The genetics of type 2 diabetes

Reiling, H.W.

Chapter 2

Genetic association analysis of *LARS2* with type 2 diabetes

E. Reiling[1*], B. Jafar-Mohammadi[2,3*], E. van 't Riet[4,5], M.N. Weedon[6,7], J.V. van Vliet-Ostaptchouk[8], T. Hansen[9,24], R. Saxena[10], T.W. van Haeften[12], P.A. Arp[13], S. Das[2], G. Nijpels[4,14], M.J. Groenewoud[1], E.C. van Hove[1], A.G Uitterlinden[13], J.W.A. Smit[15], A.D. Morris[23], A.S.F. Doney[23], C.N.A. Palmer[23], C. Guiducci [10], A.T. Hattersley[6,7], T.M. Frayling[6,7], O. Pedersen[9,16,17], P.E. Slagboom[11], D.M. Altshuler[10,18,19] L. Groop[20,21], J.A. Romijn[15], J.A. Maassen[1,5], M.H. Hofker[8], J.M. Dekker[4,5], M.I. McCarthy[2,3,22], Leen M. 't Hart[1]

1. Leiden University Medical Center, Department of Molecular Cell Biology, Leiden, the Netherlands
2. Oxford Center for Diabetes, Endocrinology and Metabolism, University of Oxford, Churchill Hospital, Oxford, OX3 7LJ, United Kingdom
3. National Institute for Health Research, Oxford Biomedical Research Centre, University of Oxford, Old Road, Headington, Oxford OX3 7LJ, United Kingdom
4. VU University Medical Center, EMGO Institute for Health and Care Research, Amsterdam, the Netherlands
5. VU University Medical Center, Department of Epidemiology and Biostatistics, Amsterdam, the Netherlands
6. Genetics of Complex Traits, Institute of Biomedical and Clinical Science, Peninsula Medical School, Exeter, United Kingdom.
7. Diabetes Genetics Institute of Biomedical and Clinical Science, Peninsula Medical School, Exeter, United Kingdom
8. Molecular Genetics, Medical Biology Section, Dept. of Pathology & Medical Biology, University Medical Centre Groningen and University of Groningen, Groningen, the Netherlands
9. Steno Diabetes Center and Hagedorn Research Institute, Gentofte, Denmark
10. Program in Medical and Population Genetics, Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts
11. Leiden University Medical Center, Department of Molecular Epidemiology, the Netherlands
12. University Medical Center Utrecht, Department of Internal Medicine, Utrecht, the Netherlands
13. Erasmus University Medical Center, Department of Internal Medicine, Rotterdam, the Netherlands
14. Department of General Practice, VU University Medical Center, Amsterdam, the Netherlands
15. Leiden University Medical Center, Department of Endocrinology, the Netherlands
16. Aarhus University, Faculty of Health Science, Aarhus, Denmark
17. University of Copenhagen, Faculty of Health Science, Copenhagen, Denmark
18. Center for Human Genetic Research and Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts
19. Department of Genetics, Harvard Medical School, Boston, Massachusetts
20. Department of Clinical Sciences, University Hospital Malmö, Clinical Research Center, Lund University, Malmö, Sweden
21. Department of Medicine, Helsinki University Hospital, University of Helsinki, Helsinki, Finland
22. Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, United Kingdom
23. Diabetes Research Group, Ninewells Hospital and Medical School, University of Dundee, Dundee, United Kingdom
24. University of Southern Denmark, Faculty of Health Science, Denmark

* Both authors contributed equally to this work

Diabetologia 2010 Jan 53 (1): 103-111

Abstract

Aims/hypothesis

LARS2 has been previously identified as a potential type 2 diabetes susceptibility gene through the low frequency (LF) H324Q (rs71645922) variant (MAF 3.0%). However, this association did not achieve genome-wide levels of significance. The aim of this study was to establish the true contribution of this variant and common variants in *LARS2* (MAF >5%) to type 2 diabetes risk.

Methods

We combined genome-wide association study (GWAS) data (n=10128) from the DIAGRAM consortium, with independent data derived from a tagging SNP approach in Dutch individuals (n=999), and took forward two SNPs of interest for replication in up to 11163 Dutch subjects. In addition, because inspection of the GWAS data identified a cluster of LF variants with evidence of type 2 diabetes association, we attempted replication of rs9825041 (a proxy for this group) and the previously identified H324Q variant in up to 35715 subjects of European descent.

Results

No association of the common SNPs in *LARS2* with type 2 diabetes was found. Our replication studies for the 2 LF variants, rs9825041 and H324Q failed to confirm an association with type 2 diabetes in Dutch, Scandinavian and UK samples (OR 1.03 (0.95-1.12), p=0.45, n=31962 and OR 0.99 (0.90-1.08), p=0.78, n=35715 respectively).

Conclusion

In this study, the largest study examining the association of sequence variants in *LARS2* with type 2 diabetes susceptibility we find no evidence to support previous data indicating a role in type 2 diabetes susceptibility.

Introduction

Changes in mitochondrial function are observed in patients with type 2 diabetes and their first degree relatives. Previous studies have indicated that genes involved in oxidative phosphorylation are down regulated in the muscle cells of type 2 diabetes patients (1). Furthermore, the muscle mitochondria from patients with type 2 diabetes have an impaired bioenergetic capacity (2). Mitochondria also play an important role in insulin secretion and sensitivity (3;4). Previously, our group has shown that a mutation in the mitochondrial DNA encoded tRNA-Leu(UUR) gene is associated with maternally inherited diabetes and deafness (5). In addition, an H324Q (rs71645922) variant in the nuclear encoded mitochondrial *LARS2* gene has shown an association with type 2 diabetes in work previously carried out by our group (6). The *LARS2* gene encodes for the mitochondrial leucyl tRNA synthetase (EC 6.1.1.4), which catalyzes the aminoacylation of both mitochondrial leucyl tRNAs with leucine and is therefore essential for mitochondrial protein synthesis. By analyzing the coding region for the *LARS2* gene we found an H324Q (rs71645922) variant, and demonstrated an association with type 2 diabetes susceptibility in a meta-analysis of four independent cohorts from the Netherlands and Denmark (OR = 1.40 (95% CI 1.12-1.76, P = 0.004, n = 7836) (6).

In recent years the advent of genome-wide association studies (GWAS) and the accumulation of large data sets capable of detecting associations to levels of genome wide significance appropriate for such studies ($p<5*10^{-8}$) has identified close to 20 loci impacting on type 2 diabetes susceptibility. However, low frequency variants such as H324Q are generally poorly captured by such studies. We set out therefore to re-evaluate the possible contribution of this LF variant to type 2 diabetes susceptibility in appropriately sized samples. We also used a combination of publicly available (DIAGRAM consortium) and newly derived tagging SNP data to undertake the most comprehensive assessment of the *LARS2* locus yet performed.

Materials and methods

Study samples

The first part of our study was aimed at the identification of common alleles associated with increased type 2 diabetes susceptibility using DIAGRAM consortium data and a tagging SNP approach. For this we genotyped several European samples.

The first sample we included was from the Hoorn Study (here designated as NL1) (7). From this Dutch population based study from the city of Hoorn, in North-Western Netherlands, we selected 519 normal glucose tolerant (NGT) subjects and 480 type 2 diabetes subjects. Glucose tolerance was assessed using a fasting oral glucose tolerance test (OGTT), according to 1999 World Health Organization (WHO) criteria (8). This sample was used for the analysis of common variation in *LARS2* with a tagging SNP approach. Variants in *LARS2* identified from the DIAGRAM meta-analysis and the tagging SNP approach were then taken forward for replication in three Dutch samples, designated NL2, 3 and 4 respectively.

The second sample from the Netherlands (designated NL2) included 1517 controls and 821 cases (9;10). The1517 controls were randomly selected from the New Hoorn Study (NHS), which is an ongoing, population based study from the city of Hoorn, which does not overlap with the original Hoorn Study (NL1). We included 147 cases from the NHS and the remainder of the cases (n=674) were recruited from the diabetes clinics of the Leiden University Medical Centre (LUMC, Leiden) and from the Vrije Universiteit medical centre (VUmc, Amsterdam). All subjects in this replication sample were Dutch Caucasians and all NGT subjects underwent an OGTT according to WHO criteria(8).

The third replication sample was ascertained from the Breda study (NL3) (11;12). This is a case control study from the city of Breda, in Southern Netherlands. The 920 controls were from the Dutch blood bank and self reported a non-diabetic state. The 501 cases had type 2 diabetes based on WHO criteria (8).

For the fourth replication sample we selected 5183 NGT subjects and 1222 type 2 diabetes subjects from the population based ERGO study from the city of Rotterdam in the South-western region of the Netherlands (NL4) (13).

In total 8139 controls and 3024 type 2 diabetes cases were included in our replication study in the Netherlands.

The second part of this study was focused on the follow up of 2 low frequency variants in *LARS2* and for this we carried out replication in samples from the Netherlands (NL1-4) as well as samples from the UK (UK1,2), Denmark (DK1), Finland (FI1,2) and Sweden (SE1).

We included one replication sample from Denmark (designated DK1) (14). This sample consists of 514 NGT controls which are randomly selected from public registers at the Steno Diabetes Center and the Research Centre for Prevention and Health, Copenhagen, Denmark. The 706 cases were recruited from the Steno Diabetes Center. NGT subjects underwent an OGTT according to WHO criteria (8).

Two UK samples were included. The first (UK1) was the UKT2DGC (United Kingdom Type 2 Diabetes Genetics Consortium) case-control sample comprising 4124 type 2 diabetes cases and 5126 controls ascertained in Tayside, Scotland. Details of the ascertainment scheme and recruitment criteria for this sample have been described elsewhere (15;16): the enlarged sample used here represents continuing recruitment to this resource under precisely the same criteria. The second sample (UK2) consists of 1853 type 2 diabetes cases ascertained as part of the BDA Warren 2 collection (Exeter, London, Oxford, Norwich and Newcastle) and 10220 control samples. The latter represent the full British 1958 Birth Cohort (n=7133) and The United Kingdom Blood Services Collection of Common Controls (UKBS) (n= 3087), a subset of which featured in the WTCCC genome wide association scan (both samples were collected throughout the UK) (15;16).

Finally, we included samples from Finland and Sweden. One was a case-control sample from the Botnia region of Finland, here designated as FI1. This sample consisted of 353 controls and 402 cases. The second sample originated from Sweden (Skara and Malmö), here designated as the Swedish case-control study (SE1) and consisted of 468 controls and 480 cases (17;18). Furthermore, we included a set of trios originating from the Botnia region of Finland (FI2). This sample consisted of 211 probands (multiple diabetic sibs) and 370 parents. All study samples are summarized in table 1.

In total 25191 controls and 10800 type 2 diabetes cases were included for the follow up of the LF variants.

All studies were approved by the appropriate medical ethical committees and were in accordance with the principles of the Declaration of Helsinki. All participants provided written, informed consent for this study.

Common SNP selection

Common SNPs (MAF >5%) in the *LARS2* locus were selected for follow-up based on the data from the DIAGRAM meta-analysis (gene boundaries chr3: 45373001…45698001) (22). SNPs with a P<0.05 were genotyped in the Dutch replication samples (NL1-4). Furthermore, tagging SNPs in *LARS2* were selected for genotyping in the NL1 sample using the HapMap database and Tagger software (19;20) (selection criteria and SNPs shown in supplementary table S1).

Table 1. Description of study samples.

Study	Subjects (%male)		Mean Age (years)(SD)		Mean BMI (Kg/m^2)(SD)	
	Controls	cases	controls	cases	controls	cases
NL1	519 (55)	480 (52)	65 (8)	67 (8)	26.4 (4.5)	28.8 (4.6)
NL2	1517 (44)	821 (50)	53 (7)	61 (11)	25.5 (3.6)	29.0 (4.6)
NL3	920 (61)	501 (46)	48 (13)	71 (10)	n.a.	27.8 (4.1)
NL4	5183 (41)	1222 (39)	69 (9)	73 (9)	26.0 (3.9)	27.4 (4.0)
DK1	514 (46)	706 (48)	57 (10)	59 (10)	25.9 (3.8)	29.3 (5.1)
UK1	5126 (51)	4124 (55)	60 (13)	66 (6)	26.9 (11.4)	31.2 (13.8)
UK2	10220 (50)	1853 (61)	42 (7)	57 (9)	27.2 (6.4)[a]	31.8 (6.7)
FI1	353 (53)	402 (55)	60 (10)	61 (10)	26.1 (3.6)	28.7 (4.5)
FI2	370 (50)[b]	211 (47)[c]	n.a	40 (9)	28.5 (5.5)	n.a.
SE1	468 (52)	480 (53)	66 (12)	67 (11)	27.5 (4.1)	27.9 (4.1)

n.a. not available

a. Based on the British 1958 Birth Cohort (7133) and Panel 2 of the United Kingdom Blood Services Collection of Common Controls (n=1643)

b. parents

c. probands

Genotyping and quality control

SNPs selected for follow-up in our replication samples were genotyped using Taqman SNP genotyping assays (Applied Biosystems, Foster City, USA). Tagging SNPs were genotyped in the NL1 sample using the Sequenom platform (Sequenom, San Diego, USA). Assays showing overlapping clusters, success rates below 95% or not obeying Hardy Weinberg Equilibrium (HWE) ($p<0.05$) were excluded from analysis. Duplicate samples (~5%) showed complete concordance.

Statistical analysis

Differences in genotype distribution and allele frequencies were analyzed using a chi-squared test. ORs were calculated using an additive model, which was the best fit for the data. Homogeneity of ORs between the different samples was calculated with a Tarone's test after which a common OR was calculated with a Mantel-Haenszel test using a fixed effects model. Results from OGTT (only normal glucose tolerant subjects) were analyzed with univariate analysis of variance, using additive, recessive and dominant models and correction for age, BMI and gender as possible confounders. Association in the Botnia trios was assessed by the transmission disequilibrium test (TDT). All general statistics were calculated using SPSS 16.0 (SPSS Inc, Chicago, USA). For statistics involving the geographical distribution of the H324Q (rs71645922) variant in the UK population (described below) we used StatXact v 6.0 (Cytel software corps, Cambridge, MA, USA). Power calculations were performed using Quanto (21). From the DIAGRAM consortium meta-analysis of common variants we selected for replication all common SNPs with a $p<0.05$. At this alpha the DIAGRAM consortium meta-analysis had at least 80% power to detect a variant with OR ≥1.20 (MAF >0.05) (22). Combined with our Dutch replication sample we had at least 80% power to replicate the association of a variant with an OR ≥1.09 at the observed MAFs of 0.19 (rs952621) and 0.24 (rs17637703) respectively (alpha = 0.05) or OR ≥1.12 at alpha = 10^{-4}). Power of the tagging SNP approach in NL1 was limited (80% power to detect a variant with an OR≥1.6 (alpha=0.05, MAF=0.05) or OR≥1.45 at the observed lowest MAF of 0.10) therefore we replicated in NL2-4 only our strongest signal from the NL1 sample (rs17637703, $p = 0.07$).

Whilst extensive GWAS have indicated that the effect sizes of common variants influencing type 2 diabetes risk are modest, the potential remains for low-frequency variants to have effects on type 2 diabetes risk that are more substantial which was corroborated by our previous observation regarding the H324Q variant (6). Power calculations at the start of the project demonstrated that we had at least 99% power to detect an effect size similar to our initial finding concerning H324Q (rs71645922) (OR 1.4) and at least 80% power to detect an OR of 1.13 (α=0.05) (6). From the DIAGRAM meta-analysis we used an alpha of 0.05 to select other LF SNPs for replication. At this alpha the power in DIAGRAM was 80% to detect association for variants with ORs ranging from 1.24 (MAF = 0.03) to 1.45 (MAF = 0.01 and alpha = 0.05). For replication of the two low frequency variants (observed MAFs ~0.03(H324Q, rs71645922) and ~0.05 (rs9825041) respectively) we had in our complete replication sample at least 80% power to detect an OR ≥1.13 (25191 controls and 10800 type 2 diabetes cases and alpha = 0.05).

Results

Common LARS2 variants in available DIAGRAM *GWAS data*

We analyzed the data from the DIAGRAM GWAS meta-analysis (22) for the *LARS2* gene (100% coverage (MAF > 5%), according to HapMap phase 2, April 2007, CEU population) and observed 1 common SNP (rs952621, directly typed) showing weak evidence of association with type 2 diabetes (OR = 1.11 (1.02 – 1.20), p = 0.01 for the T allele). This SNP was also captured in our complementary tagging SNP approach (NL1) and we found an OR of 1.13 (0.89 – 1.43), p = 0.33 for the same allele. However, additional genotyping in the Dutch samples (NL2,4) and meta-analysis of all data resulted in a common OR of 1.05 (0.99 – 1.11), p = 0.13 (n=19870). As there was no convincing evidence of association in our samples, this SNP was not analysed further.

No other common SNP in *LARS2* showed evidence for association with type 2 diabetes in the GWAS data. The same was true of the tagging SNP analysis conducted in the NL1 sample (supplementary table s1). In the latter analysis, rs17637703 showed weak evidence of association (OR = 1.22 (95% C.I. 0.99 – 1.50), p = 0.07) but this was not confirmed in the Dutch replication samples (the

common OR was 0.98 (0.91 – 1.06), p = 0.62 (n = 10087), in line with the DIAGRAM result for this SNP (OR 1.02 (0.94 – 1.10)).

Low frequency variants in LARS2

In addition to the common variants, the DIAGRAM meta-analysis also captured fourteen LF SNPs (0.01<MAF<0.05) within the *LARS2* gene, ten of which are in high LD with each other (r^2 > 0.95 according to HapMap, supplementary figure S1) and showed some evidence for association with type 2 diabetes (ORs 1.17 – 1.21; p 0.02 – 0.05). We selected rs9825041 (OR = 1.20 (1.03 – 1.39), p = 0.02) as a proxy for the group for genotyping in the replication samples but no association with type 2 diabetes was observed (table 2). Homogeneity of ORs was tested with a Tarone's test (p = 0.67) and we calculated a common OR across all studies of 1.03 (0.95 – 1.12), p = 0.45, (8959 cases, 23003 controls).

Follow up of the H324Q (rs71645922) variant in LARS2

Finally, we examined the association of the H324Q (rs71645922) variant with type 2 diabetes in our replication samples from the Netherlands (NL2-4), UK (UK1, 2), Sweden (SE1) and Finland (FI1). This variant was not captured by the GWAS and was not captured by any of the SNPs mentioned above (r^2 < 0.17). Subjects in the Dutch replication samples that were included in our original study of this variant were excluded from analysis (n = 914 from the NL4 study). The replication samples did not confirm our previously observed association. A meta-analysis of all available studies including our previous data from the Netherlands (NL1) and Denmark (DK1, 2) (6) resulted in an overall OR of 0.99 (0.90 – 1.08), p = 0.78, n = 35715 subjects (10399 type 2 diabetes subjects, table 2). In addition, we did not observe a significant excess of transmission of the risk allele in the Botnia trios (FI2, transmitted / untransmitted = 18 / 14, OR 1.29 (0.64 – 2.59), p = 0.48).
To investigate possible heterogeneity between the studies we performed several analyses. For age stratification we created, based on the age distribution in the Dutch samples the following age strata; ≤60 years, 61-70 years and >70 years. A decreased frequency of the risk allele was observed in type 2 diabetes subjects with increasing age in most but not all samples (data not shown, available on

request), but this did not reach statistical significance. Furthermore we looked at age at diagnosis of type 2 diabetes and allele frequencies in those with early onset diabetes (≤45 years) and those with an age at diagnosis above 45 years. Although the allele frequency was slightly higher in those with early onset diabetes this was not statistically significant nor was the age at diagnosis in carriers and non-carriers (all $p > 0.05$, data not shown). Stratification for gender and BMI (where data available) did not affect the outcome in the Dutch studies (NL1-NL4, data not shown) and was therefore not further investigated.

H324Q (rs71645922) shows marked variation in MAF between the various European-descent samples examined (control MAF ranges from 1.9% to 4.8%). In the two large UK control samples, for example, there was a highly significant ($p = 5 \cdot 10^{-7}$, using an exact implementation of the Cochran Armitage trend test) difference in allele frequencies between UK1 (recruited exclusively in Scotland) and UK2 (recruited throughout the UK) that made us consider the possibility that this variant was showing variation in allele frequency along the south-north cline previously described in the WTCCC study and others (16;23-26).To test this, we made use of information on the region of ascertainment available for the UK 1958 Birth Cohort and UK Blood Service and analysed genotype frequencies based on subdivisions of the UK into 4 major regions, namely (1) Scotland, (2) Northern England(Yorkshire and the Humber, North East, North West), (3) UK Midlands (East Midlands, West Midlands, Wales, East of England) and (4) South of England (South East, Greater London, and South West). We found some evidence (supplementary figure S2) for a North - South gradient across the UK. (MAF 4.66 %, 3.41%, 3.31%, and 3.28% respectively) with $p = 0.038$ calculated using the Jonckheere-Terpstra Test (StatXact v 6.0; Cytel software corporation, Cambridge, MA, USA). No such MAF gradient was observed in other European samples (supplementary figure S3)

Table 2. Genotyping results for rs9825041 and H324Q (rs71645922).

Study	rs9825041				H324Q (rs71645922)			
	Controls	Cases	OR 95% CI	P_{Add}	Controls	Cases	OR 95% CI	P_{Add}
NL1	5.5	7.3	1.35 (0.92 – 1.98)	0.15	1.9	4.3	2.26 (1.10 – 4.66)[a]	0.04
NL2	6.0	6.6	1.11 (0.85 – 1.44)	0.46	3.4	2.7	0.78 (0.49 – 1.25)	0.18
NL3	5.7	6.8	1.21 (0.88 – 1.66)	0.25	3.3	3.0	0.90 (0.57 – 1.41)	0.74
NL4	6.1	5.5	0.89 (0.74 – 1.08)	0.27	3.1	3.5	1.11 (0.84 – 1.47)[a]	0.47
DK1	4.2	4.7	1.12 (0.75 – 1.68)	0.62	2.2	2.8	1.24 (0.74 – 2.09)[a]	0.44
DK2	n.m.	n.m.	n.m.	n.m.	2.8	3.7	1.33 (0.97 – 1.82)[a]	0.08
UK1	4.9	4.9	1.00 (0.87 – 1.14)	1.00	4.8	4.4	0.91 (0.79 – 1.05)	0.19
UK2	4.8	5.1	1.05 (0.88 – 1.27)	0.57	3.6	3.4	0.94 (0.77 – 1.15)	0.59
FI1	n.m.	n.m.	n.m.	n.m.	4.2	4.9	1.18 (0.72 – 1.93)	0.51
SE1	n.m.	n.m.	n.m.	n.m.	4.5	4.2	0.94 (0.59 – 1.48)	0.79
M-A			1.03 (0.95 – 1.12)	0.48			0.99 (0.90 – 1.08)	0.82
DIAGRAM[b]			1.20 (1.03 – 1.39)	0.02			n.m.	n.m.

Minor allele frequencies are shown for cases and controls.
n.m.: not measured
M-A: meta-analysis
a. Data taken from reference (6). Results from the NL4 study are partially from this previous research (n = 914).
b. Meta analysis of the DIAGRAM consortium GWAS, n = 10128 (4549 cases and 5579 controls) (22).

Discussion

We found no evidence of common SNPs in *LARS2* being associated with type 2 diabetes in our samples. We therefore conclude that it is unlikely that common SNPs in *LARS2* are associated with type 2 diabetes susceptibility.

Several low frequency SNPs, which are all in high LD with each other ($r^2 > 0.95$) showed nominal evidence of association with type 2 diabetes in the DIAGRAM meta-analysis. However, we have been unable to confirm this association in our large replication samples from the Netherlands, Denmark and the UK (n = 31962) and therefore conclude that the nominal p-values observed in the GWAS are most likely consistent with statistical noise.

The previously observed association of the H324Q (rs71645922) variant in *LARS2* with type 2 diabetes was not confirmed in our replication samples (table 2). Power

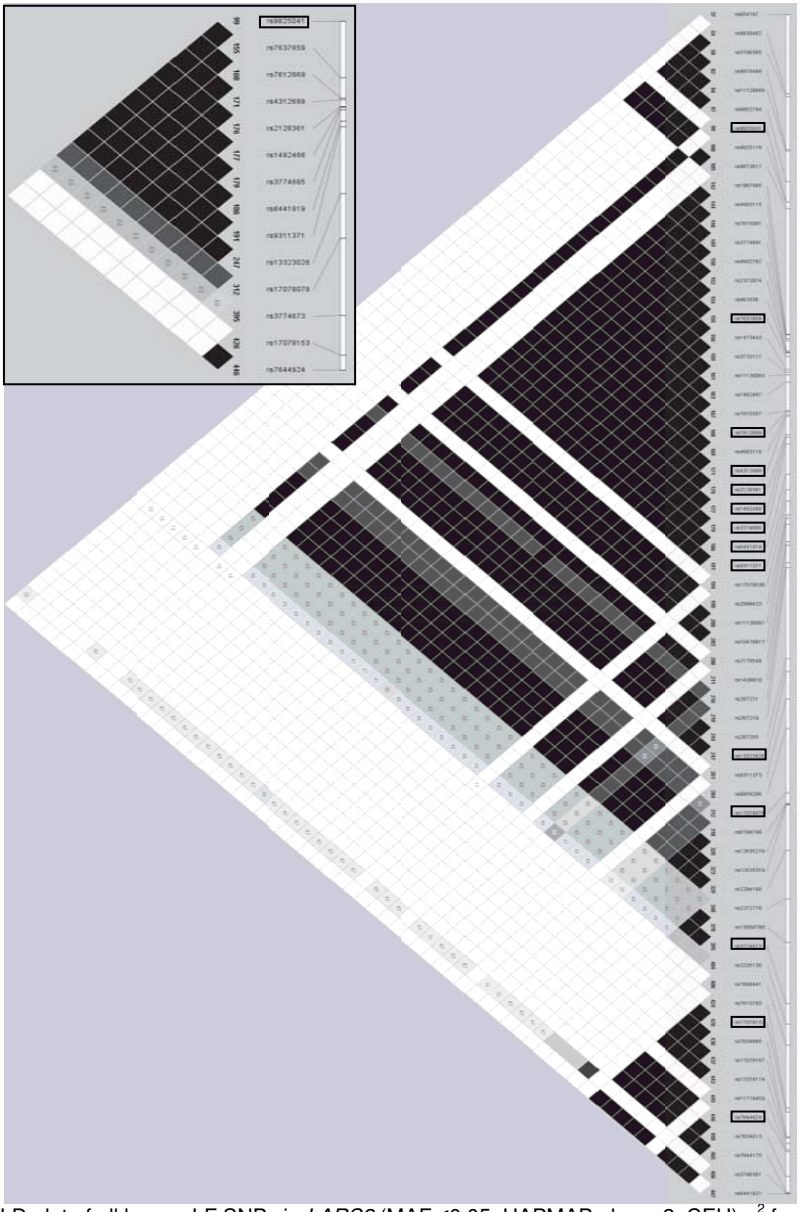
calculations at the start of the project demonstrated that we had at least 99% power to detect an effect size similar to our initial finding concerning H324Q (rs71645922) (OR 1.4) and at least 80% power to detect an OR of 1.13. Before excluding the previous association as false, we considered the possibility of heterogeneity but found no evidence that age, age at diagnosis, BMI and gender were responsible. Another possibility, raised by the evidence for variation in H324Q MAF across the UK is that the previous association in Dutch and Danish subjects reflected the effects of hidden population structure. It seems, however, unlikely that population stratification effects were responsible for the original reports of H324Q (rs71645922) associations as the cases and controls in that study were recruited from the same relatively narrow geographic regions within the Netherlands and Denmark (6). Also in our additional Dutch replication cohorts we could not detect a MAF gradient across the country. However, there are differences in MAF between different countries (supplementary Figure S2, S3). Migration patterns in the UK appear to reflect an increase in the MAF of H324Q (rs71645922) and therefore this may be a potential migration marker, however this needs to be demonstrated in other populations. Since stratification for BMI and gender did not affect our result, we can exclude that these variables confounded our observation. The reason for the discrepancy between our first and current study is likely to reflect chance. Three other low frequency non-synonymous SNPs are present in the *LARS2* locus; K727N (rs36054230), E831D (rs9827689) and E868K (rs34965084). However, according to dbSNP and our own sequencing efforts (6) these SNPs are only identified in the African population and not polymorphic in the European population. Therefore, these additional non-synonymous variants were not analysed in this study. As our study does not include a thorough resequencing of the complete *LARS2* locus we cannot fully exclude that other yet unknown LF variants are present and associated with type 2 diabetes. Results from the 1000 genomes project should facilitate a thorough investigation of LF SNPs in *LARS2* in the future. In conclusion, our findings do not support the hypothesis that common variants in *LARS2* are major type 2 diabetes susceptibility factors. We have also conducted one of the largest (up to 35715 subjects) replication studies for two low frequency variants for type 2 diabetes and also these data do not support a significant role as

type 2 diabetes susceptibility variants. We therefore conclude that currently known genetic variation in *LARS2* does not play an important role in type 2 diabetes susceptibility.

Acknowledgements

This study was supported by the Netherlands Organization for Scientific Research, ZonMW RIDE program and the Dutch Diabetes Research Foundation. The Danish study was supported by grants from the Lundbeck Foundation Centre of Applied Medical Genomics for Personalized Disease Prediction, Prevention and Care (LUCAMP), the Danish Health Research Council, and The European Union (EUGENE2, grant no. LSHM-CT-2004-512013). We acknowledge use of genotype data from the British 1958 Birth Cohort DNA collection, funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. We thank the staff and senior management of the UK Blood Services responsible for the UK Blood Services Collection. Diabetes UK funded the collection of the Warren 2 resource and the UK Type 2 Diabetes Genetics Consortium collection was supported by the Wellcome Trust (Biomedical Collections Grant GR072960). We also acknowledge funding from Diabetes UK (Grant RD04/0002809), Wellcome Trust (076113 and GR072960), UK Medical Research Council (G0601261) and the Oxford NIHR Biomedical Research Centre. The DIAGRAM consortium is acknowledged for sharing data. Bahram Jafar-Mohammadi is a Diabetes UK clinical training fellow. The authors would like to acknowledge the participants of all study cohorts for their cooperation.

Supplementary figure S1. LD plot of all known LF SNPs in the LARS2 locus.

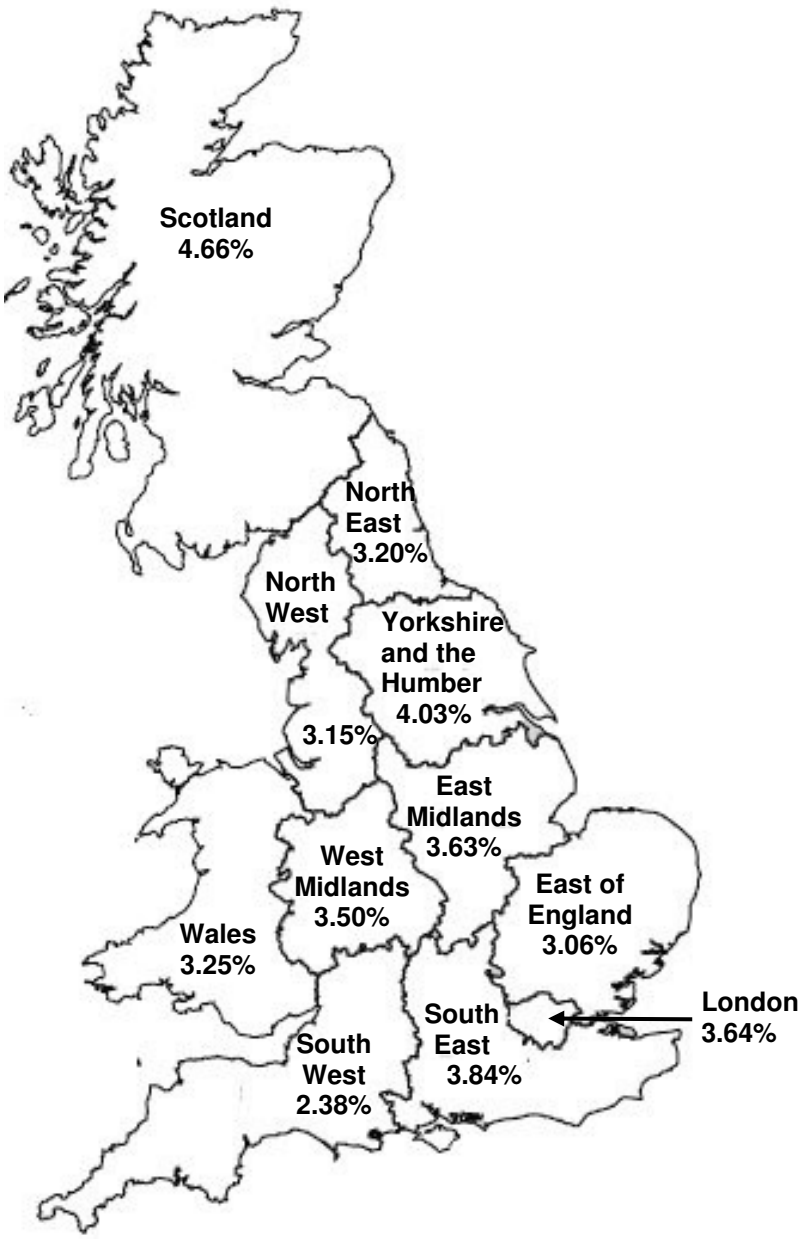


LD plot of all known LF SNPs in *LARS2* (MAF ≤0.05; HAPMAP phase 2, CEU). r^2 for each of the SNP pairs is given (black r^2=1, shades of gray 0< r^2<1 and white r^2=0).

The 14 LF SNPs directly genotyped in DIAGRAM are boxed. In the insert only the 14 LF SNPs included in the DIAGRAM scan are shown.

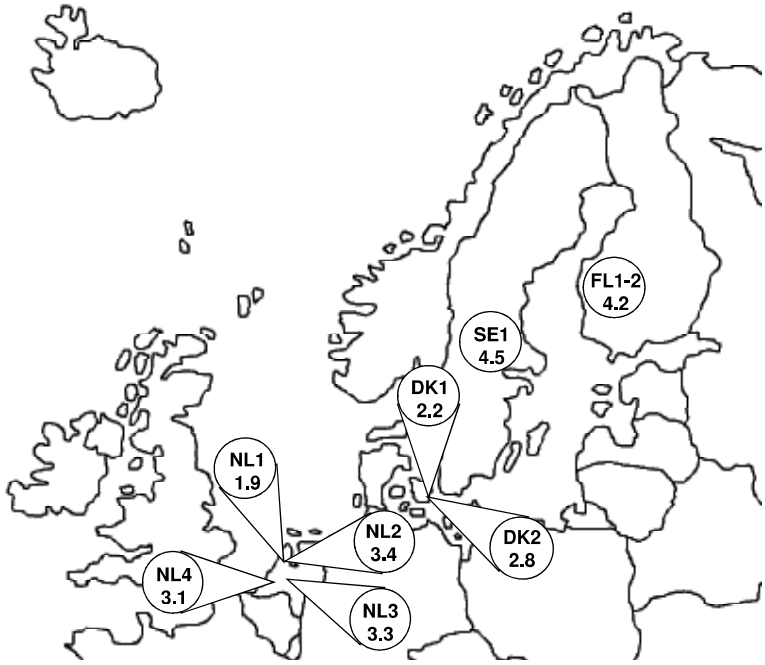
rs9825041 was used as a proxy for the block and is shown in the box.

Supplementary figure S2. Geographic stratification in the UK for the H324Q (rs71645922) variant.



The numbers represent minor allele frequencies of the H324Q (rs71645922) variant in each of the geographic regions.

Supplementary figure S3. Geographic stratification in other European samples.



The numbers is the circles represent minor allele frequencies of H324Q (rs71645922) in each of the study samples.

Supplementary table S1. Overview of tagging SNPs, including results from the NL1 study.

rs number	Position	Allele 1	Allele 2	MAF controls	MAF cases	OR 95% C.I.	p-value
rs854191	45409713	T	C	0.19	0.16	0.86 0.68 – 1.08	0.22
rs17637580	45411477	G	A	0.33	0.35	1.08 0.89 – 1.30	0.45
rs17077759	45411945	A	C	0.23	0.20	0.86 0.69 – 1.07	0.17
rs17637703	45415725	A	G	0.22	0.25	1.22 0.99 – 1.50	0.07
rs864391	45422021	G	A	0.23	0.20	0.88 0.71 – 1.09	0.23
rs6802884	45423962	T	C	0.19	0.17	0.91 0.72 – 1.14	0.42
rs17576051	45424984	A	G	0.11	0.10	0.94 0.71 – 1.25	0.72
rs3774696	45425473	A	C	0.24	0.27	1.16 0.95 – 1.42	0.17
rs9878643	45426336	G	A	n.a.	n.a.	n.a.	n.a.
rs6764923	45429693	A	G	0.40	0.37	0.88 0.69 – 1.13	0.32
rs17576289	45433737	A	G	0.14	0.12	0.87 0.67 – 1.12	0.29
rs4682783	45445335	A	G	n.a.	n.a.	n.a.	n.a.
rs4683113	45450337	A	G	0.42	0.42	1.00 0.80 – 1.25	1.00
rs9825692	45451366	G	A	0.28	0.31	1.17 0.97 – 1.42	0.12
rs952621	45451698	C	T	0.18	0.20	1.13 0.89 – 1.43	0.33
rs11926964	45464453	A	T	0.18	0.16	0.89 0.70 – 1.13	0.34
rs1877960	45471148	T	C	0.29	0.28	0.95 0.73 – 1.24	0.74
rs1949304	45475954	T	C	0.28	0.28	0.97 0.74 – 1.26	0.84
rs3774694	45481450	C	T	0.31	0.32	1.02 0.84 – 1.23	0.89
rs12487705	45492763	A	G	0.36	0.38	1.12 0.93 – 1.34	0.26
rs2128358	45524032	A	G	0.13	0.13	1.04 0.80 – 1.34	0.79
rs11706785	45546084	A	G	0.43	0.42	0.95 0.75 – 1.19	0.64
rs2118749	45547261	C	T	0.28	0.30	1.09 0.84 – 1.42	0.54
rs9825041	45502488	G	A	0.06	0.07	1.35 0.92 – 1.98	0.15

The *LARS2* locus was also investigated using a tagging SNP approach in the NL1 sample.
Common SNPs (MAF > 0.05), in *LARS2* were identified using HAPMAP (gene boundaries set at 45,405,079 - 45,565,332)[19].
Tagging SNPs were selected using the Tagger option in Haploview (NCBI build 36, phase II, April 2007, population CEU, LOD threshold 3.0 and $r^2 > 0.8$)[20].
Allele 1 is the major and allele 2 is the minor allele.
n.a.: not available.

Reference list

1. Mootha VK, Lindgren CM, Eriksson KF, et al (2003) PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat.Genet. 34: 267-273

2. Kelley DE, He J, Menshikova EV, Ritov VB (2002) Dysfunction of mitochondria in human skeletal muscle in type 2 diabetes. Diabetes. 51: 2944-2950

3. Matschinsky FM, Glaser B, Magnuson MA (1998) Pancreatic beta-cell glucokinase: closing the gap between theoretical concepts and experimental realities. Diabetes. 47: 307-315

4. Maassen JA, Romijn JA, Heine RJ (2007) Fatty acid-induced mitochondrial uncoupling in adipocytes as a key protective factor against insulin resistance and beta cell dysfunction: a new concept in the pathogenesis of obesity-associated type 2 diabetes mellitus. Diabetologia. 50: 2036-2041

5. van den Ouweland JM, Lemkes HH, Ruitenbeek W, et al (1992) Mutation in mitochondrial tRNA(Leu)(UUR) gene in a large pedigree with maternally transmitted type II diabetes mellitus and deafness. Nat.Genet. 1: 368-371

6. 't Hart LM, Hansen T, Rietveld I, et al (2005) Evidence that the mitochondrial leucyl tRNA synthetase (LARS2) gene represents a novel type 2 diabetes susceptibility gene. Diabetes 54: 1892-1895

7. Mooy JM, Grootenhuis PA, de VH, et al (1995) Prevalence and determinants of glucose intolerance in a Dutch caucasian population. The Hoorn Study. Diabetes Care. 18: 1270-1273

8. World Health Organization: Definition, diagnosis and classification of Diabetes Mellitus, Report of a WHO Consultation, Part 1: Diagnosis and classification of Diabetes Mellitus, WHO/NCD/NCS/99.2, Geneva. 1-1-1999.

9. van 't Riet E, Rijkelijkhuizen JM, Nijpels G, Dekker JM (2008) Limited agreement between HbA1c and glucose in the general Dutch population: The New Hoorn Study. Diabetologia 51: S164-S164 (abstract)

10. Reiling E, van Vliet-Ostaptchouk JV, van 't RE, et al (2009) Genetic association analysis of 13 nuclear-encoded mitochondrial candidate genes with type II diabetes mellitus: the DAMAGE study. Eur.J.Hum.Genet 17: 1056-1062

11. van Tilburg JH, Sandkuijl LA, Strengman E, et al (2003) A genome-wide scan in type 2 diabetes mellitus provides independent replication of a susceptibility locus on 18p11 and suggests the existence of novel Loci on 2q12 and 19q13. J.Clin.Endocrinol.Metab. 88: 2223-2230

12. Monsuur AJ, de Bakker PI, Alizadeh BZ, et al (2005) Myosin IXB variant increases the risk of celiac disease and points toward a primary intestinal barrier defect. Nat.Genet. 37: 1341-1344

13. Hofman A, Grobbee DE, de Jong PT, van den Ouweland FA (1991) Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. Eur.J.Epidemiol. 7: 403-422

14. Drivsholm T, Ibsen H, Schroll M, Davidsen M, Borch-Johnsen K (2001) Increasing prevalence of diabetes mellitus and impaired glucose tolerance among 60-year-old Danes. Diabet.Med. 18: 126-132

15. Zeggini E, Weedon MN, Lindgren CM, et al (2007) Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science. 316: 1336-1341

16. Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 447: 661-678

17. Florez JC, Sjogren M, Agapakis CM, et al (2007) Association testing of common variants in the insulin receptor substrate-1 gene (IRS1) with type 2 diabetes. Diabetologia. 50: 1209-1217

18. Saxena R, Gianniny L, Burtt NP, et al (2006) Common single nucleotide polymorphisms in TCF7L2 are reproducibly associated with type 2 diabetes and reduce the insulin response to glucose in nondiabetic individuals. Diabetes. 55: 2890-2895

19. Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P (2005) A haplotype map of the human genome. Nature. 437: 1299-1320

20. de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D (2005) Efficiency and power in genetic association studies. Nat.Genet. 37: 1217-1223

21. Gauderman WJ (2002) Sample size requirements for association studies of gene-gene interaction. Am.J.Epidemiol. 155: 478-484

22. Zeggini E, Scott LJ, Saxena R, et al (2008) Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat.Genet. 40: 638-645

23. Bersaglieri T, Sabeti PC, Patterson N, et al (2004) Genetic signatures of strong recent positive selection at the lactase gene. Am.J.Hum.Genet. 74: 1111-1120

24. Coelho M, Luiselli D, Bertorelle G, et al (2005) Microsatellite variation and evolution of human lactase persistence. Hum.Genet. 117: 329-339

25. Sabeti PC, Schaffner SF, Fry B, et al (2006) Positive natural selection in the human lineage. Science. 312: 1614-1620

26. Todd JA, Walker NM, Cooper JD, et al (2007) Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. Nat.Genet. 39: 857-864