nature COMMUNICATIONS

ARTICLE

Received 15 Jun 2015 | Accepted 16 Dec 2015 | Published 1 Feb 2016

DOI: 10.1038/ncomms10495 OPEN

# New loci for body fat percentage reveal link between adiposity and cardiometabolic disease risk

Yingchang Lu *et al.*#

To increase our understanding of the genetic basis of adiposity and its links to cardiometabolic disease risk, we conducted a genome-wide association meta-analysis of body fat percentage (BF%) in up to 100,716 individuals. Twelve loci reached genome-wide significance ($P<5\times10^{-8}$), of which eight were previously associated with increased overall adiposity (BMI, BF%) and four (in or near *COBLL1/GRB14*, *IGF2BP1*, *PLA2G6*, *CRTC1*) were novel associations with BF%. Seven loci showed a larger effect on BF% than on BMI, suggestive of a primary association with adiposity, while five loci showed larger effects on BMI than on BF%, suggesting association with both fat and lean mass. In particular, the loci more strongly associated with BF% showed distinct cross-phenotype association signatures with a range of cardiometabolic traits revealing new insights in the link between adiposity and disease risk.

Correspondence and requests for materials should be addressed to R.J.F.L. (email: ruth.loos@mssm.edu).
#A full list of authors and their affiliations appears at the end of the paper.

Large-scale meta-analyses of genome-wide association studies (GWAS) for adiposity traits and obesity risk have identified at least 160 loci that contribute to body weight and fat distribution in adults and children of diverse ancestry[1–20]. Studies of overall adiposity, assessed by body mass index (BMI), have mainly implicated genes that provide support for a role of the central nervous system (CNS) in obesity susceptibility[1–6,10,19], whereas genetic loci associated with body fat distribution, assessed by waist-to-hip ratio (WHR), seem enriched for genes involved in adipocyte metabolism[9,11,20]. Although these commonly studied adiposity traits are easily collected in large populations and thus allow statistically well-powered meta-analyses, they represent heterogeneous phenotypes, for example, people with the same BMI or WHR may vary in BF%, translating in differences in cardiometabolic risk[21,22].

To assess the genetic contribution to adiposity, we previously performed the first GWAS for BF% in nearly 40,000 individuals and identified two new loci (near *IRS1* and *SPRY2*), not identified in earlier large-scale GWAS for BMI[13]. Follow-up analyses of these loci provided strong evidence for *IRS1* to be involved in tissue-specific body fat storage and subsequent effects on cardiometabolic disease, such as type 2 diabetes (T2D) and coronary artery disease (CAD)[13]. While little is known about *SPRY2*, the Spry1 homolog in mice has been implicated in adipose tissue differentiation[23]. Taken together, these loci for BF% pointed towards new mechanisms involved in adipocyte metabolism that differ from the BMI-associated loci that suggested a role for the CNS[13,19].

Here, we have extended our study to include more than 100,000 individuals and continue to discover novel genetic loci associated with BF% that have not been identified before for any of the commonly studied adiposity traits[1–20]. Through an in-depth integrative characterization, including cross-trait association analyses, expression quantitative trait loci (eQTL), pathway and network analyses, regulome analyses and transgenic drosophila models, we show that these loci provide new insights into the biology that underlies adiposity and related cardiometabolic health, by specifically highlighting peripheral physiological mechanisms.

## Results

**Analyses in >100,000 individuals identify 12 loci for BF%.** In our primary meta-analysis, we combined results of genetic associations with BF% for up to 100,716 individuals from 43 GWAS ($n$ up to 76,137) and 13 MetaboChip studies ($n$ up to 24,582), predominantly of European ancestry ($n$ up to 89,297), but also of non-European ancestry ($n$ up to 11,419) populations (Supplementary Table 1 and Supplementary Fig. 1). As women have on average a higher BF% than men, we also stratified meta-analyses by sex ($n_{men}$ up to 52,416; $n_{women}$ up to 48,956). In secondary meta-analyses, we combined data from European-ancestry populations only ($n$ up to 89,297; $n_{men}$ up to 44,429; $n_{women}$ up to 45,525) to reduce genotypic and phenotypic heterogeneity that may have been introduced in the overall analyses by combining diverse ancestries.

In our primary meta-analysis of men and women combined, single-nucleotide polymorphisms (SNPs) in 10 independent loci reached genome-wide significance (GWS, $P<5\times10^{-8}$; Table 1 and Supplementary Fig. 2), including the three loci that we identified before[13]. Two additional loci, near *PLA2G6* and in *CRTC1*, were identified in men-specific and women-specific analyses, respectively (Table 1 and Supplementary Fig. 3). The European-ancestry-only analyses revealed the same loci, but no additional ones (Supplementary Tables 4–6, Supplementary Figs 4 and 5). We did not identify evidence of secondary signals at any of the 12 loci.

Two (near *IRS1* and *SPRY2*) of the 12 loci had been first identified in our previous genome-wide screen for BF% (ref. 13), and six loci (in/near *FTO*, *MC4R*, *TMEM18*, *TOMM40/APOE*, *TUFM/SH2B1* and *SEC16B*) had been first reported for association with BMI[1–6,10]. Four of the 12 loci, in or near *COBLL1/GRB14*, *IGF2BP1*, *PLA2G6* and *CRTC1*, have not been associated with an overall adiposity trait (such as BMI, BF%, obesity risk) before (Fig. 1 and Supplementary Fig. 6). Of note, the *COBLL1/GRB14* locus was previously established as a locus for body fat distribution independent of overall adiposity, assessed by $WHR_{adjBMI}$[11], and the *CRTC1* locus has been first reported for its association with age at menarche[24] (Table 2, Supplementary Table 7, See also 'Cross-phenotype association' section).

**Effect sizes and explained variance.** Index SNPs in the 12 established loci increase BF% by 0.024 to 0.051 s.d. per allele (equivalent to 0.16 to 0.33% in BF%, Table 1, Fig. 2). Given the high correlation between BF% and BMI, the BF% increasing alleles of each of the 12 loci are associated with increased

**Table 1 | Loci reaching genome-wide significance ($P<5\times10^{-8}$) for body fat percentage in all ancestry analyses, sorted according to significance in the overall analysis.**

| SNP | Chr. | Position (bp) | Nearest gene | Other nearby genes of interest | Fat% increasing allele | Fat% increasing allele frequency* | Other allele | All ancestry | | | | | All ancestry-men | | | | | All ancestry-women | | | | | Sex difference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Per allele change in body fat %* | | P | Explained variance | N | Per allele change in body fat %† | | P | Explained variance | N | Per allele change in body fat %† | | P | Explained variance | N | P |
| | | | | | | | | β | s.e. | | | | β | s.e. | | | | β | s.e. | | | | |
| rs1558902 | 16 | 52,361,075 | *FTO* | | A | 40% | T | 0.051 | 0.005 | 3.8E−27 | 0.125% | 99,328 | 0.051 | 0.0064 | 3.8E−15 | 0.122% | 51,498 | 0.050 | 0.0067 | 7.2E−14 | 0.120% | 48,486 | 0.96 |
| rs2943652 | 2 | 226,816,690 | *IRS1*‡ | | C | 36% | T | 0.034 | 0.005 | 1.5E−12 | 0.052% | 99,323 | 0.046 | 0.0065 | 1.3E−12 | 0.098% | 51,492 | 0.023 | 0.0068 | 5.6E−04 | 0.025% | 48,487 | 0.013 |
| rs6567160 | 18 | 55,980,115 | *MC4R* | | C | 25% | T | 0.034 | 0.005 | 1.3E−10 | 0.044% | 100,642 | 0.042 | 0.0072 | 6.1E−09 | 0.065% | 52,380 | 0.029 | 0.0076 | 1.1E−04 | 0.032% | 48,918 | 0.23 |
| rs6755502 | 2 | 625,721 | *TMEM18* | | C | 83% | T | 0.039 | 0.006 | 1.4E−10 | 0.043% | 99,855 | 0.027 | 0.0084 | 1.6E−03 | 0.020% | 51,778 | 0.052 | 0.0087 | 2.7E−09 | 0.075% | 48,733 | 0.034 |
| rs6738627 | 2 | 165,252,696 | *COBLL1* | *GRB14* | A | 37% | G | 0.030 | 0.005 | 5.7E−09 | 0.043% | 80,196 | 0.035 | 0.0073 | 1.9E−06 | 0.057% | 39,698 | 0.026 | 0.0072 | 3.8E−04 | 0.031% | 41,153 | 0.36 |
| rs693839 | 13 | 79,856,289 | *SPRY2*‡ | | C | 32% | T | 0.028 | 0.005 | 6.6E−09 | 0.035% | 100,190 | 0.034 | 0.0067 | 3.6E−07 | 0.050% | 51,906 | 0.021 | 0.0069 | 2.4E−03 | 0.019% | 48,940 | 0.17 |
| rs6857 | 19 | 50,084,094 | *TOMM40* | *APOE, APOC1 SH2B1, APOB48R,* | C | 83% | T | 0.048 | 0.008 | 6.8E−09 | 0.065% | 68,857 | 0.035 | 0.0112 | 1.8E−03 | 0.035% | 35,868 | 0.058 | 0.0118 | 7.3E−07 | 0.096% | 33,644 | 0.15 |
| rs4788099 | 16 | 28,763,228 | *TUFM* | *ATXN2L, SBK1, SULT1A2* | G | 38% | A | 0.027 | 0.005 | 1.2E−08 | 0.034% | 100,659 | 0.032 | 0.0064 | 6.7E−07 | 0.048% | 52,385 | 0.024 | 0.0067 | 3.6E−04 | 0.027% | 48,929 | 0.37 |
| rs9906944 | 17 | 44,446,419 | *IGF2BP1* | | C | 67% | T | 0.033 | 0.006 | 2.9E−08 | 0.049% | 74,338 | 0.025 | 0.0083 | 2.9E−03 | 0.027% | 38,242 | 0.036 | 0.0084 | 1.5E−05 | 0.059% | 36,751 | 0.31 |
| rs543874 | 1 | 176,156,103 | *SEC16B* | | G | 19% | A | 0.032 | 0.006 | 4.5E−08 | 0.031% | 100,705 | 0.028 | 0.0079 | 3.7E−04 | 0.024% | 52,410 | 0.037 | 0.0081 | 5.8E−06 | 0.042% | 48,951 | 0.43 |
| *Loci identified in sex-specific all-ancestry analyses* | | | | | | | | | | | | | | | | | | | | | | | |
| rs3761445 | 22 | 36,925,357 | *PLA2G6* | *PICK1* | G | 41% | A | 0.024 | 0.005 | 1.7E−07 | 0.029% | 99,614 | 0.037 | 0.0063 | 2.5E−09 | 0.068% | 51,687 | 0.017 | 0.0066 | 0.013 | 0.013% | 48,114 | 0.020 |
| rs757318 | 19 | 18,681,308 | *CRTC1* | | C | 50% | A | 0.024 | 0.005 | 2.1E−07 | 0.030% | 98,814 | 0.012 | 0.0064 | 0.054 | 0.008% | 51,484 | 0.037 | 0.0067 | 4.8E−08 | 0.067% | 47,986 | 0.0075 |

Chr., chromosome; positions (bp) according to Build 36; and allele coding based on the positive strand.
*Based on all-ancestry sex-combined analyses.
†Effects sizes are expressed in s.d., based on inverse normally transformed outcomes (mean 0, s.d. 1).
‡Loci first reported in the previous genome-wide association study of body fat percentage[13] (PMID:21706003).

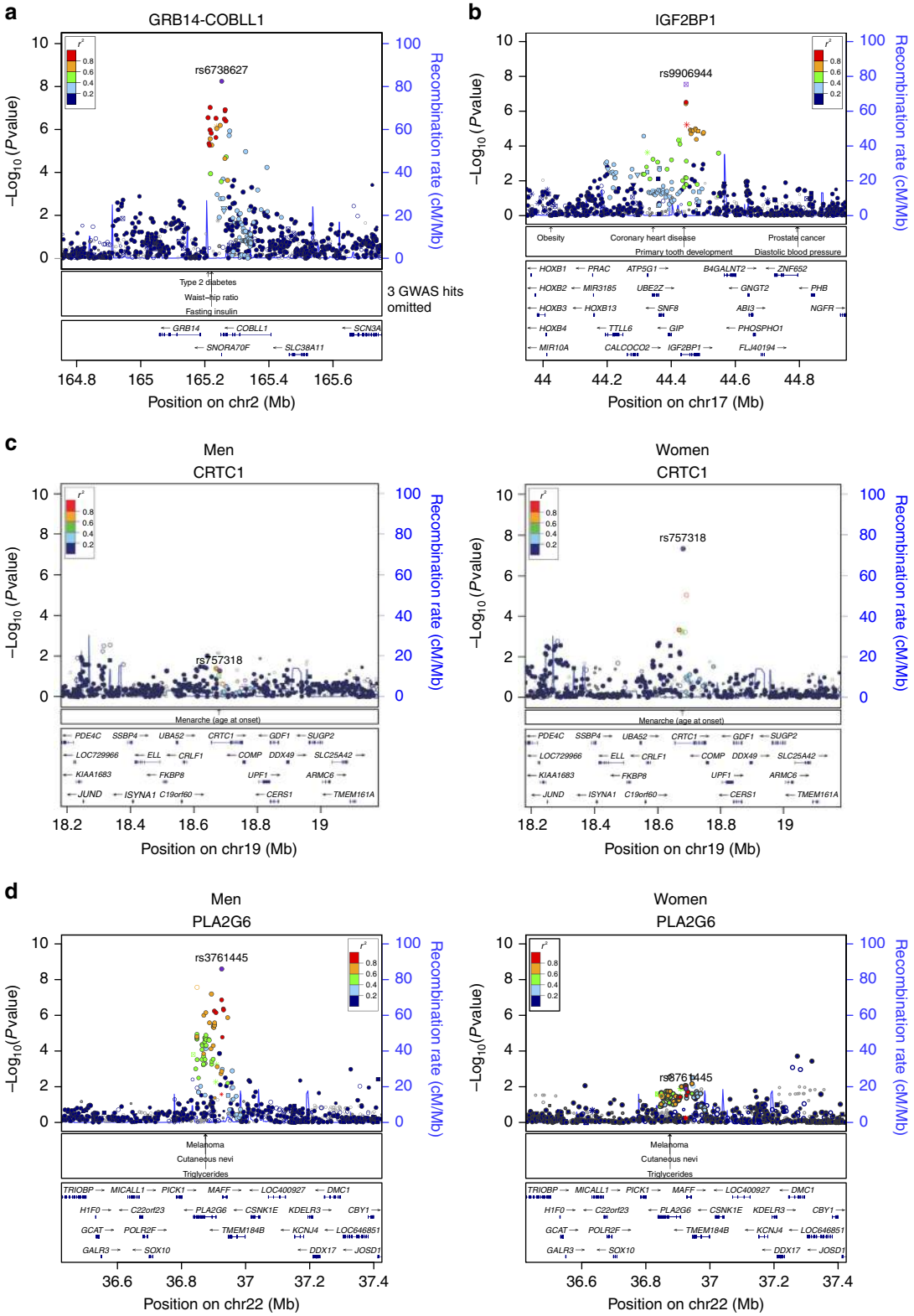


**Figure 1 | Regional plots of the four newly identified loci that reached genome-wide significant association with body fat percentage.** Regional plots of the four newly identified loci that reached genome-wide significant association with body fat percentage in all-ancestry analyses, in men and women combined for the *COBLL1/GRB14* and *IGF2BP1* loci (**a**,**b**), and separately for the *CRTC1* and *PLA2G6* (**c**,**d**). Each symbol represents the significance (*P* value on a −log10 scale) of a SNP with BF% as a function of the SNP's genomic position (NCBI Build 36). For each locus, the index SNP is represented in the purple colour. The colour of all other SNPs indicates LD with the index SNP (estimated by CEU $r^2$ from the HapMap Project data Phase II CEU). Recombination rates are also estimated from International HapMap Project data, and gene annotations are obtained from the UCSC Genome Browser. GWAS catalogues SNPs with *P* value $<5\times10^{-8}$ are shown in the middle panel. Different shapes denote the different categories of the SNPs: up-triangle for framestop or splice SNPs, down-triangle for nonsynonymous SNPs, square for coding or untranslated region (UTR) SNPs; star for SNPs in tfbscons region, square filled with 'X' symbol for SNPs located in mcs44placental region and circle for SNPs with no annotation information.

**Table 2 | Cross-phenotype associations: associations signatures of 12 established body fat percentage loci for anthropometric and cardiometabolic traits through look-ups in large-scale genetics consortia.**

| Nearby gene | | *FTO* | | *IRS1* | | *MC4R* | | *TMEM18* | | *COBLL1/GRB14* | | *SPRY2* | | *TOMM40/APOE* | | *TUFM/SH2B1* | | *IGF2BP1* | | *SEC16B* | | *PLA2G6/PICK1*† | | | | *CRTC1*† | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SNP | | rs1558902 | | rs2943652† | | rs6567160 | | rs6755502 | | rs6738627† | | rs693839 | | rs6857 | | rs4788099 | | rs9906944† | | rs543874 | | rs3761445 | | | | rs757318 | | | |
| Fat%-increasing allele (frequency %) | | A (40%) | | C (36%) | | C (25%) | | C (83%) | | A (37%) | | C (32%) | | C (83%) | | G (38%) | | C (67%) | | G (19%) | | G (41%) | | | | C (50%) | | | |
| Trait | Consortium (Max. *N*)‡ | Effect per fat% increasing allele | *P* | Effect per fat% increasing allele | *P* | Effect per fat% increasing allele | *P* | Effect per fat% increasing allele | *P* | Effect per fat% increasing allele | *P* | Effect per fat% increasing allele | *P* | Effect per fat% increasing allele | *P* | Effect per fat% increasing allele | *P* | Effect per fat% increasing allele | *P* | Effect per fat% increasing allele | *P* | Men | | Women | | Men | | Women | |
| | | | | | | | | | | | | | | | | | | | | | | Effect per fat% increasing allele | *P* | Effect per fat% increasing allele | *P* | Effect per fat% increasing allele | *P* | Effect per fat% increasing allele | *P* |
| Body fat percentage (INV) | Current study (100,705) | 0.051 | 3.8E−27 | 0.034 | 1.5E−12 | 0.034 | 1.3E−10 | 0.039 | 1.4E−10 | 0.030 | 5.7E−09 | 0.028 | 6.6E−09 | 0.048 | 6.8E−09 | 0.0269 | 1.2E−08 | 0.0333 | 2.9E−08 | 0.0315 | 4.5E−08 | 0.0374 | 2.5E−09 | 0.017 | 0.01 | 0.0123 | 0.054 | 0.0366 | 4.8E−08 |
| BMI (INV) | GIANT (339,148) | 0.081 | 1E−156 | 0.014 | 2.4E−06 | 0.056 | 6.7E−59 | 0.060 | 2.0E−53 | 0.011 | 6.1E−04 | 0.010 | 3.1E−03 | 0.021 | 1.0E−04 | 0.031 | 1.1E−24 | 0.010 | 0.018 | 0.050 | 2.3E−40 | 0.012 | 3.2E−03 | 0.006 | 0.14 | 0.013 | 2.2E−03 | 0.021 | 1.7E−07 |
| Circulating leptin (Ln ng ml$^{-1}$) | LEPgen (32,158) | 0.033 | 1.8E−07 | 0.020 | 1.9E−03 | 0.027 | 3.9E−04 | 0.026 | 1.7E−03 | 0.036 | 8.3E−07 | 0.012 | 0.097 | 0.026 | 0.016 | 0.027 | 3.9E−05 | 0.013 | 0.072 | 0.009 | 0.28 | 0.016 | 0.12 | 0.016 | 0.051 | −0.009 | 0.41 | −0.001 | 0.92 |
| Subcutaneous adipose tissue | SAT-VAT (10,557) | + | 6.2E−07 | + | 9.2E−04 | + | 0.093 | + | 6.1E−05 | + | 0.022 | + | 0.009 | + | 0.0037 | + | 0.036 | + | 0.70 | + | 0.097 | + | 0.13 | + | 0.22 | − | 0.26 | + | 0.11 |
| Visceral adipose tissue | SAT-VAT (10,557) | + | 4.6E−04 | + | 0.60 | + | 0.13 | + | 0.05 | + | 0.51 | + | 0.077 | + | 2.1E−04 | + | 0.36 | + | 0.24 | + | 0.53 | + | 0.017 | + | 0.09 | − | 0.66 | + | 0.12 |
| WHR$_{adjBMI}$ (INV) | GIANT (209,997) | 0.004 | 0.26 | 0.000 | 0.99 | −0.003 | 0.54 | −0.008 | 0.07 | −0.021 | 2.2E−08 | 0.002 | 0.55 | 0.024 | 1.2E−04 | 0.002 | 0.49 | 0.010 | 0.037 | −0.002 | 0.69 | 0.003 | 0.60 | −0.005 | 0.29 | −0.002 | 0.73 | −0.008 | 0.08 |
| Height (Z) | GIANT (253,217) | −0.010 | 1.2E−03 | −0.003 | 0.38 | 0.025 | 2.0E−12 | 0.006 | 0.16 | 0.001 | 0.72 | 0.007 | 0.038 | −0.006 | 0.17 | 0.002 | 0.42 | −0.016 | 1.1E−06 | 0.006 | 0.091 | 0.016 | 8.0E−04 | 0.010 | 0.015 | 0.005 | 0.31 | 0.003 | 0.53 |
| Triglycerides (INV) | GLGC (177,828) | 0.018 | 2.3E−06 | −0.027 | 1.3E−13 | 0.012 | 8.4E−04 | 0.008 | 0.027 | −0.017 | 3.3E−05 | −0.002 | 0.46 | −0.054 | 4.6E−19 | −0.002 | 0.57 | 0.003 | 0.84 | 0.004 | 0.21 | − | 2.5E−03 | − | 0.018 | + | 0.36 | + | 0.71 |
| HDL-Cholesterol (INV) | GLGC (187,131) | −0.018 | 2.7E−07 | 0.032 | 8.2E−17 | −0.026 | 2.9E−09 | −0.013 | 0.008 | 0.019 | 4.9E−05 | −0.001 | 0.91 | 0.067 | 2.6E−17 | −0.012 | 5.4E−04 | −0.012 | 0.025 | −0.011 | 0.018 | + | 0.054 | + | 0.044 | − | 0.32 | + | 0.56 |
| LDL-Cholesterol (INV) | GLGC (173,055) | −0.002 | 0.45 | −0.006 | 0.14 | 0.001 | 0.86 | −0.010 | 0.024 | −0.012 | 0.035 | 0.005 | 0.20 | −0.192 | 5.1E−110 | −0.003 | 0.41 | 0.003 | 0.39 | −0.010 | 0.068 | − | 0.43 | − | 0.51 | − | 0.40 | − | 0.76 |
| Fasting glucose (mmol l$^{-1}$) | MAGIC (120,901) | 0.006 | 0.004 | −0.004 | 0.084 | 0.006 | 0.030 | 0.006 | 0.031 | −0.001 | 0.58 | −0.001 | 0.83 | 0.010 | 0.012 | 0.000 | 0.92 | 0.002 | 0.59 | 0.005 | 0.044 | −0.002 | 0.50 | 0.000 | 0.89 | 0.003 | 0.26 | 0.003 | 0.37 |
| Fasting insulin (Ln pmol l$^{-1}$) | MAGIC (85,501) | 0.019 | 1.8E−12 | −0.015 | 3.8E−08 | 0.008 | 0.018 | 0.007 | 0.062 | −0.009 | 0.004 | 0.001 | 0.80 | 0.003 | 0.49 | 0.008 | 0.003 | 0.000 | 0.97 | 0.012 | 5.1E−04 | −0.009 | 0.02 | 0.000 | 0.95 | 0.007 | 0.08 | 0.008 | 0.021 |
| Type 2 diabetes (OR) | DIAGRAM (86,195) | 1.120 | 4.4E−21 | 0.920 | 4.7E−12 | 1.070 | 6.0E−07 | 1.040 | 0.005 | 0.940 | 2.3E−05 | 0.980 | 0.09 | 1.088 | 0.0014 | 1.020 | 0.11 | 1.051 | 7.7E−05 | 1.020 | 0.20 | 0.968 | 0.03 | 0.972 | 0.09 | 1.008 | 0.60 | 1.028 | 0.12 |
| Coronary artery disease (OR) | CARDIoGRAM plusC4D (213,938) | 1.025 | 0.008 | 0.971 | 8.8E−04 | 1.031 | 1.9E−03 | 1.028 | 0.018 | 0.982 | 0.063 | 0.994 | 0.51 | 0.899 | 5.9E−11 | 1.010 | 0.33 | 1.045 | 2.2E−06 | 0.999 | 0.96 | 1.011 | 0.33 | 0.998 | 0.88 | 1.010 | 0.50 | 1.029 | 0.17 |

CARDIoGRAMplusC4D, Coronary ARtery DIsease Genome-wide Replication and Meta-analysis (CARDIoGRAM) plus The Coronary Artery Disease (C4D) Genetics consortium; DIAGRAM, DIAbetes Genetics Replication And Meta-analysis consortium; GIANT, Genetic Investigation of ANthropometric Traits consortium; GLGC, Global Lipids Genetics Consortium; INV, inverse-normal transformation (mean of 0, s.d. of 1); LEPgen, circulating leptin consoritum (Kilpeläinen *et al.*, in preparation); Ln, natural logarithm-transformation; MAGIC, the Meta-Analyses of Glucose and Insulin-related traits Consortium; OR, odds ratio; SAT-VAT, subcutaneous adipose tissue (SAT)-visceral adipose tissue (VAT) consortium; WHRadjBMI, waist-to-hip ratio adjusted by BMI; Z, z-score transformation (mean of 0, s.d. of 1). The fat percentage (Fat%) increasing allele frequency was based on all-ancestry sex-combined analysis. The '+/−' in effect stands for increasing or decreasing phenotypes. The threshold for a statistically significant association with Bonferroni correction for 13 traits is $P = 0.00385$ (0.05/13). Colour coding of cells: BF%-increasing shows risk-increasing association with respective cardiometabolic traits at nominal (faded red) or multiple-testing corrected (solid red) significance. BF%-increasing shows risk-reducing association with respective cardiometabolic traits at nominal (faded green) or multiple-testing corrected (solid green) significance.
*Results of men and women combined are presented in Supplementary Table 8.
†The SNP of rs2943646 was used as a proxy for rs2943652 regarding coronary artery disease ($R^2 = 1$ and $D' = 1$); the SNP of rs2075650 was used as a proxy for rs6857 regarding CAD ($R^2 = 0.88$ and $D' = 1$); the SNP of rs4794018 was used as a proxy for rs9906944 regarding coronary artery disease and type 2 diabetes ($R^2 = 0.9$ and $D' = 1$).
‡The maximum sample size invoved in the 12 SNP assocation testing was reported from each respective consortium.

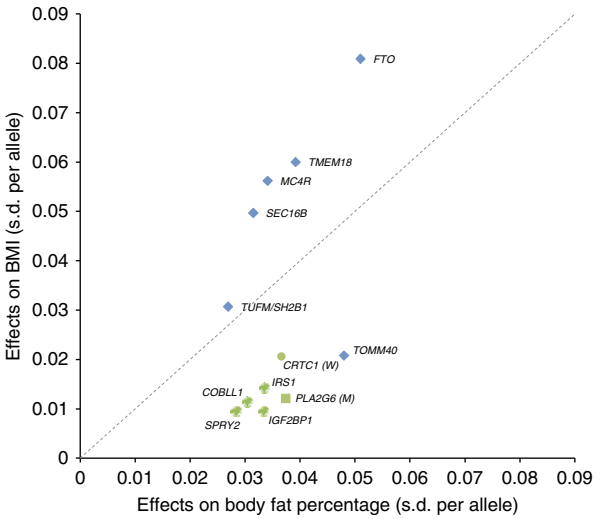


**Figure 2 | Comparison of effects of the 12 loci on body fat percentage (*x* axis) and on BMI (*y* axis).** Both outcomes (BMI and BF%) were inverse normally transformed (mean 0, s.d. 1) such that effects sizes are at the same scales and directly comparable. Effect sizes for BMI were obtained from Locke *et al.*[19]. The allele effects for the *PLA2G6* (square) and *CRTC1* (round) loci were derived, respectively, from the men- and women-based meta-analyses. Six loci had first been identified for BMI (blue), whereas six others were first identified for BF% (green).

BMI (Fig. 2, Table 2, and Supplementary Table 7). However, loci that had been previously identified for BMI, have larger effects (expressed in s.d. per allele) on BMI than on BF%, except the *TOMM40/APOE* locus, which has a substantially more pronounced effect on BF% than on BMI[25] (Fig. 2). The *TOMM40/APOE* locus, together with the loci previously (*IRS1* and *SPRY2*) and newly (*COBLL1/GRB14*, *IGF2BP1*, *PLA2G6* and *CRCT1*) identified for BF% all have larger effects on BF% than on BMI (Fig. 2). This division based on effect sizes, illustrated in Fig. 2, suggests that *IRS1, SPRY2, COBLL1/GRB14, TOMM40/APOE, IGF2BP1, PLA2G6* and *CRTC1* affect adiposity in particular, which is not fully captured by BMI (which represents both lean and fat mass).

Of the 12 loci, four showed significant sex-specific effects. For the loci near *IRS1* and *PLA2G6*, the effect in men was twice as large as in women, whereas for the *TMEM18* and *CRTC1* loci the effect was two- to threefold larger in women than in men (Table 1). As the European-ancestry-only populations represent the vast majority (90%) of the total sample, effects sizes from European only and all-ancestry analyses were similar (Supplementary Tables 5 and 8).

In aggregate, the 12 loci explained 0.58% of the variance in BF% in men and women combined. Because of the sex-specific effects of four loci, the explained variance was slightly higher, when estimated in men (0.62%) and women (0.61%) separately. Individually, the *FTO* locus explained the most variance of all identified loci (0.12%) (Table 1).

**Cross-phenotype association with cardiometabolic traits**. To gain insight in how the BF% loci affect anthropometric and cardiometabolic traits and comorbidities, we performed look-ups in the most recent large-scale GWAS meta-analyses from the GIANT (Genetic Investigation of ANthropometric Traits) consortium (WHR$_{adjBMI}$ and height)[20,26], the SAT-VAT consortium (abdominal visceral adipose tissue (VAT) and subcutaneous adipose tissue (SAT))[27], the LEPgen consortium (circulating leptin), the GLGC (high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C) and triglycerides (TG))[28], the MAGIC (fasting glucose and fasting insulin)[29], DIAGRAM (T2D)[30] and CARDIoGRAMplusC4D (CAD)[31]. To account for multiple testing, associations were considered statistically significant if *P* values were $< 5.2 \times 10^{-4}$ (Bonferroni-corrected $P = 0.05/96$ (12 SNP * eight trait groups)).

*Associations with anthropometric and adiposity traits*. The BF% increasing alleles for 11 of the 12 loci were associated with

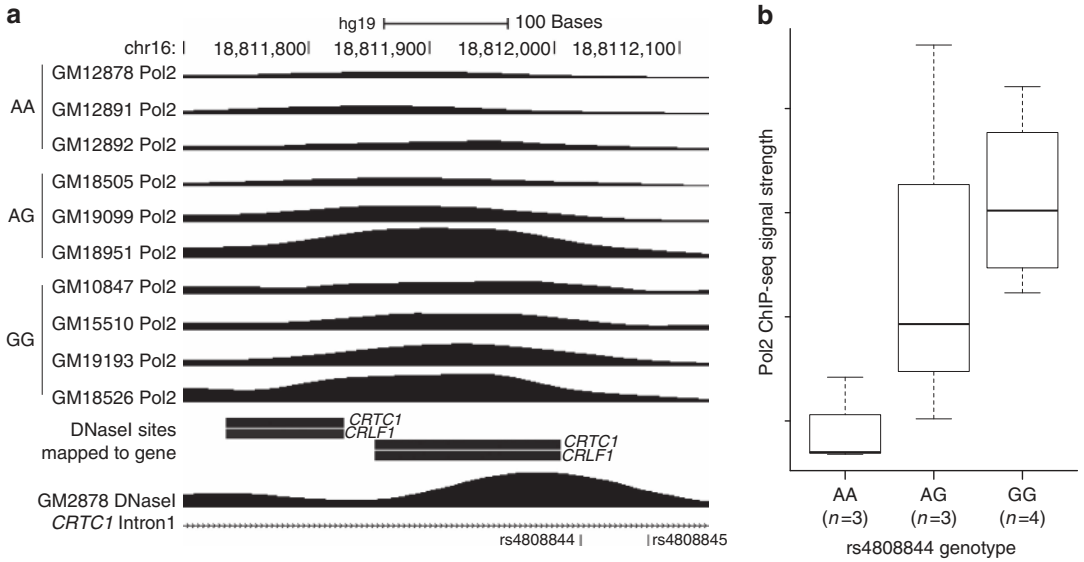


**Figure 3 | Genotype influences Pol2 binding at rs4808844.** (**a**) UCSC Genome browser track (hg19) of chromosome 16 displaying Pol2 binding signal in 10 lymphoblastoid cell lines (LCLs), grouped by genotype and correlations between DNaseI hypersensitivity and nearby gene transcription. (**b**) Binding signals from Pol2 ChIP-seq from 10 LCLs, grouped by genotype.

increased circulating leptin levels ($P_{binomial} = 0.006$), of which four reached statistical significance and another four were nominally significant (Table 2, Supplementary Table 7). These results are consistent with the notion that leptin is secreted by adipocytes proportional to adipose tissue mass.

The BF% increasing alleles of all 12 loci were associated with increased SAT and VAT ($P_{binomial} = 0.0005$), two (*FTO* and *TMEM18*) of which reached significance for association with SAT, and two (*FTO* and *TOMM40/APOE*) with VAT. The BF% increasing allele of the locus near *IRS1* was associated with a lower VAT/SAT ratio, indicative of a proportionally greater subcutaneous than visceral fat storage, as we have shown previously[13] (Table 2, Supplementary Table 7).

As expected, most of the identified BF% loci showed no association with $WHR_{adjBMI}$, as this trait, because of the adjustment for BMI, does not correlate with overall adiposity. Nevertheless, associations with $WHR_{adjBMI}$ for two loci (*COBLL1/GRB14* and *TOMM40/APOE*) did reach statistical significance. The *COBLL1/GRB14* locus was previously identified as a $WHR_{adjBMI}$ locus[11]. We show that it is the BF% increasing allele that is associated with lower $WHR_{adjBMI}$, suggestive of a preferential gluteal rather than abdominal fat storage. Although the *COBLL1/GRB14* association with $WHR_{adjBMI}$ is five times stronger in women than in men[11], we observed no sex difference for association with BF% (Table 1). For the *TOMM40/APOE* locus, it is the BF% increasing allele that is also associated with increased $WHR_{adjBMI}$, suggesting that the *TOMM40/APOE* locus increases abdominal and overall fat accumulation, at least in part, in an additive and independent manner. Furthermore, the BF% increasing allele was also significantly associated with increased VAT (Table 2, Supplementary Table 7) and liver fat storage ($P = 3.4 \times 10^{-4}$, $n = 5{,}550$, Methods section).

SNPs in three loci (*MC4R*, *PLA2G6* and *IGF2BP1*) showed significant association with height, two of which (*PLA2G6* and *IGF2BP1*) have not been reported in large GWAS studies before. Similar to the *MC4R* locus, the BF% increasing allele of the *PLA2G6* (rs3761445) was associated with greater adult height ($P = 6.7 \times 10^{-5}$; Table 2, Supplementary Table 7). Following up this variant in data from the Early Growth Genetics Consortium, we found that the BF% increasing allele was associated with higher birth weight ($P = 0.003$, $n$ up to 26,836; ref. 32) and greater prepubertal height ($P = 0.007$, $n = 13{,}948$; ref. 33), yet not with growth during or timing of puberty (Supplementary Table 10)[33]. In contrast, the BF% increasing allele in *IGF2BP1* (rs9906944) was associated with shorter height (Table 2, Supplementary Table 7), a cross-phenotype association pattern that is consistent with the effects of the GH/IGF1 axis[34]. SNPs in *IGF2BP1*, in linkage disequilibrium (LD) with rs9906944 ($r^2_{EUR} = 0.47$), have been previously implicated with primary tooth development in infancy[35]. Consistently, the BF% increasing allele of *IGF2BP1* (rs9906944) showed association with a later eruption of the first tooth ($\beta = 0.16$ months per allele; $P = 3.1 \times 10^{-8}$) and reduced number of teeth at 1 year ($\beta = -0.14$ number of teeth at age 1 year per allele; $P = 1.1 \times 10^{-7}$; ref. 35). Even though this suggests a role in maturation, we found no evidence for association with pre-pubertal height or pubertal growth and timing (Supplementary Table 10)[33] or age at menarche ($\beta = 0.01$ age of menarche (years) per allele; $P = 0.11$; ref. 24). Although this locus harbours a number of genes, data in rodents suggest that *IGF2BP1* might be a potential candidate gene driving the associations observed here, as Igf2bp1 knockout mice demonstrate fetal and postnatal growth retardation[36].

Taken together, alleles of each of the 12 loci are associated with increased BF%, yet their associations with other anthropometric traits differ, which in turn might result in varying impacts on cardiometabolic health.

*Associations with cardiometabolic traits*. Although phenotypic correlations observed in epidemiological studies have shown that increased adiposity is associated with increased cardiometabolic risk, the BF% increasing alleles of identified loci do not always associate with poorer health outcomes (Table 2 and Supplementary Table 11). For some loci, the BF% increasing allele may even have significant protective effects, as we have shown previously for the locus near-*IRS1* (ref. 13).

For the loci in/near *FTO, MC4R, TMEM18, TUFM/SH2B1 and SEC16B,* which were all five previously established for BMI, the observed cross-phenotype associations with cardiometabolic traits are generally directionally consistent with the phenotypic correlations. Specifically, their BF% increasing allele is typically associated with an unfavourable lipid profile and increased insulin resistance (Table 2, Supplementary Tables 12 and 13). These cross-phenotype associations translate in increased risk of

T2D and CAD and higher CRP levels, at least for the *FTO, TMEM18 and MC4R* loci (Fig. 2, Table 2, Supplementary Tables 9,12 and 13).

For the remaining seven loci, which all have a larger effect on BF% than on BMI (Fig. 2), the cross-phenotype associations are not always consistent with the phenotypic correlation between BF% and cardiometabolic traits. For example, the *COBLL1/GRB14* locus was previously identified for its association with fasting insulin[29], TG[37], HDL-C[37] and,T2D risk[30] (Table 2, Supplementary Tables 12 and 13). However, we show for the first time that it is the BF% increasing allele that is associated with a protective effect on cardiometabolic health; that is, with significantly lower TG levels and higher HDL-C levels, and a reduced risk of T2D (Table 2, Supplementary Tables 12 and 13). This association signature of the *COBLL1/GRB14* locus is consistent with the observation that its BF% increasing allele is associated with a lower $WHR_{adjBMI}$, corresponding to a proportionally lower abdominal and higher gluteal fat accumulation and, at nominal significance, with SAT but not with the metabolically more harmful VAT. The *COBLL1/GRB14* association signature is similar to that of the near-*IRS1* locus (Table 2, Supplementary Tables 12 and 13), and suggest that the beneficial cardiometabolic effects of the loci near *COBLL1/GRB14* and *IRS1* might be mediated through a favourable influence on body fat distribution, despite increased adiposity.

The BF% increasing allele of rs6857 near *TOMM40/APOE* is significantly associated with increased overall adiposity (BMI), abdominal adiposity ($WHR_{adjBMI}$), visceral adipose tissue (VAT) and liver fat storage, which may be mediating the nominally significant association with increased fasting glucose and risk of T2D (Table 2 and Supplementary Table 13). However, most notably, the BF% increasing allele was also highly significantly associated with a favourable lipid profile and reduced risk of CAD (Table 2 and Supplementary Table 12). The associations with lipid levels seem to be only partially driven by the nearby *APOE* locus for which previously highly significant associations with LDL-C[37], CRP[38] (both rs4420638), HDL-C and TG (rs439401; ref. 37) levels have been reported (Supplementary Fig. 7). These two SNPs (rs4420638, rs439401) are in low LD with each other ($r^2_{EUR}=0.13$, $D'_{EUR}=0.96$), and with the here-identified *TOMM40*-rs6857 ($r^2_{EUR}=0.39$, $D'_{EUR}=0.72$ and $r^2_{EUR}=0.06$, $D'_{EUR}=0.77$, respectively). Although the *APOE*-rs4420638 allele shows evidence of association with BF% ($P=3.9\times10^{-5}$), the association is completely abolished ($P=1.00$) after conditioning for *TOMM40*-rs6857 (Supplementary Table 14). The *APOE*-rs439401 SNP, previously associated with HDL-C levels, was not associated with BF% ($P=0.72$). Conversely, the *TOMM40*-rs6857 associations with TG ($P=4.5\times10^{-19}$; $P_{conditional}=3.6\times10^{-5}$) and HDL-C ($P=2.6\times10^{-17}$; $P_{conditional}=8.4\times10^{-14}$) remain significant after conditioning for the lipid-associated *APOE* SNPs (rs4420638, rs439401), whereas its association with LDL-C ($P=5.1\times10^{-110}$; $P_{conditional}=0.97$) is completely abolished after adjusting for the *APOE*-rs4420638 (Supplementary Table 14). Taken together, these observations show that associations of *TOMM40*-rs6857 are independent from the HDL-C and TG-associated *APOE*-rs439401 and partially independent from the LDL-C-associated *APOE*-rs4420638 (Supplementary Table 14). Another SNP (rs2075650) in this region, in high LD ($r^2_{EUR}=0.77$, $D'_{EUR}=0.96$) with the *TOMM40*-rs6857 and associated with BF% ($P=1.4\times10^{-7}$), has been previously identified for its association with Alzheimer's disease[39], cognitive function[40] and ageing[41], with the BF% increasing allele being associated with reduced risk of Alzheimer's disease, slower cognitive decline and increased longevity.

Although we do not observe association of *IGF2BP1*-rs9906944 with circulating lipid levels or glycemic traits, interestingly, the BF% increasing allele is significantly associated with increased risk of T2D and CAD, and with higher CRP levels (Table 2, Supplementary Tables 9,12 and 13).

The sex-specific effect of *PICK1/PLA2G6*-rs3761445 does not translate in sexual dimorphic associations with other traits (Table 2, Supplementary Tables 12 and 13). Interestingly, the BF% increasing allele is associated with a favourable lipid profile; in particular with lower TG levels ($P=8.1\times10^{-12}$) and higher HDL-C levels ($P=3.9\times10^{-6}$, Supplementary Table 12), but no association with CAD risk was observed (Supplementary Table 12). The *PICK1/PLA6G2*-rs3761445 is in moderate LD with SNPs identified before for nevus count (rs2284063, $r^2_{EUR}=0.67$, $D'_{EUR}=0.90$; ref. 42) and melanoma risk (rs738322, $r^2_{EUR}=0.77$, $D'_{EUR}=0.98$; refs 42,43). Consistently, the rs3761445 BF% increasing allele is associated with a lower number of cutaneous nevi ($-0.067$ nevi/allele, $P=9.4\times10^{-6}$; ref. 43) and reduced melanoma risk (OR = 0.86 per allele, $P=5.3\times10^{-10}$; ref. 44).

The BF% increasing allele of *CRTC1*-rs757318, which showed a significantly stronger association in women than men, was not associated with any of the cardiometabolic traits in either sex-stratified or sex-combined results. Rs757318 is in moderate LD ($r^2_{EUR}=0.57$, $D'_{EUR}=1$) with another *CRTC1* SNP (rs10423674) that was previously established for age at menarche[24] and, consistently, also the rs757318 BF% increasing allele was significantly associated with earlier age at menarche ($\beta=-0.03$ years per allele; $P=2.4\times10^{-10}$; ref. 24).

**Functional annotation of genome-wide significant loci**. The causal genes and/or variants underlying most of the BF% associated loci remain unknown. For the 12 genome-wide significant loci, and also for putative loci ($P<1\times10^{-5}$), we used multiple complementary approaches to prioritize candidate genes and/or variants and to elucidate the mechanisms involved in body fat regulation. These approaches include identification of nearby coding variants or copy-number variants (CNVs), *cis*-eQTL analysis, epigenetic marker and functional regulatory genomic element analysis, pathway and tissue enrichment analysis, and a transgenic *Drosophila* model.

*Coding variants and CNV analysis*. Among the 12 index SNPs, only rs4788099 near *SH2B1* was in high LD with seven coding variants ($r^2_{EUR}>0.7$) in nearby genes (*APOBR, SH2B1* and *ATP2A1*; Supplementary Table 15, Methods section). Two of these seven variants were non-synonymous, of which, one, Thr484Ala (rs7498665) in *SH2B1*, was in perfect LD with our index SNP. Thr484Ala shows a high degree of conservation, but was predicted to be functionally benign by PolyPhen and tolerated by SIFT. None of the other 11 index SNPs were in high LD with coding or CNVs.

*eQTL analysis*. We examined *cis*-associations between each index SNP and gene expression of transcripts within 1 Mb-region flanking the respective SNP (Supplementary Tables 16 and 17, Methods section). As shown previously[13], the BF% increasing allele of rs2943652 near *IRS1* is associated with increased IRS1 expression in omental and subcutaneous fat. SNPs within the same locus (LD $r^2_{EUR}>0.95$) have also been shown to be associated with increased IRS1 expression in skeletal muscle[45]. We also identified significant ($P<1\times10^{-5}$ or 5% FDR) eQTLs for other BF% associated loci, even after conditioning for the most significant SNP-transcript association in the regions. The BF% increasing allele of *COBLL1/GRB14*-rs6738627 is associated with lower expression of GRB14, whereas there is no evidence of association with COBLL1 expression. The BF% increasing allele

for *PLA2G6*/*MAFF*-rs3761445 is associated with lower expression of MAFF and TMEM184B in omental and subcutaneous fat. *TUFM*/*SH2B1*-rs4788099 is associated with the expression of a number of genes, such as TUFM (blood), APOBR (blood), SBK1 (blood), SULT1A2 (omental and subcutaneous fat) and SH2B1 (omental fat).

*Epigenetic marker and functional regulatory genomic element analysis*. We examined the overlap of 746 variants in LD ($r^2_{CEU} > 0.70$) with the 12 index SNPs with regulatory elements in brain, blood, liver, adipose and pancreatic islets from the ENCODE Consortium and Roadmap Epigenomic Projects (Supplementary Table 18). Across loci, 179 (24%) variants showed evidence of being located in a regulatory element as defined by overlapping variants in two or more data sets from the same tissue (Supplementary Table 19). Promoter variants, located within 2 kb of a transcription start site, overlapped with an average of 22 regulatory elements, while more distal variants (>2 kb) overlapped with an average of nine elements.

Two of the distal variants with the greatest amount of regulatory overlap were rs4808844 and rs4808845 (43 and 41 elements, respectively; Supplementary Table 19). These variants are located 58 bp apart in intron 1 of *CRTC1* and overlap evidence of open chromatin, histone marks that are characteristic of active transcription regulation and Pol2 binding (Fig. 3a). We found that rs4808844 was significantly associated ($P = 0.036$) with Pol2 binding signal strength (Fig. 3b). In addition, DNaseI hypersensitivity signal in this region has been shown to negatively correlate with *CRTC1* and *CRLF1* transcription levels across many cell types[46]. These data suggest that rs4808844 and rs4808845, both in high LD ($r^2_{CEU} = 0.76$ and 0.79, respectively) with our index SNP (rs757318), may influence the transcription of these and/or other nearby genes.

We further characterized variants overlapping with regulatory elements at each of the 12 loci using RegulomeDB, and two loci stood out. In the *TUFM-SH2B1* region, three SNPs (rs4788084, rs1074631 and rs149299) in LD ($r^2_{CEU} = 0.82$, 0.76 and 0.75, respectively) with rs4788099 are located in an EBF1-binding protein ChIP-seq signal in lymphoblastoid cells. In addition, rs4788084 is located within an EBF1-binding motif. EBF1 is involved in the thalamic axon projection into the neocortex[47] and the genetic variants around rs4788099 might affect the regulation of EBF1 of the nearby *SH2B1* (ref. 48). In the *PLA2G6/PICK1* region, rs4384 in LD with rs3761445 ($r^2_{EUR} = 0.73$) overlapped with more elements (50 elements in four tissues, Supplementary Table 19) than any other distal variant. This variant is located in a HEN1-binding motif with evidence of a DNase footprint in multiple cell types (Supplementary Fig. 8). HEN1 is a transcription factor potentially involved in the CNS development[49].

*Pathway, network and tissue-enrichment analysis*. To test for enrichment and define pathways and networks between the genes harboured by the 12 GW-significant loci and 31 loci with putative evidence ($P < 1 \times 10^{-5}$) of association with BF%, we applied a number of approaches (see Methods section). Neither DEPICT (data-driven enrichment prioritized integration for complex traits)[50] nor Ingenuity IPA identified pathways, tissues or networks that were significantly enriched among the genes across the 43 loci (Supplementary Tables 20–22). Also, GRAIL (Gene Relationships Among Implicated Loci), which searches the published literature to identify relationships between genes, and DAPPLE (Disease Association Protein–protein Link Evaluator), which tests for protein–protein interactions, did not identify significant connection between any of the genes in the identified loci. Their limited power may be due to the relatively small number of loci identified in this meta-analyses or to limited knowledge related to adipogenesis[51].

**Experimental follow-up of candidate genes in *Drosophila***. We used *Drosophila* as a fast and inexpensive model to help prioritize which genes within the identified loci are the most likely candidates to underlie the observed associations.

To gain first insights in the potential candidacy of the genes located within the 12 BF% associated loci, we performed a look-up in data from a genome-wide transgenic RNAi screen for fat content in adult *Drosophila*[52]. In that screen, whole-body TG, also in *Drosophila* the major lipid storage form, were used as a direct measure of fly adiposity upon activation of a heat shock-inducible Hsp70-GAL4 system. As such, transgenic fly lines were made to test the adiposity regulating potential of 10,489 of the ~14,000 annotated *Drosophila* protein coding genes. Of the 80 genes located within a 1 Mb-window of each of the 12 index SNPs, 44 *Drosophila* orthologues were available, yet, 12 of these 44 transgenic RNAi fly lines were too weak to be screened. Of the remaining 32 fly lines, 15 fly lines had substantially lower (>2 s.d. less) whole-body TG than the wild-type flies, whereas five fly lines showed higher TG (>2 s.d. more) (Supplementary Table 23). Next, we selected one to three candidate genes within each of the 12 loci based on their potential role in adipocyte metabolism. We knocked down their corresponding orthologues in *Drosophila* that were subsequently exposed to a high-sugar diet (Supplementary Table 24), as described before[53]. Both *Drosophila* experiments pinpoint the *SPRY2* (or *sty*) as the potential causal gene within the locus; that is, knockdown flies for *sty* have significantly lower whole-body TG levels than wild-type flies. While the genome-wide transgenic RNAi screen pointed towards the *CRTC1* gene in the *CRTC1* locus, we could not confirm a role for *CRTC1* in the knockdown experiment.

**Established loci and body fat percentage**. The most recent GWAS meta-analysis for BMI, including nearly 340,000 individuals, identified 97 loci that reached GWS[19]. Each of the 97 BMI-associated SNPs showed directionally consistent association with BF% ($P_{binomal} < 1 \times 10^{-4}$), 71 of which also reached nominal statistical significance (Supplementary Table 25). One of the reasons for the non-significance for the remaining loci might be insufficient power as the current final meta-analysis sample size for BF% was only one-third of that for BMI.

Of the 12 loci previously identified through GWAS for extreme and early-onset obesity[7,12,54,55], 11 showed directionally consistent association with BF% ($P_{binomal} < 0.006$), of which five also reached nominal statistical significance (Supplementary Table 25).

## Discussion

Our meta-analysis of data from more than 100,000 individuals identified 12 loci significantly associated with BF%. While a recent GWAS including more than 340,000 individuals reported nearly 100 loci associated with BMI, a commonly used proxy measure for overall adiposity, four (*SPRY2*, *IGF2BP1*, *PLA2G6* and *CRTC1*) of the 12 BF% associated loci did not reach GWS for BMI, despite the enormous sample size[19]. This observation most likely reflects the heterogeneity of BMI as a marker of overall adiposity and emphasizes the increased statistical power of more precisely measured phenotypes.

The 12 BF% associated loci divide into two distinct groups. The first group comprises the five loci (*FTO*, *MC4R*, *TMEM18*, *SEC16B* and *SH2B1*) of which the association is stronger with BMI than with BF%, suggesting that they affect both fat mass and lean mass. All five loci have been identified and described in detail before in relation with BMI[5,10,19]. Their associations with cardiometabolic outcomes are predictable, reflecting the phenotypic correlations with BF%; that is, their BF% increasing

alleles are associated with an unfavourable glycemic and lipid profile and with an increased risk of T2D and CVD.

The second group, comprising the remaining seven loci (*IRS1*, *SPRY2*, *TOMM40*/*APOE*, *CRCT1*, *PLA2G6*, *IGB2BP1* and *COBLL1*/*GRB14*), all show a more pronounced effect on BF% than on BMI, suggesting a specific effect on adiposity rather than on overall body mass. Most notably, the association patterns with cardiometabolic traits of this group of loci, as opposed to the first group, often do not reflect the phenotypic correlations. For example, as we have described before, the BF% increasing allele of the index SNP 500 kb upstream of *IRS1*, which affects IRS1 expression, is associated with a favourable cardiometabolic risk profile, including a reduced risk of T2D and CVD[13]. We showed that this association signature, which goes against the phenotypic correlations, could be explained by an effect on fat distribution, as the BF% increasing allele was associated with increased subcutaneous, but not with the metabolically more harmful visceral fat[13]. The locus between *GRB14* and *COBLL1* shows a similar association signature. In fact, this locus was first described for its association with a lower $WHR_{adjBMI}$[11] and reduced risk of T2D[30]. Here, we show that the same allele is associated with increased BF%, suggesting that the association with $WHR_{adjBMI}$ likely reflects a proportionally greater fat accumulation at hip and thighs rather than at the waist. Although this locus requires further experimental follow-up, current observations point towards *GRB14* as the candidate gene in this locus. *GRB14* encodes a protein that binds directly to the insulin receptor (IR), and the BF% increasing allele of the index SNP is associated with reduced GRB14 expression in adipose tissue. This is consistent with previous observations showing that Grb14/GRB14 expression is increased in adipose tissue of insulin-resistant rodents and in obese patients with T2D[56]. Furthermore, Grb14-deficient mice show improved glucose homeostasis and enhanced insulin action through increased IR-mediated IRS1 phosphorylation in the liver and skeletal muscle[57]. The similar cross-phenotype association signatures of the *IRS1* and *GRB14/COBLL1* loci might be a reflection of the close interaction between IRS1 and GRB14 in the IR-signalling pathway.

The BF% increasing allele of the *PLA2G6* locus is associated with lower insulin and TG levels and reduced T2D risk, particularly in men. *PLA2G6* is the nearest gene and encodes a calcium-independent phospholipase A2 involved in the hydrolysis of phospholipids. However, this locus harbours a number of other genes that would make plausible candidates for driving the cross-phenotype associations, including *PICK1*, which is membrane sculpting BAR domain protein. PICK1-deficient mice and flies display marked growth retardation, which at least in mice, might be due to impaired storage and secretion of growth hormone from the pituitary and possibly insulin from the pancreas[58]. PICK1-deficient mice, despite their smaller size, demonstrate increased body fat and reduced lean mass, reduced TG levels and impaired insulin secretion, which was compensated by increased insulin sensitivity[58]. Given the locus' association with nevus count, SOX10, which encodes a member of the SOX (SRY-related HMG-box) family of transcription factors, is another candidate gene in this locus. SOX genes are involved in the regulation of embryonic development and *SOX10* in particular is important for the development of neural crest and peripheral nervous system. Mutations in *SOX10* have been implicated in uveal melanoma and Waardenburg syndrome, which presents with pigmentation abnormalities and hearing loss, and Kallmann syndrome, which presents with failure to start or complete puberty and hypogonadotropic hypogonadism (short stature, absence of puberty and sex hormones, among others) and absence of smell[59,60]. The phenotype similarity of these syndromes and the association signature may suggest that *SOX10* could be driving the associations observed for the *PLA2G6* locus.

The *TOMM40*/*APOE* locus is another locus with an intriguing association signature; while the BF% increasing allele has an unfavourable effect on glycemic traits and T2D risk, it is associated with a favourable lipid profile and reduced risk of CVD. The high LD in this region poses a major challenge to elucidate whether the association with lipid traits is due to a 'spillover' effect from nearby lipid-associated loci in *APOE*. Using conditional analyses, we provide evidence suggesting that at least the association with lower TG and high HDL-C levels might be distinct from previously reported loci. Of interest is that the BF% increasing allele seems to be associated with markers of increased longevity[41].

The *CRTC1* locus is another gene-rich locus, but given the epigenetic marks in this gene and data from animal models, *CRTC1* poses to be a good candidate gene. *CRTC1* is primarily expressed in the brain, and it may affect leptin anorexic effect in the hypothalamus[61]. *CRTC* knockout mice demonstrated hyperphagia, increased white adipose tissue and infertility[61].

Our meta-analysis was limited by the fact that participating studies all had imputed HapMap reference panels for autosomal chromosomes and that the analysis model assumed additive effects. Future discovery efforts based on genome-wide imputation of 1000 Genomes reference panels, that include X- and Y-chromosomes and that also test recessive and dominant inheritance, will allow for the discovery of more and lower-frequency variants and for refining association signatures of already established BF%-associated loci.

Taken together, our expanded genome-wide meta-analyses of BF% has identified a number of loci with distinct cross-phenotype association signature that, together with our functional follow-up analyses, facilitated the identification of strong positional candidates. Particularly striking is that two of the 12 loci harbour genes (*IRS1*, *GRB14*) that influence insulin receptor signalling, and two other loci contain genes (*IGF2BP1*, *PICK1)* that are involved in the GH/IGF1 pathway, that in turn also relates to insulin receptor signalling.

## Methods

**Discovery of new loci.** *Study design*. A two-stage meta-analysis was performed to identify loci associated with BF%. In Stage 1, we conducted two parallel meta-analyses; one meta-analysis combined summary statistics from 43 GWAS, totalling up to 76,137 adult individuals (65,831 European ancestry, 7,557 South Asian ancestry, 2,333 East Asian ancestry and 416 African Americans), and the other meta-analysis combined summary statistics from 13 additional studies genotyped using the Metabochip, totalling up to 24,582 individuals (23,469 Europeans and 1,113 African Americans). In Stage 2, we combined the GWAS meta-analysis results and Metabochip meta-analysis results from Stage 1 (Supplementary Table 1 and Supplementary Figs 1 and 2) in one final meta-analysis, including 100,716 individuals from 56 studies. All the studies were approved by their local institutional review boards and written consent was obtained from all the study participants.

Although our primary analysis, described above, combined all the data available to us, in the secondary analyses, we conducted stratified analyses for (1) all-ancestry men-only, (2) all-ancestry women-only, (3) European ancestry, (4) European ancestry men-only and (5) European ancestry women-only (Supplementary Tables 4–6 and Supplementary Figs 2–5).

*Phenotype*. BF% in each cohort was measured either with bioimpedance analysis (BIA) or dual energy X-ray absorptiometry (DEXA) as described in detail before[13]. For each study, BF% was adjusted for age, $age^2$ and study-specific covariates (for example, genotype-based principle components, study centre and others), if necessary. For studies of unrelated individuals, the residuals were calculated separately in men and women, and in cases and controls. For studies of family-based design, the residuals were calculated in men and women together, and sex was additionally adjusted in the model. The residuals were then inverse normally transformed for association testing. For studies of family-based design, the family relatedness was additionally adjusted in the association testing.

*Sample quality control, imputation and association*. Each study did the study-specific quality control (QC) (Supplementary Table 2). The GWAS common SNPs were imputed in each study using the respective HapMap Phase II (Release 22)

reference panels (EUR for studies of European-ancestry populations, CHB + JPT for studies of Eastern Asian ancestry populations, and CEU + YRI + CHB + JPT for studies of Indian Asian ancestry populations and African American populations). Individual SNPs were associated with inverse normally transformed BF% residuals using linear regression with an additive model. All the SNPs with low imputation scores (MACH $r^2$-hat $<0.3$, IMPUTE proper_info $<0.4$ or PLINK info $<0.8$) and a MAC $\leq 3$ were removed. The EasyQC software was used for detailed QC of study level analyses and meta-level analysis, as described elsewhere[62].

*Meta-analysis*. Meta-analyses were performed using inverse variance-weighted fixed-effect method in METAL. Inflation before genomic control (GC)-correction was generally low in all-ancestry ($\lambda_{men+women}=1.13$; $\lambda_{men}=1.07$; $\lambda_{women}=1.09$) and European-only ($\lambda_{men+women}=1.13$; $\lambda_{men}=1.07$; $\lambda_{women}=1.10$) analyses. To reduce the inflation of the test statistics from potential population structure, individual GWAS results and GWAS meta-analysis results were corrected for GC using all SNPs. Individual Metabochip results and Metabochip meta-analysis results were GC-corrected using 4,425 SNPs, which are derived from pruning of QT-interval replication SNPs within 500 kb of an anthropometry replication SNP on the Metabochip. The GC-corrected GWAS and Metabochip meta-analysis results were finally meta-analysed (Supplementary Fig. 1).

Using the LD score regression method in the European-only meta-analyses suggests that the observed inflation is not due to population substructure[63]. The regression intercept, which estimates inflation after removing polygenic signals, was 1.0045 (with $\lambda_{GC}=1.136$ and mean $\chi^2=1.16$) for sex-combined, 0.999 ($\lambda_{GC}=1.062$ and mean $\chi^2=1.079$) for men-only and 1.014 ($\lambda_{GC}=1.105$ and mean $\chi^2=1.112$) for women-only analyses. Using these regression intercepts, rather than the $\lambda_{GC}$, to correct our meta-analyses, results in more significant associations (for example, for the rs1558902-*FTO* SNP, $P=3.24\times10^{-27}$ in the modified European sex-combined meta-analysis compared with $P=1.1\times10^{-25}$ (Supplementary Table 6)). Overall, however, the less stringent correction did not result in the identification of novel loci.

*Identification of novel loci*. Each unique locus was defined as $\pm$500 kb on either side of the most significant SNP that reached a GWS threshold ($P<5\times10^{-8}$) in the meta-analysis. These GWS-index SNP loci from the primary analysis as well as from secondary analyses were highlighted for further analyses (Table 1 and Supplementary Tables 4–6). The genotype data for the genome-wide significant SNPs was of high quality with a median imputation score of $\geq$0.95 (Supplementary Table 26). The fifth percentile for all SNPs was $\geq$0.80, except for the previously established *TOMM40* SNP (P5 = 0.52).

*Joint and conditional multiple SNP association analysis*. We used the GCTA approach to identify potential additional signals in regions of GWS-index SNP. This approach uses summary meta-analysis statistics and a LD matrix from an ancestry-matched sample to perform approximate joint and conditional SNP association analysis. Although our primary analyses were based on all ancestry populations, the 12 GWS-index SNPs were strongly associated with BF% in European populations, 6 of them reaching the GWS (Supplementary Table 5). The estimated LD matrix based on 6,654 unrelated individuals of European ancestry in ARIC cohort was used in the analysis.

*Heterogeneity among studies*. The potential heterogeneity in the effect estimates for our GWS-index SNPs were investigated between men and women in all-ancestry populations and in European populations, and between individuals of European ancestry and individuals of all ancestry. We also tested for heterogeneity between results from studies that used BIA for BF% assessment and that used DEXA. Heterogeneity was assessed using a *t*-statistic, $t=(\beta_1-\beta_2)/(se_1^2+se_2^2-2*r*se_1*se_2)^{1/2}$ to account for relatedness, where $\beta_1$ and $\beta_2$ are the effect size estimates, $se_1$ and $se_2$ are the corresponding standard errors and $r$ is Spearman's correlation coefficient of beta values between men and women or between European ancestry and all ancestry.

*Variance explained*. The variance explained by each GWS-index SNP was calculated using the effect allele frequency ($f$) and beta ($\beta$) from the respective meta analyses using the formula[6] of Explained variance $=2f(1-f)\beta^2$.

**Cross-trait association lookups**. *Cardiometabolic consortia*. To explore the relationship between BF% and an array of cardiometabolic traits and diseases, the association results for the 12 GWS-index SNPs were requested from seven primary cardiometabolic genetic consortia: the LEPgen consortium (circulating leptin, Kilpeläinen *et al.*, in preparation), VATGen consortium[27], GIANT (BMI, height and $WHR_{adjBMI}$)[19,20,26], GLGC (HDL-C, LDL-C, TG, TC)[28], MAGIC[29], DIAGRAM (T2D)[30] and CARDIoGRAMplusC4D (CAD)[31]. On the basis of known correlations among these cardiometabolic traits, we considered circulating leptin levels, abdominal adipose tissue storage, height, $WHR_{adjBMI}$, plasma lipid levels, plasma glycemic traits, T2D and CAD as eight independent trait groups. In addition, the associations for these 12 SNPs were also looked up in four consortia that examined phenotypes more distantly related to BF%: ADIPOGen (BMI-adjusted adiponectin)[64], ReproGen (age at menarche)[24], liver enzyme meta-analysis[65] and CRP meta-analysis[38]. For certain GWAS-index SNPs, we also did specific lookups: rs6857 association in liver fat storage, rs3761445 associations in cutaneous nevi and melanoma risk meta-analysis[42–44], early growth genetics (birth weight[32] and pubertal height[33]), insulin-like growth factor 1 meta-analysis (Teumer *et al.* under review) and CHARGE testosterone meta-analysis[66], and rs9906944 associations in tooth development meta-analysis[35] and Early Growth Genetics Consortium (birth weight[32] and pubertal height[33]).

*NHGRI GWAS catalogue lookups*. We manually curated and searched the National Human Genome Research Institute (NHGRI) GWAS Catalogue (www.genome.gov/gwastudies) for previously reported associations for SNPs within 500 kb and $r^2>0.7$ (1000 Genomes Pilot1 EUR population based on SNAP: http://www.broadinstitute.org/mpg/snap/ldsearch.php) with each of the 12 GWS-index SNPs. All previously reported associations that reached $P<5\times10^{-8}$ were retained (Supplementary Table 11).

**Coding variants and CNVs.** To determine whether any of our 12 GWS-index SNPs might be tagging potentially functional variants, we identified all variants within 500 kb and in LD ($r^2>0.7$, HapMap release 22/1000 Genomes Pilot1 EUR) with our GWS-index SNPs. As such, we identified 776 variants and annotated each of them using Annovar (http://www.openbioinformatics.org/annovar/). The predicted functional impacts for coding variants were accessed via the Exome Variant Server (http://evs.gs.washington.edu/EVS/) for PhastCon, Grantham, GERP and PolyPhen, and were also from SIFT (http://sift.jcvi.org/). To determine whether any of the 12 GWS-index SNPs tagged ($r^2>0.7$) CNVs, all genetic variants (SNV, Indel and SVS) within a 1 Mb window of the index SNPs from the 1000 Genomes Project EUR population (Phase 1) were downloaded. The LD indexes were calculated between each of the 12 GWS-index SNPs and any nearby CNV variants.

**Analyses of eQTLs.** The *cis*-associations between 12 GWS-index SNPs and expression of nearby genes ($\pm$500 kb of the index-SNP) were examined in the whole blood ($n=2{,}360$) from the eQTL meta-analysis study[67], the abdominal fat tissue ($n=742$ for omental fat and $n=610$ for subcutaneous fat) from the bariatric surgery study[68], the abdominal subcutaneous fat tissue ($n=54$) and gluteal subcutaneous fat tissue ($n=65$) from the MolOBB study[69], and the brain tissue from the cortical brain study ($n=193$; ref. 70). Conditional analyses were conducted by including both GWS-index SNP and the most significant *cis*-associated SNP for the given transcript in the model to examine whether observed associations were driven by our GWS-index SNP or by other nearby variants. Conditional analyses were conducted for all tissues except the brain tissue.

**Regulatory annotation using ENCODE and Roadmap.** *Regulatory element overlap*. We identified variants in LD ($r^2>0.7$, 1000 Genomes Project Pilot, EUR) with each of the 12 GWS-index SNPs and tested for overlap between these variants and elements from regulatory datasets. In total, 746 variants at the 12 GWS-index loci were examined for overlap with regulatory elements in 181 data sets (Supplementary Tables 18 and 19) from five tissues (blood, brain, liver, adipose tissue and pancreatic islets). These data sets, downloaded from the ENCODE Consortium and Roadmap Epigenomics Projects, identify regions of open chromatin (DNase-seq, FAIRE-seq), histone modification signal enrichment (H3K4me1, H3K27ac, H3K4me3, H3K9ac and H3K4me2), and transcription factor binding in cell lines and tissues believed to influence BF%. When available, we downloaded data processed as a part of the ENCODE Integrative Analysis. Roadmap Epigenomics sequencing data were processed with MACS2 and the same irreproducible discovery rate pipeline used in the ENCODE Integrative analysis when multiple data sets were available, or MACS2 alone when only a single replicate was available.

*Pol2 binding*. We tested for correlation between Pol2 binding strength and genotype in lymphoblastoid cell lines at two SNPs, rs4808844 and rs4808845 that are in LD with GWS-index SNP of rs757318 in *CRTC1*. Pol2 binding data uniformly processed as part of the ENCODE Integrative analysis were download for 10 lymphoblastoid cell lines (GM10847, GM12878, GM12891, GM12892, GM15510, GM18505, GM18526, GM18951, GM19099, GM19193). We examined the alleles present at these variants in Pol2 ChIP-seq alignment BAM files to determine sample genotypes and compared these with genotypes generated by the 1000 Genomes Project for the same samples. For the eight samples also genotyped by the 1000 Genomes Project, genotype calls were 100% concordant. Correlation between genotype and Pol2 binding signal at each SNP was calculated in R using a linear model (signal ~ genotype).

*RegulomeDB annotation*. We further characterized the variants at selected loci using the web-based tool RegulomeDB (http://regulomedb.org/). The reference sequence identifiers of variants that overlap two or more regulatory elements in the same tissue were used to conduct the RegulomeDB search.

**Pathway, network and tissue-enrichment analysis.** To define pathways, networks and tissue enrichment, we extended the list of genome-wide significant loci to also include loci that showed putative ($P<1\times10^{-5}$) association with BF% (using the same criteria described above to define independent loci). As such loci, represented by 43 index SNPs, were considered for gene prioritization, pathway enrichment (DEPICT, Ingenuity Pathway Analyses), gene relationship analysis (GRAIL) and protein–protein interaction analyses (DAPPLE).

*Data-driven enrichment prioritized integration for complex traits*. Details of this method are provided in Pers *et al.*[50] DEPICT is designed to systematically identify the most likely causal gene at a given locus, to test gene sets for enrichment for genetic associations, and to identify tissues and cell types in which genes from associated loci are highly expressed.

DEPICT assigned genes to the 43 associated loci if the genes resided within the associated LD region ($r^2>0.5$) of a given associated SNP. After merging overlapping regions and discarding regions that mapped within the extended major histocompatibility complex locus, we were left with 42 non-overlapping regions that covered a total of 82 genes. We then used DEPICT to test enrichment at these loci for a total of 14,461 reconstituted gene sets, and for 209 tissue and cell type annotations.

*Ingenuity pathway analyses*. We used HaploReg v2 (http://www.broadinstitute.org/mammals/haploreg/haploreg.php) and adopted a stringent LD ($r^2>0.8$ in 1000 Genome phase 1 EUR) to extract all the nearby genes (88 genes in total) of the index SNPs based on both GENCODE and RefSeq. For 65 out of them, they were successfully mapped to the Ingenuity Knowledge Base, and those unmapped genes are mainly lincRNA, miRNA, antisense or processed transcript genes derived from GENCODE. The 65 genes were incorporated into Ingenuity Canonical pathway enrichment analysis. The *P* values are calculated based on Fisher's right-tailed exact test. The default settings were used for Ingenuity Interaction network analysis.

*Gene relationships among implicated loci*. The GRAIL was used to examine relationships between genes. For each query and seed SNP, we adopted the default methods implemented in GRAIL to extract the genes around each index SNP: that is, (1) we first identified neighboring SNPs in the 3′ and 5′ direction in LD ($r^2>0.5$, CEU HapMap), proceeding outwards in each direction to the nearest recombination hotspots to define an interval region, and extracted all the genes in this interval; (2) if there are no genes in that interval region, the interval is extended an additional 250 kb in either direction. The 12 GWS-index SNP regions were input as seed regions, and the regions for the remaining 31 SNPs were input as query regions. Connections between genes were inferred from textual relationships based on published scientific text using PubMed abstracts as of December 2006. The significant gene similarity was declared based on $P_{GRAIL}<0.01$.

*Disease association protein–protein link evaluator*. The DAPPLE package was used to examine the potential encoded protein–protein interaction evidence for the genes located in the 43 associated loci. Genes from 32 of the 43 loci were annotated in the high-confidence pair-wise interaction InWeb database. Both the direct and indirect interactions were considered. The running settings were 1,000 permutation, common interactor binding degree $=2$, and 110 kb upstream and 40 kb downstream to define a gene' residence.

***Drosophila* knockdown experiments.** *Genome-wide screen*. We first identified all genes within $\pm 500$ kb of the 12 GWS-index SNPs, and subsequently identified the corresponding *Drosophila* orthologues available in the ensembl orthologue database (www.ensembl.org, Supplementary Table 23). *Drosophila* triglyceride content values were mined from a publicly available genome-wide obesity screen data set[52]. Estimated values represent fractional changes in triglyceride content in adult male flies. Data are from male progeny resulting from crosses of male UAS-RNAi flies from the VDRC and Hsp70-GAL4; Tub-GAL8ts virgins females. Two-to-five-day-old males were sorted into groups of 20 and subjected to two 1-h wet heatshocks 4 days apart. On the seventh day, flies were picked in groups of eight, manually crushed and sonicated, and the lysates heat-inactivated for 10 min in a thermocycler at 95 °C. Centrifuge-cleared supernatants were then used for triglyceride (GPO Trinder, Sigma) and protein (Pierce) determination. Triglyceride values from these adult-induced ubiquitous RNAi knockdown individuals were normalized to those obtained in parallel from non-heatshocked progeny from the very same crosses.

*Targeted follow-up*. Based on known biology, one to three potential candidate genes within $\pm 500$ kb of the 12 GWS-index SNPs were selected. Corresponding *Drosophila* orthologues were available for 11 loci, but no orthologue exists for *FTO* (Supplementary Table 24, http://www.flyrnai.org/cgi-bin/DRSC_orthologs.pl). The respective fly RNAi stocks for each *Drosophila* orthologue were acquired from the Vienna *Drosophila* Resource Center, as well as genetic background controls w1118 (for GD lines, VDRC #60000); tub-gal4/TM6 and w; tub-gal80ts/TM6 is available from the Bloomington *Drosophila* Stock Center. For fly triglyceride assay in the adult, male RNAi flies were crossed with w; tub-gal4 tub-gal80ts/TM6 virgins. Progenies were kept in 16 °C until enclosure. Adults were transferred to 25 °C for 2 weeks. Whole-animal triglycerides were measured as previously described[53]. Briefly, triglycerides were measured using the Infinity Triglycerides Reagent kit (Thermo Fisher #TR22321) on whole-animal homogenates of groups of three animals. Proteins from the same homogenates were measured using the Pierce BCA protein Assay kit (Thermo Scientific #23227). Triglycerides were normalized by proteins. Data were average of three experiments. The fractional changes in triglyceride content in adult male flies between knockdown group and the control groups were compared using the two-tailed *t*-tests in SAS version 9.2 software (SAS Institute, Cary, NC).

## References


1. Frayling, T. M. *et al.* 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).
2. Scuteri, A. *et al.* Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet.* **3,** e115 (2007).
3. Loos, R. J. *et al.* Common variants near MC4R are associated with fat mass, weight and risk of obesity. *Nat. Genet.* **40,** 768–775 (2008).
4. Chambers, J. C. *et al.* Common genetic variation near MC4R is associated with waist circumference and insulin resistance. *Nat. Genet.* **40,** 716–718 (2008).
5. Willer, C. J. *et al.* Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat. Genet.* **41,** 25–34 (2009).
6. Thorleifsson, G. *et al.* Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. *Nat. Genet.* **41,** 18–24 (2009).
7. Meyre, D. *et al.* Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. *Nat. Genet.* **41,** 157–159 (2009).
8. Lindgren, C. M. *et al.* Genome-wide association scan meta-analysis identifies three Loci influencing adiposity and fat distribution. *PLoS Genet.* **5,** e1000508 (2009).
9. Heard-Costa, N. L. *et al.* NRXN3 is a novel locus for waist circumference: a genome-wide association study from the CHARGE Consortium. *PLoS Genet.* **5,** e1000539 (2009).
10. Speliotes, E. K. *et al.* Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.* **42,** 937–948 (2010).
11. Heid, I. M. *et al.* Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat. Genet.* **42,** 949–960 (2010).
12. Scherag, A. *et al.* Two new Loci for body-weight regulation identified in a joint analysis of genome-wide association studies for early-onset extreme obesity in French and german study groups. *PLoS Genet.* **6,** e1000916 (2010).
13. Kilpelainen, T. O. *et al.* Genetic variation near IRS1 associates with reduced adiposity and an impaired metabolic profile. *Nat. Genet.* **43,** 753–760 (2011).
14. Wen, W. *et al.* Meta-analysis identifies common variants associated with body mass index in east Asians. *Nat. Genet.* **44,** 307–311 (2012).
15. Okada, Y. *et al.* Common variants at CDKAL1 and KLF9 are associated with body mass index in east Asian populations. *Nat. Genet.* **44,** 302–306 (2012).
16. Bradfield, J. P. *et al.* A genome-wide association meta-analysis identifies new childhood obesity loci. *Nat. Genet.* **44,** 526–531 (2012).
17. Monda, K. L. *et al.* A meta-analysis identifies new loci associated with body mass index in individuals of African ancestry. *Nat. Genet.* **45,** 690–696 (2013).
18. Falchi, M. *et al.* Low copy number of the salivary amylase gene predisposes to obesity. *Nat. Genet.* **46,** 492–497 (2014).
19. Locke, A. E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518,** 197–206 (2015).
20. Shungin, D. *et al.* New genetic loci link adipose and insulin biology to body fat distribution. *Nature* **518,** 187–196 (2015).
21. Heo, M., Faith, M. S., Pietrobelli, A. & Heymsfield, S. B. Percentage of body fat cutoffs by sex, age, and race-ethnicity in the US adult population from NHANES 1999-2004. *Am. J. Clin. Nutr.* **95,** 594–602 (2012).
22. Despres, J. P., Lemieux, I. & Prud'homme, D. Treatment of obesity: need to focus on high risk abdominally obese patients. *BMJ* **322,** 716–720 (2001).
23. Urs, S. *et al.* Sprouty1 is a critical regulatory switch of mesenchymal stem cell lineage allocation. *FASEB J.* **24,** 3264–3273 (2010).
24. Perry, J. R. *et al.* Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. *Nature* **514,** 92–97 (2014).
25. Guo, Y. *et al.* Gene-centric meta-analyses of 108 912 individuals confirm known body mass index loci and reveal three novel signals. *Hum. Mol. Genet.* **22,** 184–201 (2013).
26. Wood, A. R. *et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* **46,** 1173–1186 (2014).
27. Fox, C. S. *et al.* Genome-wide association for abdominal subcutaneous and visceral adipose reveals a novel locus for visceral fat in women. *PLoS Genet.* **8,** e1002695 (2012).
28. Willer, C. J. *et al.* Discovery and refinement of loci associated with lipid levels. *Nat. Genet.* **45,** 1274–1283 (2013).
29. Scott, R. A. *et al.* Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat. Genet.* **44,** 991–1005 (2012).
30. Morris, A. P. *et al.* Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat. Genet.* **44,** 981–990 (2012).
31. Nikpay *et al.* A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat. Genet.* **47,** 1121–1130 (2015).
32. Horikoshi, M. *et al.* New loci associated with birth weight identify genetic links between intrauterine growth and adult height and metabolism. *Nat. Genet.* **45,** 76–82 (2013).
33. Cousminer, D. L. *et al.* Genome-wide association and longitudinal analyses reveal genetic loci linking pubertal height growth, pubertal timing and childhood adiposity. *Hum. Mol. Genet.* **22,** 2735–2747 (2013).
34. Berryman, D. E., Glad, C. A., List, E. O. & Johannsson, G. The GH/IGF-1 axis in obesity: pathophysiology and therapeutic considerations. *Nat. Rev. Endocrinol.* **9,** 346–356 (2013).
35. Pillas, D. *et al.* Genome-wide association study reveals multiple loci associated with primary tooth development during infancy. *PLoS Genet.* **6,** e1000856 (2010).

36. Hansen, T. V. *et al.* Dwarfism and impaired gut development in insulin-like growth factor II mRNA-binding protein 1-deficient mice. *Mol. Cell. Biol.* **24,** 4448–4464 (2004).
37. Teslovich, T. M. *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466,** 707–713 (2010).
38. Dehghan, A. *et al.* Meta-analysis of genome-wide association studies in >80 000 subjects identifies multiple loci for C-reactive protein levels. *Circulation* **123,** 731–738 (2011).
39. Jun, G. *et al.* Comprehensive search for Alzheimer disease susceptibility loci in the APOE region. *Arch. Neurol.* **69,** 1270–1279 (2012).
40. Davies, G. *et al.* A genome-wide association study implicates the APOE locus in nonpathological cognitive ageing. *Mol. Psychiatry* **19,** 76–87 (2014).
41. Deelen, J. *et al.* Genome-wide association study identifies a single major locus contributing to survival into old age; the APOE locus revisited. *Aging Cell* **10,** 686–698 (2011).
42. Falchi, M. *et al.* Genome-wide association study identifies variants at 9p21 and 22q13 associated with development of cutaneous nevi. *Nat. Genet.* **41,** 915–919 (2009).
43. Nan, H. *et al.* Genome-wide association study identifies nidogen 1 (NID1) as a susceptibility locus to cutaneous nevi and melanoma risk. *Hum. Mol. Genet.* **20,** 2673–2679 (2011).
44. Barrett, J. H. *et al.* Genome-wide association study identifies three new melanoma susceptibility loci. *Nat. Genet.* **43,** 1108–1113 (2011).
45. Rung, J. *et al.* Genetic variant near IRS1 is associated with type 2 diabetes, insulin resistance and hyperinsulinemia. *Nat. Genet.* **41,** 1110–1115 (2009).
46. Sheffield, N. C. *et al.* Patterns of regulatory activity across diverse human cell types predict tissue identity, transcription factor binding, and long-range interactions. *Genome Res.* **23,** 777–788 (2013).
47. Garel, S., Yun, K., Grosschedl, R. & Rubenstein, J. L. The early topography of thalamocortical projections is shifted in Ebf1 and Dlx1/2 mutant mice. *Development* **129,** 5621–5634 (2002).
48. Ren, D., Li, M., Duan, C. & Rui, L. Identification of SH2-B as a key regulator of leptin sensitivity, energy balance, and body weight in mice. *Cell Metab.* **2,** 95–104 (2005).
49. Brown, L. & Baer, R. HEN1 encodes a 20-kilodalton phosphoprotein that binds an extended E-box motif as a homodimer. *Mol. Cell. Biol.* **14,** 1245–1255 (1994).
50. Pers, T. H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat. Commun.* **6,** 5890 (2015).
51. Rosen, E. D. & Spiegelman, B. M. What we talk about when we talk about fat. *Cell* **156,** 20–44 (2014).
52. Pospisilik, J. A. *et al.* Drosophila genome-wide obesity screen reveals hedgehog as a determinant of brown versus white adipose cell fate. *Cell* **140,** 148–160 (2010).
53. Musselman, L. P. *et al.* A high-sugar diet produces obesity and insulin resistance in wild-type Drosophila. *Dis. Model. Mech.* **4,** 842–849 (2011).
54. Wheeler, E. *et al.* Genome-wide SNP and CNV analysis identifies common and low-frequency variants associated with severe early-onset obesity. *Nat. Genet.* **45,** 513–517 (2013).
55. Berndt, S. I. *et al.* Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nat. Genet.* **45,** 501–512 (2013).
56. Cariou, B. *et al.* Increased adipose tissue expression of Grb14 in several models of insulin resistance. *FASEB J.* **18,** 965–967 (2004).
57. Cooney, G. J. *et al.* Improved glucose homeostasis and enhanced insulin signalling in Grb14-deficient mice. *EMBO J.* **23,** 582–593 (2004).
58. Holst, B. *et al.* PICK1 deficiency impairs secretory vesicle biogenesis and leads to growth retardation and decreased glucose tolerance. *PLoS Biol.* **11,** e1001542 (2013).
59. Bondurand, N. *et al.* Deletions at the SOX10 gene locus cause Waardenburg syndrome types 2 and 4. *Am. J. Hum. Genet.* **81,** 1169–1185 (2007).
60. Pingault, V. *et al.* Loss-of-function mutations in SOX10 cause Kallmann syndrome with deafness. *Am. J. Hum. Genet.* **92,** 707–724 (2013).
61. Altarejos, J. Y. *et al.* The Creb1 coactivator Crtc1 is required for energy balance and fertility. *Nat. Med.* **14,** 1112–1117 (2008).
62. Winkler, T. W. Quality control and conduct of genome-wide association meta-analyses. *Nat. Protoc.* **9,** 1192–1212 (2014).
63. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47,** 291–295 (2015).
64. Dastani, Z. *et al.* Novel loci for adiponectin levels and their influence on type 2 diabetes and metabolic traits: a multi-ethnic meta-analysis of 45,891 individuals. *PLoS Genet.* **8,** e1002607 (2012).
65. Chambers, J. C. *et al.* Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. *Nat. Genet.* **43,** 1131–1138 (2011).
66. Ohlsson, C. *et al.* Genetic determinants of serum testosterone concentrations in men. *PLoS Genet.* **7,** e1002313 (2011).
67. Fehrmann, R. S. N. *et al.* Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. *PLoS Genet.* **7,** e1002197 (2011).
68. Zhong, H., Yang, X., Kaplan, L. M., Molony, C. & Schadt, E. E. Integrating pathway analysis and genetics of gene expression for genome-wide association studies. *Am. J. Hum. Genet.* **86,** 581–591 (2010).
69. Min, J. L. *et al.* Coexpression network analysis in abdominal and gluteal adipose tissue reveals regulatory genetic loci for metabolic syndrome and related phenotypes. *PLoS Genet.* **8,** e1002505 (2012).
70. Myers, A. J. *et al.* A survey of genetic human cortical gene expression. *Nat. Genet.* **39,** 1494–1499 (2007).


## Acknowledgements

A full list of acknowledgements can be found in the Supplementary Notes. This work was supported by the following: Aase and Ejner Danielsens Foundation; Academy of Finland; Agency for Health Care Policy Research; Ahokas Foundation; ALFEDIAM; ALK-Abelló A/S (Hørsholm, Denmark); Althingi (the Icelandic Parliament); ANR ('Agence Nationale de la 359 Recherche'); American Heart Association; Ardix Medical; Arthritis Research UK; Association Diabète Risque Vasculaire, the Fédération Française de Cardiologie; AstraZeneca; Australian Research Council; Bayer Diagnostics; BBSRC; Becton Dickinson; BMBF (DEEP); Boehringer Ingelheim Foundation; Boston University School of Medicine; British Heart Foundation; British Skin Foundation; Canadian Institutes of Health Research; Cancer Research UK; Cardionics; Centers for Disease Control and Prevention/Association of Schools of Public Health; Chief Scientist Office of the Scottish Government; Cohortes Santé TGIR; CMSB; CPER ('Contrat de Projets État-Région'); Danish Agency for Science, Technology and Innovation; Danish Council for Independent Research; Danish Medical Research Council; Department of Health, UK; Deutsche Forschungsgemeinschaft; Deutshe Forschungsgemeinschaft (SFB992); DHFD (Diabetes Hilfs- und Forschungsfonds Deutschland); Diabetes UK; Dutch Dairy Association (NZO); Dutch Kidney Foundation; Dutch Inter University Cardiology Institute Netherlands (ICIN); Emil Aaltonen Foundation; ENGAGE consortium; Food Standards Agency, UK; Erasmus Medical Center; Erasmus University; Estonian Government; European Commission; European Community's Seventh Framework Programme; European Research Council; European Research Council (ERC-StG-281641); European Union framework program 6 EUROSPAN project; European Union; European Union (EU_FP7_NoE 'Epigenesys'); Faculty of Biology and Medicine of Lausanne; Federal State of Mecklenburg-West Pomerania; Federal Ministry of Education and Research (German Obesity Biomaterial Bank); Finnish Cardiovascular Research Foundation; Finnish Centre for Pensions; Finnish Cultural Foundation; Finnish Diabetes Research Foundation; Finnish Diabetes Research Society; Finnish Foundation for Cardiovascular Research; Finnish Foundation for Pediatric Research; Finnish Special Governmental Subsidy for Health Sciences; Finska Läkaresällskapet; Folkhälsan Research Foundation; German Bundesministerium fuer Forschung und Technology; German Diabetes Association; German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF); German Research Council; GlaxoSmithKline; Göran Gustafsson Foundation; Health and Safety Executive, UK; Helmholtz Zentrum München—German Research Center for Environmental Health; Hjartavernd (the Icelandic Heart Association); Illinois Department of Public Health; INSERM (Réseaux en Santé Publique, Interactions entre les déterminants de la santé); Integrated Research and Treatment Centre (IFB); Juho Vainio Foundation; John D and Catherine T MacArthur Foundation Research Networks; John W. Barton Sr Chair in Genetics and Nutrition; Kompetenznetz Adipositas; King's College London; Knut och Alice Wallenberg Foundation; Kuopio, Tampere and Turku University Hospital Medical Funds; Kuopio University Hospital; La Fondation de France; Li Ka Shing Foundation; Liv och Hälsa; Local Government Pensions Institution (KEVA); Ludwig-Maximilians-Universität; Lundbeck Foundation; Lundberg Foundation; Max-Planck Institute; Medical Research Council, UK; MEKOS Laboratories Denmark; Merck Santé; Ministry for Health, Welfare and Sports, The Netherlands; Ministry of Cultural Affairs and Social Ministry of the Federal State of Mecklenburg-West Pomerania; Ministry of Economic Affairs, Agriculture and Innovation, The Netherlands; Ministry of Education, Culture and Science, The Netherlands; Ministry of Science, Education and Sport of the Republic of Croatia; MRC Centre for Causal Analyses in Translational Epidemiology (MRC CAiTE); MRC-GlaxoSmithKline; MRC Human Genetics Unit; Munich Center of Health Sciences (MC Health); Municipality of Rotterdam; National Center for Advancing Translational Sciences, CTSI; National Center for Research Resources (NCRR); National Genome Research Institute, Korean Center for Disease Control and Prevention; National Heart Lung and Blood Institute (NHLBI); National Heart, Lung and Blood Institute's Framingham Heart Study; National Institute for Health and Welfare (THL); National Institute for Health Research (NIHR); National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS); National Institutes of Health; National Center for Advancing Translational Sciences; National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC); National Institute of Neurological Disorders and Stroke (NINDS); National Institute on Aging (NIA); Netherlands Consortium Healthy Ageing (NCHA); Netherlands Genomics Initiative (NGI); Netherlands Heart Foundation; Netherlands Organisation for Scientific Research (NWO); Netherlands Organization for the Health Research and Development (ZonMw); Novo Nordisk; Novo Nordisk Foundation; ONIVINS; Orion-Farmos Research Foundation; Paavo Nurmi Foundation; Pierre Fabre; Research Centre for Prevention and Health, the Capital Region of Denmark; Research Institute for Diseases in the Elderly (RIDE); Research into Ageing, UK; Roche; Royal Swedish Academy of Science; Russian Foundation for Basic Research; Sanger Institute; Samfundet Folkhälsan; SFD ('Société

Francophone du 358 Diabète'); Siemens Healthcare; Signe and Ane Gyllenberg Foundation; Sigrid Jusélius Foundation; Social Insurance Institution of Finland (KELA); State of Bavaria; Stroke Association, UK; Swedish Diabetes Foundation; Swedish Foundation for Strategic Research; Swedish Heart-Lung Foundation; Swedish Research Council; Swedish Research Council for Infrastructures; Swiss National Science Foundation; Sylvia & Charles Viertel Charitable Foundation; Tampere Tuberculosis Foundation; Timber Merchant Vilhelm Bangs Foundation; Topcon; Torsten and Ragnar Söderberg's Foundation; Translational Genomics Research Institute; Unilever UK; University Cancer Research Fund at UNC Chapel Hill; University of Eastern Finland; University of Maryland General Clinical Research Center; Uppsala University; Uppsala University Hospital; USDA National Institute of Food and Agriculture; VA Clinical Science Research and Development; Velux Foundation; VU University Medical Center; Wageningen University; Wellcome Trust.

## Author contributions

R.J.F.L. oversaw all aspects of this study and chaired the writing group that consisted of Y.L., F.R.D., S.G., M.L.B., J.N., K.L.M., M.C.Z., J.A.P., C.L. and T.O.K.. Data cleaning and preparations were performed by F.R.D., S.G., R.J.F.L., Y.L. and R.W.W.; and meta-analyses were performed by F.R.D., S.G. and Y.L. A full list of author contributions can be found in the Supplementary Notes.

## Additional information

**Supplementary Information** accompanies this paper at http://www.nature.com/naturecommunications

**Competing financial interests:** R.E. received honoraria for lectures from the following companies (within the past 2 years): MSD Sharp&Dohme and Pfizer. J.M.J. disclosed Trinity Partners, Inc., Osteoarthritis Research Society International, Chronic Osteoarthritis Management Initiative of US Bone and Joint Initiative, Samumed, Interleukin Genetics, Inc. and Algynomics, Inc. K.S. is a full-time employee of GlaxoSmithKline. D.S. received honoraria for lectures from the following companies (within the past 2 years): MSD Sharp&Dohme. E.S.-T. received honoraria for lectures, research grants and consultancy fees from the following companies (within the past 2 years): Aegerion, Fresenius, MSD Sharp&Dohme, Pfizer and Sanofi. Peter Vollenweider received an unrestricted grant from GSK to build the CoLaus study.

**Reprints and permission** information is available online at http://npg.nature.com/reprintsandpermissions/

**How to cite this article:** Lu, Y. *et al.* New loci for body fat percentage reveal link between adiposity and cardiometabolic disease risk. *Nat. Commun.* 7:10495 doi: 10.1038/ncomms10495 (2016).



Yingchang Lu[1,2], Felix R. Day[3], Stefan Gustafsson[4,5], Martin L. Buchkovich[6], Jianbo Na[7], Veronique Bataille[8,9], Diana L. Cousminer[10], Zari Dastani[11], Alexander W. Drong[12], Tõnu Esko[13,14,15,16], David M. Evans[17,18], Mario Falchi[9,19], Mary F. Feitosa[20], Teresa Ferreira[12], Åsa K. Hedman[4,5,12], Robin Haring[21,22], Pirro G. Hysi[9], Mark M. Iles[23], Anne E. Justice[24], Stavroula Kanoni[25,26], Vasiliki Lagou[12,27], Rui Li[11], Xin Li[28], Adam Locke[29], Chen Lu[30], Reedik Mägi[12,13], John R.B. Perry[3], Tune H. Pers[15,16,31,32,33], Qibin Qi[34], Marianna Sanna[9,19], Ellen M. Schmidt[35], William R. Scott[36,37], Dmitry Shungin[38,39,40], Alexander Teumer[41,42], Anna A.E. Vinkhuyzen[43], Ryan W. Walker[1,2], Harm-Jan Westra[44,45,46], Mingfeng Zhang[47], Weihua Zhang[36,37], Jing Hua Zhao[3], Zhihong Zhu[43], Uzma Afzal[36,37], Tarunveer Singh Ahluwalia[31,48,49,50], Stephan J.L. Bakker[51], Claire Bellis[52], Amélie Bonnefond[53,54,55], Katja Borodulin[56], Aron S. Buchman[57], Tommy Cederholm[58], Audrey C. Choh[59], Hyung Jin Choi[60], Joanne E. Curran[61], Lisette C.P.G.M. de Groot[62], Philip L. De Jager[44,63,64], Rosalie A.M. Dhonukshe-Rutten[62], Anke W. Enneman[65], Elodie Eury[53,54,55], Daniel S. Evans[66], Tom Forsen[67], Nele Friedrich[21], Frédéric Fumeron[68,69,70,71], Melissa E. Garcia[72], Simone Gärtner[73], Bok-Ghee Han[74], Aki S. Havulinna[56], Caroline Hayward[75], Dena Hernandez[76], Hans Hillege[77], Till Ittermann[41], Jack W. Kent[52], Ivana Kolcic[78], Tiina Laatikainen[56,79,80], Jari Lahti[81,82], Irene Mateo Leach[77], Christine G. Lee[83,84], Jong-Young Lee[74], Tian Liu[85,86], Youfang Liu[87], Stéphane Lobbens[53,54,55], Marie Loh[36,88], Leo-Pekka Lyytikäinen[89,90], Carolina Medina-Gomez[65,91,92], Karl Michaëlsson[93], Mike A. Nalls[76], Carrie M. Nielson[94,95], Laticia Oozageer[37], Laura Pascoe[96], Lavinia Paternoster[18], Ozren Polašek[78,97], Samuli Ripatti[10,26,98], Mark A. Sarzynski[99], Chan Soo Shin[100], Nina Smolej Narančić[101], Dominik Spira[102,103], Priya Srikanth[94,95], Elisabeth Steinhagen-Thiessen[102,103], Yun Ju Sung[104], Karin M.A. Swart[105,106], Leena Taittonen[107,108], Toshiko Tanaka[109], Emmi Tikkanen[10,98], Nathalie van der Velde[65], Natasja M. van Schoor[105,106], Niek Verweij[77], Alan F. Wright[75], Lei Yu[57], Joseph M. Zmuda[110], Niina Eklund[56], Terrence Forrester[111], Niels Grarup[31], Anne U. Jackson[29], Kati Kristiansson[10,56], Teemu Kuulasmaa[112], Johanna Kuusisto[112,113,114], Peter Lichtner[115], Jian'an Luan[3], Anubha Mahajan[12], Satu Männistö[56], Cameron D. Palmer[14,15], Janina S. Ried[116], Robert A. Scott[3], Alena Stancáková[117], Peter J. Wagner[10,56], Ayse Demirkan[118], Angela Döring[119,120], Vilmundur Gudnason[121,122], Douglas P. Kiel[123,124], Brigitte Kühnel[116,120,125], Massimo Mangino[9], Barbara Mcknight[126,127,128], Cristina Menni[9],

Jeffrey R. O'Connell[129], Ben A. Oostra[118], Alan R. Shuldiner[129,130], Kijoung Song[131], Liesbeth Vandenput[132], Cornelia M. van Duijn[91,118,133], Peter Vollenweider[134], Charles C. White[30], Michael Boehnke[29], Yvonne Boettcher[135,136], Richard S. Cooper[137], Nita G. Forouhi[3], Christian Gieger[116,120,125], Harald Grallert[120,125,138], Aroon Hingorani[139], Torben Jørgensen[140,141,142], Pekka Jousilahti[56], Mika Kivimaki[143], Meena Kumari[143], Markku Laakso[112,113,114], Claudia Langenberg[3,143], Allan Linneberg[144], Amy Luke[137], Colin A. Mckenzie[111], Aarno Palotie[10,26,145], Oluf Pedersen[31], Annette Peters[120,125], Konstantin Strauch[116,146], Bamidele O. Tayo[137], Nicholas J. Wareham[3], David A. Bennett[57], Lars Bertram[147,148], John Blangero[61], Matthias Blüher[135,136], Claude Bouchard[99], Harry Campbell[97], Nam H. Cho[149], Steven R. Cummings[66], Stefan A. Czerwinski[59], Ilja Demuth[102,150], Rahel Eckardt[102], Johan G. Eriksson[56,67,81], Luigi Ferrucci[109], Oscar H. Franco[91,92], Philippe Froguel[53,54,55], Ron T. Gansevoort[51], Torben Hansen[31,151], Tamara B. Harris[72], Nicholas Hastie[75], Markku Heliövaara[56], Albert Hofman[91,92], Joanne M. Jordan[87], Antti Jula[56], Mika Kähönen[152,153], Eero Kajantie[56,154,155], Paul B. Knekt[56], Seppo Koskinen[56], Peter Kovacs[135], Terho Lehtimäki[89,90], Lars Lind[156], Yongmei Liu[157], Eric S. Orwoll[95], Clive Osmond[158], Markus Perola[10,13,56], Louis Pérusse[159,160], Olli T. Raitakari[161,162], Tuomo Rankinen[99], D.C. Rao[20,104,163], Treva K. Rice[104,163], Fernando Rivadeneira[65,91,92], Igor Rudan[97], Veikko Salomaa[56], Thorkild I.A. Sørensen[18,31,164], Michael Stumvoll[135,136], Anke Tönjes[136], Bradford Towne[59], Gregory J. Tranah[66], Angelo Tremblay[159], André G. Uitterlinden[65,91,92], Pim van der Harst[77,165,166], Erkki Vartiainen[56], Jorma S. Viikari[167], Veronique Vitart[75], Marie-Claude Vohl[160,168], Henry Völzke[41,169,170], Mark Walker[44,171], Henri Wallaschofski[21,169], Sarah Wild[172], James F. Wilson[75,97], Loïc Yengo[53,54,55], D. Timothy Bishop[23], Ingrid B. Borecki[20,173], John C. Chambers[36,37,174], L. Adrienne Cupples[30,175], Abbas Dehghan[176], Panos Deloukas[25,26,177], Ghazaleh Fatemifar[18], Caroline Fox[63,175], Terrence S. Furey[6,178], Lude Franke[77,166], Jiali Han[179], David J. Hunter[14,28,180,181], Juha Karjalainen[166], Fredrik Karpe[27,182], Robert C. Kaplan[34], Jaspal S. Kooner[37,174,183], Mark I. McCarthy[12,27,182], Joanne M. Murabito[184,185], Andrew P. Morris[12,186], Julia A.N. Bishop[23], Kari E. North[187], Claes Ohlsson[132], Ken K. Ong[3,188,189], Inga Prokopenko[12,19,27], J. Brent Richards[11,190,191,192], Eric E. Schadt[193,194], Tim D. Spector[9], Elisabeth Widén[10], Cristen J. Willer[35,195,196], Jian Yang[43], Erik Ingelsson[4,5,197], Karen L. Mohlke[6], Joel N. Hirschhorn[14,15,16], John Andrew Pospisilik[198], M. Carola Zillikens[65,91], Cecilia Lindgren[12,14,199], Tuomas Oskari Kilpeläinen[3,31] & Ruth J.F. Loos[1,2,3,200,201]

[1] The Charles Bronfman Institute for Personalized Medicine, The Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. [2] The Department of Preventive Medicine, The Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. [3] MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Institute of Metabolic Science, University of Cambridge, Cambridge Biomedical Campus, Cambridge CB2 0QQ, UK. [4] Science for Life Laboratory, Uppsala University, 750 85 Uppsala, Sweden. [5] Department of Medical Sciences, Molecular Epidemiology, Uppsala University, 751 85 Uppsala, Sweden. [6] Department of Genetics, University of North Carolina, Chapel Hill, North Carolina 27599, USA. [7] Department of Developmental and Regenerative Biology, The Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. [8] West Herts NHS Trust, Herts HP2 4AD, UK. [9] Department of Twin Research and Genetic Epidemiology, King's College London, London SE1 7EH, UK. [10] Institute for Molecular Medicine Finland, University of Helsinki, FI-00290 Helsinki, Finland. [11] Department Epidemiology, Biostatistics and Human Genetics, Lady Davis Institute, Jewish General Hospital, McGill University, Montréal, Quebec, Canada H3T1E2. [12] Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK. [13] Estonian Genome Center, Univeristy of Tartu, Tartu, 51010, Estonia. [14] Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge 2142, USA. [15] Divisions of Endocrinology and Genetics and Center for Basic and Translational Obesity Research, Boston Children's Hospital, Boston, Massachusetts 02115, USA. [16] Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. [17] University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, Queensland 4102, Australia. [18] MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Bristol BS82BN, UK. [19] Department of Genomics of Common Disease, School of Public Health, Imperial College London, London W12 0NN, UK. [20] Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine, St Louis, Missouri 63108, USA. [21] Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, 17475 Greifswald, Germany. [22] European University of Applied Sciences, Faculty of Applied Public Health, 18055 Rostock, Germany. [23] Leeds Institute of Cancer and Pathology, Cancer Research UK Leeds Centre, University of Leeds, Leeds LS9 7TF, UK. [24] Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [25] William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London EC1M 6BQ, UK. [26] Wellcome Trust Sanger Institute, Human Genetics, Hinxton, Cambridge CB10 1SA, UK. [27] Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Churchill Hospital, Oxford OX3 7LJ, UK. [28] Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts 02115, USA. [29] Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, Michigan 48109, USA. [30] Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts 02118, USA. [31] Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, Faculty of Health and Medical Sciences, University of Copenhagen,

2100 Copenhagen, Denmark. [32] Medical and Population Genetics Program, Broad Institute of MIT and Harvard, Cambridge 02142, USA. [33] Department of Epidemiology Research, Statens Serum Institut, 2100 Copenhagen, Denmark. [34] Department of Epidemiology and Popualtion Health, Albert Einstein College of Medicine, Bronx, New York 10461, USA. [35] Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan 48109, USA. [36] Department of Epidemiology and Biostatistics, Imperial College London, London W2 1PG, UK. [37] Ealing Hospital NHS Trust, Middlesex UB1 3HW, UK. [38] Lund University Diabetes Centre, Department of Clinical Science, Genetic and Molecular Epidemiology Unit, Skåne University Hosptial, 205 02 Malmö, Sweden. [39] Department of Public Health and Clinical Medicine, Unit of Medicine, Umeå University, 901 87 Umeå, Sweden. [40] Department of Odontology, Umeå University, 901 85 Umeå, Sweden. [41] Institute for Community Medicine, University Medicine Greifswald, 17475 Greifswald, Germany. [42] Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, 17475 Greifswald, Germany. [43] Queensland Brain Institute, The University of Queensland, Brisbane 4072, Australia. [44] Program in Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA. [45] Divisions of Genetics and Rheumatology, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02446, USA. [46] Partners Center for Personalized Genetic Medicine, Boston, Massachusetts 02446, USA. [47] Department of Dermatology, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. [48] Copenhagen Prospective Studies on Asthma in Childhood, Faculty of Health and Medical Sceinces, University of Copenhagen, 2200 Copenhagen, Denmark. [49] Danish Pediatric Asthma Center, Gentofte Hospital, The Capital Region, 2200 Copenhagen, Denmark. [50] Steno Diabetes Center A/S, DK-2820 Gentofte, Denmark. [51] University of Groningen, University Medical Center Groningen, Department of Medicine, 9700 RB Groningen, The Netherlands. [52] Department of Genetics, Texas Biomedical Research Institute, San Antonio, Texas 78245, USA. [53] CNRS UMR 8199, F-59019 Lille, France. [54] European Genomic Institute for Diabetes, 59000 Lille, France. [55] Université de Lille 2, 59000 Lille, France. [56] National Institute for Health and Welfare, FI-00271 Helsinki, Finland. [57] Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, Illinois 60612, USA. [58] Department of Public Health and Caring Sciences, Clinical Nutrition and Metabolism, Uppsala University, 751 85 Uppsala, Sweden. [59] Lifespan Health Research Center, Wright State University Boonshoft School of Medicine, Dayton, Ohio 45420, USA. [60] Department of Anatomy, Seoul National University College of Medicine, Seoul 03080, Korea. [61] South Texas Diabetes and Obesity Institute, University of Texas Rio Grande Valley, Brownsville, Texas 78520, USA. [62] Department of Human Nutrition, Wageningen University, 6700 EV Wageningen, The Netherlands. [63] Harvard Medical School, Boston, Massachusetts 02115, USA. [64] Program in Translational NeuroPsychiatric Genomics, Department of Neurology, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. [65] Department of Internal Medicine, Erasmus Medical Center, 3015GE Rotterdam, The Netherlands. [66] California Pacific Medical Center Research Institute, San Francisco, California 94107, USA. [67] Department of General Practice and Primary Health Care, University of Helsinki, FI-00014 Helsinki, Finland. [68] INSERM, UMR_S 1138, Centre de Recherche des Cordeliers, F-75006 Paris, France. [69] Sorbonne Universités, UPMC Univ Paris 06, UMR_S 1138, Centre de Recherche des Cordeliers, F-75006 Paris, France. [70] Université Paris Descartes, Sorbonne Paris Cité, UMR_S 1138, Centre de Recherche des Cordeliers, F-75006 Paris, France. [71] Univ Paris Diderot, Sorbonne Paris Cité, UMR_S 1138, Centre de Recherche des Cordeliers, F-75006 Paris, France. [72] Laboratory of Epidemiology and Population Sciences, National Institute on Aging, Bethesda, Maryland 20892, USA. [73] Department of Medicine A, University Medicine Greifswald, 17475 Greifswald, Germany. [74] Center for Genome Science, National Institute of Health, Osong Health Technology Administration Complex, Chungcheongbuk-do 370914, Korea. [75] MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK. [76] Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, Maryland 20892, USA. [77] University of Groningen, University Medical Center Groningen, Department of Cardiology, 9700 RB Groningen, The Netherlands. [78] Department of Public Health, Faculty of Medicine, University of Split, Split 21000, Croatia. [79] Hospital District of North Karelia, FI-80210 Joensuu, Finland. [80] Institute of Public Health and Clinical Nutrition, University of Eastern Finland, FI-70211 Kuopio, Finland. [81] Folkhälsan Research Centre, FI-00290 Helsinki, Finland. [82] Institute of Behavioural Sciences, University of Helsinki, FI-00014 Helsinki, Finland. [83] Department of Medicine, Oregon Health and Science University, Portland, Oregon 97239, USA. [84] Research Service, Veterans Affairs Medical Center, Portland, Oregon 97239, USA. [85] Max Planck Institute for Molecular Genetics, Department of Vertebrate Genomics, 14195 Berlin, Germany. [86] Max Planck Institute for Human Development, 14194 Berlin, Germany. [87] Thurston Arthritis Research Center, University of North Carolina at Chapel Hill, Chaper Hill, North Carolina 27599-7280, USA. [88] Translational Laboratory in Genetic Medicine (TLGM), Agency for Science, Technology and Research (A*STAR), 8A Biomedical Grove, Immunos, Level 5, Singapore 138648, Singapore. [89] Department of Clinical Chemistry, University of Tampere School of Medicine, FI-33014 Tampere, Finland. [90] Department of Clinical Chemistry, Fimlab Laboratories and School of Medicine, University of Tampere, FI-33520 Tampere, Finland. [91] Netherlands Genomics Initiative (NGI)-sponsored Netherlands Consortium for Healthy Aging (NCHA), Rotterdam The Netherlands. [92] Department of Epidemiology, Erasmus Medical Center, 3015GE Rotterdam, The Netherlands. [93] Department of Surgical Sciences, Orthopedics, Uppsala University, 751 85 Uppsala, Sweden. [94] School of Public Health, Oregon Health & Science University, Portland, Oregon 97239, USA. [95] Bone & Mineral Unit, Oregon Health & Science University, Portland, Oregon 97239, USA. [96] Institute of Cell & Molecular Biosciences, Newcastle University, Newcastle NE1 7RU, UK. [97] Centre for Global Health Research, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Teviot Place, Edinburgh EH8 9AG, UK. [98] Hjelt Institute, University of Helsinki, FI-00014 Helsinki, Finland. [99] Human Genomics Laboratory, Pennington Biomedical Research Center, Baton Rouge, Los Angeles 70808, USA. [100] Department of Internal Medicine, Seoul National University College of Medicine, Seoul 03080, Korea. [101] Institute for Anthropological Research, Zagreb 10000, Croatia. [102] The Berlin Aging Study II; Research Group on Geriatrics; Charité—Universitätsmedizin Berlin, 13347 Berlin, Germany. [103] Lipid Clinic at the Interdisciplinary Metabolism Center, Charité-Universitätsmedizin Berlin, 13353 Berlin, Germany. [104] Division of Biostatistics, Washington University School of Medicine, St Louis, Missouri 63110, USA. [105] EMGO Institute for Health and Care Research, VU University Medical Center, 1081 BT Amsterdam, The Netherlands. [106] VUMC, Department of Epidemiology and Biostatistics, 1081 BT Amsterdam, The Netherlands. [107] Department of Pediatrics, University of Oulu, FI-90014 Oulu, Finland. [108] Department of Pediatrics, Vaasa Central Hospital, FI-65100 Vaasa, Finland. [109] Translational Gerontology Branch, National Institute on Aging, Baltimore, Maryland 21225, USA. [110] Department of Epidemiology; University of Pittsburgh, Pittsburgh, Pennsylvania 15261, USA. [111] Tropical Metabolism Research Unit, Tropical Medicine Research Institute, University of the West Indies, Mona JMAAW15, Jamaica. [112] Faculty of Health Sciences, Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland, 70210 Kuopio, Finland. [113] Department of Medicine, University of Eastern Finland, 70210 Kuopio, Finland. [114] Kuopio University Hospital, 70029 Kuopio, Finland. [115] Institute of Human Genetics, Helmholtz Zentrum München—German Research Center for Environmental Health, 85764 Neuherberg, Germany. [116] Institute of Genetic Epidemiology, Helmholtz Zentrum München—German Research Center for Environmental Health, 85764 Neuherberg, Germany. [117] Department of Medicine, University of Eastern Finland and Kuopio University Hospital, 70210 Kuopio, Finland. [118] Genetic Epidemiology Unit, Department of Epidemiology, Erasmus University Medical Center, 3015GE Rotterdam, The Netherlands. [119] Institute of Epidemiology I, Helmholtz Zentrum München—German Research Center for Environmental Health, 85764 Neuherberg, Germany. [120] Institute of Epidemiology II, Helmholtz Zentrum München—German Research Center for Environmental Health, 85764 Neuherberg, Germany. [121] Icelandic Heart Association, Kopavogur 201, Iceland. [122] University of Iceland, Faculty of Medicine, Reykjavik 101, Iceland. [123] Department of Medicine Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts 02115, USA. [124] Institute for Aging Research Hebrew Senior Life, Boston, Massachusetts 02131, USA. [125] Research Unit of Molecular Epidemiology, Helmholtz Zentrum München—German Research Center for Environmental Health, 85764 Neuherberg, Germany. [126] Cardiovascular Health Research Unit, University of Washington, Seattle, Washington 98101, USA. [127] Program in Biostatistics and Biomathematics, Divison of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, USA. [128] Department of Biostatistics, University of Washington, Seattle, Washington 98195, USA. [129] Program for Personalized and Genomic Medicine, Division of Endocrinology, Diabetes and Nutrition,

Department of Medicine, University of Maryland School of Medicine, Baltimore, Maryland 21201, USA. [130] Geriatric Research and Education Clinical Center, Vetrans Administration Medical Center, Baltimore, Maryland 21042, USA. [131] Genetics, Projects Clinical Platforms and Sciences, GlaxoSmithKline, Philadelphia, Pennsylvania 19112, USA. [132] Centre for Bone and Arthritis Research, Department of Internal Medicine and Clinical Nutrition, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, 413 45 Gothenburg, Sweden. [133] Center for Medical Systems Biology, 2300 Leiden, The Netherlands. [134] Department of Internal Medicine, University Hospital Lausanne (CHUV) and University of Lausanne, 1011 Lausanne, Switzerland. [135] University of Leipzig, IFB Adiposity Diseases, 04103 Leipzig, Germany. [136] University of Leipzig, Department of Medicine, 04103 Leipzig, Germany. [137] Department of Public Health Sciences, Stritch School of Medicine, Loyola University Chicago, Maywood, Illinois 61053, USA. [138] German Center for Diabetes Research (DZD), 85764 Neuherberg, Germany. [139] Institute of Cardiovascular Science, University College London, London WC1E 6BT, UK. [140] Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, 2200 Copenhagen, Denmark. [141] Faculty of Medicine, University of Aalborg, 9220 Aalborg, Denmark. [142] Research Centre for Prevention and Health, DK2600 Capital Region of Denmark, Denmark. [143] Department of Epidemiology and Public Health, UCL, London WC1E 6BT, UK. [144] Research Centre for Prevention and Health, Glostrup Hospital, 2600 Glostrup, Denmark. [145] Massachusetts General Hospital, Center for Human Genetic Research, Psychiatric and Neurodevelopmental Genetics Unit, Boston, Massachusetts 02114, USA. [146] Institute of Medical Informatics, Biometry and Epidemiology, Chair of Genetic Epidemiology, Ludwig-Maximilians-Universität, 81377 Munich, Germany. [147] School of Public Health, Faculty of Medicine, Imperial College London, London W6 8RP, UK. [148] Lübeck Interdisciplinary Platform for Genome Analytics, Institutes of Neurogenetics and Integrative and Experimental Genomics, University of Lübeck, 23562 Lübeck, Germany. [149] Ajou University School of Medicine, Department of Preventive Medicine, Suwon Kyoung-gi 443-721, Korea. [150] Institute of Medical and Human Genetics, Charité—Universitätsmedizin Berlin, 13353 Berlin, Germany. [151] Faculty of Health Sciences, University of Southern Denmark, 5000 Odense, Denmark. [152] Department of Clinical Physiology, Tampere University Hospital, FI-33521 Tampere, Finland. [153] Department of Clinical Physiology, University of Tampere School of Medicine, FI-33014 Tampere, Finland. [154] Children's Hospital, Helsinki University Hospital and University of Helsinki, FI-00029 Helsinki, Finland. [155] Department of Obstetrics and Gynecology, MRC Oulu, Oulu University Hospital and University of Oulu, FI-90029 Oulu, Finland. [156] Department of Medical Sciences, Uppsala University, 751 85 Uppsala, Sweden. [157] Center for Human Genetics, Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina 27157, USA. [158] MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton General Hospital, Southampton SO16 6YD, UK. [159] Department of Kinesiology, Laval University, Québec City, Quebec, Canada G1V 0A6. [160] Institute of Nutrition and Functional Foods, Laval University, Québec City, Quebec, Canada G1V 0A6. [161] Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, FI-20521 Turku, Finland. [162] Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, FI-20520 Turku, Finland. [163] Department of Psychiatry, Washington University School of Medicine, St Louis, Missouri 63110, USA. [164] Institute of Preventive Medicine, Bispebjerg and Frederiksberg Hospital, The Capital Region, 2000 Frederiksberg, Denmark. [165] Durrer Center for Cardiogenetic Research, Interuniversity Cardiology Institute Netherlands-Netherlands Heart Institute, 3501 DG Utrecht, The Netherlands. [166] Department of Genetics, University Medical Center Groningen, University of Groningen, 9700 RB Groningen, The Netherlands. [167] Department of Medicine, University of Turku, FI-20521 Turku, Finland. [168] School of Nutrition, Laval University, Québec City, Quebec, Canada G1V 0A6. [169] DZHK (German Centre for Cardiovascular Research), partner site Greifswald, 17475 Greifswald, Germany. [170] DZD (German Centre for Diabetes Research), partner site Greifswald, 17475 Greifswald, Germany. [171] Institute of Cellular Medicine, Newcastle University, Newcastle NE2 4HH, UK. [172] Centre for Population Health Sciences, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh EH8 9AG, UK. [173] Analytical Genetics Group, Regeneron Genetics Center, Regeneron Pharmaceuticals, Inc., Tarrytown, New York 10591, USA. [174] Imperial College Healthcare NHS Trust, London W12 0HS, UK. [175] National Heart, Lung, and Blood Institute, the Framingham Heart Study, Framingham, Massachusetts 01702, USA. [176] Department of Epidemiology, Erasmus Medical Center, 3000CA Rotterdam/Zuidholland, The Netherlands. [177] Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah 21589, Saudi Arabia. [178] Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599, USA. [179] Department of Epidemiology, Richard M. Fairbanks School of Public Health, Melvin and Bren Simon Cancer Center, Indianapolis, Indiana 46202, USA. [180] Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA. [181] Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts 02115, USA. [182] Oxford NIHR Biomedical Research Centre, Oxford OX3 7LJ, UK. [183] National Heart and Lung Institute, Imperial College London, London W12 0NN, UK. [184] Boston University School of Medicine, Department of Medicine, Section of General Internal Medicine, Boston, Massachusetts 02118, USA. [185] NHLBI's and Boston University's Framingham Heart Study, Framingham, Massachusetts 01702, USA. [186] Department of Biostatistics, University of Liverpool, Liverpool L69 3GA, UK. [187] Carolina Center for Genome Sciences and Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7400, USA. [188] MRC Unit for Lifelong Health and Ageing at UCL, London WC1B 5JU, UK. [189] Department of Paediatrics, University of Cambridge, Cambridge CB2 0QQ, UK. [190] Department of Medicine, Lady Davis Institute, Jewish General Hospital, McGill University, Montréal, Quebec, Canada H3T1E2. [191] Department of Twin Research, King's College London, London SE1 1E7, UK. [192] Division of Endocrinology, Lady Davis Institute, Jewish General Hospital, McGill University, Montréal, Quebec, Canada H3T1E2. [193] Icahn Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. [194] Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. [195] Department of Human Genetics, University of Michigan, Ann Arbor, Michigan 48109, USA. [196] Department of Internal Medicine, Division of Cardiovascular Medicine, University of Michigan, Ann Arbor, Michigan 48109, USA. [197] Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, California 94305, USA. [198] Department of Epigenetics, Max Planck Institute of Immunobiology and Epigenetics, D-76108 Freiburg, Germany. [199] The Big Data Institute, University of Oxford, Oxford OX3 7LJ, UK. [200] The Genetics of Obesity and Related Metabolic Traits Program, The Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. [201] The Mindich Child Health and Development Institute, The Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA.