	D10Mgh6>1371645_at, D10Mgh6>1394228_at, Abpa>1372064_at, Abpa>1372503_at, Abpa>1370930_at	D10Cebr44s3>1370901_at, D10Cebr44s3>1373490_at, D10Cebr44s3>1388784_at, D10Rat161>1374891_at, D10Rat26>1375994_at, D10Mit7>1390034_at, D10Mit7>1370930_at, D10Mit7>1384517_at, D10Cebrp1016s5>1371645_at, D10Cebr23s8>1385808_at, D10Cebr23s8>1372355_at, D10Cebr23s8>1371676_at, D10Rat93>1370930_at, D10Rat145>1374323_at, D10Rat28>1370930_at
83	Bp71	70363	10	-	D10Mgh6 - D10Rat106	SS	congenic	-	0.04000		D10Rat26>1375994_at, D10Cebrp1016s5>1371645_at, D10Cebr23s8>1385808_at, D10Cebr23s8>1372355_at, D10Cebr23s8>1371676_at, D10Rat93>1370930_at, D10Rat145>1374323_at, D10Rat28>1370930_at

84	Liver Cholesterol (Bottger et al., 1998)	-	10	-	---	BN.Lx	RI strains	3.30	-	-	-
85	BpQTLcluster10	70180	11	-	D11Mit1 - D11Mit5	WKY, DRY	F2	3.19	-	-	-
86	Bp104	631506	11	Sst	D11Mgh4 - D11Wox6	wild	backcross	2.80	-	-	Sst>1376011_at, D11Rat47>1376011_at, D11Rat94>1369242_at
87	Irg2	61334	12	-	D12Mgh1 - D12Mit8 (D12Mgh10)	WKY	F2	6.30	-	D12Mit5>1376550_at, D12Rat36>1372415_at, D12Rat36>1372532_at, D12Mit7>1375958_at	Eln>1377244_at, Eln>1377244_at, Epo>1369349_a_at, Epo>1368585_at, Epo>1377199_at, Epo>1373816_at, Hsp27>1375928_at, Hsp27>1368007_at, Hsp27>1371907_at, Lsn2>1372415_at, Lsn2>1372532_at, D12Mit5>1375636_at, D12Mit7>1398960_at
88	Irg2	-	12	-	D12Mgh1 - D12Mit8 (D12Mgh10)	WKY	backcross	1.90	-	D12Mit5>1376550_at, D12Rat36>1372415_at, D12Rat36>1372532_at, D12Mit7>1375958_at	Eln>1377244_at, Eln>1377244_at, Epo>1369349_a_at, Epo>1368585_at, Epo>1377199_at, Epo>1373816_at, Hsp27>1375928_at, Hsp27>1368007_at, Hsp27>1371907_at, Lsn2>1372415_at, Lsn2>1372532_at, D12Mit5>1375636_at, D12Mit7>1398960_at
89	Heart weight (Hamet et al., 1996)	-	12	Hsbp1	---	BN.Lx	RI strains	-	0.01120	D12Mit5>1376550_at, D12Mit7>1375958_at	Eln>1377244_at, Eln>1377244_at, Epo>1369349_a_at, Epo>1368585_at, Epo>1377199_at, Epo>1373816_at, Hsp27>1375928_at, Hsp27>1368007_at, Hsp27>1371907_at, Lsn2>1372415_at, Lsn2>1372532_at, D12Mit5>1375636_at, D12Mit7>1398960_at
90	Trs2	61443	12	-	D12Mit1 - Hspb1	BN.Lx	RI strains	3.30	-	D12Mit7>1375958_at	Epo>1369349_a_at, Epo>1368585_at, Epo>1377199_at, Epo>1373816_at, Lsn2>1372415_at, Lsn2>1372532_at, D12Mit7>1398960_at
91	Bp142	634310	12	D12Mit3	---	Lew	F2	4.20	-	D12Rat36>1372415_at, D12Rat36>1372532_at, D12Mit7>1375958_at	Eln>1377244_at, Eln>1377244_at, Epo>1369349_a_at, Epo>1368585_at, Epo>1377199_at, Epo>1373816_at, Hsp27>1375928_at, Hsp27>1368007_at, Hsp27>1371907_at, Lsn2>1372415_at, Lsn2>1372532_at, D12Mit7>1398960_at
92	Bp (Pravenec et al., 1991)	-	13	Renin	---	BN.Lx	RI strains	-	-	D13Cebr5s3>1374123_at	D13Rat88>1387671_at, D13Rat126>1372397_at, D13Rat126>1384402_a_at, D13Rat126>1372616_at, D13Cebrp149s11>1373171_at, D13Cebrp149s11>1394124_at, D13Cebr5s4>1371520_at
93	Bp55	70220	13	D13Mit2	---	WKY	F2	5.75	-	-	Lca>1373661_a_at, Lca>1389244_x_at, D13Rat126>1372397_at, D13Rat126>1384402_a_at, D13Rat126>1372616_at, D13N2>1367700_at, D13N2>1383577_at, D13N2>1372953_at
94	Bp31	61349	13	D13Mit2	---	WKY	F2	5.75	-	-	Lca>1373661_a_at, Lca>1389244_x_at, D13Rat126>1372397_at, D13Rat126>1384402_a_at, D13Rat126>1372616_at, D13N2>1367700_at, D13N2>1383577_at, D13N2>1372953_at
95	Bp80	619615	13	D13Uwm1	D3Mit1 - Atp1a2	BB/OK	-	0.04	-	D13Mit3>1368304_at, D13Rat131>1371732_at, D13Cebr5s3>1374123_at	D13Rat88>1387671_at, Lca>1373661_a_at, Lca>1389244_x_at, D13Rat126>1372397_at, D13Rat126>1384402_a_at, D13Rat126>1372616_at, D13N2>1367700_at, D13N2>1383577_at, D13N2>1372953_at, D13Cebrp149s11>1373171_at, D13Cebrp149s11>1394124_at, D13Cebr5s4>1371520_at
96	Scl11	631528	14	Gc	Afp - D14Wox4	wild	backcross	4.90	-	D14Utr6>1387314_at, D14Rat13>1377181_at, D14Rat52>1375194_at	D14Cebrp312s2>1390032_at, D14Cebrp312s2>1398262_at, D14Cebrp312s2>1375910_at, D14Cebrp312s2>1375932_at, D14Cebrp312s2>1383734_at, D14Mit3>1393759_at, D14Rat13>1374561_at, D14Rat13>1369784_at, H>1377463_at, H>1371820_at, H>1387279_at, H>1373833_at, D14Rat52>1393759_at, D14Rat52>1373737_at, D14Rat52>1373332_at, D14Rat52>1388820_at, D14Rat52>1376900_at, D14Rat52>1377506_at
97	Bg1	631516	15	Ednrb	D15Mgh3 - D15Mgh5	wild	backcross	7.00	-	-	D15Rat68>1390205_at, D15Rat68>1398935_at, D15Ucsf1>1369512_at, D15Ucsf1>1375557_at
98	Bp23	61338	16	-	D16Mit2 - D16Mit5	BN	F2	4.30	-	D16Mit2>1390435_at, D16Rat87>1390435_at, D16Rat87>1386352_at	D16Cebr10s10>1398450_at, D16Cebr10s10>1387550_a_at, D16Rat54>1392746_x_at, D16Rat54>1367769_at, D16Ucsf1>1373644_at
99	BpQTLcluster13	70183	16	Glud1	D16Mit2 - D16Mgh1	WKY, DRY	F2	3.33	-	D16Mit2>1390435_at, D16Rat87>1390435_at, D16Rat87>1386352_at	D16Mit5>1372162_at, D16Mit5>1373834_at, D16Mit5>1399080_at, D16Mit5>1377959_at, D16Mit5>1372136_at, D16Mit5>1388508_at, D16Mit5>1389793_at, D16Mit5>1374553_at, D16Rat77>1390029_at, D16Rat77>1388549_at, D16Rat77>1388909_at, D16Rat77>1387294_at, D16Rat77>1392746_x_at, D16Rat77>1367769_at, D16Rat77>1390435_at, D16Rat72>1389722_at, D16Cebr10s10>1398450_at, D16Cebr10s10>1387550_a_at, D16Rat75>1369594_at, D16Rat75>1387002_at, D16Rat75>1367769_at, D16Rat54>1392746_x_at, D16Rat54>1367769_at, D16Ucsf1>1373644_at, D16Mit1>1373773_at, D16Mit1>1375966_at, D16Mit1>1387061_at, D16Mit1>1368928_at, D16Mit1>1376729_at
100	Kw (Hamet et al.,1998)	-	17	D1a	---	BN.Lx	RI strains	-	0.01000	-	Drd1a>1375511_at
101	Heart weight (Pravenec et al., 1995)	-	17	Drd1a	---	BN.Lx	RI strains	-	0.00038	-	Drd1a>1375511_at
102	Sll1	631499	17	-	D17Wox3 - Aldoc	wild	backcross	3.60	-	D17Rat144>1368225_at, D17Rat144>1388908_at, D17Rat144>1388617_at, D17Mit2>1374959_at	D17Cebrp203s2>1371954_at, Prl>1370622_at, Prl>1388001_at, Prl>1381263_at, D17Rat144>1372011_at, D17Mit2>1390947_at
103	Lvm6/Cm26	631502	17	D17Rat52	D17Rat39 - D17Rat52	Lew	F2	3.71	-	D17Rat62>1376259_at, D17Rat62>1389867_at, D17Rat151>1393902_at	-
104	Bw11	631518	18	D18Rat36	Spink1 - D18Wox18 (D18Rat36)	wild	backcross	2.80	-	-	D18Cebrp97s6>1390365_at
105	Bp & Ss (Pravenec et al., 2001)	-	18	-	D18Rat32 - D18Rat12	BN.Lx	congenic	-	-	D18Rat55>1374560_at, D18Rat89>1390128_at, D18Rat13>1380293_at, D18Rat13>1388630_at, D18Rat13>1375343_at, D18Rat41>1374075_at	D18Rat97>1372191_at, D18Rat89>1367912_at, D18Cebr19s1>1386470_at, D18Cebrp97s6>1390365_at, D18Rat41>1368072_at
106	Scl22	631264	18	Tlr	Tlr - Nr3c1	BB/OK	F2	6.20	-	-	D18Cebr19s1>1386470_at, D18Cebrp97s6>1390365_at
107	Bp46	61382	18	-	Tlr - Nr3c1	BB/OK	F2	18.80	-	-	D18Cebr19s1>1386470_at, D18Cebrp97s6>1390365_at
108	Bp47	61383	18	-	Nr3c1 - Gja1 (D18Wox13)	BB/OK	F2	17.30	-	D18Rat55>1374560_at, D18Rat41>1374075_at	D18Rat97>1372191_at, D18Rat41>1368072_at

109	Spl3	631274	18	D18Mit17	Nr3c1 - D18Mit9	BB/OK	F2	5.30	-	D18Rat55>1374560_at, D18Rat89>1390128_at, D18Rat13>1380293_at, D18Rat13>1388630_at, D18Rat13>1375343_at, D18Rat41>1374075_at	D18Rat97>1372191_at, D18Rat89>1367912_at, D18Rat41>1368072_at
110	Arr11	738005	18	D18Rat55	- - -	BN.Lx	RI strains	3.40	-	D18Rat55>1374560_at, D18Rat89>1390128_at, D18Rat41>1374075_at	D18Rat89>1367912_at, D18Rat41>1368072_at
111	Bp48	61384	18	-	Ofl - D18Mit9	BB/OK	F2	19.40	-		
112	Cm9/Bp111	631678	19	D19Rat18	- - -	SS/JrRkb	F2	4.27	-		
113	Cm12/Bp114	631681	19	D19Rat18	- - -	SS/JrRkb	F2	3.33	-		
114	Scl12	61407	19	D19Mit2	Es2 - Ednra	BN.Lx	RI strains	-	0.00006	D19Rat14>1386699_at	D19Rat14>1387214_at, D19Rat14>1370352_at, D19Rat11>1387214_at, D19Rat11>1386908_at, D19Rat11>1368605_at, D19Rat11>1370352_at
115	Km3	724546	19	D19Uia1	D19Rat29 - D19Uia1	Dahl	backcross	3.10	-		D19Rat11>1387214_at, D19Rat11>1386908_at, D19Rat11>1368605_at, D19Rat11>1370352_at, D19Rat72>1388919_at
116	Bp32	61350	19	D19Mit7	Tat - Agt	BN.Lx	RI strains	-	0.01200	D19Utr7>1371870_at	D19Cebr204s28>1390032_at, D19Cebr204s28>1399156_at, D19Cebr204s28>1398262_at, D19Cebr204s28>1375932_at
117	Ir2	-	19	D19Mit7	D19Mit7 - D19Rat57	BN.Lx	RI strains	4.10	-	D19Utr7>1371870_at, D19Cebrp150s1>1368891_at	D19Cebr204s28>1390032_at, D19Cebr204s28>1399156_at, D19Cebr204s28>1398262_at, D19Cebr204s28>1375932_at
118	Bp144	724519	20	-	D20Utr1 - Tnf	BN.Lx	RI strains	-	0.00100		
119	Scl13	61409	20	-	- - -	BN.Lx	RI strains	-	0.00035		
120	Heart weight (Kunes et al., 1990)	-	20	-	- - -	BN.Lx	-	-	-		

Footnotes

[a]pQTL identifier as reported in Figures 3.7 and 3

[b]pQTL names from Rat Genome Database (www.rgd.mcw.edu). Where pQTL names have not been assigned the publication reference is giv

[c]pQTL identifier from Rat Genome Database (www rgd.mcw.edu

[d]*cis* acting eQTLs that overlap with each individual pQTL are listed as "marker at the peak of eQTL linkage > Affimetrix probe set identyfing the eQ

[e]*trans* acting eQTLs that overlap with each individual pQTL are listed as "marker at the peak of eQTL linkage > Affimetrix probe set identyfing the eQ

Key to QTLs

arr	anxiety related response
bg	blood glucose
bp	blood pressure
bw	body weight
cm	cardiac mass
crg	compensatory renal growth
iddm	insulin dependant diabetes mellitus
Inf	infertility severity
ir	insulin resistance
irg	insulin resistance gene
km	kidney mass
kw	kidney weight
lvm	left ventricular mass
map	mean arterial pressure
scl	serum cholestrol level
ss	salt sensitivity
stl	serum triglyceride
str	sensitivity to stroke
tc	total cholestrol
trs	thermal response to stress

References

Aitman T. et al. 1997. Nat Genet 16:197-201

Bottger A. et al. 1996. J Clin Invest 98:856-862

Bottger A. et al. 1998. Biochem Biophys Res Commun 246:272-275

Hamet P. et al. 1996. Hypertension 28:1112-1117

Hamet P. et al. 1998. Kidney Int 53:1488-1492

Huang H. et al. 1995. J Hypertens 13:1247-51

Kunes J. et al. 1990. J Hypertens 1990:1091-1095

Pravenec M. et al. 1991. Genomics 9:466-472

Pravenec M. et al. 1995. J Clin Invest 96:1973-1978

Pravenec M. et al. 1997. Mamm Genome 8:387-389

Pravenec M. et al. 2001. Am J Hypertens 14:82A

Pravenec M. et al. 2002. Mamm Genome 13:253-258

Appendix III Trans gene primers

Name on tube	Probeset	Primer name	Primer sequence
FM1	1369784_at	1369784F1	ACTCTGATTTGGCCTCCAGA
FM2		1369784R1	AGAAGATGTGCATCGCTCAA
FM3		1369784F2	TTCGAGGAGAAGGACTCTGG
FM4		1369784R2	ATGCCAAGGCAGATTGACA
FM5		1369784F3	TCACCTCACCTCCCTTCATC
FM6		1369784R3	CTTGCAGACAAGAACCTACCA
FM7		1369784F4	AATGTTTGGCTGTGGGTCTC
FM8		1369784R4	TGCCCACTCAAGCTTAGACA
FM9		1369784F5	TGAGTGGCACCTTTGTTCTG
FM10		1369784R5	TTCAGGTGAAACAGAAGATCCA
FM11	1372923_at	1372923F1	CCGAAAGGCAGATCAGAGTC
FM12		1372923R1	TCTGGTGCCTAGCCCAAT
FM13		1372923F2	CTGGGTCTGTGCCTCATTTT
FM14		1372923R2	TCTTTGCATCCTTCCCATTC
FM15		1372923F3	CTCATAGGTTCCTGGGTTGT
FM16		1372923R3	CAGGAGGTACCTGGGAGTT
FM17		1372923F4	ATAGCAGCCTCCCTCCCTTA
FM18		1372923R4	CTCCTCCCGAATAGAGGAT
FM19		1372923F5	ACCAATGCTGACCTGAAAGG
FM20		1372923R5	GCTTGGCTCTGAGCACTAAAG
FM21	1373773_at	1373773F1	catgttgacaataagttggtg
FM22		1373773R1	GAACTCACAGGGAACCCAAA
FM23		1373773F2	gagatgggctgaaacagagg
FM24		1373773R2	ttagcaatcgtgctcaacct
FM25		1373773F3	GACACAGCTTTGACAGACAG
FM26		1373773R3	ACAGTGGGAAAGCCCAGATT
FM27		1373773F4	TGAATCTGACTGGGTGAGCA
FM28		1373773R4	CACAGTACCTTTTTGTGTCTGTCC
FM29	1385823_s_at	1385823F1	ACAAACAGTACCCCACTTCCT
FM30		1385823R1	TTCCTGGGCTGCTAATATGG
FM31		1385823F2	GATCCGAGAATCTCTCCACT
FM32		1385823R2	TGCACAGTGCCCATAATACC
FM33		1385823F3	CGAGGCCCTTAAATCTCCAT
FM34		1385823R3	acggggatcctccataatg
FM35		1385823F4	GAATCTATGCCTGCGTAAAGG
FM36		1385823R4	AGAAGCGAGGAGGGAAGGT
FM37		1385823f4a	tccagggcggagtttcct
FM38		1385823r4a	ctgggctgcgaacaataag
FM39	1390346_at	1390346F1	gtgcctgtagaagcctaa
FM40		1390346R1	gaaagaccagcacagcc
FM41		1390346F2	cagttctcagtggactgaat
FM42		1390346R2	cagtctaggctactacagtg
FM43		1390346F3	acacttgaagtggccccagtc
FM44		1390346R3	atttcgggcacagatgaagat
FM45		1390346F4	gtctgtttgattggacagtcgt
FM46		1390346R4	tcgttgttttgcattttcca
FM47		1390346f4a	aaaccccatgctcaatgaag
FM169	1369273_a_at	NPR3F1	cgagaaagtgaacagcagca
FM170		NPR3R1	cacccgttgcagagacaag
FM171		NPR3F2	ggatccatagccggtgttta
FM172		NPR3R2	accaggtctgcaaatgaacc
FM173		NPR3F3	tgttgcttcggatgttgggt
FM174		NPR3R3	ggcagctggtgaagaatctc
FM175		NPR3F4	atttggctgggcgtcagt
FM176		NPR3R4	cccagactcttccagtgagc
FM177		NPR3F5	catgaaggactggggtttcct
FM178		NPR3R5	tgacatccttcagcacagcc
FM179		NPR3F6	ggtcggtaagcaaagggatt
FM180		NPR3R6	ctcttctggaaagcctgaaa
MH33	1369171_at	MH33gMSTF1	gttcctcagttcacaggaaga
MH34		MH34gMSTF1	gaactccctcctgaaggcaa
MH35		MH35gMSTF1	gactaagtctgggtggccaa
MH36		MH36gMSTF1	ctactgacacttctatgccaaa
MH37		MH37gMSTF1	gagatgtcacactaagcccaa
MH38		MH38gMSTF1	ctggcctgacaggtcctat
mh181	1370806_at	rmt7gF1	gcctatcagctaacctggca
mh182		rmt7gr1	ccagggccttgtgcttgcta
mh183		rmt7gf2	gagtaggggatccaaggttt
mh184		rmt7gf3	ccctcttaccggaacgct
mh185		rmt7gr3	ggcttgtgtggccccatggt
mh186		rmt7gf4	catgccctggatgagcagaa