Diabetes Publish Ahead of Print, published online November 23, 2009

Examination of all type 2 diabetes GWAS loci reveals *HHEX-IDE* as a locus influencing pediatric BMI

Jianhua Zhao[1], Jonathan P. Bradfield[2], Haitao Zhang[2], Kiran Annaiah[2], Kai Wang[2], Cecilia E. Kim[2], Joseph T. Glessner[2], Edward C. Frackelton[2], F. George Otieno[2], James Doran[2], Kelly A. Thomas[2], Maria Garris[2], Cuiping Hou[2], Rosetta M. Chiavacci[2], Mingyao Li[3], Robert I. Berkowitz[4,5], Hakon Hakonarson[1,2,6*] and Struan F.A. Grant[1,2,6*]

[1]Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA; [2]Center for Applied Genomics, Abramson Research Center, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA; [3]Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA; [4]Behavioral Health Center and Department of Child and Adolescent Psychiatry, The Children's Hospital of Philadelphia, Philadelphia PA 19104, USA; [5]Center for Weight and Eating Disorders, Department of Psychiatry, University of Pennsylvania, Philadelphia PA 19104, USA; [6]Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA

***To whom correspondence should be addressed**.
Struan F.A. Grant
E-mail: grants@chop.edu

Submitted 2 July 2009 and accepted 13 November 2009.

Additional information for this article can be found in an online appendix at http://diabetes.diabetesjournals.org

Objective: A number of studies have found that body mass index (BMI) in early life influences the risk of developing type 2 diabetes (T2D) later in life. Our goal was to investigate if any T2D variants uncovered through genome wide association studies (GWAS) impact BMI in childhood.

Design and Methods: Utilizing data from an ongoing GWAS of pediatric BMI in our cohort, we investigated the association of pediatric BMI with 20 SNPs at 18 T2D loci uncovered through GWAS, consisting of *ADAMTS9*, *CDC123-CAMK1D*, *CDKAL1*, *CDKN2A/B*, *EXT2*, *FTO*, *HHEX-IDE*, *IGF2BP2*, the intragenic region on 11p12, *JAZF1*, *KCNQ1*, *LOC387761*, *MTNR1B*, *NOTCH2*, *SLC30A8*, *TCF7L2*, *THADA* and *TSPAN8-LGR5*. We randomly partitioned our cohort exactly in half in order to have a 'discovery' cohort (n=3592) and a 'replication' cohort (n=3592).

Results: Our data show that the major, T2D-risk conferring G allele of rs7923837 at the *HHEX-IDE* locus was associated with higher pediatric BMI in both the discovery (P=0.0013; and survived correction for 20 tests) and replication (P=0.023) sets (combined P=1.01x10^{-4}). Association was not detected with any other known type 2 diabetes loci uncovered to date through GWAS except for the well established *FTO*.

Conclusions: Our data show that the same genetic *HHEX-IDE* variant which is associated with type 2 diabetes from previous studies also influences pediatric BMI.

Diabetes mellitus affects an estimated 194 million adults worldwide and more than 18 million in the United States, with the chronic complications including microvascular disease and accelerated development of cardiovascular disease. Approximately 90 to 95 percent of those affected with diabetes have the type 2 diabetes (T2D) form of the disease. Hyperglycemia is a key feature of T2D and occurs through two possible mechanisms: a) abnormal insulin secretion due to pancreatic □-cell defects or b) insulin resistance in skeletal, muscle, liver and adipose tissue.

T2D has been the focus of more genome wide association studies (GWAS) than any other disorder studied to date; such analyses have revealed a number of loci(1-9). The strongest association in European populations has been with a gene established in 2006, namely the Wnt-signaling pathway member transcription factor 7–like 2 (*TCF7L2*)(10), while in China and Japan the strongest association has been to the gene encoding potassium channel, voltage-gated, KQT-like subfamily, member 1 *(KCNQ1)* (8; 9). The first batch of such studies(1-6), revealed new loci and with a recent meta-analysis(7) of T2D genome wide SNP genotype data producing another six loci, there are now 17 genes established in the disease, including *CDKAL1*, *SLC30A8* and *JAZF1*. *MNTR1B* which was first implicated in multiple GWAS of the related trait of fasting glucose and was subsequently associated with T2D within the same studies(11-13).

All the T2D genes uncovered by GWAS to date have been implicated in primarily impacting insulin secretion, with the exception of the fat mass and obesity associated gene (*FTO*), which was uncovered as a consequence of a T2D GWAS but turned out to be operating through insulin resistance and was therefore primarily an obesity risk factor(14).

A question therefore arises: if specific genomic variants can impact insulin resistance or insulin secretion, can this in turn impact BMI earlier on in life? As such, we aimed at examining these T2D GWAS findings in a large pediatric cohort with BMI measures and to determine the relative impact of these variants on the trait of interest. We utilized data from an ongoing GWAS in a cohort of 7,184 European American children with recorded heights and weights randomly partitioned precisely in half in order to have a 'discovery' cohort and a subsequent 'replication' cohort.

Loci selected had been discovered directly from published T2D GWAS. We therefore queried for known variants at the 18 T2D-associated loci of *ADAMTS9*, *CDC123-CAMK1D*, *CDKAL1*, *CDKN2A/B*, *EXT2*, *FTO*, *HHEX-IDE*, *IGF2BP2*, the intragenic region on 11p12, *JAZF1*, *KCNQ1*, *LOC387761*, *MTNR1B*, *NOTCH2*, *SLC30A8*, *TCF7L2*, *THADA* and *TSPAN8-LGR5* with respect to their correlation with pediatric BMI.

MATERIAL AND METHODS

Research Subjects: Childhood European American Cohort from Philadelphia: All subjects were consecutively recruited from the Greater Philadelphia area from 2006 to 2009 at the Children's Hospital of Philadelphia. Our study cohort consisted of 7,184 singleton children of European ancestry with systematically recorded height and weight. All subjects were consecutively and randomly recruited from the greater metropolitan area of Philadelphia from 2006 to 2009 at the Children's Hospital of Philadelphia i.e. participants were not specifically targeted for obesity-related traits. The study was approved by the Institutional Review Board of the Children's Hospital of

Philadelphia. Parental informed consent was given for each study participant for both the blood collection and subsequent genotyping.

Genotyping: Illumina Infinium™ assay*:* We performed high throughput genome-wide SNP genotyping, using the Illumina Infinium™ II HumanHap550 or Human 610 BeadChip technology(Illumina, San Diego), at the Children's Hospital of Philadelphia's Center for Applied Genomics, as described previously (15). The overall genomic control value was 1.036. The SNPs analyzed survived the filtering of the genome wide dataset for SNPs with call rates <95%, minor allele frequency <1%, missing rate per person >2% and Hardy-Weinberg equilibrium $P < 10^{-5}$.

Most loci described from GWA studies published to date have been found using either the Affymetrix or Illumina platform. In the event a locus was reported using both the Illumina and Affymetrix arrays, we used the SNPs present on the Illumina array. In the event of a signal only being described on the Affymetrix array, we either already had the SNP on our Illumina array or we identified and used the best surrogate SNP available based on the CEU HapMap (Supplementary Table 1 which can be found in an online appendix at http://diabetes.diabetesjournals.org). We utilized two SNPs at the *CDKAL1* (rs4712523 and rs7756992; r^2 = 0.677) and *HHEX-IDE* (rs1111875 and rs7923837; r^2= 0.698) loci as the association with T2D from various GWA studies reported different SNPs which were in imperfect LD with each other. rs3751812 at *FTO* was included as a positive control as we have previously reported association with this SNP and both pediatric obesity and pediatric BMI previously(16; 17).

Analysis: Normalization of BMI*:* BMI percentiles were defined using the standard Center for Disease control (CDC) growth chart z-scores that take in to account age and gender. All subjects were biologically unrelated and were aged between 2 and 18 years old. All subjects were between -3 and +3 standard deviations of CDC corrected BMI i.e. outliers (n=356) were excluded to avoid the consequences of potential measurement error or Mendelian causes of extreme obesity.

Association: We queried the data for the SNPs of interest in our pediatric sample. All statistical analyses were carried out using the software package PLINK version 1.05(18). We applied the PLINK to the generation of genome-wide IBS estimates between all subjects and then generated multi-dimensional scaling (MDS) plots for visual examination of population outliers. To help interpret the population genetic analysis, we have included 924 HapMap3 individuals from 11 populations as positive controls into the MDS analysis. The individuals of European ancestry were selected by the principal component 1 of more than 0.04 and principal component 2 of more than 0.01. Comparing self-identified ancestry with the MDS-inferred ancestry confirmed the reliability of MDS to identify genetically inferred individuals of European ancestry.

By treating the normalized BMI z-score as a quantitative trait, association analysis for each SNP was carried out using linear regression (additive model) with the SNP included as an independent variable (coded as 0, 1, and 2). With 3592 subjects in the discovery cohort, the powers to detect 0.2%, 0.3%, 0.4%, 0.5%, 0.6%, 0.8% and 1% variation at the α=0.0025 level were 27.0%, 49.0%, 68.2%, 82.0%, 90.6%, 97.9% and 99.6%, respectively.

RESULTS

In our analysis, twenty SNPs corresponding to the eighteen T2D loci previously discovered in GWAS of the disorder were investigated, namely *ADAMTS9*, *CDC123-CAMK1D*, *CDKAL1*, *CDKN2A/B*, *EXT2*, *FTO*, *HHEX-IDE*, *IGF2BP2*, the intragenic region on 11p12,

JAZF1, *KCNQ1*, *LOC387761*, *MTNR1B*, *NOTCH2*, *SLC30A8*, *TCF7L2*, *THADA* and *TSPAN8-LGR5* (Table 1).

We randomly partitioned our cohort exactly in half in order to have a 'discovery' cohort (n=3592) and a 'replication' cohort (n=3592). Five of these twenty SNPs yielded at least nominally significant association to BMI ($P < 0.05$) in the discovery cohort, representing four different independent loci.

Of these four loci, the minor allele of rs3751812 at the *FTO* locus yielded the strongest association with $P=3.81\text{x}10^{-5}$ and tracked with higher BMI. The direction of effect was also readily replicated in the additional cohort ($P=5.56\text{x}10^{-6}$) yielding a combined $P=1.05\text{x}10^{-9}$.

The major T2D-conferring G allele of rs7923837 at the *HHEX-IDE* locus was associated with higher pediatric BMI in both the discovery (unadjusted $P=0.0013$; Bonferroni correction for 20 variants threshold $P\leq0.0025$) and replication (unadjusted $P=0.023$) sets (combined unadjusted $P=1.01\text{x}10^{-4}$). The major C allele of rs1111875 at the same locus was also trending with higher pediatric BMI but did not survive the Bonferroni correction for multiple testing in the discovery cohort.

As for the other two nominally significant loci in the discovery cohort, rs4402960 at *IGF2BP2* ($P=0.05$) and rs11257622 at *CDC123-CAMK1D* ($P=0.024$), they failed to replicate in the additional cohort. Association was not detected at all with any of the other T2D loci uncovered to date through GWAS.

We also analyzed males and females separately but the effect of the G allele rs7923837 at the *HHEX-IDE* locus on pediatric BMI did not vary by gender (Supplementary Table 2). However, we did look at different age bins and found that the variant was associated with higher pediatric BMI most strongly in the 2-6 years old age bin (Supplementary Table 3). Breaking the ages down further in to individual years, nominal significant association for this *HHEX-IDE* variant in the same direction was observed at ages 3, 7, 14 and 16 years old (Supplementary Table 4). However we did not observe an overall statistical interaction with age, with the interaction *P*-values for rs1111875 and rs7923837 being 0.2507 and 0.1076 respectively.

DISCUSSION

If a genomic variant is well established to be associated with a trait which is the consequence of a defect of recognition of insulin by the body or by a fault in the amount of insulin released for the pancreatic islets, i.e. type 2 diabetes, then if these defects are operating at all in childhood one might expect there to be an impact on BMI in childhood.

With this notion in mind, we queried the existing dataset from our ongoing GWAS of pediatric BMI if any of the T2D loci uncovered in GWAS to date play a role in our trait of interest; it should be noted that *PPARG*, *KCNJ11* and of *WFS1* were not included as their discovery with respect to being T2D loci pre-dates GWA studies and thus have already been more extensively investigated. Our data in fact do show that the same genetic *HHEX-IDE* variant which is significantly associated with T2D from previous studies also influences pediatric BMI. Indeed, the major G allele of rs7923837 at the *HHEX-IDE* locus was associated with higher pediatric BMI in both the 'discovery' and 'replication' cohorts, which is the same allele that has been reported to confer risk of T2D. This mirrors very well what is seen with the much more established *FTO* gene, seen here and in other studies.

SNP rs7923837 yielded the fourth strongest association to T2D in a Canadian/French GWAS carried out on the Illumina HumanHap platform(1). SNPs rs1111875 and rs7923837 yielded the

strongest association at the *HHEX-IDE* locus but it should be noted that they are far from being in perfect LD with each other (r^2= 0.698) thus the inclusion of both in the current study; however, despite the lack of complete concordance and the large sample size, we were unable to separate the effects of these SNPs as they cannot be considered to be totally independent signals either.

One hypothesis could be that the fetal genotype for rs7923837 is primarily associated with birth weight, as reduced birth weight is often reported to be associated with increased BMI and T2D later in life. However, this doesn't appear to be the case as we have already investigated and reported the role of these T2D loci in the context of birth weight in our cohort. Although we have agreed with previous studies that *CDKAL1* is a birth weight associated gene, we have not observed such an association with *HHEX-IDE*(19). Further, although there is no CDC categorization for the under 2 year old age group, following our own normalization we do not observe association between rs7923837 and BMI in this age category (data not shown). The correlation between birth weight and BMI in later childhood in less correlated than in earlier stages, suggesting the *HHEX-IDE* variant exerts its physiological influence directly rather than as a consequence of a 'knock-on' effect from a primary impact on birth weight. However, we do acknowledge, of the age bins studied, the strongest effect was observed in the 2-6 years age bin (effect size(SE)=0.12(+/-0.04)) (Supplementary Table 3) but is not the whole story due to the fact that at the individual age level, although more limited in terms of power, the impact continues to be observed in to the mid-teens (Supplementary Table 4).

The assumption in this study is that deficient insulin secretion mediates the effect on childhood BMI but it is also possible that higher childhood BMI results in impaired insulin secretion later in life. There could indeed be pleiotropic associations from multiple independent mechanisms; however we were not able to address this as we do not have insulin secretion/sensitivity measures in our study.

From our analysis, apart from *FTO* it is clear only one of the loci previously reported from T2D GWAS plays a role in our phenotype of interest i.e. pediatric BMI. While this recently discovered locus unveils a new biomolecular pathway not previously studied in the context of T2D and obesity, it is also important to note this and other genetic associations with childhood obesity explain very little of the genetic risk for the pathogenesis of the trait(17); indeed, an estimate of the explained variance of the *HHEX-IDE* and *FTO* loci combined is only 0.98%, suggesting the existence of additional loci whose number and effect size remain mainly unknown. Current knowledge concerning the impact of genetic factors in the determination of pediatric BMI may still be very limited due to both the lack of availability of large pediatric cohorts with GWAS data and methodological difficulties in the analysis of the phenotype that changes with age and depends on many other contributing factors. Once our GWAS is complete, we will have the opportunity to look for other variants in the genome associated with BMI in childhood.

ACKNOWLEDGEMENTS

We would like to thank all participating subjects and families. Elvira Dabaghyan, Hope Thomas, Kisha Harden, Andrew Hill, Kenya Fain, Crystal Johnson-Honesty, Cynthia Drummond, Shanell Harrison and Sarah Wildrick provided expert assistance with genotyping or data collection and management. We would also like to thank Smari Kristinsson, Larus Arni Hermannsson and Asbjörn Krisbjörnsson of Raförninn ehf for their extensive software design and contribution. This research was financially

supported by the Children's Hospital of Philadelphia.

The study is supported in part by a Research Development Award from the Cotswold Foundation (H.H. & S.F.A.G.) and NIH grant 1R01HD056465-01A1.

REFERENCES

1. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, Balkau B, Heude B, Charpentier G, Hudson TJ, Montpetit A, Pshezhetsky AV, Prentki M, Posner BI, Balding DJ, Meyre D, Polychronakos C, Froguel P: A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 445:881-885, 2007

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

3. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, Hughes TE, Groop L, Altshuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Bengtsson Bostrom K, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nilsson P, Orho-Melander M, Rastam L, Speliotes EK, Taskinen MR, Tuomi T, Guiducci C, Berglund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjogren M, Sterner M, Surti A, Svensson M, Svensson M, Tewhey R, Blumenstiel B, Parkin M, Defelice M, Barry R, Brodeur W, Camarata J, Chia N, Fava M, Gibbons J, Handsaker B, Healy C, Nguyen K, Gates C, Sougnez C, Gage D, Nizzari M, Gabriel SB, Chirn GW, Ma Q, Parikh H, Richardson D, Ricke D, Purcell S: Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 316:1331-1336, 2007

4. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, Timpson NJ, Perry JR, Rayner NW, Freathy RM, Barrett JC, Shields B, Morris AP, Ellard S, Groves CJ, Harries LW, Marchini JL, Owen KR, Knight B, Cardon LR, Walker M, Hitman GA, Morris AD, Doney AS, Burton PR, Clayton DG, Craddock N, Deloukas P, Duncanson A, Kwiatkowski DP, Ouwehand WH, Samani NJ, Todd JA, Donnelly P, Davison D, Easton D, Evans D, Leung HT, Spencer CC, Tobin MD, Attwood AP, Boorman JP, Cant B, Everson U, Hussey JM, Jolley JD, Knight AS, Koch K, Meech E, Nutland S, Prowse CV, Stevens HE, Taylor NC, Walters GR, Walker NM, Watkins NA, Winzer T, Jones RW, McArdle WL, Ring SM, Strachan DP, Pembrey M, Breen G, St Clair D, Caesar S, Gordon-Smith K, Jones L, Fraser C, Green EK, Grozeva D, Hamshere ML, Holmans PA, Jones IR, Kirov G, Moskvina V, Nikolov I, O'Donovan M C, Owen MJ, Collier DA, Elkin A, Farmer A, Williamson R, McGuffin P, Young AH, Ferrier IN, Ball SG, Balmforth AJ, Barrett JH, Bishop DT, Iles MM, Maqbool A, Yuldasheva N, Hall AS, Braund PS, Dixon RJ, Mangino M, Stevens S, Thompson JR, Bredin F, Tremelling M, Parkes M, Drummond H, Lees CW, Nimmo ER, Satsangi J, Fisher SA, Forbes A, Lewis CM, Onnie CM, Prescott NJ, Sanderson J, Mathew CG, Barbour J, Mohiuddin MK, Todhunter CE, Mansfield JC, Ahmad T, Cummings FR, Jewell DP, Webster J, Brown MJ, Lathrop GM, Connell J, Dominiczak A, Braga Marcano CA, Burke B, Dobson R, Gungadoo J, Lee KL, Munroe PB, Newhouse SJ, Onipinla A, Wallace C, Xue M, Caulfield M, Farrall M, Barton A, Bruce IN, Donovan H, Eyre S, Gilbert PD, Hider SL, Hinks AM, John SL, Potter C, Silman AJ, Symmons DP, Thomson W, Worthington J, Dunger DB, Widmer B, Newport M, Sirugo G, Lyons E, Vannberg F, Hill AV, Bradbury LA, Farrar C, Pointon JJ, Wordsworth P, Brown MA, Franklyn JA, Heward JM, Simmonds MJ, Gough SC, Seal S, Stratton MR, Rahman N, Ban M, Goris A, Sawcer SJ, Compston A, Conway D, Jallow M, Rockett KA, Bumpstead SJ, Chaney A, Downes K, Ghori MJ, Gwilliam R, Hunt SE, Inouye M, Keniry A, King E, McGinnis R, Potter S, Ravindrarajah R, Whittaker P, Widden C, Withers D, Cardin NJ, Ferreira T, Pereira-Gale J, Hallgrimsdottir IB, Howie BN, Su Z, Teo YY, Vukcevic D, Bentley D, Compston A, Ouwehand NJ, Samani MR, Isaacs JD, Morgan AW, Wilson GD, Ardern-Jones A, Berg J, Brady A, Bradshaw N, Brewer C, Brice G, Bullman B, Campbell J, Castle B, Cetnarsryj R, Chapman C, Chu C, Coates N, Cole T, Davidson R, Donaldson A, Dorkins H, Douglas F, Eccles D, Eeles R,

Elmslie F, Evans DG, Goff S, Goodman S, Goudie D, Gray J, Greenhalgh L, Gregory H, Hodgson SV, Homfray T, Houlston RS, Izatt L, Jackson L, Jeffers L, Johnson-Roffey V, Kavalier F, Kirk C, Lalloo F, Langman C, Locke I, Longmuir M, Mackay J, Magee A, Mansour S, Miedzybrodzka Z, Miller J, Morrison P, Murday V, Paterson J, Pichert G, Porteous M, Rahman N, Rogers M, Rowe S, Shanley S, Saggar A, Scott G, Side L, Snadden L, Steel M, Thomas M, Thomas S, McCarthy MI, Hattersley AT: Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 316:1336-1341, 2007
5. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, Erdos MR, Stringham HM, Chines PS, Jackson AU, Prokunina-Olsson L, Ding CJ, Swift AJ, Narisu N, Hu T, Pruim R, Xiao R, Li XY, Conneely KN, Riebow NL, Sprau AG, Tong M, White PP, Hetrick KN, Barnhart MW, Bark CW, Goldstein JL, Watkins L, Xiang F, Saramies J, Buchanan TA, Watanabe RM, Valle TT, Kinnunen L, Abecasis GR, Pugh EW, Doheny KF, Bergman RN, Tuomilehto J, Collins FS, Boehnke M: A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 316:1341-1345, 2007
6. Steinthorsdottir V, Thorleifsson G, Reynisdottir I, Benediktsson R, Jonsdottir T, Walters GB, Styrkarsdottir U, Gretarsdottir S, Emilsson V, Ghosh S, Baker A, Snorradottir S, Bjarnason H, Ng MC, Hansen T, Bagger Y, Wilensky RL, Reilly MP, Adeyemo A, Chen Y, Zhou J, Gudnason V, Chen G, Huang H, Lashley K, Doumatey A, So WY, Ma RC, Andersen G, Borch-Johnsen K, Jorgensen T, van Vliet-Ostaptchouk JV, Hofker MH, Wijmenga C, Christiansen C, Rader DJ, Rotimi C, Gurney M, Chan JC, Pedersen O, Sigurdsson G, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K: A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat Genet* 39:770-775, 2007
7. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, de Bakker PI, Abecasis GR, Almgren P, Andersen G, Ardlie K, Bostrom KB, Bergman RN, Bonnycastle LL, Borch-Johnsen K, Burtt NP, Chen H, Chines PS, Daly MJ, Deodhar P, Ding CJ, Doney AS, Duren WL, Elliott KS, Erdos MR, Frayling TM, Freathy RM, Gianniny L, Grallert H, Grarup N, Groves CJ, Guiducci C, Hansen T, Herder C, Hitman GA, Hughes TE, Isomaa B, Jackson AU, Jorgensen T, Kong A, Kubalanza K, Kuruvilla FG, Kuusisto J, Langenberg C, Lango H, Lauritzen T, Li Y, Lindgren CM, Lyssenko V, Marvelle AF, Meisinger C, Midthjell K, Mohlke KL, Morken MA, Morris AD, Narisu N, Nilsson P, Owen KR, Palmer CN, Payne F, Perry JR, Pettersen E, Platou C, Prokopenko I, Qi L, Qin L, Rayner NW, Rees M, Roix JJ, Sandbaek A, Shields B, Sjogren M, Steinthorsdottir V, Stringham HM, Swift AJ, Thorleifsson G, Thorsteinsdottir U, Timpson NJ, Tuomi T, Tuomilehto J, Walker M, Watanabe RM, Weedon MN, Willer CJ, Illig T, Hveem K, Hu FB, Laakso M, Stefansson K, Pedersen O, Wareham NJ, Barroso I, Hattersley AT, Collins FS, Groop L, McCarthy MI, Boehnke M, Altshuler D: Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet* 40:638-645, 2008
8. Unoki H, Takahashi A, Kawaguchi T, Hara K, Horikoshi M, Andersen G, Ng DP, Holmkvist J, Borch-Johnsen K, Jorgensen T, Sandbaek A, Lauritzen T, Hansen T, Nurbaya S, Tsunoda T, Kubo M, Babazono T, Hirose H, Hayashi M, Iwamoto Y, Kashiwagi A, Kaku K, Kawamori R, Tai ES, Pedersen O, Kamatani N, Kadowaki T, Kikkawa R, Nakamura Y, Maeda S: SNPs in KCNQ1 are associated with susceptibility to type 2 diabetes in East Asian and European populations. *Nat Genet* 40:1098-1102, 2008
9. Yasuda K, Miyake K, Horikawa Y, Hara K, Osawa H, Furuta H, Hirota Y, Mori H, Jonsson A, Sato Y, Yamagata K, Hinokio Y, Wang HY, Tanahashi T, Nakamura N, Oka Y, Iwasaki N, Iwamoto Y, Yamada Y, Seino Y, Maegawa H, Kashiwagi A, Takeda J, Maeda E, Shin HD, Cho

YM, Park KS, Lee HK, Ng MC, Ma RC, So WY, Chan JC, Lyssenko V, Tuomi T, Nilsson P, Groop L, Kamatani N, Sekine A, Nakamura Y, Yamamoto K, Yoshida T, Tokunaga K, Itakura M, Makino H, Nanjo K, Kadowaki T, Kasuga M: Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus. *Nat Genet* 40:1092-1097, 2008
10. Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J, Helgason A, Stefansson H, Emilsson V, Helgadottir A, Styrkarsdottir U, Magnusson KP, Walters GB, Palsdottir E, Jonsdottir T, Gudmundsdottir T, Gylfason A, Saemundsdottir J, Wilensky RL, Reilly MP, Rader DJ, Bagger Y, Christiansen C, Gudnason V, Sigurdsson G, Thorsteinsdottir U, Gulcher JR, Kong A, Stefansson K: Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet* 38:320-323, 2006
11. Prokopenko I, Langenberg C, Florez JC, Saxena R, Soranzo N, Thorleifsson G, Loos RJ, Manning AK, Jackson AU, Aulchenko Y, Potter SC, Erdos MR, Sanna S, Hottenga JJ, Wheeler E, Kaakinen M, Lyssenko V, Chen WM, Ahmadi K, Beckmann JS, Bergman RN, Bochud M, Bonnycastle LL, Buchanan TA, Cao A, Cervino A, Coin L, Collins FS, Crisponi L, de Geus EJ, Dehghan A, Deloukas P, Doney AS, Elliott P, Freimer N, Gateva V, Herder C, Hofman A, Hughes TE, Hunt S, Illig T, Inouye M, Isomaa B, Johnson T, Kong A, Krestyaninova M, Kuusisto J, Laakso M, Lim N, Lindblad U, Lindgren CM, McCann OT, Mohlke KL, Morris AD, Naitza S, Orru M, Palmer CN, Pouta A, Randall J, Rathmann W, Saramies J, Scheet P, Scott LJ, Scuteri A, Sharp S, Sijbrands E, Smit JH, Song K, Steinthorsdottir V, Stringham HM, Tuomi T, Tuomilehto J, Uitterlinden AG, Voight BF, Waterworth D, Wichmann HE, Willemsen G, Witteman JC, Yuan X, Zhao JH, Zeggini E, Schlessinger D, Sandhu M, Boomsma DI, Uda M, Spector TD, Penninx BW, Altshuler D, Vollenweider P, Jarvelin MR, Lakatta E, Waeber G, Fox CS, Peltonen L, Groop LC, Mooser V, Cupples LA, Thorsteinsdottir U, Boehnke M, Barroso I, Van Duijn C, Dupuis J, Watanabe RM, Stefansson K, McCarthy MI, Wareham NJ, Meigs JB, Abecasis GR: Variants in MTNR1B influence fasting glucose levels. *Nat Genet* 41:77-81, 2009
12. Bouatia-Naji N, Bonnefond A, Cavalcanti-Proenca C, Sparso T, Holmkvist J, Marchand M, Delplanque J, Lobbens S, Rocheleau G, Durand E, De Graeve F, Chevre JC, Borch-Johnsen K, Hartikainen AL, Ruokonen A, Tichet J, Marre M, Weill J, Heude B, Tauber M, Lemaire K, Schuit F, Elliott P, Jorgensen T, Charpentier G, Hadjadj S, Cauchi S, Vaxillaire M, Sladek R, Visvikis-Siest S, Balkau B, Levy-Marchal C, Pattou F, Meyre D, Blakemore AI, Jarvelin MR, Walley AJ, Hansen T, Dina C, Pedersen O, Froguel P: A variant near MTNR1B is associated with increased fasting plasma glucose levels and type 2 diabetes risk. *Nat Genet* 41:89-94, 2009
13. Lyssenko V, Nagorny CL, Erdos MR, Wierup N, Jonsson A, Spegel P, Bugliani M, Saxena R, Fex M, Pulizzi N, Isomaa B, Tuomi T, Nilsson P, Kuusisto J, Tuomilehto J, Boehnke M, Altshuler D, Sundler F, Eriksson JG, Jackson AU, Laakso M, Marchetti P, Watanabe RM, Mulder H, Groop L: Common variant in MTNR1B associated with increased risk of type 2 diabetes and impaired early insulin secretion. *Nat Genet* 41:82-88, 2009
14. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, Perry JR, Elliott KS, Lango H, Rayner NW, Shields B, Harries LW, Barrett JC, Ellard S, Groves CJ, Knight B, Patch AM, Ness AR, Ebrahim S, Lawlor DA, Ring SM, Ben-Shlomo Y, Jarvelin MR, Sovio U, Bennett AJ, Melzer D, Ferrucci L, Loos RJ, Barroso I, Wareham NJ, Karpe F, Owen KR, Cardon LR, Walker M, Hitman GA, Palmer CN, Doney AS, Morris AD, Smith GD, Hattersley AT, McCarthy MI: A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 316:889-894, 2007
15. Hakonarson H, Grant SFA, Bradfield JP, Marchand L, Kim CE, Glessner JT, Grabs R, Casalunovo T, Taback SP, Frackelton EC, Lawson ML, Robinson LJ, Skraban R, Lu Y,

Chiavacci RM, Stanley CA, Kirsch SE, Rappaport EF, Orange JS, Monos DS, Devoto M, Qu H-Q, Polychronakos C: A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. *Nature* 448:591-594, 2007
16. Grant SF, Li M, Bradfield JP, Kim CE, Annaiah K, Santa E, Glessner JT, Casalunovo T, Frackelton EC, Otieno FG, Shaner JL, Smith RM, Imielinski M, Eckert AW, Chiavacci RM, Berkowitz RI, Hakonarson H: Association analysis of the FTO gene with obesity in children of Caucasian and African ancestry reveals a common tagging SNP. *PLoS ONE* 3:e1746, 2008
17. Zhao J, Bradfield JP, Li M, Wang K, Zhang H, Kim CE, Annaiah K, Glessner JT, Thomas K, Garris M, Frackelton EC, Otieno FG, Shaner JL, Smith RM, Chiavacci RM, Berkowitz RI, Hakonarson H, Grant SF: The Role of Obesity-associated Loci Identified in Genome-wide Association Studies in the Determination of Pediatric BMI. *Obesity (Silver Spring)*, 2009
18. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC: PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559-575, 2007
19. Zhao J, Li M, Bradfield JP, Wang K, Zhang H, Sleiman P, Kim CE, Annaiah K, Glaberson W, Glessner JT, Otieno FG, Thomas KA, Garris M, Hou C, Frackelton EC, Chiavacci RM, Berkowitz RI, Hakonarson H, Grant SF: Examination of type 2 diabetes loci implicates CDKAL1 as a birth weight gene. *Diabetes* 58:2414-2418, 2009

Table 1. Quantitative association results for the known type 2 diabetes risk alleles with pediatric BMI in the European American cohort (n=3,592) followed by a replication effort (n=3,592), sorted by chromosomal location

					DISCOVERY COHORT					REPLICATION COHORT					
CHR	SNP	T2D associated allele	BP	Nearby Gene	N	Effect Size	SE	Test Statistic	P	N	Effect Size	SE	Test Statistic	P	Combined P
1	rs2793831	C	120235944	*NOTCH2*	3592	0.03508	0.04637	0.7565	0.449	3592	0.02797	0.04603	0.6076	0.544	0.3353
2	rs7578597	T*	43644474	*THADA*	3592	0.01896	0.04632	0.4094	0.682	3592	0.007494	0.04521	0.1658	0.868	0.6785
3	rs4411878	C*	64678705	*ADAMTS9*	3591	-0.02394	0.03145	-0.7611	0.447	3592	0.01207	0.03082	0.3917	0.695	0.8086
3	rs4402960	T	186994389	*IGF2BP2*	3587	-0.05843	0.0298	-1.961	0.05	3592	-0.004747	0.02886	-0.1645	0.869	0.1375
6	rs4712523	G	20765543	*CDKAL1*	3592	0.01223	0.02991	0.4087	0.683	3592	-0.02724	0.02996	-0.909	0.363	0.7294
6	rs7756992	G	20787688	*CDKAL1*	3591	0.02923	0.03128	0.9344	0.350	3592	-0.01428	0.03104	-0.4599	0.646	0.7371
7	rs1635852	C*	27962651	*JAZF1*	3590	0.009886	0.02783	0.3552	0.722	3592	-0.01975	0.02778	-0.7108	0.477	0.8058
8	rs13266634	C*	118253964	*SLC30A8*	3590	0.003039	0.03004	0.1012	0.919	3588	-0.01446	0.03039	-0.4759	0.634	0.7949
9	rs2383207	A*	22105959	*CDKN2A/B*	3591	0.04088	0.02787	1.467	0.142	3592	-0.04318	0.02783	-1.552	0.121	0.9482
10	rs11257622	C	12335345	*CDC123-CAMK1D*	3580	-0.08373	0.03703	-2.261	0.0238	3591	0.08785	0.03706	2.37	0.0178	0.9463
10	rs1111875	C*	94452862	*HHEX-IDE*	3592	0.08005	0.02839	2.82	0.00483	3592	0.05527	0.02823	1.957	0.0504	$7.14x10^{-4}$
10	****rs7923837**	**G***	**94471897**	***HHEX-IDE***	**3592**	**0.0913**	**0.02845**	**3.209**	**0.00134**	**3592**	**0.06523**	**0.02865**	**2.277**	**0.0229**	**$1.01x10^{-4}$**
10	rs7903146	T	114748339	*TCF7L2*	3592	-0.01407	0.03025	-0.465	0.642	3592	-0.00646	0.02988	-0.2162	0.829	0.636
11	rs163171	C*	2777641	*KCNQ1*	3588	-0.03288	0.0328	1.002	0.316	3588	-0.09764	0.03347	-2.917	0.00355	0.19
11	rs9300039	C*	41871942	*Intragenic*	3585	-0.08931	0.05007	1.784	0.0746	3592	0.03837	0.04824	0.7953	0.427	0.07334
11	rs7480010	G	42203294	*LOC387761*	3591	0.02035	0.03049	0.6673	0.505	3592	-0.01925	0.0299	-0.6437	0.520	0.9935
11	rs729287	C*	44236666	*EXT2*	3592	-0.0188	0.03215	0.5849	0.559	3591	0.0261	0.03194	0.8173	0.414	0.3223
11	rs1387153	T	92313476	*MTNR1B*	3592	-0.003922	0.0308	-0.1273	0.899	3592	-0.01709	0.03052	-0.5598	0.576	0.6223
12	rs1353362	C	69899543	*TSPAN8-LGR5*	3581	-0.01765	0.03074	-0.5743	0.566	3589	-0.01196	0.03008	-0.3977	0.691	0.4916
16	****rs3751812**	**T**	**52375961**	***FTO***	**3587**	**0.1159**	**0.0281**	**4.124**	**$3.81x10^{-5}$**	**3592**	**0.1273**	**0.02799**	**4.549**	**$5.56x10^{-6}$**	**$1.05x10^{-9}$**

BP: Base pair position (dbSNP build 125); N: number of individuals tested; Effect Size: regression coefficient for the test SNP; SE: standard error of the regression coefficient; Test statistic: additive model; P: unadjusted two-sided trend test *P*-value. The direction of effect is shown for the T2D-risk allele in each case. *The T2D-risk allele is the major allele; **P≤0.0025 in the discovery cohort i.e. surivive Bonferroni correction for number of variants tested.