Chapter 3

Investigation of lipin family genes for impact on syndromes of severe insulin resistance and metabolic traits in UK populations

3.1 Summary

Loss of Lpin1 activity causes lipodystrophy and insulin resistance in the fld (fatty liver dystrophy) mouse, and *LPIN1* expression and common genetic variation were recently suggested to influence adiposity and insulin sensitivity in humans. I conducted a comprehensive association study to clarify the influence of *LPIN1* common variation on adiposity, insulin sensitivity, and other metabolic traits previously associated with *LPIN1*, in UK populations. Twenty-two SNPs tagging *LPIN1* common variation were genotyped in the MRC Ely (N = 1709) and Hertfordshire (N = 2901) population-based cohorts. Where possible, data were meta-analysed with other in-house and publicly available datasets to increase power to detect modest effect sizes. No association was found between *LPIN1* SNPs and fasting insulin, but I report a nominal association between rs13412852 and BMI (P = 0.042) in a meta-analysis of 8504 samples. I also detected nominal associations between *LPIN1* SNPs and traits underlying metabolic syndrome, but these require replication in additional large cohorts.

To investigate the putative role of lipin family mutations in insulin resistance syndromes I sequenced *LPIN1*, *LPIN2*, and *LPIN3* exons, exon/intron boundaries and 3'UTR in 158 patients with idiopathic severe insulin resistance (including 23 lipodystrophic patients), and controls. Three rare nonsynonymous *LPIN1* variants (A353T, R552K and G582R) were detected but these did not co-segregate with disease in affected families and Lipin1 protein expression and phosphorylation in patient fibroblasts was indistinguishable from controls. Two rare nonsynonymous changes in *LPIN2* were predicted benign and not prioritised for further analysis. Two rare nonsynonymous *LPIN3* variants (G41S and W110C) were detected within the conserved N-terminal lipin domain in single individuals. W110C was also detected in a Druze control and was therefore considered unlikely to be pathogenic. G41S was

absent from controls but no family DNA was available for co-segregation analysis. Functional work will be required to evaluate pathogenicity.

In summary, large scale association and re-sequencing studies do not support a major effect of *LPIN1* common variation on metabolic traits and suggest that mutations in lipin family genes are not a common cause of lipodystrophy and insulin resistance in humans. *LPIN1* data was published (Fawcett et al. 2008).

3.2 Introduction

3.2.1 Lipin 1 null mutations cause lipodystrophy and insulin resistance in fatty liver dystrophy (fld) mice

Lipin 1 was identified by positional cloning as the gene responsible for two independent mutant mouse models, the fld and fld^{2J} models, both characterised by a triglyceride-filled fatty liver, lipodystrophy, insulin resistance, and progressive peripheral neuropathy (Langner et al. 1989; Peterfy et al. 2001; Reue et al. 2000). The fld mouse strain carries two copies of Lpin1 with gross structural abnormalities, while the phenotype of the fld^{2J} mouse strain results from a point mutation leading to substitution of arginine for glycine at residue 84 of the lipin 1 protein (Peterfy et al. 2001). The phenotype of these mouse models shares features with human lipodystrophies, and therefore *LPIN1* is a good candidate gene for human lipodystrophies. However, to my knowledge there has only been one study screening *LPIN1* for pathogenic mutations in human lipodystrophic patients (N=15), with no pathogenic mutation being reported (Cao and Hegele 2002).

3.2.2 Lipin 1 is required for the development of mature adipocytes

Using primary mouse embryonic fibroblasts (MEFs) from fld mice it was shown that Lpin1 is required for induction of adipogenic genes, peroxisome proliferator-activated receptor-γ (*PPARγ*) and CCAAT enhancer-binding protein-α (*C/EBPα*), and for adipocyte differentiation (Phan et al. 2004). The primary defect in fld mice is therefore likely to be reduced adipose tissue mass with ectopic deposition of lipids and/or aberrant adipokine signalling causing secondary characteristics such as insulin resistance. Lpin1 expression is induced at two time points during differentiation of the 3T3-L1 preadipocyte cell line. There is a transient spike at 10 hours into the differentiation process, which precedes induction of *PPARγ* expression at 20 hours. Lipin 1 levels then return to baseline and are induced at 2 days,

reaching a peak in mature lipid-loaded adipocytes (Peterfy et al. 2005; Phan et al. 2004).

3.2.3 Lipin 1 isoforms reveal distinct roles during adipocyte development

Two lipin 1 protein isoforms are generated by alternative splicing of Lpin1 mRNA (Peterfy et al. 2005). Lipin-1A levels diminish during the differentiation process whereas Lipin-1B, which includes 33 extra amino acids, is the predominant isoform during the transient spike at 10 hours and in mature adipocytes. Reintroduction of Lipin-1A to lipin-1-deficient MEFs induces adipogenic genes, whereas Lipin-1B expression leads to stronger induction of lipid synthesis and storage genes compared to Lipin-1A (Peterfy et al. 2005). The biphasic expression of lipin 1 and the different functions of its two isoforms suggest that lipin 1 has two distinct roles in adipocyte development: the induction of the adipogenic gene expression program, and lipid accumulation in mature adipocytes.

3.2.4 Lpin1 overexpression causes obesity in transgenic mice

Transgenic mice with adipose tissue-specific overexpression of Lpin1B exhibit diet-induced obesity and enhanced insulin sensitivity compared to wild-type littermates (Phan and Reue 2005). In these mice adipose tissue expression of lipid synthesis and storage genes diacylglycerol acyltransferase (*DGAT*), acetyl-CoA carboxylase-1 (*ACC-1*), and phospoenolpyruvate (*PEPCK*) is elevated, supporting a role for Lpin1 in lipid accumulation of mature adipocytes. Interestingly, Lpin1 appears to have distinct roles in different tissues as skeletal muscle-specific overexpression of Lpin1 results in more pronounced obesity, insulin resistance, and changes in whole-body energy expenditure and fuel utilisation (Phan and Reue 2005; Xu et al. 2006). This occurs with decreased expression of fatty acid oxidation genes such as carnitine palmitoyl transferase 1 (*CPT-1*) and acyl-CoA oxidase (*AOX*) in skeletal muscle.

3.2.5 Lipin 1 is a phospatidate phosphatase

The mechanism through which lipin 1 influences adipocyte development and fat accumulation is not entirely known. However, recent data shows that lipin 1 is a magnesium-dependent phospatidate phosphatase (PAP) responsible for catalysing the conversion of phosphatidate (PA) to DAG, the penultimate step in triacylglyceride synthesis (Figure 3.1). This could explain why Lpin1 deficient fld mice cannot accumulate fat in their limited number of mature adipocytes (Han et al. 2006; O'Hara et al. 2006). This is also the penultimate step in the synthesis of phospholipids, important components of cell membranes, which might explain how the budding yeast homolog of lipin 1 (Smp2) regulates nuclear membrane growth during the cell cycle by controlling phospholipid biosynthesis (Santos-Rosa et al. 2005) and mutation of the fission yeast homolog of lipin 1 (Ned1) causes aberrant nuclear shape (Tange et al. 2002).

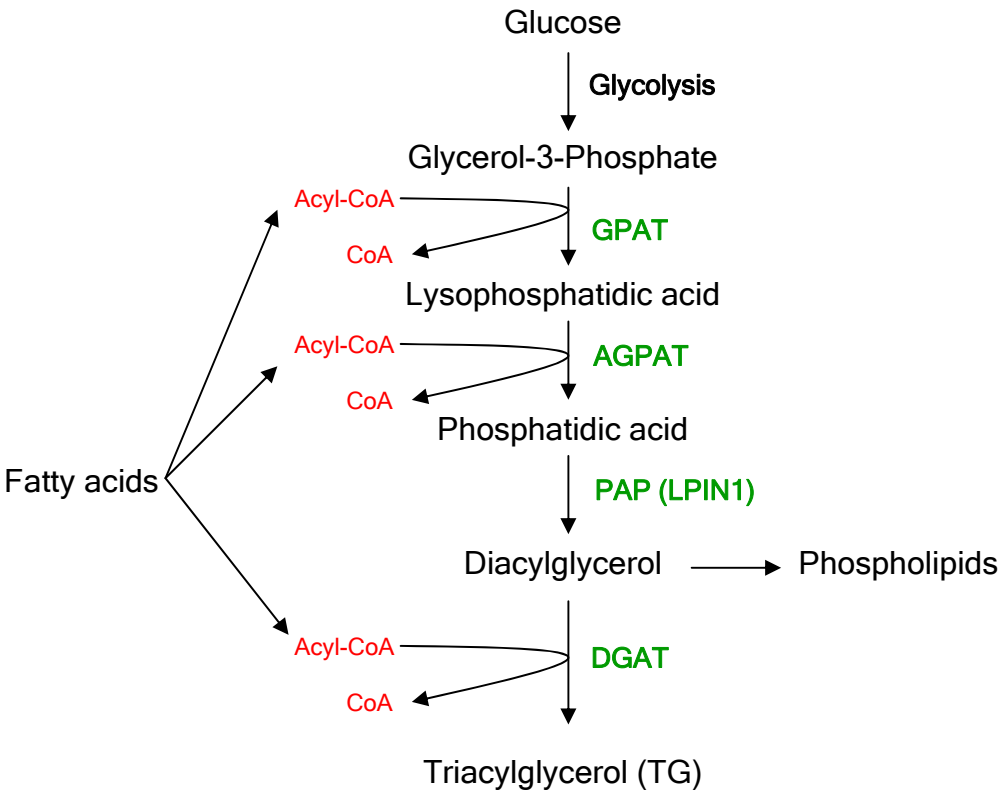

Figure 3.1 The role of *LPIN1* in triacylglycerol (or triglyceride) and phospholipid synthesis
Enzymes are shown in green and acylation in red. GPAT = glycerol-3-phosphatate acyltransferase, AGPAT = 1-acylglycerol-3-phosphate acyltransferase, PAP = phosphatidic acid phosphatase, and DGAT = diacylglycerol acyltransferase.

3.2.6 Lipin 1 is an inducible transcriptional coactivator

In addition to its role in TAG synthesis, lipin 1 appears to function as a transcriptional coactivator of *PGC1α/PPARα* target genes during fasting (Finck et al. 2006). Lipin 1 interacts directly with PGC1α and PPARα through an α-helical leucine-rich motif (LXXIL) to form a complex which then modulates gene transcription. Overexpression of lipin 1 in murine liver increases expression of *PGC1α/PPARα* target genes such as fatty acid oxidation genes and suppresses expression of genes involved in *de novo* fatty acid and TAG synthesis. Lipin 1 may therefore increase hepatic capacity for β-oxidation and help maintain hepatic lipid balance during increased lipid delivery under fasting conditions and diabetes (Finck et al. 2006).

3.2.7 Regulation of lipin 1

Regulation of lipin 1 activity occurs at the levels of mRNA transcription, mRNA splicing, protein phosphorylation, and subcellular localisation. Recent studies in mouse and rat cells have shown that glucocorticoids can induce Lpin1 transcription, increase Lpin1 expression and increase PAP acitivity (Manmontri et al. 2008; Zhang et al. 2008). Also, multiple sites on the lipin 1 protein are phosphorylated in response to insulin and amino acids and dephosphorylated in response to epinephrine and fatty acids (Harris et al. 2007). Insulin- or amino acid-stimulated phosphorylation of lipin 1 occurs in an mTOR (mammalian target of rapamycin) dependent manner (Huffman et al. 2002) in rat adipocytes. A study of murine lipin 1 showed that the phosphorylation status of lipin 1 does not appear to have any effect on its intrinsic PAP activity but does correlate with its subcellular localisation (Harris et al. 2007). Insulin increases the amount of lipin 1 and PAP activity in the soluable fraction of the cell, and decreases the amount of lipin 1 at intracellular membranes where PA is synthesised (Harris et al. 2007).

3.2.8 *LPIN1* expression and human adiposity and insulin resistance

In humans, *LPIN1* expression in adipose tissue appears to be inversely correlated with measures of adiposity such as BMI, and positively correlated with insulin sensitivity (Croce et al. 2007; Donkor et al. 2007b; Lindegaard et al. 2007; Suviolahti et al. 2006; van Harmelen et al. 2007; Yao-Borengasser et al. 2006). For example, *LPIN1* transcript levels from 19 Finnish human fat biopsies were inversely correlated with fasting plasma insulin, glucose, triglycerides and HOMA-IR (Suviolahti et al. 2006). In a US study, lipin 1 mRNA levels in adipose tissue from 36 women and 3 men were inversely correlated with obesity, BMI, percentage body fat and waist circumference, and positively correlated with insulin sensitivity and were higher in normal glucose tolerant subjects than in impaired glucose tolerant (IGT) subjects matched for BMI (Yao-Borengasser et al. 2006). In the same study lipin expression

was inversely correlated with intramyocellular lipids (IMCLs) independent of the fatty acid oxidative capacity of muscle. Therefore, lipin deficiency may lead to partitioning of lipids into muscle and insulin resistance in humans. Subjects treated with the drug pioglitazone, which increases adipose tissue mass and decreases IMCLs, also demonstrated increased expression of lipin and improved insulin sensitivity (Yao-Borengasser et al. 2006). These data suggest *LPIN1* genetic variation that influences expression of lipin 1 and/or lipin 1 function might impact upon human adiposity and insulin sensitivity.

3.2.9 *LPIN1* genetic variation and human adiposity and insulin resistance

There have been a number of studies evaluating the role of common variation in *LPIN1* on human metabolic phenotypes (Loos et al. 2007; Suviolahti et al. 2006; Wiedmann et al. 2007), but the results have been inconsistent across studies, and sometimes within the same study. For example, rs2716610 and a SNP in high linkage disequilibrium, rs2716609, were associated with BMI in a Finnish obesity case-control and in the Quebec Family Study (Loos et al. 2007; Suviolahti et al. 2006) but not in a German population-based cohort (the MONICA study) (Wiedmann et al. 2007). Moreover, *LPIN1* haplotypes were strongly associated with traits underlying metabolic syndrome in the MONICA study but these haplotypes often had the opposite effect on the same traits in a replication cohort (Wiedmann et al. 2007). This inconsistency suggests that further studies are needed to clarify the role of *LPIN1* variation on human metabolic traits.

3.2.10 *LPIN* family of genes

Multiple sequence alignments of lipin-related proteins in a broad range of eukaryotic organisms reveal some strongly conserved domains (Figure 3.2): the amino-terminal lipin domain (NLIP), a nuclear localisation signal (NLS), and a carboxy-terminal lipin domain (CLIP) which carries the DXDXT motif characteristic of a superfamily of magnesium-dependent phosphatases (Donkor et al. 2007a; Han et al. 2006; Peterfy

et al. 2001) and the LXXIL motif that mediates interaction with nuclear receptors (Finck et al. 2006) (see Figure 3.3). As expected, deletion of the CLIP domain abolishes PAP activity. However, insulin-induced phosphorylation of lipin 1 is more pronounced in the NLIP domain and this domain is also required for full PAP activity (Harris et al. 2007). These domains are shared by two other mammalian lipin 1-related proteins, lipin 2 and lipin 3. All three lipin family members in mouse and human possess PAP activity and exhibit unique but overlapping tissue distributions (Donkor et al. 2007a). Lipin 1 is the predominant form in mammalian adipose tissue and skeletal muscle. This is supported by the absence of PAP activity in adipose tissue and skeletal muscle of fld mice (Donkor et al. 2007a). Whereas lipin 2 predominates in brain and liver where fld mice have comparable PAP activity to wild-type (Donkor et al. 2007a). However, lipin 1 appears to account for most or all of the fasting- and glucocorticoid-induced PAP activity in liver (Finck et al. 2006; Harris et al. 2007; Manmontri et al. 2008). In humans *LPIN2* expression is also high in adipose tissue though this needs to be confirmed in other studies (Donkor et al. 2007a). In mice and humans lipin 3 is the main lipin family member in the gastro-intestinal tract (Donkor et al. 2007a).

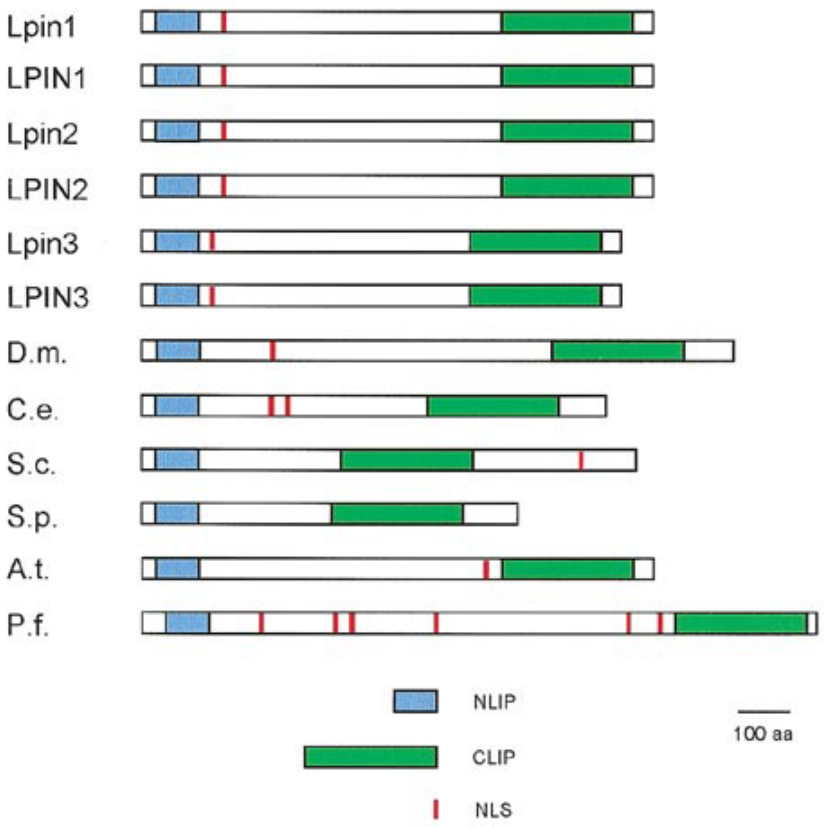

Figure 3.2 Evolutionary conservation of the lipin-protein family, (from (Peterfy et al. 2001)). Lipin homologs in mouse, human, *Drosophila* (D.m.), *C. elegans* (C.e.), *S. cerevisiae* (S.c.), *S. pombe* (S.p.), *A. thaliana* (A.t.) and *P. falciparum* (P.f.). Lpin1, Lpin2 and Lpin3 protein sequences were deduced from full-length cDNAs obtained by RACE cloning in this study. The *LPIN1*, *LPIN2*, *LPIN3*, *Drosophila* , *C. elegans*, *S. cerevisiae*, *S. pombe*, *A. thaliana* and *P. falciparum* protein sequences are based on predictions from EST and genomic sequences. NLIP (blue) and CLIP (green) domains, and predicted nuclear localization signals (NLS - red) are indicated.

Figure 3.3 Known functional motifs and disease mutations in lipin proteins, (adapted from (Reue and Zhang 2008)). G84R causes the phenotype of fld^{2J} mice and S734L is a *LPIN2* Majeed syndrome mutation. NLIP = amino-terminal lipin domain, NLS = nuclear localisation signal, and CLIP = carboxy-terminal lipin domain. DXDXT and LXXIL are peptide sequence motifs.

3.2.11 Genetic studies of *LPIN2* in humans

Mutations in *LPIN2* have been shown to cause Majeed Syndrome, a rare, autosomal recessive autoinflammatory disorder characterised by chronic recurrent multifocal osteomyelitis (CRMO), congenital dyserythropoietic anaemia (CDA), and inflammatory dermatosis (MIM no. #609628) (Al-Mosawi et al. 2007; Ferguson et al. 2006; Ferguson et al. 2005). There have been no reports of abnormalities in fat or sugar metabolism in these patients. However, given that different mutations in the same gene have been known to cause different phenotypes, for example, *LMNA* mutations lead to at least 11 different clinical syndromes (Worman and Bonne 2007), *LPIN2* may still be a good candidate for human insulin resistance and lipodystrophy. Indeed, a recent study reported a SNP within the 3'UTR of *LPIN2* associated with type 2 diabetes and measures of adiposity and insulin sensitivity in a Dutch population-based cohort (Aulchenko et al. 2007). This SNP requires replication but given time constraints I was unable to pursue this during the course of my work.

To my knowledge no genetic studies of *LPIN3* have yet been carried out so, given its likely role as a PAP in the gastro-intestinal tract, it may be worth screening *LPIN3* for pathogenic mutations causing dyslipidemia and lipodystrophy.

3.2.12 Aims of this study

1. To investigate, in UK populations (N=4610), the role of common genetic variation in *LPIN1* on insulin sensitivity, BMI, and other metabolic traits previously associated with *LPIN1* variation (Section 3.3.1);
2. To identify potentially pathogenic mutations in *LPIN1*, *2* and *3* in idiopathic insulin resistant subjects (N=158), including 23 lipodystrophic patients (Sections 3.3.2 and 3.3.3).

3.3 Results

3.3.1 Association studies of *LPIN1* tagging SNPs and metabolic traits

3.3.1.1 Identification and selection of common variants for association testing

To identify *LPIN1* common genetic variants for association studies I took two approaches (Figure 3.4). First, I used data available from the International HapMap consortium (Rel 20, PhaseII) (http://www.hapmap.org) to identify SNPs that were genotyped in CEPH Utah residents of Northern and Western European ancestry (CEU) within *LPIN1* and its flanking ~4 kb regions (Chromosome 2, coordinates 11800212 -11889941 (NCBI B36 coordinates). There were 46 SNPs with MAF>0.01 genotyped in HapMap within this region. Second, I sequenced *LPIN1* exons, exon-intron boundaries, and 3'UTR in 31 unrelated CEPH samples that overlapped with HapMap CEU samples. This approach identified 13 SNPs (including 4 novel variants - Table 3.1). Five of these SNPs overlapped with those in HapMap and two had call rates from my sequencing data <80% leaving 52 SNPs for association testing in UK population-based cohorts.

To avoid genotyping redundant SNPs I evaluated the extent of linkage disequilibrium (LD) between these SNPs (Figure 3.5) and used pairwise LD measures to select tagging SNPs (Section 3.4.2) using an r^2 cutoff of 0.8. Twenty -five SNPs selected using this processs tagged all SNPs in the full HapMap CEU panel except one, rs17603350, which was added to the tagging SNP set (Figure 3.4).

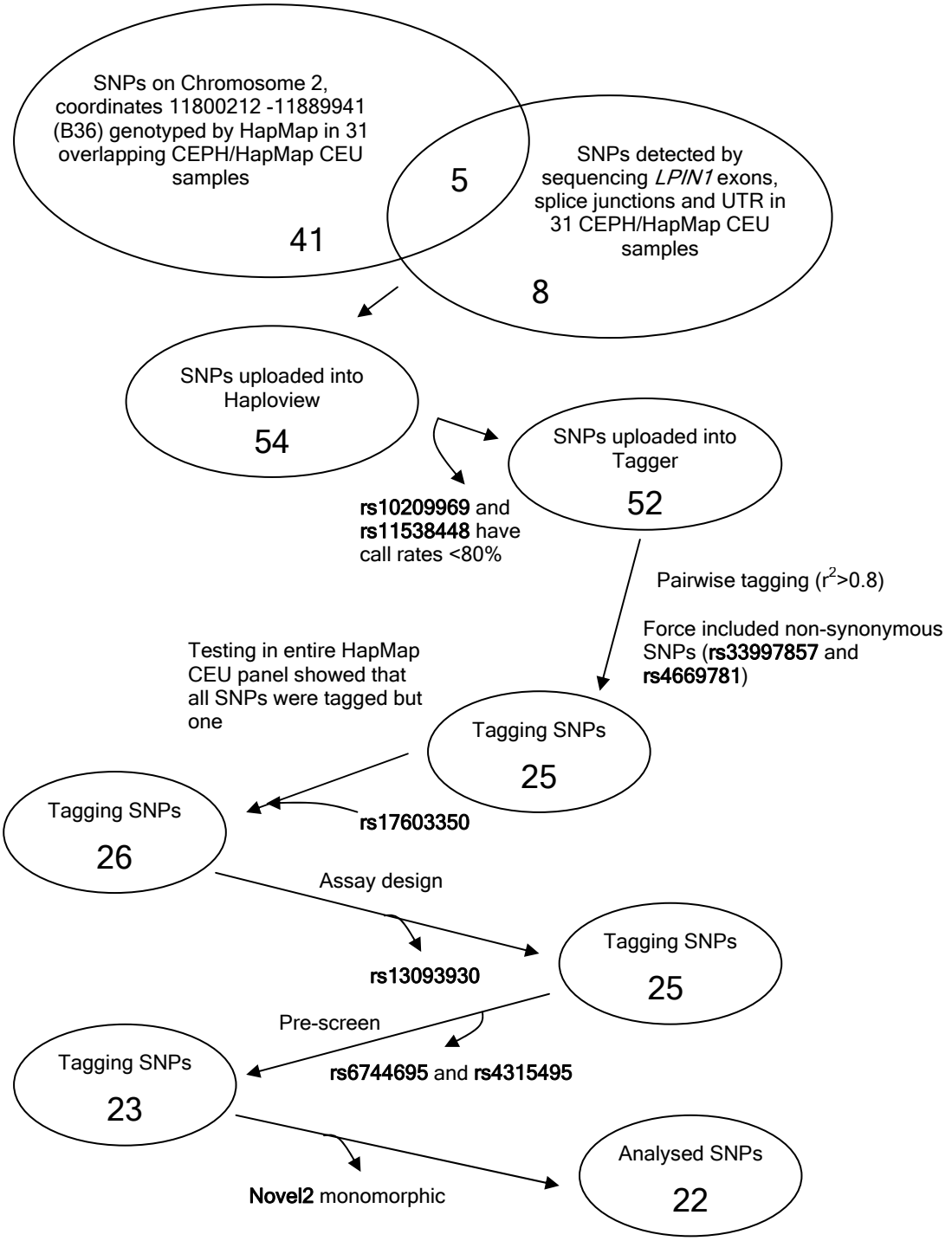

Figure 3.4 Flow chart showing the process of SNP identification and selection

Table 3.1 *LPIN1* sequence variants detected in 48 CEPH controls

Genic position	Genomic position	Nucleotide substitution	Protein consequence	MAF in CEPH	Rs number
Upstream	11789942-3	InsT		0.27	rs3214670
Intron 2	11823327	T>C		0.35	rs10209969
	11825293	C>T		0.28	Novel1
Exon 4	11829212	C>T	I184I	0.14	rs11538448
Intron 5	11836949	G>A		0.11	rs2289193
	11841434	C>T	D395D	0.016	Novel2
Intron 8	11842401	C>T		0.28	rs3795974
Exon 10	11844689	G>A	**V494M**	0.02	rs33997857
Intron 11	11849392-3	InsT		0.12	Indel1
Intron 12	11849624	G>T		0.44	rs7561070
Exon 14	11860533	C>T	**P610S**	0.04	rs4669781
Intron 15	11862198	DelG		0.14	Indel2
Exon 20	11882248	G>A	P851P	0.016	Novel3
3`UTR	11883265	C>T		0.3	rs1050800
	11883768	T>C		0.17	rs11524
	11884454	C>G		0.16	Novel4

Genomic coordinates correspond to NCBI Build 36. Non-synonymous changes are highlighted in bold. Ins = insertion. Del = deletion.

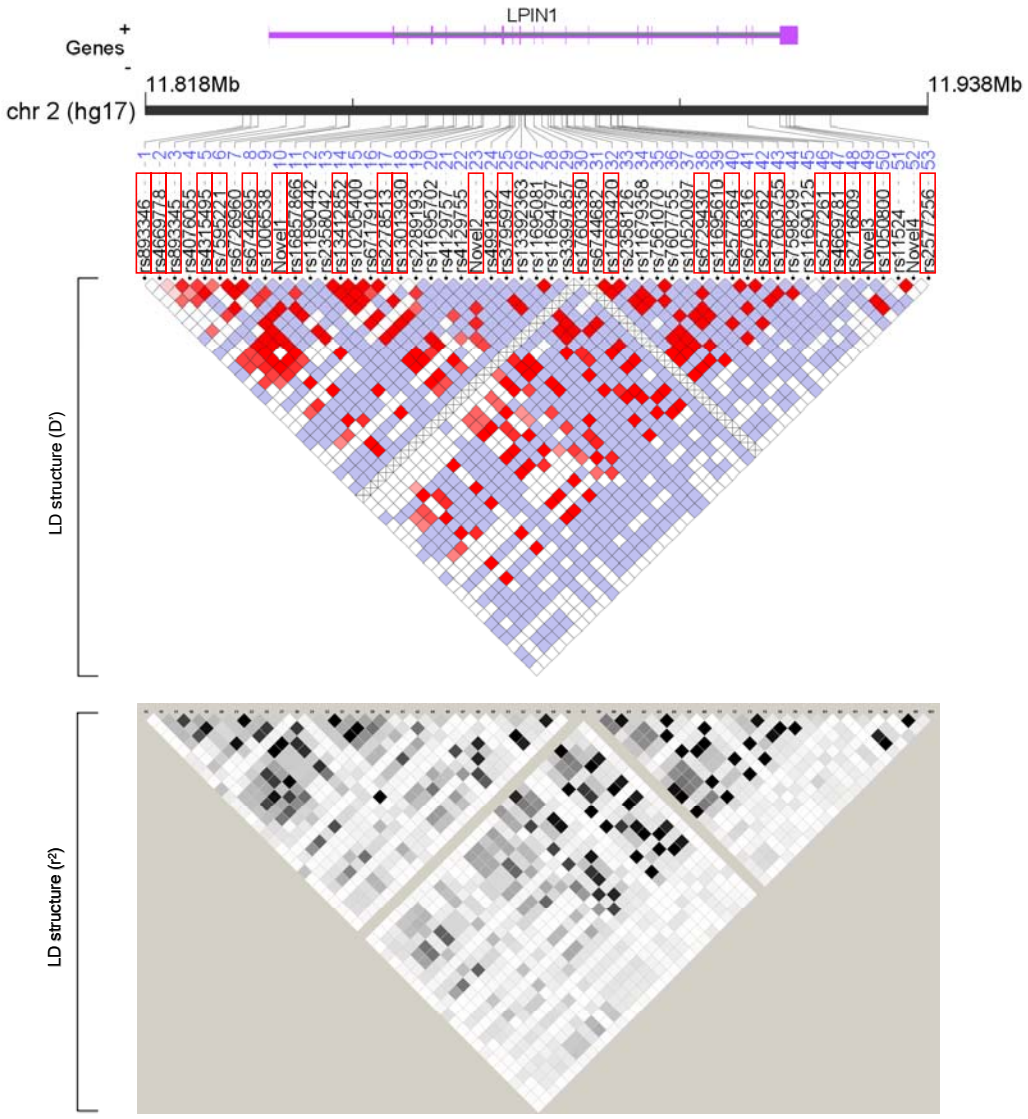

Figure 3.5 Pairwise linkage disequilibrium (LD) between SNPs in LPIN1 and the surrounding sequence

SNPs displayed are a combination of those detected from sequencing of 31 CEPH samples and those genotyped in overlapping samples of CEPH Utah trios in HapMap (release 20, phaseII Jan06, on NCBI B35 assembly, dbSNP B125) as well as one SNP (rs17603350) in CEU trios that was not present in the 31 overlapping samples. The upper plot presents LD as D' - see figure key for details. This figure was generated using Locusview (T. Petryshen, A. Kirby, M. Ainscow, unpublished software). In the lower plot LD is represented as r^2. Black diamonds represent an r^2 value of 1, grey diamonds represent intermediate values of r^2, and white diamonds indicate r^2 values of 0. This plot was generated using Haploview (Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005 Jan 15). SNPs in red boxes were selected as tagSNPs for this study using Tagger based on a pairwise r^2 threshold ≥0.8.

3.3.1.2 Association results of *LPIN1* tagging SNPs with fasting insulin and BMI

Of the 26 SNPs selected to tag the gene, one failed assay design, two failed pre-screening and one was monomorphic in the samples tested (section 3.5.4), leaving 22 SNPs for association testing (Figure 3.4 and Table 3.2). Of these 22 tagging SNPs with genotype data one, rs17603755, deviated from Hardy-Weinberg Equilibrium ($P > 0.01$) and was not tested for association. Tagging SNPs were successfully genotyped in >85% of samples from the MRC Ely and Hertfordshire cohort studies (call rates for each SNP are presented in Table 3.2) and the average call rate was 92.1%.

Association of tagging SNPs with fasting insulin levels and BMI was tested by linear regression analysis in the MRC Ely and Hertfordshire cohorts separately (Tables 3.3 and 3.4). In the MRC Ely cohort, the minor allele of rs13412852 is nominally associated with lower fasting insulin levels ($P = 0.041$) and the minor allele of rs17603350 is nominally associated with higher BMI ($P = 0.031$) (Table 3.3) but these associations are not replicated in the Hertfordshire cohort (Table 3.4). Conversely, in the Hertfordshire cohort, rs17603420 and rs2577261 are nominally associated with BMI ($P = 0.01$ and $P = 0.006$ respectively) (Table 3.4), but are not associated with BMI in the MRC Ely cohort (Table 3.3). To increase the statistical power of this study to detect modest effects of SNPs on insulin levels and BMI, I pooled data from Ely and Hertfordshire cohorts and performed a joint analysis (Table 3.5). No SNPs were associated with fasting insulin levels but rs13412852, rs17603420 and rs2577261 were nominally associated with BMI ($P \leq 0.05$). I performed 10,000 permutations of BMI to test for empirical significance of rs13412852, rs17603420 and rs2577261, which was confirmed in every case (P = 0.028, 0.006, and 0.005 respectively).

Table 3.2 Tagging SNPs in *LPIN1* analysed for association with metabolic traits

TagSNPs	Genomic position	Genic position	Protein consequence	Alleles (major/minor)	MAF in CEPHs	Call rate in Ely and Hertfordshire (%)
rs893346	11800212	5'		A/G	0.065	94
rs4669778	11801431	5'		T/C	0.371	92
rs893345	11802567	5'		G/A	0.387	86
rs7595221	11814426	Intron 1		A/G	0.371	93
Novel1	11825293	Intron 2		C/T	0.204	91
rs16857866	11828169	Intron 3		C/T	0.016	93
rs13412852	11832392	Intron 5		C/T	0.371	93
rs2278513	11835356	Intron 5		C/T	0.419	92
rs3795974	11842401	Intron 8		G/A	0.29	92
rs33997857	11844689	Exon 10	494V>M	G/A	0.016	94
rs17603350	11845874	Intron 10		C/T	0	91
rs17603420	11847191	Intron 11		G/A	0.5	93
rs6729430	11852207	Intron 12		C/T	0.016	94
rs2577264	11854249	Intron 13		C/T	0.306	93
rs2577262	11856397	Intron 13		G/A	0.339	93
rs17603755	11856469	Intron 13		G/A	0.125	87
rs2577261	11858238	Intron 13		G/T	0.145	91
rs4669781	11860533	Exon 14	610P>S	C/T	0.016	95
rs2716609	11877356	Intron 18		T/C	0.129	88
Novel3	11882248	Exon 20	851P>P	G/A	0.016	94
rs1050800	11883265	3'UTR		C/T	0.145	92
rs2577256	11889941	3'		A/G	0.468	93

Minor Allele Frequency (MAF) is provided for a subset of 31 CEPH samples that overlapped between the HapMap CEU trios and samples used for my own resequencing efforts. These frequencies compare well to those reported in HapMap. Genomic coordinates correspond to NCBI Build 36.

Table 3.3 Mean fasting insulin levels and mean BMI of study participants by *LPIN1* tagSNP genotype in the MRC Ely cohort

	Ely							
	Insulin (pmol/l)				BMI (kg/m^2)			
SNP	0	1	2	P value	0	1	2	P value
rs893346	49.18 ± 1.02	47.65 ± 1.05	57.57 ± 1.08	*0.609*	27.23 ± 0.13	27.5 ± 0.39	26.3 ± 1.7	*0.303*
rs4669778	49.99 ± 1.03	48.88 ± 1.02	47.23 ± 1.03	*0.216*	27.42 ± 0.24	27.41 ± 0.16	27.11 ± 0.27	*0.109*
rs893345	48.17 ± 1.03	48.69 ± 1.02	49.76 ± 1.03	*0.424*	27.27 ± 0.28	27.75 ± 0.32	25.87 ± 1.57	*0.967*
rs7595221	47.75 ± 1.03	50.08 ± 1.02	48.59 ± 1.04	*0.55*	27.11 ± 0.22	26.99 ± 0.24	26.09 ± 0.92	*0.815*
Novel1	48.46 ± 1.02	49.79 ± 1.03	49.67 ± 1.08	*0.409*	27.15 ± 0.15	27.49 ± 0.21	27.01 ± 0.64	*0.393*
rs16857866	49.1 ± 1.01	53.35 ± 1.08	47.7 ± 0	*0.398*	27.23 ± 0.12	28.02 ± 0.64	29.65 ± 0	*0.215*
rs13412852	50.18 ± 1.02	48.63 ± 1.02	45.1 ± 1.05	***0.042***	27.42 ± 0.18	27.36 ± 0.18	26.44 ± 0.35	*0.054*
rs2278513	47.41 ± 1.03	49.45 ± 1.02	48.82 ± 1.04	*0.37*	27.06 ± 0.22	27.28 ± 0.16	27.45 ± 0.26	*0.467*
rs3795974	48.59 ± 1.02	49.57 ± 1.02	49.93 ± 1.04	*0.441*	27.21 ± 0.19	27.76 ± 0.96	27.86 ± 3.19	*0.844*
rs33997857	48.95 ± 1.01	46.11 ± 1.1	56.18 ± 1.9	*0.627*	27.24 ± 0.12	27.7 ± 0.67	28.76 ± 0.57	*0.401*
rs17603350	48.99 ± 1.02	47.89 ± 1.05	99.5 ± 1.37	*0.986*	27.34 ± 0.13	26.3 ± 0.36	27.52 ± 1.28	***0.031***
rs17603420	50.63 ± 1.03	47.75 ± 1.02	49.14 ± 1.03	*0.271*	27.33 ± 0.2	27.39 ± 0.18	26.9 ± 0.25	*0.269*
rs6729430	48.95 ± 1.02	45.77 ± 1.1	49.99 ± 1.47	*0.61*	27.24 ± 0.12	27.39 ± 0.34	29.79 ± 1.28	*0.508*
rs2577264	48.61 ± 1.02	48.91 ± 1.02	50.6 ± 1.04	*0.368*	27.19 ± 0.2	27.37 ± 0.18	27.21 ± 0.3	*0.733*
rs2577262	49.25 ± 1.02	48.96 ± 1.02	48.72 ± 1.05	*0.639*	27.24 ± 0.17	27.61 ± 0.28	27.99 ± 0.86	*0.857*
rs2577261	49.52 ± 1.02	46.67 ± 1.03	48.09 ± 1.12	*0.15*	27.18 ± 0.12	27.38 ± 0.18	27.23 ± 0.3	*0.256*
rs4669781	48.91 ± 1.02	49.56 ± 1.04	43.59 ± 1.21	*0.844*	27.22 ± 0.12	27.19 ± 0.18	27.25 ± 0.29	*0.780*
rs2716609	48.99 ± 1.02	48.25 ± 1.05	44.85 ± 1.2	*0.61*	27.13 ± 0.15	27.34 ± 0.17	26.82 ± 0.24	*0.073*
Novel3	48.11 ± 1.02	50.44 ± 1.04	52.4 ± 1.17	*0.117*	27.25 ± 0.12	27.25 ± 1.26	0 ± 0	*0.987*
rs1050800	49.01 ± 1.01	47.15 ± 1.19	0 ± 0	*0.728*	27.21 ± 0.14	27.34 ± 0.23	26.55 ± 0.62	*0.993*
rs2577256	48.81 ± 1.02	49.32 ± 1.03	49.09 ± 1.09	*0.862*	27.02 ± 0.24	27.41 ± 0.19	26.94 ± 0.38	*0.208*

Data are means ± standard error. The *P* value indicates the results of a regression analysis assuming an additive model of gene action (nominally significant values, p<0.05, are highlighted in bold). For fasting insulin the analysis was performed on log-transformed data, and the table shows geometric means and standard errors. 0 = homozygous for the major allele (refer to Supplementary table 2), 1 = heterozygous, 2 = homozygous for the minor allele.

Table 3.4 Mean fasting insulin levels and mean BMI of study participants by *LPIN1* tagSNP genotype in the Hertfordshire cohort study

	Hertfordshire							
	Insulin (pmol/l)				BMI (kg/m^2)			
SNP	0	1	2	P value	0	1	2	P value
rs893346	70.6 ± 1.01	72.33 ± 1.04	78.17 ± 1.2	*0.486*	27.36 ± 0.09	27.43 ± 0.24	26.14 ± 1.08	*0.993*
rs4669778	71.28 ± 1.02	72.11 ± 1.02	69.06 ± 1.03	*0.409*	27.41 ± 0.16	27.37 ± 0.12	27.33 ± 0.16	*0.719*
rs893345	71.15 ± 1.03	71.73 ± 1.02	69.14 ± 1.03	*0.406*	27.21 ± 0.17	27.38 ± 0.12	27.49 ± 0.17	*0.234*
rs7595221	70.01 ± 1.02	71.45 ± 1.02	70.3 ± 1.03	*0.708*	27.34 ± 0.14	27.44 ± 0.12	27.23 ± 0.19	*0.734*
Novel1	71.71 ± 1.02	70.63 ± 1.02	71.25 ± 1.06	*0.575*	27.34 ± 0.1	27.52 ± 0.16	27.72 ± 0.45	*0.172*
rs16857866	70.96 ± 1.01	72.11 ± 1.06	0 ± 0	*0.734*	27.36 ± 0.09	27.8 ± 0.4	0 ± 0	*0.306*
rs13412852	70.57 ± 1.02	71.63 ± 1.02	69.18 ± 1.04	*0.89*	27.47 ± 0.13	27.36 ± 0.12	27.15 ± 0.2	*0.249*
rs2278513	69.47 ± 1.02	71.46 ± 1.02	72.34 ± 1.03	*0.234*	27.45 ± 0.14	27.32 ± 0.12	27.42 ± 0.2	*0.753*
rs3795974	71.04 ± 1.02	70.16 ± 1.02	72.91 ± 1.04	*0.602*	27.31 ± 0.13	27.41 ± 0.13	27.42 ± 0.22	*0.633*
rs33997857	70.82 ± 1.01	78.33 ± 1.07	73.83 ± 0	*0.166*	27.36 ± 0.08	27.91 ± 0.44	25.2 ± 0	*0.259*
rs17603350	71.27 ± 1.01	67.78 ± 1.05	60.16 ± 2.08	*0.282*	27.4 ± 0.09	27.15 ± 0.36	30.9 ± 4.96	*0.725*
rs17603420	70.43 ± 1.02	71.41 ± 1.02	70.26 ± 1.03	*0.971*	27.58 ± 0.16	27.42 ± 0.12	26.92 ± 0.17	***0.01***
rs6729430	70.98 ± 1.01	71.68 ± 1.08	0 ± 0	*0.849*	27.34 ± 0.08	28.37 ± 0.53	0 ± 0	*0.053*
rs2577264	71.27 ± 1.02	69.8 ± 1.02	73.51 ± 1.04	*0.644*	27.28 ± 0.13	27.39 ± 0.12	27.48 ± 0.23	*0.462*
rs2577262	70.76 ± 1.02	71.58 ± 1.02	69.77 ± 1.04	*0.95*	27.47 ± 0.12	27.37 ± 0.13	26.98 ± 0.25	*0.148*
rs2577261	71.08 ± 1.02	71.39 ± 1.03	73.5 ± 1.08	*0.719*	27.23 ± 0.09	27.67 ± 0.22	29.06 ± 0.72	***0.006***
rs4669781	71.03 ± 1.01	69.44 ± 1.03	66.11 ± 1.11	*0.426*	27.41 ± 0.09	27.02 ± 0.24	26.81 ± 2.07	*0.126*
rs2716609	71.19 ± 1.01	69.03 ± 1.04	78.39 ± 1.22	*0.597*	27.35 ± 0.1	27.42 ± 0.17	27.82 ± 0.66	*0.465*
Novel3	71.65 ± 1.02	68.38 ± 1.03	76.59 ± 1.08	*0.34*	27.38 ± 0.08	25.71 ± 0.72	0 ± 0	*0.103*
rs1050800	71.02 ± 1.01	69.3 ± 1.15	0 ± 0	*0.848*	27.4 ± 0.1	27.34 ± 0.16	27.14 ± 0.55	*0.585*
rs2577256	70.48 ± 1.02	71.77 ± 1.02	74.09 ± 1.07	*0.386*	27.19 ± 0.16	27.42 ± 0.12	27.47 ± 0.18	*0.219*

Data are means ± standard error. The *P* value indicates the results of a regression analysis assuming an additive model of gene action (nominally significant values, p<0.05, are highlighted in bold). For fasting insulin the analysis was performed on log-transformed data, and the table shows geometric means and standard errors. 0 = homozygous for the major allele (refer to Supplementary table 2), 1 = heterozygous, 2 = homozygous for the minor allele.

Table 3.5 Joint analysis of the association between *LPIN1* tagSNPs and fasting insulin or BMI in combined Ely and Hertfordshire datasets

	Fasting insulin		BMI	
SNP	β ± SE	P value	β ± SE	P value
rs893346	0.0086 ± 0.03	*0.755*	0.1339 ± 0.2	*0.494*
rs4669778	-0.0191 ± 0.01	*0.163*	-0.131 ± 0.1	*0.18*
rs893345	-0.0042 ± 0.01	*0.768*	0.084 ± 0.1	*0.401*
rs7595221	0.0085 ± 0.01	*0.532*	0.0005 ± 0.1	*0.996*
Novel1	0.0002 ± 0.02	*0.988*	0.1893 ± 0.12	*0.123*
rs16857866	0.0374 ± 0.05	*0.442*	0.5632 ± 0.35	*0.107*
rs13412852	-0.0183 ± 0.01	*0.203*	-0.2242 ± 0.1	***0.028***
rs2278513	0.0207 ± 0.01	*0.137*	0.0281 ± 0.1	*0.777*
rs3795974	0.0116 ± 0.01	*0.408*	0.0602 ± 0.1	*0.544*
rs33997857	0.0386 ± 0.05	*0.446*	0.4933 ± 0.36	*0.171*
rs17603350	-0.0299 ± 0.04	*0.428*	-0.4412 ± 0.27	*0.1*
rs17603420	-0.0091 ± 0.01	*0.502*	-0.2653 ± 0.1	***0.006***
rs6729430	-0.0121 ± 0.06	*0.839*	0.8156 ± 0.42	*0.053*
rs2577264	0.0121 ± 0.01	*0.39*	0.0873 ± 0.1	*0.381*
rs2577262	-0.0049 ± 0.01	*0.738*	-0.1236 ± 0.1	*0.237*
rs2577261	-0.0128 ± 0.02	*0.593*	0.4756 ± 0.17	***0.005***
rs4669781	-0.0237 ± 0.03	*0.434*	-0.1977 ± 0.21	*0.356*
rs2716609	0.003 ± 0.02	*0.881*	0.2383 ± 0.14	*0.093*
Novel3	-0.043 ± 0.12	*0.71*	-1.0324 ± 0.83	*0.214*
rs1050800	0.0144 ± 0.02	*0.447*	-0.0461 ± 0.13	*0.731*
rs2577256	-0.0094 ± 0.01	*0.503*	0.1761 ± 0.1	*0.074*

β (regression coefficient) is the mean change (± standard error) in fasting insulin or BMI per minor allele. Analysis for fasting insulin was performed on log-transformed data and all analyses included an adjustment for cohort term. Nominally significant SNPs are highlighted in bold.

3.3.1.3 Use of publicly available datasets to increase power

Two of these SNPs, rs13412852 and rs2577261, overlapped with SNPs on the Affymetrix 500k and Illumina 300k SNP chips, and rs17603420 could be imputed. Consequently I was able to further increase the power of my study to detect modest effects of these SNPs on BMI by performing meta-analyses with in-house data (EPIC-Obesity study, N = 2415) and, in the case of rs13412852 and rs2577261, data deposited by the WTCCC and WTSI and published online from the British 1958 DNA collection (N = 1479) (http://www.b58cgene.sgul.ac.uk/, accessed January 2008). SNPs rs2577261 and rs17603420 were not associated with BMI in the meta-analysis (*P* = 0.114 and 0.071 respectively). However, the association between rs13412852 and BMI remained marginally statistically significant (*P* = 0.042) in the meta-analysis (Figure 3.6).

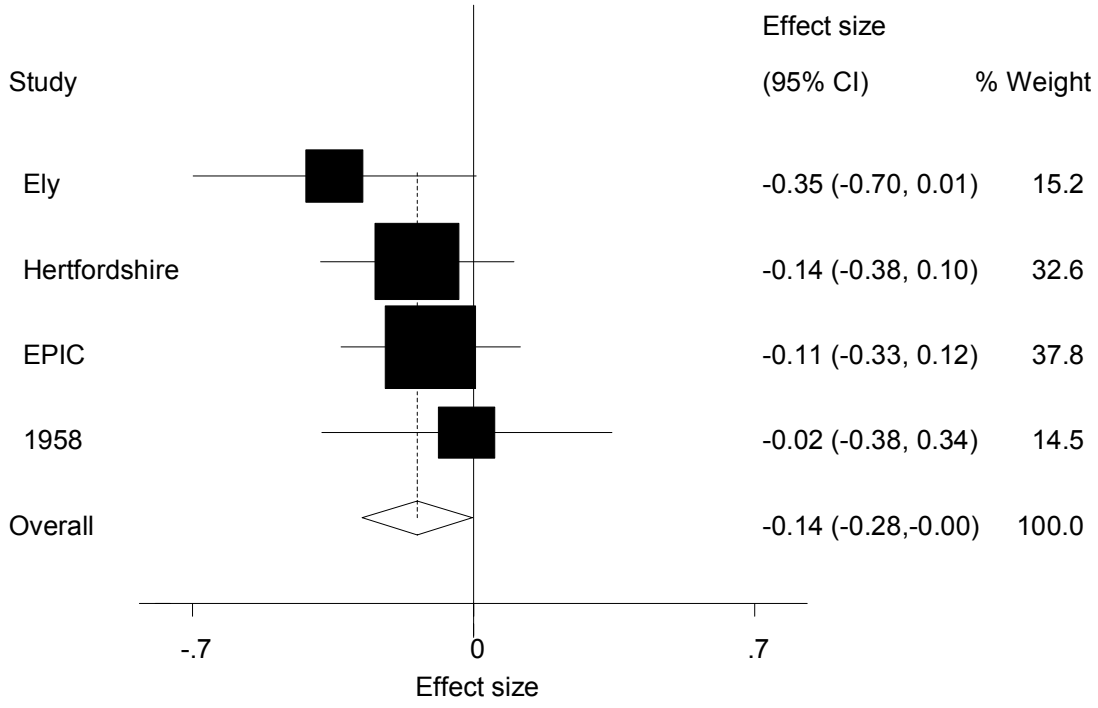

Figure 3.6 Association between rs13412852 and BMI in studies within the meta-analysis
Ely = MRC Ely study, Hertfordshire = Hertfordshire cohort study, EPIC = EPIC Obesity cohort, and 1958 = 1958 British Birth cohort. Overall = combined effect size, *P* = 0.045.

3.3.1.4 Testing for association of tagging SNPs with additional metabolic quantitative traits

In pooled data from Ely and Hertfordshire cohorts, linear regression analysis was used to test for association between *LPIN1* tagging SNPs and metabolic quantitative traits found to be associated with *LPIN1* variation in the MONICA study (Table 3.6A). A number of nominal associations (P <0.05) were detected, namely between *LPIN1* variation and systolic and diastolic blood pressure, low-density lipoprotein (LDL) and high-density lipoprotein (HDL) cholesterol, triglycerides and, in Ely only, glycosylated haemoglobin (HbA1c) levels. Two SNPs, rs2278513 and rs2577256, are present on the Affymetrix SNP chip and have been analysed as part of the DGI study at the Broad Institute. Consequently I was able to use publicly available summary statistics (http://www.broad.mit.edu/diabetes/scandinavs/metatraits.html) to perform meta-analyses with my data (see Table 3.6B for P-values).

3.3.1.5 Testing for association of tagging SNPs with hypertension and diabetes

In addition to continuous metabolic traits, the MONICA study reported association between *LPIN1* variation and risk of diabetes and hypertension. These traits were also tested in my pooled Ely and Hertfordshire data by logistic regression analysis. One SNP, rs2577256, was associated with diabetes and several SNPs were associated with hypertension status (Table 3.7). For rs2577256 meta-analysis was performed with publicly available WTCCC diabetes and hypertension data (see WTCCC and meta-analysis summary statistics in Table 3.7).

The MONICA study detected associations between metabolic traits and haplotypes of three *LPIN1* SNPs, rs33997857, rs6744682, and rs6708316. Only one of these SNPs, rs33997857, is within my SNP set but rs2577262 is highly correlated with the two other SNPs in the haplotype, rs6744682 and rs6708316 (r^2 = 1.0 and 0.96 respectively in HapMap CEU trios). To attempt replication of the MONICA study data

I tested haplotypes of rs33997857 and rs2577262 against metabolic traits in Ely and Hertfordshire but only found nominal associations with hypertension (Table 3.8).

Table 3.6A Statistically significant associations between traits underlying metabolic syndrome and *LPIN1* tagSNPs in Ely and Hertfordshire cohorts

SNP	Ely				Hertfordshire				Ely + Hertfordshire	
	0	1	2	P	0	1	2	P	β ± SE	P
	Systolic blood pressure (mmHg)									
rs2278513	130.30 ± 0.76	131.20 ± 0.59	133.33 ± 0.95	0.01	132.81 ± 0.67	133.09 ± 0.52	134.62 ± 0.83	0.074	0.009 ± 0.003	0.004
rs2716609	131.34 ± 0.51	130.17 ± 0.97	134.90 ± 3.29	0.824	132.56 ± 0.42	135.56 ± 0.76	132.49 ± 2.60	0.005	0.01 ± 0.004	0.023
	Diastolic blood pressure (mmHg)									
rs2278513	77.63 ± 0.46	78.62 ± 0.37	79.73 ± 0.61	0.004	70.39 ± 0.38	70.76 ± 0.30	71.24 ± 0.51	0.067	0.01 ± 0.003	0.001
rs2716609	78.64 ± 0.32	78.41 ± 0.62	80.29 ± 1.89	0.689	70.32 ± 0.25	72.03 ± 0.44	69.29 ± 1.51	0.011	0.012 ± 0.005	0.008
	HbA1c (%)									
rs4669778	5.45 ± 0.03	5.4 ± 0.02	5.34 ± 0.03	0.007	Not tested	Not tested	Not tested	N/A	-0.009 ± 0.004	0.007
rs13412852	5.44 ± 0.03	5.37 ± 0.02	5.33 ± 0.04	0.004	Not tested	Not tested	Not tested	N/A	-0.01 ± 0.004	0.004
rs3795974	5.36 ± 0.02	5.41 ± 0.02	5.46 ± 0.05	0.016	Not tested	Not tested	Not tested	N/A	0.008 ± 0.003	0.016
rs17603420	5.44 ± 0.03	5.4 ± 0.02	5.32 ± 0.03	0.004	Not tested	Not tested	Not tested	N/A	-0.01 ± 0.003	0.004
rs2577264	5.35 ± 0.02	5.42 ± 0.02	5.46 ± 0.05	0.004	Not tested	Not tested	Not tested	N/A	0.01 ± 0.003	0.004
Novel3	5.4 ± 0.01	5.05 ± 0.09		0.016	Not tested	Not tested	Not tested	N/A	-0.07 ± 0.029	0.016
	LDL cholesterol (mmol/l)									
rs2577256	3.43 ± 0.05	3.5 ± 0.03	3.57 ± 0.05	0.036	3.91 ± 0.04	3.93 ± 0.03	4.02 ± 0.04	0.02	0.019 ± 0.006	0.002
	HDL cholesterol (mmol/l)									
rs7595221	1.43 ± 0.02	1.49 ± 0.01	1.48 ± 0.02	0.011	1.53 ± 0.01	1.51 ± 0.01	1.57 ± 0.02	0.368	0.012 ± 0.005	0.019
rs893345	1.47 ± 0.02	1.48 ± 0.02	1.42 ± 0.02	0.032	1.56 ± 0.02	1.51 ± 0.01	1.54 ± 0.02	0.332	-0.011 ± 0.006	0.038
	Triglycerides (mmol/l)									
rs2577256	1.43 ± 0.04	1.41 ± 0.03	1.53 ± 0.05	0.04	1.59 ± 0.03	1.66 ± 0.02	1.67 ± 0.04	0.066	0.029 ± 0.01	0.007

Data for Ely and Hertfordshire cohorts separately are presented as means ± standard error. 0 = homozygous for the major allele (refer to Supplementary table 2), 1 = heterozygous, 2 = homozygous for the minor allele. β (regression coefficient) is the mean change (± standard error) in the metabolic trait per minor allele. The *P* values indicate the results of a regression analysis on log-transformed data that assumed an additive model of gene action. It should be noted that there is a high prevalence of treatment for hypertension in the Hertfordshire cohort (36%) which might render measured blood pressure values of limited use.

Table 3.6B Meta-analysis of Ely, Hertfordshire and Broad continuous trait data

SNP	Trait	Broad β ± SE	Meta P
rs2278513	Systolic BP	-0.570 ± 0.543	0.110
rs2278513	Diastolic BP	-0.860 ± 0.416	0.157
rs2577256	LDL cholesterol	-0.01 ± 0.027	0.051
rs2577256	Triglycerides	0 ± 0.27	0.006

Regression coefficients ± SE of publicly available data from the Broad Institute are represented in the penultimate column and *P* values for the meta-analysis of my UK data and the Broad data are represented in the last column.

Table 3.7 Statistically significant associations between *LPIN1* tagSNPs and risk of diabetes or hypertension in Ely and Hertfordshire cohorts, and in a pooled analysis of both cohorts and publicly available WTCCC data

SNP	Ely		Hertfordshire		Ely and Hertfordshire		WTCCC		Meta (+WTCCC)	
	OR ± SE	P value	OR ± SE	P value	OR ± SE	P value	OR ± SE	P value	OR ± SE	P value
	Diabetes (502 cases, 3688 controls)									
rs2577256	1.06 ± 0.17	0.716	1.29 ± 0.10	0.001	1.24 ± 0.09	0.002	1.02 ± 0.04	0.589	1.08 ± 0.04	0.038
	Hypertension (1603 cases, 2991 controls)									
rs3795974	1.05 ± 0.09	0.562	1.14 ± 0.06	0.023	1.11 ± 0.05	0.024				
rs17603420	1.09 ± 0.09	0.283	0.84 ± 0.05	0.002	0.91 ± 0.04	0.044				
rs2577264	1.06 ± 0.09	0.503	1.16 ± 0.07	0.009	1.13 ± 0.05	0.010				
rs2577262	0.90 ± 0.08	0.250	0.89 ± 0.05	0.058	0.90 ± 0.04	0.029				
rs4669781	0.73 ± 0.14	0.093	0.78 ± 0.10	0.052	0.76 ± 0.08	0.011				
rs1050800	0.84 ± 0.10	0.141	0.84 ± 0.07	0.025	0.84 ± 0.05	0.009				
rs2577256	1.10 ± 0.09	0.248	1.20 ± 0.07	0.001	1.17 ± 0.05	0.001	1.00 ± 0.04	0.998	1.07 ± 0.03	0.021

Data are odds ratios (giving the increase in disease risk per minor allele) ± standard error. *P* values indicate the outcome of logistic regression analyses. Included as diabetes cases were individuals with previously diagnosed diabetes, and also newly diagnosed diabetes based on glucose measurements taken 2 hours after an oral glucose tolerance test. The two columns headed WTCCC provide summary statistics for the publicly available Wellcome Trust Case Control Consortium (WTCCC). The two columns headed Meta (+WTCCC) describe effect size and significance of the association between rs2577256 and diabetes or hypertension in meta-analyses of Ely, Hertfordshire and the publicly available WTCCC data. The meta-analyses are based on diabetes data for 2426 cases and 6619 controls, and hypertension data for 3486 cases and 5796 controls.

Table 3.8 Statistical association of haplotypes of rs33997857 and rs2577262 with metabolic traits in Ely and Hertfordshire cohorts

	P values		
Trait	Haplotype 11	Haplotype 12	Haplotype 21
Systolic blood pressure	0.1421	0.1956	0.5611
Diastolic blood pressure	0.1492	0.2575	0.6469
HbA1c *	0.617	0.781	0.448
LDL cholesterol	0.9382	0.937	0.9806
HDL cholesterol	0.1022	0.08027	0.8017
Waist	0.8637	0.7763	0.6769
Triglycerides	0.9778	0.6594	0.1002
Diabetes	0.8743	0.9038	0.904
Hypertension	*0.02982*	*0.02752*	0.9864

P values are the result of meta-analysis of Ely and Hertfordshire summary statistics using the METAL program, except in the case of HbA1c data which is based on Ely only (*).

3.3.2 *LPIN1* mutation screening in the SIR cohort

3.3.2.1 *LPIN1* variation in SIR

A total of 44 variants were detected in insulin resistant or lipodystrophic patients (Table 3.9), eight of which were present in the coding sequence (Figure 3.7). Coding sequence variants that did not alter the amino acid sequence (shown below the schematic in Figure 3.7) and/or that were also present in CEPH controls were considered unlikely to be pathogenic. This left three rare nonsynonymous variants (A353T, R552K, and G582R - underlined above the schematic in Figure 3.7) that did not fall within any known functional domains within *LPIN1*.

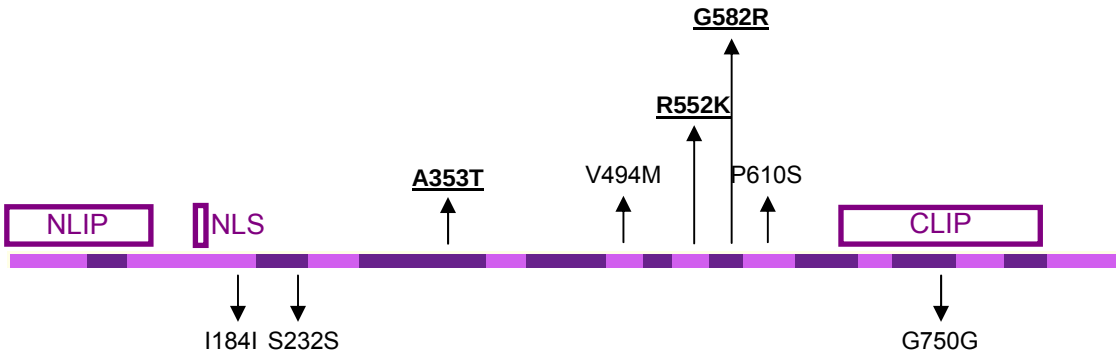

Figure 3.7 Coding *LPIN1* variants in the SIR cohort
Schematic of the lipin 1 protein showing exons in alternating bright and dark purple and known domains among lipin family proteins in boxes. Arrows indicate the location of coding SNPs detected in *LPIN1* by sequencing 23 patients with partial lipodystrophy and 135 patients with other syndromes of severe insulin resistance. Non-synonymous variants are above the schematic and synonymous variants are below the schematic. Non-synonymous mutations absent from control samples were considered potentially pathogenic (underlined). NLIP (amino acids 1-114) = N-terminal lipin domain, NLS (amino acids 153-158) = nuclear localisation signal, and CLIP (amino acids 674-830) = C-terminal lipin domain, also referred to as the LNS2 (Lipin/Ned1/Smp2) domain.

Table 3.9 *LPIN1* sequence variants detected in a cohort of severe insulin resistant patients and 48 CEPH controls

Genic position	Genomic position	Nucleotide substitution	Protein consequence	MAF in SIR	MAF in CEPH	Rs number
Upstream	11789942-3	InsT		0.27	0.27	rs3214670
Intron 2	11823327	T>C		0.25	0.35	rs10209969
	11823350	T>C		<0.01	0	
	11825297	G>T		<0.01	0	
	11825293	C>T		0.19	0.28	Novel1
Exon 4	11829212	C>T	I184I	0.14	0.14	rs11538448
Exon 5	11831296	G>C	S232S	<0.01	0	
Intron 5	11836949	G>A		0.12	0.11	rs2289193
Exon 7	11839985	G>A	**A353T**	<0.01	0	
	11841434	C>T	D395D	0	0.016	Novel2
Intron 8	11841548	A>T		<0.01	0	
	11842376	C>T		<0.01	0	
	11842401	C>T		0.43	0.28	rs3795974
Exon 10	11844689	G>A	**V494M**	0.03	0.02	rs33997857
Intron 10	11845874	C>T		0.02	0	rs17603350
	11845928	C>T		<0.01	0	
Intron 11	11846054	G>C		<0.01	0	
	11849353	C>T		<0.01	0	
	11849392-3	InsT		0.1	0.12	
	11849396	C>T		<0.01	0	
Exon 12	11849540	G>A	**R552K**	<0.01	0	
Intron 12	11849624	G>T		0.49	0.44	rs7561070
	11852911	A>G		<0.01	0	
	11852932	G>T		<0.01	0	
Exon 13	11853030	G>A	**G582R**	<0.01	0	
Intron 13	11853126	T>C		<0.01	0	
Exon 14	11860533	C>T	**P610S**	0.05	0.04	rs4669781
Intron 14	11860690	G>A		<0.01	0	
Intron 15	11862198	DelG		0.07	0.14	
	11862564	T>G		<0.01	0	
	11862577	T>G		<0.01	0	
Intron 16	11872591	T>C		<0.01	0	
Exon 17	11872773	G>C	G750G	0.01	0	
Intron 17	11872940	G>A		<0.01	0	
Exon 20	11882248	G>A	P851P	0	0.016	Novel3
3`UTR	11883265	C>T		0.23	0.3	rs1050800
	11883450	DelG		<0.01	0	
	11883469	C>T		<0.01	0	
	11883631	C>T		<0.01	0	
	11883712	DelATT		<0.01	0	
	11883768	T>C		0.13	0.17	rs11524
	11883798	C>T		<0.01	0	
	11884301	G>C		<0.01	0	
	11884454	C>G		0.08	0.16	Novel4
	11884670	G>T		0.02	0	
	11884739	T>C		<0.01	0	

Genomic coordinates correspond to NCBI Build 36. Non-synonymous changes are highlighted in red bold. Ins = insertion. Del = deletion.

3.3.2.2 Investigation of the A353T variation

A353T was detected in a female patient with a Pakistani father and British white mother. She presented with clinical features of severe insulin resistance at 8 years old, which worsened with weight gain in the second decade, before improving dramatically with weight loss in adult life. She had no evidence of lipodystrophy. Three methods were used to investigate whether A353T caused disease in the patient. First, I used a web-based program, PANTHER, which uses information on evolutionary sequence conservation to predict whether an amino acid substitution is likely to have a functional impact on the protein. An alanine to threonine change was predicted to have no functional impact on the lipin 1 protein. Figure 3.8 shows ClustalW alignments of orthologous lipin 1 peptide sequences from a variety of organisms. Highly conserved residues are more likely to have been subject to purifying selection and their substitution is more likely to have a functional impact on the protein (Miller and Kumar 2001). The alanine 353 residue is conserved in primates, mouse, rat, frog and chicken (Figure 3.8) but not opossum or zebrafish. Second, to determine if the variant co-segregates with disease in the family, DNA from the patient's mother, maternal aunts, and maternal grandparents was sequenced. Co-segregation analysis demonstrated that the A353T variant did not segregate with the hallmarks of insulin resistance in the family (Figure 3.9). Finally, A353T was also genotyped in 1064 participants of the HGDP panel to check for its presence in unaffected individuals but was not detected (data not shown).

```
H. sapiens          TLGAAAPLLPMI | KKGGRWWFSWRG | AHSTGEQPPQLSLATRVKHE
P. troglodytes      TLGAAAPLLPMI | KKGGRWWFSWRG | AHSTGEQPPQLSLATRVKHE
M. musculus         SLGAAAPPSPVA | KKGGRWWFSWRG | GHNTGEQPAQLGLATRIKHE
R. norvegicus       ALGAAAPPLSVA | KKGGRWWFSWRG | GHNTGEQPAQLGLTTRIKHE
M. domesticus       GQGGGGQALPGA | RKGGRWWFSWRG | GLVIGEQPAKLSIGTRMKEE
G. gallus           SAGAAVPSLPAN | KKGGRWWFSWRG | SRLKGEDSSQMTMANRIKDE
X. tropicalis       SLGAAAPPLPYD | KKGGRWWFSWRG | GPYSGGQPVGSSLENRIKDE
D. rerio            PISVVAH-LITE | KKGGRWWFSWRG | ESIRTG-----SVSSRLKDE
```

Figure 3.8 Multiple sequence alignments (using ClustalW) showing conservation of *LPIN1* amino acids A353, R552, and G582
Straight lines indicate hidden sequence.

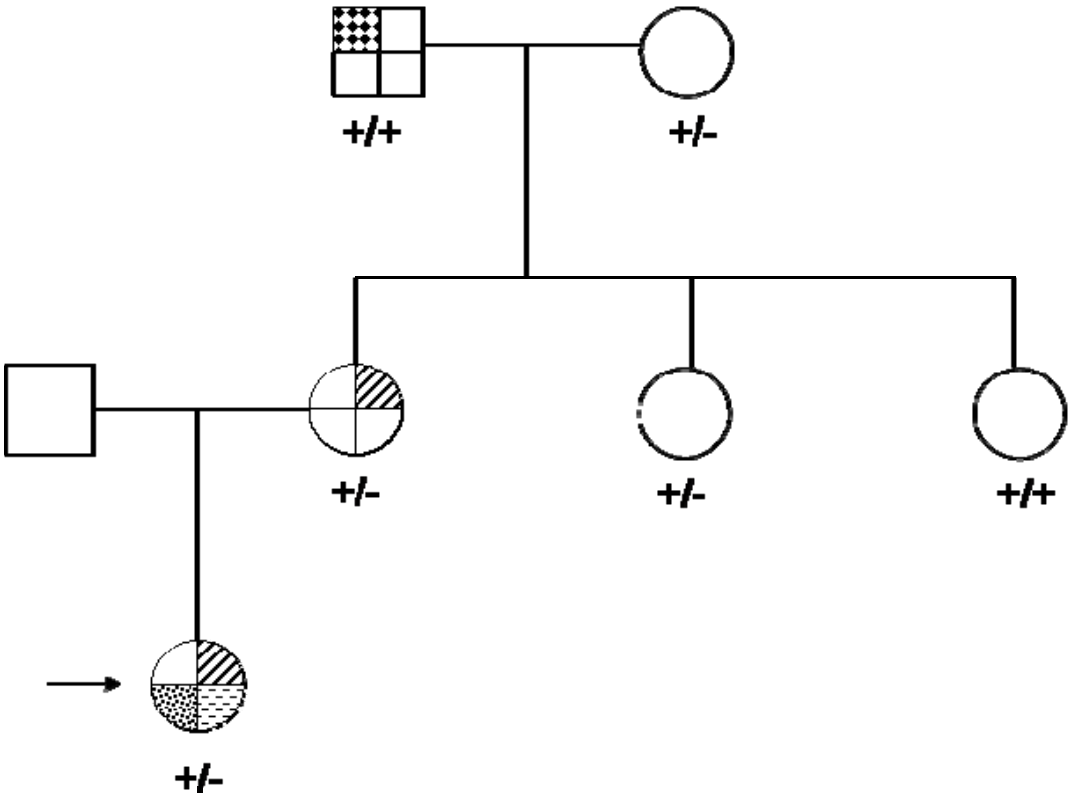

Figure 3.9 Family pedigree demonstrating that the A353T mutation does not segregate with disease in a fully penetrant manner
+/- represents a heterozygous genotype and +/+ represents the wild-type genotype. The patient (indicated by the arrow) has hyperinsulinaemia (diagonal stripes), hirsutism (spots) and acanthosis nigricans (dashes). The grandfather has diabetes (diamonds).

3.3.2.3 Investigation of the R552K variation

R552K was detected as a heterozygous change in two unrelated white European females, but not in 1064 controls from the Diversity Panel. The first proband presented with severe insulin resistance and femorogluteal lipodystrophy at 15 years old. The lipodystrophy progressed to become generalised in conjunction with the appearance of aggressive haemolytic anaemia and autoimmune liver disease. Liver failure led to her demise at 24 years old. The other proband was diagnosed with insulin resistant diabetes at 32 years old, and subsequently required in excess of 4U/day exogenous insulin to maintain satisfactory glycaemic control. She had no clinical evidence of lipodystrophy, and her BMI was sustained above 30 kg/m^2. R552 is within a highly conserved tract (Figure 3.8) and mutation to lysine is predicted by PANTHER to have deleterious effects on lipin 1 function. Family DNA was not available for co-segregation analysis for either patient.

3.3.2.4 Investigation of the G582R variation

G582R was identified as a heterozygous change in a white, European male with a complex syndrome. This included severe insulin resistance and severe, early onset sensorimotor neuropathy which confined him to a wheelchair, a combination highly reminiscent of the combined lipodystrophy/insulin resistance and neuropathy of the fld mouse. This patient also underwent allogeneic bone marrow transplantation in childhood for acute lymphoblastic leukaemia, and had a cerebral cavernous hemangioma. All genomic analyses were undertaken on DNA extracted from cultured skin fibroblasts. G582 is a well conserved residue within the protein (Figure 3.8), and mutation to arginine is predicted by PANTHER to have deleterious effects on lipin 1 function. Co-segregation analysis was performed using DNA from first-degree relatives of the patient (Figure 3.10). The father also carried the variant but although diagnosed with diabetes at age 69 years, he had no peripheral neuropathy, nor clinical or biochemical evidence of insulin resistance/lipodystrophy. This could be

because the wild-type copy of *LPIN1* can compensate for the effects of the G582R. As the patient is also heterozygous for G582R I sequenced patient cDNA to exclude the possibility that the patient only expresses the mutant copy of *LPIN1*. However, the patient's cDNA contained both alleles (Figure 3.11). Subsequently, the G582R variant was genotyped in the HGDP panel and detected in a Bedouin control individual from Nedev, Israel.

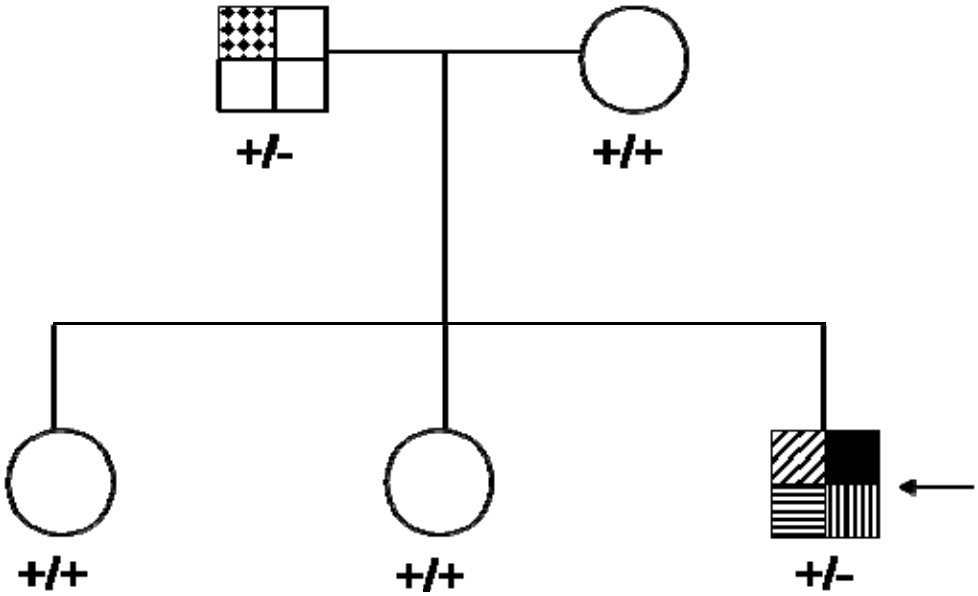

Figure 3.10 Family pedigree demonstrating that the G582R mutation does not segregate with disease in a fully penetrant manner
+/- represents a heterozygous genotype and +/+ represents the wild-type genotype. The patient (indicated by the arrow) has hyperinsulinaemia (diagonal stripes), severe peripheral neuropathy (black), previous bone marrow transplant for AML (horizontal stripes) and an intracerebral cavernous haemangioma (vertical stripes). His father has diabetes (diamonds).

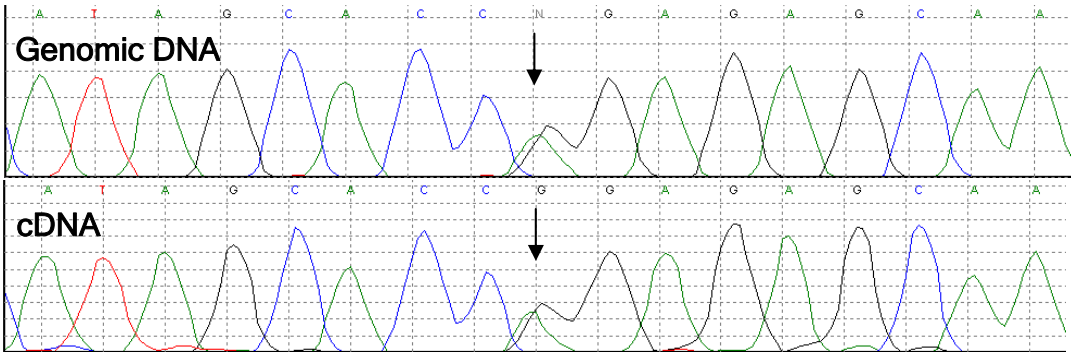

Figure 3.11 Sequences of genomic DNA and cDNA from the patient carrying the G582R variant (indicated by arrow)
This experiment shows that both alleles of G582R are expressed in patient fibroblasts.

In summary, I identified 3 rare missense variants in *LPIN1* in a cohort of patients with extreme insulin resistance. Two of these, G582R and R552K, are predicted to be deleterious to function by the PANTHER algorithm. Furthermore, the proband carrying the G582R variant had a syndrome including severe neuropathy and insulin resistance as seen in the fld mouse. Thus, despite the failure of this variant to co-segregate with the disease phenotype in the wider kindred, and despite the absence of available family members from the R552K kindred, Neil Grimsey and Symeon Siniossoglou at the Cambridge Institute for Medical Research investigated the impact of G582R and R552K on Lipin 1 function in primary skin fibroblasts from the probands. As A353T was predicted to be a benign change based on PANTHER and family co-segregation analysis, and as no fibroblasts were available, this variant was not investigated further.

3.3.2.5 Investigation of R552K and G582R in patient fibroblasts

To investigate the impact of R552K and G582R mutations on lipin 1 protein levels and phosphorylation status, total cell extracts from patient fibroblasts were probed with lipin 1, lipin 2, a nuclear pore marker (Mab414), and lamin B-specific primary antibodies to detect protein levels. This work was carried out by Neil Grimsey and Symeon Siniossoglou at the Cambridge Institute for Medical Research. The resulting

Western blot shown in Figure 3.12 shows similar intensities of all four proteins in patient fibroblasts compared to control cells. Immunocytochemistry was employed to detect changes in nuclear membrane morphology by staining a nuclear pore marker (Figure 3.13). There was no discernable difference in morphology between patient and control fibroblasts.

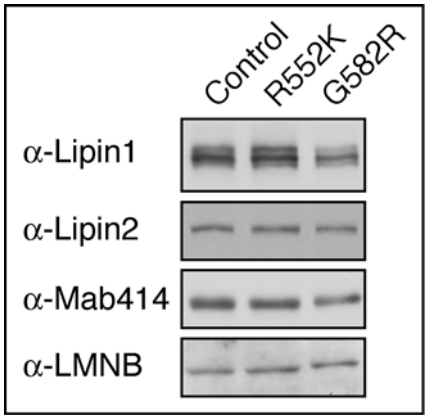

Figure 3.12 Western blot analysis of total cell extracts from cultured patient fibroblasts carrying the R552K and G582R *LPIN1* mutations (lanes 2 and 3 respectively) and control fibroblasts (lane 1)
Data from Neil Grimsey and Symeon Siniossoglou. The blots were probed with primary antibodies specific to lipin 1 (α-Lipin1), lipin 2 (α-Lipin2), a nuclear pore marker (α-Mab414) and lamin B (α-LMNB), and species-specific secondary antibodies. See Materials and Methods for details.

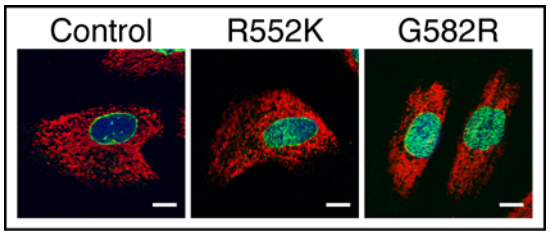

Figure 3.13 Immunofluoresence of control and patient fibroblasts showing Mab414, a nuclear pore marker (green), calreticulin, an ER calcium binding protein (red), and DNA (blue)
Data from Neil Grimsey and Symeon Siniossoglou
Bar = 5μm.

3.3.3 *LPIN2* and *LPIN3* mutation screening in the SIR cohort

3.3.3.1 *LPIN2* variation in SIR

Ninety-two *LPIN2* variants were detected in insulin resistant or lipodystrophic patients (Appendix Table A5), nine of which were present in the coding sequence (Table 3.10 and Figure 3.14). Coding sequence variants that did not alter the amino acid sequence (shown below the schematic in Figure 3.14) and/or that were also present in controls were considered unlikely to be pathogenic. This left two rare nonsynonymous variants (E497K and P626S - underlined above the schematic in Figure 3.14) that did not fall within any known functional domains within *LPIN2*, and were predicted benign by PANTHER. These were investigated no further.

Table 3.10 *LPIN2* coding variants detected in a cohort of severe insulin resistant patients and 11 Indian and 23 CEPH controls

Genic position	Genomic position	Minor|Major allele	Protein consequence	MAF in SIR	Detected in controls?
Exon 5	2930695	G|A	A202A	0.0027	
Exon 5	2930693	A|G	S203F	0.0081	Yes
Exon 10	2919124	T|C	**E497K**	0.0026	No
Exon 13	2916779	G|A	S579P	0.0026	Yes
Exon 14	2915359	T|C	E601K	0.0111	Yes
Exon 14	2915284	A|G	**P626S**	0.0027	No
Exon 16	2913790	A|G	L719L	0.0052	
Exon 20	2910387	G|A	S865S	0.0029	
Exon 20	2910372	A|G	S870S	0.0026	

Genomic coordinates correspond to NCBI Build 36.

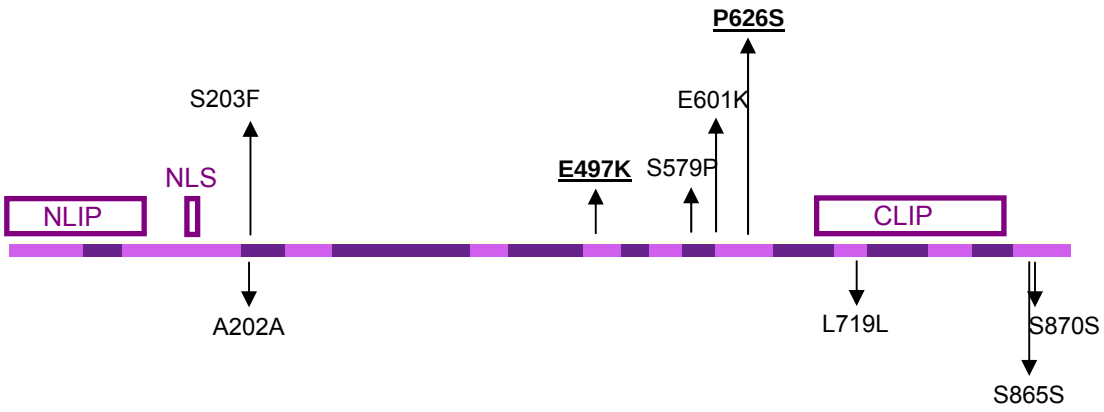

Figure 3.14 Coding *LPIN2* variants in the SIR cohort
Schematic of the lipin 2 protein showing exons in alternating bright and dark purple and known domains among lipin family proteins in boxes. Arrows indicate the location of coding SNPs detected in *LPIN2* by sequencing 23 patients with partial lipodystrophy and 135 patients with other syndromes of severe insulin resistance. Non-synonymous variants are above the schematic and synonymous variants are below the schematic. Non-synonymous mutations absent from control samples were considered potentially pathogenic (underlined). NLIP (amino acids 1-114) = N-terminal lipin domain, NLS (amino acids 153-158) = nuclear localisation signal, and CLIP (amino acids 685-841) = C-terminal lipin domain, also referred to as the LNS2 (Lipin/Ned1/Smp2) domain.

3.3.3.2 *LPIN3* variation in SIR

A total of 54 *LPIN3* variants were detected in the SIR cohort (Appendix Table A6), 17 of which were present in the coding sequence (Table 3.11 and Figure 3.15). Coding sequence variants that did not alter the amino acid sequence (shown below the schematic in Figure 3.15) and/or that were also present in controls were considered unlikely to be pathogenic. This left three rare nonsynonymous variants (G41S, W110C and E539K - underlined above the schematic in Figure 3.14), two of which fell within the highly conserved N-terminal domain (residues 1-114) common to all three members of the lipin family.

Table 3.11 *LPIN3* coding variants detected in a cohort of severe insulin resistant patients and 11 Indian and 23 CEPH controls

Genic position	Genomic position	Minor|Major allele	Protein consequence	MAF in SIR	Detected in controls?	rs ID
Exon 2	39407866	T|C	Y3Y	0.0027		
Exon 2	39407896	T|G	G13G	0.0026		rs16985673
Exon 2	39407916	A|G	R20Q	0.0132	Yes	
Exon 2	39407947	T|C	G30G	0.0026		
Exon 2	39407978	A|G	**G41S**	0.0026	No	
Exon 2	39407998	G|C	P47P	0.0165		
Exon 2	39408038	T|C	R61W	0.0239	Yes	
Exon 4	39410714	C|G	**W110C**	0.0027	No	
Exon 7	39412412	C|G	V355L	0.0161	Yes	
Exon 8	39413957	A|G	A395A	0.0026		
Exon 9	39414271	T|C	L454L	0.0052		
Exon 12	39416786	A|G	**E539K**	0.0026	No	
Exon 12	39416791	T|G	E540D	0.0026	Yes	
Exon 17	39419954	A|G	L686L	0.0315		rs2072969
Exon 17	39419963	A|G	S689S	0.0026		rs41277020

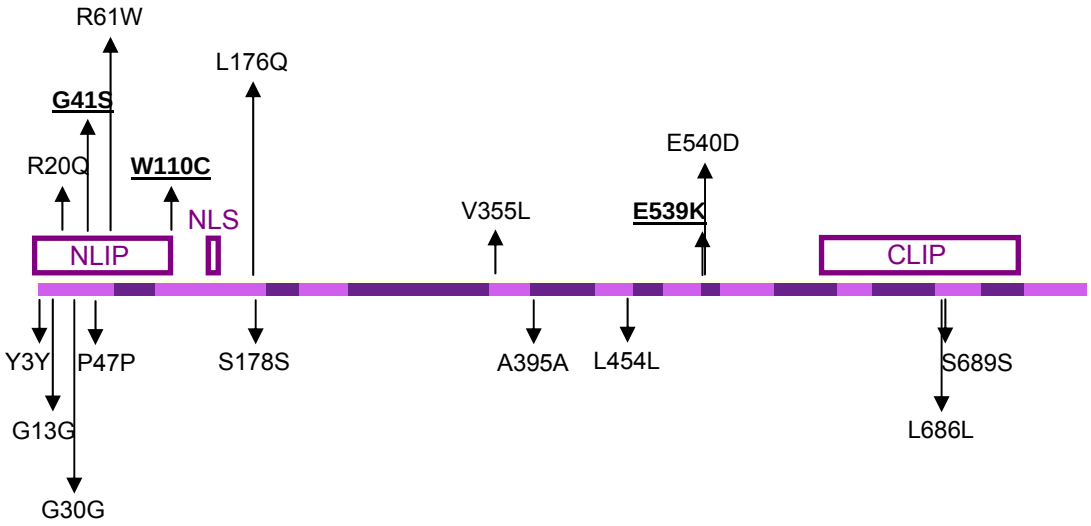

Figure 3.15 Coding *LPIN3* variants in the SIR cohort
Schematic of the lipin 3 protein showing exons in alternating bright and dark purple and known domains among lipin family proteins in boxes. Arrows indicate the location of coding SNPs detected in *LPIN3* by sequencing 23 patients with partial lipodystrophy and 135 patients with other syndromes of severe insulin resistance. Non-synonymous variants are above the schematic and synonymous variants are below the schematic. Non-synonymous mutations absent from control samples were considered potentially pathogenic (underlined). NLIP (amino acids 1-114) = N-terminal lipin domain, NLS (amino acids 141-148) = nuclear localisation signal, and CLIP (amino acids 640-796) = C-terminal lipin domain, also referred to as the LNS2 (Lipin/Ned1/Smp2) domain. HAD = haloacid dehalogenase domain.

3.3.3.3 Investigation of *LPIN3* G41S and W110C variants

I decided to prioritise those *LPIN2* and *LPIN3* variants most likely to have a functional impact upon the protein for further investigation. *LPIN3* non-synonymous variants G41S and W110C were selected as both these variants were present in single severe insulin resistant individuals, absent from controls, and mapped within a known functional domain. Both were also predicted to be highly likely to have a deleterious effect on protein function by PANTHER.

G41S was detected in an Asian female with partial lipodystrophy. The mutation was not present in 47 Indian and 47 white European controls on the CIN panel (Appendix Table 4) (data not shown). Unfortunately, family DNA was not available preventing further co-segregation analysis.

W110C was detected in an Arabic female with pseudoacromegaly and acanthosis nigricans. To investigate whether this variant was present in any unaffected individual of the same ethnic background, *LPIN3* exon 4 was sequenced in 173 Arabic controls from the HGDP panel. W110C was detected in one Druze control from Carmel, Israel, and is therefore unlikely to be pathogenic (data not shown).

3.4 Discussion

3.4.1 *LPIN1*

In this study I performed a comprehensive analysis of *LPIN1* variants and their effects on metabolic quantitative traits and syndromes of insulin resistance (including lipodystrophy). Analysis of *LPIN1* polymorphisms (MAF>0.01) in two UK population-based cohorts (N = 4610) revealed nominal significant associations with BMI, and rs13412582 remained marginally associated with BMI (P = 0.042) in a meta-analysis of UK population-based samples from in-house and publicly available genome-wide studies (N = 8504). I also detected nominal associations between my tagSNPs and metabolic traits previously reported to be associated with *LPIN1* variation (Wiedmann et al. 2007). Sequencing of 23 patients with lipodystrophy and 135 patients with syndromes of insulin resistance revealed that mutations in *LPIN1* are unlikely to be a common cause of these diseases in humans. This study has been published (Fawcett et al. 2008).

To my knowledge, neither rs13412582 nor any highly correlated SNPs have been tested in other association studies published to date. Further replication in larger cohorts will be required to confirm the BMI association.

Seven of my tagSNPs that passed quality control were directly genotyped in at least one of the other association studies (Loos et al. 2007; Suviolahti et al. 2006; Wiedmann et al. 2007). All analyses, including my own, found no association between rs4669781, rs1050800, and rs2577256 and insulin levels and measures of adiposity. However, results for the other four SNPs are inconsistent between studies. For example, rs2716610 was associated with obesity in a Finnish case-control population (Suviolahti et al. 2006). In contrast, I detected no association between rs2716609, which is in complete LD with rs2716610 (r^2 = 1) in HapMap CEU trios,

and obesity based on analysis of 1128 obese and 3601 lean individuals from Ely and Hertfordshire population-based cohorts (OR = 1.11 ± 0.08, P = 0.149).

SNP rs2716610 was also associated with BMI in lean Finnish men (Suviolahti et al. 2006). This association with quantitative measures of adiposity is supported by the Quebec Family study (Loos et al. 2007), which analysed parents (N = 335) and offspring (N = 377) from 172 French-Canadian families. Here, the highly correlated SNP rs2716609 was associated with skinfolds and waist circumference, and BMI showed the same trend. Given my sample size of 4130 individuals with full rs2716609 genotype and BMI data I had >80% power to detect the effect size observed in the Quebec Family study. Nevertheless, I did not replicate the association between rs2716609 and BMI in Ely and Hertfordshire cohorts (Table 3.3). Nor did I replicate the association with waist circumference, though the direction of the effect was consistent with the Quebec study (data not shown). My results agree with the MONICA study Augsburg (N = 1416), a German population-based cohort, which found no association between rs2716610 and BMI in men or women (Wiedmann et al. 2007).

Two other SNPs, rs893346 and rs2577262, were associated with BMI in lean Finnish men (Suviolahti et al. 2006) but showed no statistical association with BMI in 1873 lean men from Ely and Hertfordshire cohort studies (P = 0.631 and 0.253 respectively). Similarly, rs2278513 and rs2577262 were associated with BMI in obese Finnish men but not in obese men from the UK (P = 0.780 and 0.676 respectively). My data agree with the MONICA study which found no association of rs893346 and two SNPs highly correlated with rs2577262 in HapMap CEU trios (r^2 = 1.0 and 0.96 for rs6744682 and rs6708316 respectively) with BMI in men (Wiedmann et al. 2007).

In the Quebec Family study, rs2577262 and rs2577256 were associated with resting metabolic rate (RMR) in parents but not offspring. The MRC Ely study included data for mean resting energy expenditure which was not statistically

associated with rs2577262 or rs2577256 (P = 0.895 and 0.923 respectively) despite having >80% power to detect the magnitude of the effect size described previously.

The MONICA study reported strong associations between haplotypes of markers rs33997857, rs6744682 and rs6708316 and quantitative traits underlying metabolic syndrome including hypertension-, obesity-, and diabetes-related traits (Wiedmann et al. 2007). Several of the traits were also statistically associated with the same haplotypes in a replication study, but the effect was always in the opposite direction compared to the original cohort. I tested all my individual tagging SNPs as well as haplotypes of rs33997857 and rs2577262, which is highly correlated with rs6744682 and rs6708316, for association with systolic and diastolic blood pressure, HDL and LDL cholesterol, plasma triglycerides, waist circumference, HbA1c levels, and hypertension and diabetes status. I detected a number of nominal associations between these traits and individual SNPs in my study (Table 3.6, 3.7 and 3.8) but none of these associations reached statistical significance after adjustment of the P value threshold for multiple testing using the Bonferroni correction and thus require further replication.

There are several possible reasons why I could not replicate some previously reported associations between *LPIN1* variants and metabolic quantitative traits. Firstly, I may have reported false negative results. However, where effect sizes were reported in previously published studies I was able to calculate that my study had >80% power to detect them. Secondly, previous studies might have reported false positive results. In particular, as a consequence of multiple testing, detection of false positive associations becomes more likely when analyses are performed in subsets of samples and on many traits. Alternatively, the discrepancy in results between studies may be due to genetic and/or environmental differences between the populations genotyped. For example, the degree of linkage disequilibrium between *LPIN1* tag SNPs and the putative unmeasured true functional variant(s) may vary

between the cohorts. Also, *LPIN1* SNPs may interact with other genetic and/or environmental risk factors in different studies. Therefore, I cannot rule out the possibility of population-specific effects of *LPIN1* genotype on metabolic quantitative traits and hypertension, diabetes and obesity risk.

In the fld mouse model Lpin1 null mutations cause lipodystrophy, insulin resistance and peripheral neuropathy (Peterfy et al. 2001). However, of the three rare (MAF<0.01) nonsynonymous *LPIN1* variants detected within the cohort of patients with syndromes of severe insulin resistance, none are likely to be pathogenic in isolation in heterozygous form: family co-segregation analysis showed that A353T and G582R did not segregate with disease in a fully penetrant manner and G582R was also detected in one Bedouin control.

Western blotting of patient fibroblasts showed that G582R and R552K had no discernable impact on lipin 1 protein levels. Proteins orthologous to lipin 1 in yeast are proposed to be involved in nuclear membrane growth and morphology (Santos-Rosa et al. 2005; Siniossoglou et al. 1998; Tange et al. 2002). However, staining of a nuclear pore marker in patient fibroblasts with R552K and G582R variants revealed no abnormalities in membrane morphology compared to control fibroblasts.

To date, my study (N=23) and previously published work (N=15) (Cao and Hegele 2002) have demonstrated that *LPIN1* coding mutations are unlikely to be a common cause of human lipodystrophy. However, I cannot exclude the possibility that *LPIN1* mutations are rarer causes of these disorders or that rare variants in *LPIN1* interact with other genetic defects to cause severe insulin resistance. The methods I used to screen for mutations would not have detected copy number variations affecting large regions, such as whole exon deletions and duplications nor would they detect potential mutations affecting regulatory regions, therefore I cannot exclude these types of *LPIN1* variation as causes of human lipodystrophy and insulin resistance.

Furthermore, the *in vitro* assays used to assess the functional impact of *LPIN1* non-synonymous variants might have missed some functional effects, such as phosphatidic acid phosphatase (PAP) activity.

In conclusion, coding variants in *LPIN1* are not a common cause of lipodystrophy and severe insulin resistance in humans, and polymorphisms in *LPIN1* are unlikely to importantly contribute to insulin sensitivity in UK populations. SNP rs13412852 is nominally associated with BMI in UK cohorts but the effect size is very modest and requires confirmation. Nominal associations between *LPIN1* variants and blood pressure, cholesterol, triglycerides, HbA1c, and risk of hypertension need replicating in larger cohorts.

3.4.2 *LPIN2* and *LPIN3*

As far as I am aware I have performed the first screening of *LPIN2* and *LPIN3* for potentially pathogenic mutations in patients with syndromes of severe insulin resistance. Two rare, non-synonymous variants were detected in the N-terminal conserved lipin domain of *LPIN3*. One of these, W110C, was considered unlikely to be pathogenic as it was present in a Druze control. However, G41S is present in a patient with partial lipodystrophy and severe insulin resistance without family available for co-segregation analysis. The role of this mutation, if any, in the syndrome will need to be investigated further.

Two non-synonymous variants present in patients with syndromes of insulin resistance and absent from control samples were detected in *LPIN2*. However, these were not prioritised for further analysis as they were predicted to be benign changes and did not fall within any known functional domains. In future work more controls could be tested for these variants. If absent from controls and if DNA from family members is available, co-segregation analysis could be carried out.

Given that my sequencing approach would not detect large insertions and deletions, I cannot exclude the possibility that copy number variants affecting *LPIN2* and *LPIN3* impact insulin resistance. It is also plausible that combinations of rare variants in these genes cause syndromes of insulin resistance. Furthermore, I did not screen putative regulatory regions and therefore may have missed pathogenic mutations affecting *LPIN2* and *LPIN3* gene regulation.

In my study I did not evaluate the role of common variants in *LPIN2* and *LPIN3* on human metabolic traits. In this context it would be of interest to design a study to attempt replication of the recently published association between rs3745012 in the 3'UTR of *LPIN2* (Aulchenko et al. 2007) and type 2 diabetes and related metabolic traits. This SNP has not been genotyped in HapMap samples and is not present on Affymetrix or Illumina SNP chips, therefore genome-wide association study data will not provide enough data to replicate this association. Still, other SNPs within *LPIN2* and *LPIN3* are available on genome-wide arrays and may not need to be genotyped in extra cohorts.

In conclusion, coding variants in *LPIN2* and *LPIN3* are unlikely to be a common cause of lipodystrophy and severe insulin resistance in humans. A potentially pathogenic variant in *LPIN3* requires further investigation to establish whether it is causing insulin resistance and lipodystrophy in the patient.

3.5 Materials and Methods

3.5.1 Description of cohorts

Brief descriptions of cohorts used in this study follows. For more details see the corresponding sections of Chapter 2.

3.5.1.1 ELY Cohort

The Medical Research Council (MRC) Ely Study is a population-based cohort study of the aetiology and pathogenesis of type 2 diabetes and related metabolic disorders in the UK (Wareham et al. 1998). It comprises white men and women aged 35-79 years without diagnosed diabetes. Measurements of anthropometric and metabolic data analysed in this study have been described previously (Ekelund et al. 2007).

3.5.1.2 Hertfordshire Cohort

The Hertfordshire Cohort Study was recruited from the cohort of people born in Hertfordshire between 1931 and 1939. The cohort details and measurements of metabolic traits analysed in this study have been described previously (Syddall et al. 2005).

3.5.1.3 EPIC-Obesity Study

The EPIC-Obesity study is nested within the EPIC-Norfolk study, a population based cohort study of 25663 white European men and women aged 39-79 recruited in Norfolk, UK between 1993 and 1997 (Day et al. 1999). Height and weight were measured using standard anthropometric techniques (Day et al. 1999). For *LPIN1*, I analysed BMI data from a random subset of the entire EPIC-Norfolk study (N = 2415) with genome-wide association study data (see 3.5.4).

3.5.1.4 HGDP-CEPH

The HGDP-CEPH Human Genome Diversity Cell Line Panel is a resource of 1064 DNA samples from individuals distributed around the world and has been described previously (Cann et al. 2002).

3.5.1.5 CEPH

48 unrelated individuals from CEPH families supplied by Coriell Cell Repositories (Dausset et al. 1990) are control individuals of North and West European origin.

3.5.1.6 CIN

This panel includes DNA samples from the European Collection of Cell Cultures (ECACC), including 47 white European individuals and 47 individuals of Asian Indian origin.

3.5.1.7 Severe Insulin Resistance Cohort

All patients had severe insulin resistance, defined as fasting insulin > 150 pmol/l, or peak insulin on oral glucose tolerance testing > 1,500 pmol/l in non-diabetic patients. In complete insulin deficiency it was defined as an insulin requirement above 3U/kg/day. Acanthosis nigricans was also used as a marker of insulin resistance. Most patients had a BMI <30 kg/m^2 and at least 58 had BMI>30.

3.5.2 PCR and sequencing

Genomic DNA from 23 patients with lipodystophy, 135 patients with other syndromes of severe insulin resistance and, in the case of *LPIN2* and *LPIN3* sequencing, 11 Indian controls from the CIN panel and 23 CEPH controls, was whole-genome amplified (Chapter 2.3.1.1) prior to amplification with gene-specific primers covering all coding exons, splice junctions, and 3'UTR (see Appendix Table A7 for sequences and cycling conditions).

PCR was performed using standard conditions (Chapter 2.3.2) and products purified using exonuclease I and shrimp alkaline phosphatase (USB Corporation, Cleveland, OH, USA).

Bi-directional sequencing was performed using the Big Dye Terminator 3.1 kit (Applied Biosystems, Foster City, CA, USA). Sequencing reactions were run on ABI3730 capillary machines (Applied Biosystems) and analysed using Mutation Surveyor version.2.20 (SoftGenetics LLC, State College, PA, USA) (Chapter 2.3.5). All non-synonymous variants with MAF<0.01 were confirmed in a second PCR and sequencing reaction using patient genomic DNA. DNA from family members used for co-segregation analysis was genomic.

3.5.3 Tagging SNP selection

Tagging SNPs were selected from 52 known and novel *LPIN1* SNPs with a pairwise $r^2 \geq 0.8$ using Tagger (de Bakker et al. 2005) as a stand-alone program in Haploview (Barrett et al. 2005). Twenty-five SNPs were selected as tags and one was force included, giving a total of 26 SNPs for association testing.

3.5.4 Genotyping

Genotyping was performed by Susannah Bumpstead and Andrew Keniry in the genotyping facility within the Genetics of Complex Traits in Humans team at the Wellcome Trust Sanger Institute. *LPIN1* mutations A353T, R552K, and G582R were genotyped on the HGDP-CEPH Human Genome Diversity panel as stand-alone assays using the Sequenom MassArray hME platform according to the manufacturer's instructions (see Chapter 2 for details and Appendix Table A8 for primer sequences). Twenty-six SNPs were selected for genotyping in the MRC Ely population-based cohort (N = 1709) and the Hertfordshire cohort study (N = 2901) using the Sequenom MassArray iPLEX platform (Chapter 2.3.7.1.2). One SNP, rs13093930, failed Sequenom assay design but because it was a rare intronic SNP

that only tagged itself a replacement could not be designed. This left 25 tagging SNPs (Figure 3.5 red boxes). Twenty-three of these tagging SNPs passed pre-screening and all except the monomorphic loci Novel02 and rs17603755 (which did not pass HWE) were analysed for association with metabolic traits (Table 3.2). Primers, probes and conditions for the remaining twenty-two *LPIN1* tagging SNPs are presented in Appendix Table A8. All genotyped SNPs were tested for deviation from Hardy-Weinberg Equilibrium ($P > 0.01$) and call rates > 85%.

Genotyping of the EPIC-Obesity study using the Affymetrix GeneChip Human Mapping 500K Array Set has been described (Loos et al. 2008). This study contained genotype information for five of my *LPIN1* tagSNPs, each of which had a call rate >90%. In total, 2415 individuals with height and weight measures and quality-controlled genotype data were available for analyses.

3.5.5 Statistical analysis

Deviation of *LPIN1* tagSNP genotype from Hardy-Weinberg equilibrium was assessed using a goodness-of-fit χ^2 test. Linear regression analysis was used to assess the association between individual SNPs and BMI, log-transformed fasting plasma insulin, and log-transformed additional metabolic traits (systolic and diastolic blood pressure, HDL and LDL cholesterol, plasma triglycerides, waist circumference, and HbA1c levels) in Ely and Hertfordshire cohorts using Stata v9 (Stata Corporation, Texas, USA). All analyses were adjusted for age and sex and, in the case of the joint analysis, study. Logistic regression in Stata was used to test for association between *LPIN1* SNPs and risk of hypertension and diabetes. Chi-squared analysis was performed to test for significant differences ($P<0.01$) in call rate between cases and controls. Where nominally significant values were found in the Ely and Hertfordshire joint analysis, I performed 10,000 permutations of the dependent variable to test for empirical significance. This was performed in Stata.

The joint Ely and Hertfordshire cohort analysis of additional traits underlying metabolic syndrome comprised 189 tests so the *P* value threshold adjusted for multiple testing using the Bonferroni correction is 0.0003.

Fixed effect meta-analysis using the inverse variance method was performed by using the metan command in Stata (Bradburn et al. 1999). Heterogeneity among studies was assessed using the Q statistic. IMPUTE software (http://www.stats.ox.ac.uk/~marchini/software/gwas/impute.html) was used to impute genotypes for rs17603420 in the EPIC cohort.

Plink (http://pngu.mgh.harvard.edu/~purcell/plink/) was used to perform hapotype analysis (Purcell et al. 2007), and Ely and Hertfordshire cohorts were meta-analysed using METAL (http://www.sph.umich.edu/csg/abecasis/metal/index.html).

Power calculations were performed using Quanto v1.1.1. I calculated that we have >80% power to detect a per allele effect on BMI of >1.33 kg/m2 with MAF=0.01, and >0.27kg/m2 with MAF=0.5. For logged fasting insulin data this range is >1.04 to >1.22.

3.5.6 Western blotting

The following analyses of patient fibroblasts carrying the *LPIN1* mutations R552K and G582R mutations were performed by Neil Grimsey and Symeon Siniossoglou at the Cambridge Institute for Medical Research.

Patient fibroblast cells were maintained in DMEM medium supplemented with 10% fetal bovine serum (FBS) and 2mM L-glutamine in a humidified 37°C incubator with 5% CO_2. All cells were routinely assessed for and protected against mycoplasma

infection using VenorGeM® mycoplasma detection kit (Minerva biolabs, CamBio, VGM-025) and BM-cyclin (Roche; 799050) respectively.

Fibroblasts were collected by trypsin EDTA release, washed with PBS and then lysed in a 50mM HEPES pH7.4 buffer containing 150mM NaCl, 1% Triton X-100, 100µM AEBSF, Protease inhibitor cocktail 1, Phosphatase inhibitor cocktail II, 1µg/ml Dnase and 4mM MgCl2 chilled to 4°C. Each sample was homogensised by passing through a 25G needle 10 times. Insoluble debris was removed by a 16,000g centrifugation step at 4°C.

Sample protein concentration was measured by a comparative Bradford protein assay. Samples were then suspended in 1 x SDS sample buffer and boiled at 95°C for 5 minutes, loaded onto 7% SDS-PAGE, transferred onto nitrocellulose, and blocked in PBS with 1% TX-100 and 5% milk. These were then probed with protein specific primary antibodies: Lipin 1, Lipin 2, anti-Mab414 (Covance; MMS-120P), anti-laminB (Santa-Cruz; sc-6217). Lipin 1 and 2 polyclonal antibody production will be described elsewhere (Grimsey and Siniossoglou, in preparation). This was followed by species specific secondary antibodies coupled to Horse Radish Peroxidase (HRP): anti-rabbit IgG (Jackson immuno research; 211-032-171), anti-goat IgG (Novus Biologicals; NB 710-H), anti-mouse IgG (H & L) highly cross-adsorbed (Molecular probes; A11029). Proteins were then detected using standard electrochemiluminescence techniques (Amersham ECL-reagents).

3.5.7 Indirect immunofluorescence by confocal microscopy

The following analyses were performed by Neil Grimsey and Symeon Siniossoglou at the Cambridge Institute for Medical Research.

Fibroblasts were fixed with 3% Formaldehyde, permiablised with 0.1% Triton X-100, and blocked with 1mg/ml BSA in PBS. Each cover slip was labelled with primary mouse $-\alpha$Mab414 (nuclear pore marker) and secondary anti-mouse conjugated to FITC (green), primary rabbit $-\alpha$Calreticulin (Endoplasmic reticulum calcium binding protein) (Calbiochem; 208910), and secondary anti-rabbit conjugated to Alexa fluor 594 (red) (Molecular probes; A11037), DNA was stained with DAPI (blue). Each slide was mounted onto glass slides and then visualised with a 63X or 100X Plan Apochromat objective (numerical aperture,1.4) on a Ziess Axiovert 200M inverted microscope with an LSM 510 confocal laser Scanning attachment.