nature genetics

Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways

Through genome-wide association meta-analyses of up to 133,010 individuals of European ancestry without diabetes, including individuals newly genotyped using the Metabochip, we have increased the number of confirmed loci influencing glycemic traits to 53, of which 33 also increase type 2 diabetes risk ($q < 0.05$). Loci influencing fasting insulin concentration showed association with lipid levels and fat distribution, suggesting impact on insulin resistance. Gene-based analyses identified further biologically plausible loci, suggesting that additional loci beyond those reaching genome-wide significance are likely to represent real associations. This conclusion is supported by an excess of directionally consistent and nominally significant signals between discovery and follow-up studies. Functional analysis of these newly discovered loci will further improve our understanding of glycemic control.

© 2012 Nature America, Inc. All rights reserved.

The Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) previously undertook meta-analyses of genome-wide association studies (GWAS) of glycemic traits in non-diabetic individuals, leading to the discovery of multiple associated loci: 16 for fasting glucose concentration, 2 for fasting insulin concentration and 5 for post-challenge glucose concentration (2-hour glucose, 2hGlu)[1–3]. These and subsequent studies highlighted important biological pathways implicated in glucose and insulin regulation[4,5]. They also showed that some but not all loci associated with glycemic traits in non-diabetic individuals also affect the risk of type 2 diabetes (T2D)[1,6]. Despite the success of these efforts, the identification of new loci was limited by *de novo* genotyping capacity and cost, such that only a limited number of promising loci from discovery analyses were taken forward to follow-up analyses (often those reaching a threshold of approximately $P < 1 \times 10^{-5}$ in the discovery phase). Therefore, it is likely that many additional associations with common, low-penetrance variants remain to be found among SNPs not previously selected for follow-up[7,8].

The Illumina CardioMetabochip (Metabochip) is a custom iSELECT array of 196,725 SNPs developed to support cost-effective large-scale follow-up studies of putative association signals for a range of cardiovascular and metabolic traits (~66,000 SNPs) and to fine map established loci (~120,000 SNPs) (**Supplementary Fig. 1**)[9]. The ~66,000 SNPs for follow-up analysis were selected to enable genotyping of the most significant association signals for each of 23 metabolic traits for which data were contributed by a range of consortia. MAGIC contributed ~5,000 top-ranking SNPs for fasting glucose concentration and ~1,000 SNPs each for fasting insulin concentration and 2hGlu that had shown nominal association in discovery analyses ($P_{discovery} < 0.02$)[1,2].

In the present study, we combined newly available samples with genotype data for these 66,000 follow-up SNPs with previous discovery meta-analyses to discover new association signals with glycemic traits. This approach identified 41 glycemic associations not previously described[1,2]: 20 for fasting glucose concentration, 17 for fasting insulin concentration and 4 for 2hGlu. This raises the number of associated loci to 36 for fasting glucose concentration, 19 for fasting insulin concentration and 9 for 2hGlu, explaining 4.8%, 1.2% and 1.7% of the variance in these traits, respectively. Of these 53 non-overlapping loci, 33 were also associated with T2D ($P < 0.05$), which, although supporting the previous assertion of an imperfect correlation between these traits, also implicates new loci in the etiology of T2D and increases the overlap between glycemic and T2D loci.

RESULTS

Approaches to identify loci associated with glycemic traits

To follow up loci showing evidence of association ($P_{discovery} < 0.02$) in discovery GWAS, we investigated the 66,000 Metabochip follow-up SNPs for association with fasting glucose concentration, fasting insulin concentration and 2hGlu. We combined in meta-analysis data from up to 133,010 (fasting glucose), 108,557 (fasting insulin) and 42,854 (2hGlu) non-diabetic individuals of European ancestry, including individuals from the previous meta-analyses[1,2], individuals from new GWAS and individuals newly genotyped on the Metabochip array (**Supplementary Fig. 2**). All study characteristics are shown in **Supplementary Table 1**. Genome-wide association data for Filipino women were available (**Supplementary Table 1**), for which we report the effect directions and allele frequencies (**Supplementary Table 2a**,**b**). Association signals at genome-wide significance ($P < 5 \times 10^{-8}$) located more than 500 kb from and not in linkage disequilibrium (LD; HapMap Utah residents of Northern and Western European ancestry (CEU) $r^2 < 0.05$) with any variant already known to be associated with the trait were considered novel. Associated loci are referred to by the name of the nearest gene, unless a more biologically plausible gene was nearby or a nearby gene was previously associated with another

A full list of authors and affiliations appears at the end of the paper.

Received 23 January; accepted 20 July; published online 12 August 2012; doi:10.1038/ng.2385

glycemic trait. In such cases, we maintain consistency with the previous naming or name the most biologically plausible gene (nearest genes are named in **Supplementary Table 2a–d**). As body mass index (BMI) is a major risk factor for T2D and is correlated with glycemic traits, we also performed analyses adjusted for BMI.

Although not the main focus of this effort, given the increased variant density available on the Metabochip for established glycemic loci, we investigated whether these data would enable fine mapping of underlying functional variants[1–3]. In these analyses, we included data from up to 53,622 individuals for fasting glucose, 42,384 for fasting insulin and 27,602 for 2hGlu from studies with Metabochip genotypes only. However, given the lack of samples from different ancestry groups and the absence of full conditional analyses, these analyses for the most part did not improve the resolution of association signals.

Beyond single-SNP investigations for each glycemic trait, we also tested the hypothesis that gene-based analyses using VEGAS[10] would identify genes that harbor multiple association signals, which individually did not reach genome-wide significance. Among the ~66,000 SNPs, we used VEGAS to pool the results for all SNPs within 50 kb of either side of gene boundaries to identify genes with more evidence of association than expected by chance (given gene size and LD structure) by simulation that was significant after Bonferroni correction for multiple testing ($P < 5 \times 10^{-6}$).

Fasting glucose concentration

In analyses of up to 133,010 individuals, we identified 20 loci with genome-wide significant associations with fasting glucose ($P < 5 \times 10^{-8}$) (**Table 1** and **Supplementary Figs. 3** and **4**) and confirmed previously established loci[1] (**Supplementary Table 2e**). Of these 20 loci, 9 (in or near *IBKAP*, *DNLZ*, *WARS*, *KL*, *TOP1*, *P2RX2*, *AMT*, *RREB1* and *GLS2*) had not previously been associated with other metabolic traits (**Box 1**). Among these, *KL* (encoding klotho) is of particular interest. In addition to being associated with fasting glucose (but not fasting insulin) concentration, the glucose-raising allele is also associated with an increased risk of T2D (odds ratio (OR) = 1.08 (1.04–1.11); $P = 1.1 \times 10^{-5}$) (**Fig. 1**). *KL* was first identified as a gene related to suppression of aging: its reduced expression was associated with reduced lifespan, as well as hypoglycemia[11]. Despite further animal studies supporting a role for *KL* in glucose metabolism[12] and insulin sensitivity[13], human studies have generally been small and inconclusive[14,15].

We also identified new associations with fasting glucose concentration in regions previously associated with other metabolic traits or disease outcomes, including T2D[6,16] (*ARAP1*, *CDKN2B*, *GRB10*, *CDKAL1*, *IGF2BP2* and *ZBED3*, which was identified in BMI-adjusted models) and 2hGlu[2] (*GIPR*), as well as confirming the recently identified signals for fasting glucose[17–19] at *FOXA2*, *PPP1R3B*, *PCSK1* and *PDX1*. FOXA2 is a forkhead transcription factor that regulates *PDX1* expression, and *PDX1* encodes a transcription factor critical for pancreatic development[20]. *PDX1* mutations have been linked to maturity-onset diabetes of the young 4 (MODY4)[21], pancreatic agenesis[22] and permanent neonatal diabetes[23], although we observed no significant association with T2D in DIAbetes Genetics Replication and Meta-analysis (DIAGRAM) Metabochip analyses[24] (**Fig. 1**).

Given the overlap between genetic loci for fasting glucose and other metabolic traits, we performed a systematic search of all glycemic loci and their associations with other metabolic traits using data available through other consortia[25–27]. In DIAGRAM Metabochip analyses[24], 22 (>60%) of the now 36 loci associated with fasting glucose at genome-wide significance showed association ($P < 0.05$; false discovery rate (FDR) $q < 0.05$) with T2D (**Fig. 1**). In all cases, the glucose-raising allele was associated with increased risk of T2D, yet fasting glucose effect sizes and T2D ORs were weakly correlated (**Fig. 2a**).

Gene-based analyses confirmed many of the loci identified in single-SNP analyses (**Supplementary Table 3a**) and identified another 9 genomic regions (containing 14 genes) with significant association signals ($P < 5 \times 10^{-6}$), including some with biological candidacy, such as the *HKDC1* gene that encodes a putative hexokinase[28].

Fasting insulin concentration

In 108,557 individuals, we identified 17 additional loci with genome-wide significant associations to fasting insulin concentration and confirmed known associations[1]. These newly identified loci include variants in or near *HIP1*, *TET2*, *YSK4*, *PEPD* and *FAM13A* (**Table 1**, **Box 1** and **Supplementary Figs. 3** and **4**), as well as SNPs near loci previously associated with other metabolic traits, including T2D[6] (*TCF7L2* and *PPARG*), BMI[29] (*FTO*), waist-hip ratio (WHR)[26] (*LYPLAL1*, *RSPO3* and *GRB14*), triglycerides[27] (*ANKRD55-MAP3K1)* and adiponectin[30] (*ARL15*). We also confirmed the recent associations with fasting insulin at *GRB14*, *PPP1R3B*, *LYPLAL1*, *IRS1*, *UHRF1BP1* and *PDGFC*[19]. The *ANKRD55-MAP3K1* association is of interest, as the MAP3K1 protein regulates expression of IRS1 (ref. 31) as well as activation of nuclear factor (NF)-κB[32,33] and the c-Jun N-terminal kinase (JNK) pathway[34], both of which are centrally implicated in insulin resistance[35,36]. Furthermore, data from DIAGRAM Metabochip analyses show that the insulin-raising allele at this SNP is strongly associated with increased risk of T2D[24].

In contrast to fasting glucose (**Supplementary Fig. 5a**), in fasting insulin analyses adjusted for BMI, we observed a systematic decrease in the standard errors of the SNP effect estimates (**Supplementary Fig. 5b**), perhaps because BMI explains more of the variance in fasting insulin levels (R^2 = 32.6%) than in fasting glucose levels (R^2 = 8.6%) or 2hGlu (R^2 = 11.0%) (data from the Fenland study). Therefore, BMI adjustment removes more variance in fasting insulin, thereby rendering genetic associations more readily detectable. This idea is supported by the identification of another five loci in BMI-adjusted models by this approach (**Table 1** and **Supplementary Figs. 3** and **4**). As expected, BMI adjustment abolished fasting insulin associations at *FTO* ($P = 0.71$; **Supplementary Table 2b**), suggesting that the association with fasting insulin is mediated entirely through association with BMI.

In total, 13 of the 19 loci associated with fasting insulin concentration also showed associations with T2D ($P < 0.05$; FDR $q < 0.05$) (**Fig. 1**), with the insulin-raising allele associated with higher risk of T2D, except at *TCF7L2* (**Fig. 2b,c**), where the allele associated with lower fasting insulin was associated with higher fasting glucose levels (**Table 1**). Notably, the loci associated with fasting insulin showed a pattern of association with lipid traits consistent with insulin resistance, which is not observed for either fasting glucose or 2hGlu loci (**Fig. 1**). Thirteen (~68%) of the 19 loci were associated with high-density lipoprotein (HDL)-cholesterol ($q < 0.05$): all insulin-raising alleles were associated with lower HDL levels, and nine of these were also associated with higher triglycerides ($q < 0.05$) (**Fig. 1**). Further, the insulin-raising alleles of four SNPs were associated with higher WHR (adjusted for BMI) ($q < 0.05$) (**Fig. 1**), another trait linked to insulin resistance, and five SNPs were also associated with BMI, although with inconsistent direction ($q < 0.05$) (**Fig. 1**).

In gene-based analyses, we focused on BMI-adjusted results to account for the variance in fasting insulin explained by BMI. Beyond those loci containing genome-wide significant SNPs, we identified 7 distinct regions (containing 22 genes) after Bonferroni correction ($P < 5 \times 10^{-6}$). Among these genes, we identified many for which previous biological evidence suggests their role in pathways involved in insulin secretion or action (**Supplementary Table 3b**). Although the

© 2012 Nature America, Inc. All rights reserved.

association for the lead SNP in *PPARD* did not reach genome-wide significance ($P = 3.9 \times 10^{-6}$), the *PPARD* gene—a regulator of adipose, hepatic and skeletal muscle metabolism[37]—reached the gene-based significance threshold ($P < 1 \times 10^{-6}$). PPARD agonists have also been shown to induce insulin-sensitizing effects in a mouse model[38]. In addition, we identified *PTEN* to be associated (**Supplementary Table 3b**), a gene previously suggested to affect glucose metabolism through regulation of insulin signaling[39], and in which a muscle-specific deletion protected mice from insulin resistance and diabetes resulting from high-fat feeding[40].

Table 1 SNPs associated with fasting glucose, fasting insulin and 2-hour glucose at genome-wide significance in Europeans

							Primary trait				I^2		FI (BMI-adjusted)				2hGlu			
Primary trait	SNP	Chr.	Position	Gene	Alleles (effect/ other)	Freq. effect allele	Effect	SE	Global analysis *P* value	Global analysis *n*	estimate	*P* value	Effect	SE	Global analysis *P* value	Global analysis *n*	Effect	SE	Global analysis *P* value	Global analysis *n*
FG	**rs10811661**	9	22124094	*CDKN2B*	T/C	0.82	0.0238	0.003	5.6×10^{-18}	128,488	0.00	1.00	−0.0065	0.003	0.019	98,880	0.0567	0.014	8.8×10^{-5}	42,801
	rs4869272	5	95565204	*PCSK1**	T/C	0.69	0.0177	0.002	1.0×10^{-15}	131,872	0.00	1.00	0.0016	0.002	0.469	103,493	−0.0322	0.012	0.006	42,848
	rs11619319	13	27385599	*PDX1*	G/A	0.23	0.0195	0.002	1.3×10^{-15}	132,996	0.00	1.00	0.0001	0.002	0.977	103,492	0.0185	0.013	0.156	42,848
	rs983309	8	9215142	*PPP1R3B**	T/G	0.12	0.0256	0.003	6.3×10^{-15}	127,470	0.14	0.32	0.0223	0.003	1.2×10^{-12}	99,024	−0.0548	0.016	0.001	42,846
	rs6943153	7	50759073	*GRB10*	T/C	0.34	0.0154	0.002	1.6×10^{-12}	131,795	0.00	1.00	0.0091	0.002	2.3×10^{-5}	103,447	0.0110	0.011	0.333	42,794
	rs11603334	11	72110633	*ARAP1*	G/A	0.83	0.0192	0.003	1.1×10^{-11}	128,139	0.00	1.00	−0.0046	0.003	0.086	99,026	0.0294	0.014	0.037	42,839
	rs6113722	20	22505099	*FOXA2*	G/A	0.96	0.0353	0.005	2.5×10^{-11}	123,665	0.04	0.78	−0.0095	0.005	0.064	103,471	0.0493	0.030	0.101	41,416
	rs16913693	9	110720180	*IKBKAP*	T/G	0.97	0.0434	0.007	3.5×10^{-11}	125,115	0.00	1.00	−0.0018	0.007	0.785	96,357	0.0639	0.034	0.062	40,522
	rs3829109	9	138376587	*DNLZ*	G/A	0.71	0.0172	0.003	1.1×10^{-10}	115,310	0.25	0.07	−0.0002	0.003	0.948	94,964	0.0343	0.014	0.013	36,803
	rs3783347	14	99909014	*WARS*	G/T	0.79	0.0168	0.003	1.3×10^{-10}	132,544	0.02	0.89	0.0017	0.003	0.515	103,339	0.0274	0.014	0.044	42,850
	rs2302593	19	50888474	*GIPR*	C/G	0.50	0.0144	0.002	9.3×10^{-10}	116,141	0.27	0.05	0.0025	0.002	0.265	96,976	−0.0322	0.012	0.006	40,781
	rs9368222	6	20794975	*CDKAL1*	A/C	0.28	0.0143	0.002	1.0×10^{-9}	128,453	0.09	0.50	−0.0047	0.002	0.037	98,894	0.0279	0.012	0.023	42,825
	rs10747083	12	131551691	*P2RX2*	A/G	0.66	0.0133	0.002	7.6×10^{-9}	127,111	0.00	1.00	−0.0006	0.002	0.785	99,895	0.0269	0.012	0.026	42,790
	rs6072275	20	39177319	*TOP1*	A/G	0.16	0.0159	0.003	1.7×10^{-8}	128,616	0.00	1.00	0.0038	0.003	0.169	99,018	−0.0110	0.014	0.435	42,853
	rs7651090	3	186996086	*IGF2BP2*	G/A	0.31	0.0128	0.002	1.75×10^{-8}	128,548	0.02	0.86	0.0003	0.002	0.900	98,924	0.0583	0.012	1.05×10^{-6}	42,814
	rs576674	13	32452302	*KL*	G/A	0.15	0.0167	0.003	2.3×10^{-8}	131,856	0.00	1.00	−0.0001	0.003	0.983	103,472	0.0308	0.016	0.060	42,849
	rs11715915	3	49430334	*AMT*	C/T	0.68	0.0120	0.002	4.9×10^{-8}	131,523	0.30	0.02	0.0059	0.002	0.006	103,398	0.0273	0.012	0.018	42,851
FG (BMI-adjusted)	rs17762454	6	7158199	*RREB1*	T/C	0.26	0.0140	0.002	9.6×10^{-9}	123,247	0.00	1.00	−0.0002	0.002	0.919	103,470	0.0007	0.013	0.953	42,848
	rs7708285	5	76461623	*ZBED3*	G/A	0.27	0.0150	0.003	1.2×10^{-8}	117,931	0.00	1.00	0.0027	0.002	0.265	98,341	0.0349	0.013	0.008	42,803
	rs2657879	12	55151605	*GLS2*	G/A	0.18	0.0157	0.003	3.9×10^{-8}	121,596	0.39	0.03	−0.0024	0.003	0.366	102,175	0.0200	0.014	0.164	42,670
							Primary trait						FG				2hGlu			
FI	rs1421085	16	52358455	*FTO*	C/T	0.42	0.0200	0.003	1.9×10^{-15}	104,062	0.00	1.00	0.0074	0.002	0.001	128,597	0.0122	0.011	0.278	42,849
	rs983309	8	9215142	*PPP1R3B**	T/G	0.12	0.0287	0.004	3.8×10^{-14}	103,030	0.04	0.77	0.0256	0.003	6.3×10^{-15}	127,470	−0.0548	0.016	0.001	42,846
	rs9884482	4	106301085	*TET2*	C/T	0.39	0.0165	0.002	1.4×10^{-11}	108,420	0.00	1.00	0.0001	0.002	0.946	132,869	0.0004	0.011	0.973	42,745
	rs7903146	10	114748339	*TCF7L2*	C/T	0.72	0.0181	0.003	6.1×10^{-11}	103,037	0.31	0.02	−0.0220	0.002	2.7×10^{-20}	127,477	−0.0885	0.013	5.6×10^{-12}	42,851
	rs10195252	2	165221337	*GRB14**	T/C	0.59	0.0159	0.003	4.9×10^{-10}	99,126	0.00	1.00	0.0053	0.002	0.014	127,005	0.0361	0.011	0.001	42,846
	rs1167800	7	75014132	*HIP1*	A/G	0.54	0.0156	0.003	2.6×10^{-9}	91,416	0.00	1.00	0.0016	0.002	0.470	118,536	−0.0133	0.012	0.272	38,884
	rs2820436	1	217707303	*LYPLAL1*	C/A	0.67	0.0153	0.003	4.4×10^{-9}	104,044	0.01	0.97	0.0077	0.002	0.001	128,580	−0.0041	0.012	0.723	42,843
	rs2745353	6	127494628	*RSPO3*	T/C	0.51	0.0143	0.002	5.5×10^{-9}	104,075	0.06	0.67	−0.0009	0.002	0.677	128,615	−0.0005	0.011	0.962	42,853
	rs731839	19	38590905	*PEPD*	G/A	0.34	0.0145	0.003	1.7×10^{-8}	104,636	0.13	0.38	0.0046	0.002	0.038	132,528	0.0142	0.012	0.220	42,846
	rs4865796	5	53308421	*ARL15*	A/G	0.67	0.0146	0.003	2.1×10^{-8}	100,001	0.03	0.81	0.0043	0.002	0.052	127,784	0.0337	0.012	0.004	42,852
	rs2972143	2	226824609	*IRS1*	G/A	0.62	0.0142	0.003	3.2×10^{-8}	99,566	0.00	1.00	0.0035	0.002	0.107	127,473	0.0195	0.011	0.082	42,853
	rs1530559	2	135472099	*YSK4*	A/G	0.52	0.0145	0.003	3.4×10^{-8}	107,281	0.19	0.18	0.0037	0.002	0.100	129,880	0.0200	0.011	0.077	42,849
FI (BMI-adjusted)	rs2943645	2	226807424	*IRS1*	T/C	0.63	0.0193	0.002	2.3×10^{-19}	99,023	0.00	1.00	0.0034	0.002	0.112	127475	0.0210	0.011	0.061	42,846
	rs10195252	2	165221337	*GRB14**	T/C	0.60	0.0174	0.002	1.3×10^{-16}	98,997	0.00	1.00	0.0053	0.002	0.014	127005	0.0361	0.011	0.001	42,846
	rs2126259	8	9222556	*PPP1R3B*	T/C	0.11	0.0238	0.003	3.3×10^{-13}	99,021	0.14	0.51	0.0213	0.003	5.4×10^{-10}	127480	−0.0877	0.017	1.8×10^{-7}	42,849
	rs4865796	5	53308421	*ARL15*	A/G	0.67	0.0154	0.002	2.2×10^{-12}	98,314	0.48	0.01	0.0043	0.002	0.052	127784	0.0337	0.012	0.004	42,852
	rs17036328	3	12365484	*PPARG*	T/C	0.86	0.0212	0.003	3.6×10^{-12}	98,984	0.21	0.31	0.0051	0.003	0.103	128567	0.0335	0.016	0.031	42,843
	rs731839	19	38590905	*PEPD*	G/A	0.34	0.0148	0.002	5.1×10^{-12}	103,252	0.13	0.55	0.0046	0.002	0.038	132528	0.0142	0.012	0.220	42,847
	rs974801	4	106290513	*TET2*	G/A	0.38	0.0139	0.002	3.3×10^{-11}	103,489	0.09	0.67	0.0012	0.002	0.582	131866	0.0052	0.011	0.643	42,849
	rs459193	5	55842508	*ANKRD55-MAP3K1*	G/A	0.73	0.0147	0.002	1.12×10^{-10}	103,378	0.27	0.17	0.0111	0.002	1.6×10^{-6}	132989	0.0276	0.012	0.023	42,849
	rs6822892	4	157954125	*PDGFC*	A/G	0.68	0.0138	0.002	2.6×10^{-10}	103,432	0.00	1.00	0.0010	0.002	0.636	132951	0.0256	0.012	0.031	42,836
	rs4846565	1	217788727	*LYPLAL1*	G/A	0.67	0.0132	0.002	1.8×10^{-9}	99,014	0.00	1.00	0.0066	0.002	0.003	127468	0.0132	0.012	0.254	42,853
	rs3822072	4	89960292	*FAM13A*	A/G	0.48	0.0116	0.002	1.8×10^{-8}	99,977	0.00	1.00	0.0025	0.002	0.236	129432	0.0161	0.011	0.143	42,850
	rs6912327	6	34872900	*UHRF1BP1*	T/C	0.80	0.0165	0.003	2.3×10^{-8}	80,010	0.04	0.91	0.0074	0.003	0.011	103826	0.0139	0.0139	0.391	34,761
							Primary trait						FG				FI (BMI-adjusted)			
2hGlu	rs6975024	7	44198411	*GCK*	C/T	0.15	0.1026	0.016	5.2×10^{-11}	42,842	0.00	1.00	0.0605	0.003	2.9×10^{-99}	103,517	0.0063	0.003	0.030	98,458
	rs11782386	8	9239197	*PPP1R3B**	C/T	0.87	0.0985	0.017	2.2×10^{-9}	42,852	0.00	1.00	−0.0167	0.003	5.5×10^{-7}	100,595	−0.0164	0.003	6.9×10^{-7}	95,565
	rs1019503	5	96280573	*ERAP2*	A/G	0.48	0.0628	0.011	8.9×10^{-9}	42,851	19.6	0.42	−0.0061	0.002	0.003	108,113	0.0004	0.002	0.851	103,448
2hGlu (BMI-adjusted)	rs7651090	3	186996086	*IGF2BP2*	G/A	0.30	0.064	0.012	4.5×10^{-8}	42,792	63.4	0.01	0.0128	0.002	1.8×10^{-8}	104,019	0.0003	0.002	0.900	98,924

Genome-wide loci for fasting glucose (FG), fasting insulin (FI), FI (adjusted for BMI) and 2hGlu are shown along with results for the other traits aligned to the trait-raising allele for the primary trait. Non-MAGIC SNPs (identified in other consortia and selected for the Metabochip to follow up on other non-MAGIC traits) are indicated in bold. Freq., allele frequency of the primary trait-raising allele. Per-allele effect (standard error, SE) for FI represents differences in natural log–transformed levels of FI. *N* represents sample size. Heterogeneity was assessed using the I^2 index[56]. The gene shown is the nearest gene to the lead SNP, except for those marked with an asterisk, for which the nearest gene is also listed in **Supplementary Table 2a–d**.

© 2012 Nature America, Inc. All rights reserved.

Box 1 Genes nearest to loci associated with glycemic traits

Fasting glucose

IKBKAP (inhibitor of κ light polypeptide gene enhancer in β cells, kinase complex–associated protein) encodes a scaffold protein that binds IKKs and NF-κB–inducing kinase (NIK), assembling them into different active complexes. Splice-site mutations in this gene lead to familial dysautonomia[57]. Also mapping to this region are ***C9orf4***, ***C9orf5***, ***C9orf6*** and ***MIR32*** (microRNA 32, unknown function), as well as ***ACTL7A*** (actin-like 7A) and ***ACTL7B*** (actin-like 7B).

WARS (tryptophanyl-tRNA synthetase) catalyzes the aminoacylation of tRNA(Trp) with tryptophan. The intronic SNP rs3783347 is associated with *WARS* expression in liver: the glucose-raising allele associated with lower mRNA expression (age- and sex-adjusted $P = 4.19 \times 10^{-5}$) and is in perfect LD ($r^2 = 1$, $D = 1$) with a 3′ UTR SNP in *SLC25A47* (rs3736952) and in modest LD ($r^2 = 0.3$, $D = 1$) with a nonsynonymous p.Arg135Leu alteration (qualified as tolerated by SIFT and probably damaging by Polyphen). Nearby, ***YY1*** (YY1 transcription factor) codes for a zinc-finger transcription factor involved in regulating a broad set of promoters. It has been suggested that YY1-regulated transcription is linked to glucose metabolism via O-GlcNAcylation[58].

KL (klotho) encodes a type I membrane protein related to β-glucosidases. rs576674 lies ~36 kb upstream of *KL*. Variation in *KL* has been associated with insulin regulation, insulin resistance phenotypes and cardiovascular disease in some studies[14,15,59,60], but *KL* variants were not associated with diabetes risk[61]. The various SNPs in these studies are all in weak LD with rs576674 ($r^2 < 0.125$). Variation in *KL* is also associated with bone metabolism and may have a role in associations of energy metabolism with bone metabolism[62,63].

TOP1 (topoisomerase (DNA) I). rs6072275 is intronic in *TOP1* and lies in a large region of high LD in Europeans, which includes the plausible biological candidate ***LPIN3*** (lipin 3). In mice, a related homolog, *Lpin1*, is associated with fatty liver dystrophy[64], a phenotype similar to human lipodystrophy (loss of body fat, fatty liver, hypertriglyceridemia and insulin resistance). *Lpin1* mRNA is expressed at high levels in adipose tissue and is induced during differentiation of preadipocytes, suggesting that lipin is required for normal adipose tissue development, whereas *LPIN2* has been suggested to be associated with T2D and glucose metabolism[65]. rs6072275 lies in the middle of a large copy-number variation (CNV) that extends from within the 3′ end of *TOP1* to the 5′ end of ***PLCG1*** (phospholipase C, γ1).

P2RX2 (purinergic receptor P2X, ligand-gated ion channel, 2). rs10747083 lies in a small CNV approximately 150 kb upstream of five protein-coding genes, including ***P2RX2***, encoding one of a family of purinoceptors for ATP; ***GALNT9*** (UDP-N-acetyl-α-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 9 (GalNAc-T9) encoding a member of the UDP-N-acetyl-α-D-galactosamine polypeptide N-acetylgalactosaminyltransferase (GalNAc-T) family of enzymes and expressed specifically in the brain; ***FBRSL1*** (fibrosin-like 1); ***PXMP2*** (peroxisomal membrane protein 2, 22 kDa); and ***PGAM5*** (phosphoglycerate mutase family member 5), and downstream within 184 kb of ***POLE*** (polymerase (DNA-directed), ε) and ***LOC100130238*** (hypothetical LOC100130238), a miscRNA.

DNLZ contains rs3829109, which is in low LD with a well-established locus for inflammatory bowel disease. Two recent publications reported that the *CARD9* SNP rs10781499 ($r^2 = 0.29$) is associated with ulcerative colitis[66], and *CARD9-SNAPC4* SNP rs4077515 ($r^2 = 0.27$) is associated with Crohn's disease and ulcerative colitis[67,68]. Several genes are located in the region, but few with high plausibility for a role in glycemia.

AMT encodes the mitochondrial aminomethyltransferase, which is a critical component of the glycine cleavage system. Depending on the *AMT* transcript, rs11715915 is located in the 3′ UTR or within coding regions, where it causes a synonymous substitution. This SNP is also located downstream of ***TCTA*** (T-cell leukemia translocation altered), which has no known metabolic function, and upstream of ***RHOA*** (ras homolog family member A). RHOA is a signaling molecule involved in actin cytoskeleton stability and reorganization[69] that binds and activates Rho kinase (ROCK), a regulator of insulin transcription[70] and action[71] that is differentially regulated in T2D[72] and is hypothesized to have a role in glucose homeostasis[71].

GLS2 encodes liver-expressed glutaminase 2, which is required for hydrolysis of glutamine. rs2657879 causes a benign (according to Polyphen) amino-acid change (p.Leu581Pro) in the GLS2 protein. The GLS2 protein is highly expressed (Human Protein Atlas) by both liver and pancreas, and it has been shown in liver tumors that alterations in the balance of the activity of GLS2:GLS1 (the kidney-specific homolog) is important for regulating glutamate metabolism[73]. The other gene in this region, ***SPRYD4*** (SPRY domain containing 4), has no known function in metabolism.

RREB1 (ras responsive element–binding protein 1) encodes a zinc-finger transcription factor, with rs17762454 lying in an intron in the gene. The protein product of *RREB1* binds to RAS-responsive elements (RREs) of gene promoters, including the *CALCA* (encoding calcitonin) gene promoter. The role of *RREB1* in energy metabolism is not known. An uncorrelated SNP at this locus (rs675209) was associated with serum urate levels ($P = 1.0 \times 10^{-9}$) in a GWAS of serum urate, gout and cardiovascular disease risk factors[74]. Another gene at this locus, ***SSR1*** (signal sequence receptor, α), encodes a glycosylated endoplasmic reticulum (ER) membrane receptor associated with protein translocation across the ER membrane. Reactome pathway analysis places this gene in a module with key roles in the synthesis and function of insulin, insulin-like growth factors and ghrelin, making this gene a plausible biological candidate at this locus (REACTOME: REACT_15380). A third gene at this locus, ***CAGE1***, encodes cancer antigen 1. *CAGE1* has no known role in metabolism.

Fasting insulin

TET2 encodes the tet oncogene family member 2, isoform b, which catalyzes the conversion of methylcytosine to 5-hydroxymethylcytosine. The enzyme is involved in myelopoiesis, and defects in this gene have been associated with several myeloproliferative disorders (NCBI RefSeq). Perhaps more relevant to glycemic regulation is ***PPA2***, which encodes the inorganic pyrophosphatase 2 isoform 1 precursor. Its protein product is localized to mitochondria; it has high homology to members of the inorganic pyrophosphatase family, including the signature sequence that is essential for its catalytic activity (NCBI RefSeq). Pyrophosphatases catalyze the hydrolysis of pyrophosphate to inorganic phosphate.

HIP1 encodes the huntingtin-interacting protein 1, a membrane-associated protein that colocalizes with huntingtin. It is ubiquitously expressed, with the highest level found in brain. Loss of normal huntingtin-HIP1 interaction in Huntington's disease may contribute to a defect in membrane-cytoskeletal integrity in the brain. Of interest to insulin action, HIP1 is involved in clathrin-mediated endocytosis and trafficking. Mice transgenic for the mutated form of huntingtin develop diabetes[75,76]; however, although mice with double knockout of *Hip1* and *Hip1r* have severe vertebral defects, suffer from dwarfism and die in early adulthood, they do not show any fasting glucose abnormalities[77]. The lead SNP (rs1167800) is only 104 bp away from a missense SNP (rs1167801), encoding a glutamic acid–to-histadine amino-acid change; however, LD between the SNPs is low ($r^2 = 0.196$).

FAM13A (family with sequence similarity 13, member A) encodes a protein with unknown function. Previous GWAS for lung function measures[78] and chronic obstructive pulmonary disease[79] described variants in *FAM13A* that affect these traits. ***SPP1***, encoding osteopontin, a secreted matrix glycoprotein and proinflammatory cytokine involved in cell-mediated immunity, is within 1 Mb. Mice exposed to a high-fat diet show increased circulating osteopontin, and

(continued)

© 2012 Nature America, Inc. All rights reserved.

Box 1 Continued

overexpression of Spp1 in the macrophages recruited into adipose tissue improved insulin sensitivity[80], and SPP1 was highly expressed in obese twins relative to their nonobese siblings[81]. Recent work linked osteopontin to β-cell function through the gastric inhibitory pathway (GIP) pathway[82]. In carriers of the *GIPR* variant associated with impaired glucose and GIP-stimulated insulin secretion, osteopontin levels were lower compared to noncarriers. In addition, both GIP and osteopontin prevented cytokine-induced apoptosis and osteopontin-stimulated cell proliferation of functional β-cell mass.

PEPD (peptidase D) encodes a member of the peptidase family. The protein forms a homodimer that hydrolyzes dipeptides or tripeptides with a C-terminal proline or hydroxyproline residue. The enzyme serves an important role in the recycling of proline and may be rate limiting for collagen production. ***CEBPA*** (CCAAT/enhancer binding protein (C/EBP) α) is ~100 kb downstream of the lead SNP and encodes a transcription factor expressed in adipose tissue that regulates a number of genes involved in lipid and glucose metabolism. A SNP in low LD with our lead SNP was previously associated with triglyceride levels[83]. Cells from *Cebpa* $^{-/-}$ mice show a complete absence of insulin-stimulated glucose transport, secondary to reduced gene expression and tyrosine phosphorylation of the insulin receptor and IRS1 (ref. 84). CEBPA also modulates expression of leptin by binding to the promoter of the gene[85], and our lead SNP showed modest association with BMI in previous GIANT meta-analyses ($P = 0.005$).

YSK4 (Sps1/Ste20-related kinase homolog) contains rs1530559 in an intron. This gene has no known function in human energy metabolism. Three other genes at this locus also have no known role in energy metabolism, including ***RAB3GAP1*** (RAB3 GTPase–activating protein subunit 1 (catalytic), encoding the catalytic subunit of a Rab GTPase–activating protein and mutated in Warburg micro syndrome; ***CCNT2*** (cyclin T2), belonging to the highly conserved cyclin family, whose members are characterized by marked periodicity in protein abundance through the cell cycle; and ***ACMSD*** (aminocarboxymuconate semialdehyde decarboxylase), involved in the *de novo* synthesis pathway of nicotinamide adenine dinucleotide (NAD) from tryptophan. *ACMSD* has been implicated in the pathogenesis of several neurodegenerative disorders.

2-hour glucose

ERAP2 (endoplasmic reticulum aminopeptidase 2) encodes an aminopeptidase that hydrolyzes N-terminal amino acids of protein or peptide substrates. The lead SNP is strongly associated with *ERAP2* expression in liver ($P = 1.1 \times 10^{-55}$) and in lymphoblastoid cell lines in individuals from the CEU ($P = 8 \times 10^{-21}$) and Yoruba from Ibadan, Nigeria (YRI) samples ($P = 2 \times 10^{-15}$). Also near to this lead SNP is ***LNPEP*** (leucyl/cystinyl aminopeptidase), which is widely expressed and well characterized in muscle and fat cells. In response to insulin, LNPEP translocates to the cell surface and colocalizes with GLUT4 (ref. 86). Although the role it has in insulin action is unknown, this translocation is impaired in individuals with T2D[86]. ***PCSK1*** is also within 500 kb of the lead SNP, although it is on the other side of a recombination hotspot (**Supplementary Fig. 4d**).

© 2012 Nature America, Inc. All rights reserved.

npg

2-h glucose

In 42,854 individuals, we identified 4 additional loci that were associated with 2hGlu (**Table 1** and **Supplementary Figs. 3** and **4**), including a signal near *ERAP2* and 3 signals near loci previously associated with fasting glucose[1] (*GCK*), HDL-cholesterol[27] (*PPP1R3B*) and T2D[6] (*IGF2BP2*), as well as confirming the 5 previous associations[2]. To determine whether these associations reflected differences in the response to a glucose challenge or were partly driven by effects on fasting glucose, we also performed analyses adjusted for fasting glucose. No additional loci were found to associate with genome-wide significance after adjustment for fasting glucose concentration, although the association of *GCK* with 2hGlu was severely attenuated ($\beta = 0.04$ (s.e.m. = 0.016) mM per allele; $P = 0.005$ versus $\beta = 0.1$ (0.016) mM per allele; $P = 5.3 \times 10^{-11}$ in the model unadjusted for fasting glucose), suggesting that the association with 2hGlu is driven, at least in part, by a primary association with fasting glucose (**Supplementary Table 2d**). The association of SNPs near *GCK* with both fasting glucose and 2hGlu suggests a generalized increase in the glucose setpoint, consistent with inactivating mutations in *GCK* that cause MODY[41]. As for fasting glucose, when 2hGlu models were adjusted for BMI, no systematic differences were observed, although, again, the rs7651090 SNP in *IGF2BP2* reached genome-wide significance (**Table 1**).

Eight of the nine SNPs associated with 2hGlu at genome-wide levels of significance were also associated with T2D ($q < 0.05$) (**Fig. 1**), although the 2hGlu-raising alleles at *PPP1R3B*, *GCKR* and *VPS13C-C2CD4A-C2CD4B* were associated with lower risk of T2D (**Fig. 2d**), consistent with their association with lower fasting glucose levels (**Table 1** and **Supplementary Table 2e**).

In addition to SNPs with associations that reached genome-wide significance in single-SNP analyses, we identified three regions (containing six genes) showing association with 2hGlu in gene-based analyses. These included the *HKDC1* gene, as well as an association signal at *CRHR1* ($P = 2 \times 10^{-6}$) (**Supplementary Table 3c**), mostly driven by the lead SNP in this gene (rs17762954), which approached genome-wide significance ($P = 7.4 \times 10^{-7}$). *CRHR1*, together with *GIPR*, belongs to the family of class B G protein–coupled receptors (GPCRs) and is highly expressed in pancreatic β-cells, where stimulation of the receptor potentiates insulin secretion in response to glucose[42].

Fine mapping of established loci

Analyses at higher SNP density around previously established loci did not generally yield stronger associations or more plausible functional variants (**Supplementary Table 4**). For fasting glucose concentration, markedly more significant SNPs or larger effect sizes than the previous lead SNP were observed for 4 of the 16 loci: *PROX1*, *GCK*, *ADRA2A* and *VPS13C-C2CD4A-C2CD4B* (**Supplementary Table 4**). Regional plots for these loci are shown in **Supplementary Figure 6**. Although the association for the new lead SNP near *ADRA2A* was not markedly more significant than the previous lead SNP, the effect size was almost double that of the previous lead SNP (**Supplementary Table 4**). However, this and other new lead SNPs lacked more plausible functionality. The new lead SNP at *VPS13C-C2CD4A-C2CD4B*, previously associated with proinsulin[43], is far more significant and of larger effect size than the previous lead SNP ($\beta = 0.0273$ (s.e.m. = 0.0035) mM per allele; $P = 4.8 \times 10^{-15}$ versus $\beta = 0.0057$ (0.0036) mM per allele; $P = 0.111$; $r^2 = 0.27$). For fasting insulin concentration, another SNP downstream of *IGF1* was found to be more significant and had a larger effect size, although with no known functionality (**Supplementary Fig. 6** and **Supplementary Table 4**). For 2hGlu, another SNP at *VPS13C-C2CD4A-C2CD4B* was again more significant than the previous lead SNP (**Supplementary Fig. 6** and **Supplementary Table 4**) and was previously associated with diabetes in Chinese individuals[44].

Pathway analysis

Next, we explored whether glycemic loci were enriched for connectivity between genes representing particular pathways or processes. To do this, we used GRAIL software[45] and investigated both an

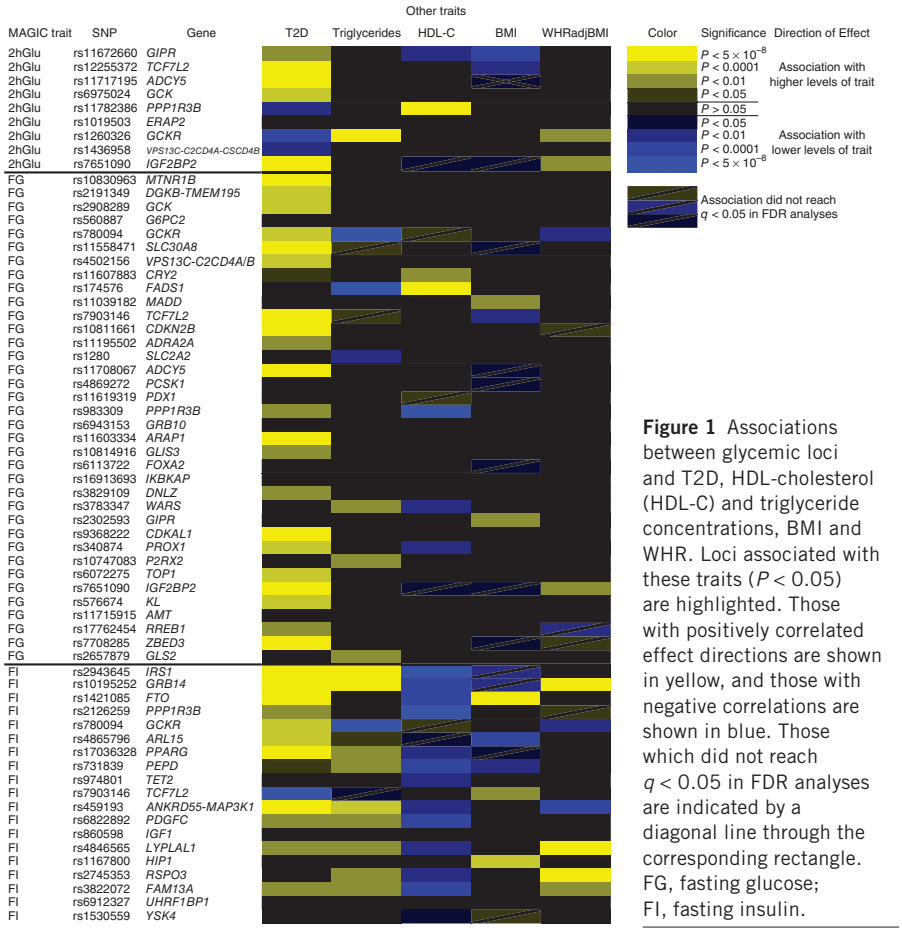

Figure 1 Associations between glycemic loci and T2D, HDL-cholesterol (HDL-C) and triglyceride concentrations, BMI and WHR. Loci associated with these traits ($P < 0.05$) are highlighted. Those with positively correlated effect directions are shown in yellow, and those with negative correlations are shown in blue. Those which did not reach $q < 0.05$ in FDR analyses are indicated by a diagonal line through the corresponding rectangle. FG, fasting glucose; FI, fasting insulin.

excess of connectivity between the established loci (that reached genome-wide significance) and then between established loci and those loci that did not reach genome-wide significance but showed a lower level of significance for association ($P < 0.0005$) (Online Methods). We aimed to establish whether there were any biologically relevant genes among this longer list of suggestively associated loci. This less stringent threshold yielded 218, 155 and 100 regions for fasting glucose, fasting insulin and 2hGlu, respectively. To further assess whether the established loci represented common biological pathways, we used MAGENTA to undertake gene set–enrichment analyses (Online Methods).

We found that genes near the 36 loci associated with fasting glucose concentration had a high degree of connectivity (see Online Methods for definition of how genes were selected). Eight genes showed highly significant similarity to genes in other associated loci at $P_{GRAIL} < 0.01$ and were connected by keywords such as 'glucose', 'insulin', 'pancreatic' and 'diabetes' (**Supplementary Fig. 7** and **Supplementary Table 5a**), at levels greater than those expected by chance ($P_{permutation} = 0.003$). We observed less connectivity among the loci that were associated with fasting insulin and 2hGlu at genome-wide significance, with no genes reaching $P_{GRAIL} < 0.01$ for fasting insulin (**Supplementary Table 5b**) and only one out of nine genes reaching this threshold for 2hGlu ($P_{permutation} = 0.07$) (**Supplementary Table 5c**).

Among the list of 218 suggestively associated loci for fasting glucose ($P < 0.0005$), we observed that 13 genes were connected to the genome-wide significant loci at $P_{GRAIL} < 0.01$, more than expected by chance ($P_{permutation} = 0.003$) (**Supplementary Table 6a**). These included genes such as *GLP1R* ($P_{GRAIL} = 3.3 \times 10^{-7}$; a glucagon receptor that mediates the GLP-1 incretin effect and stimulates insulin release), *IRS2* ($P_{GRAIL} = 6.9 \times 10^{-5}$; central to development and maintenance of β-cell mass and function[46,47]) and *INS* ($P_{GRAIL} = 2.5 \times 10^{-6}$; the insulin gene encoding proinsulin). The presence of these genes and other biologically plausible genes support our conjecture that many of the SNPs approaching genome-wide significance are likely to represent true associations. Of the 155 suggestively associated loci for fasting insulin (adjusted for BMI), we observed that 7 were connected to the genome-wide significant loci at $P_{GRAIL} < 0.01$, more than expected by chance ($P_{permutation} = 0.002$), and these genes included *INSR* ($P_{GRAIL} = 1.5 \times 10^{-4}$; encoding insulin receptor precursor), *CD36* ($P_{GRAIL} = 0.001$; previously implicated in insulin resistance[48]), *GCG* ($P_{GRAIL} = 0.008$; glucagon gene) and *HNF1A* ($P_{GRAIL} = 0.005$; mutations in the gene are associated with MODY3 (ref. 49)) (**Supplementary Table 6b**). Of the 100 suggestively associated loci for 2hGlu ($P < 0.0005$), we found that 3 reached $P_{GRAIL} < 0.01$ ($P_{permutation} = 0.014$), and the gene highlighted as most biologically connected to the genome-wide significant loci was again *HNF1A* ($P_{GRAIL} = 3.4 \times 10^{-4}$) (**Supplementary Table 6c**).

Using MAGENTA, we identified four pathways enriched for fasting glucose associations: GOTERM pathways lens development in camera-type eye ($P = 0.004$), PANTHER processes gut mesoderm development ($P = 0.009$), other steroid metabolism ($P = 0.02$) and KEGG MODY pathway ($P = 0.03$), although these were no longer significant ($P > 0.05$) after removing lead genes, all of which were known fasting glucose loci (*PROX1* for eye and gut and *G6PC2* and *GCK* for steroid and MODY pathways, respectively).

Directional consistency between discovery and follow-up

Given the wealth of biologically plausible genes in loci with associations that almost reached genome-wide significance (**Supplementary Table 6a–c**) and the deviation of the observed distribution from the expected in quantile-quantile plots, even after removing all established loci (**Supplementary Fig. 8a–d**), we hypothesized that additional loci not reaching genome-wide significance were likely to represent true associations with small effects. To establish the presence of further true associations that did not reach genome-wide significance, we compared SNP associations in discovery studies (those included in the original meta-analyses for 42,078 (fasting glucose), 34,230 (fasting insulin) and 15,252 (2hGlu) individuals)[1,2] with those in the follow-up studies (consisting of 85,710 (fasting glucose), 69,240 (fasting insulin) and 27,602 (2hGlu) individuals). We identified all SNPs that had a nominally significant association ($P < 0.05$) in the follow-up studies alone and, for these SNPs, performed a binomial test to determine whether more SNPs than expected by chance (50%) had a consistent direction of effect with that observed in the discovery analyses. We were also able to make comparisons among SNPs that were nominated for follow-up analysis by different consortia (**Supplementary Fig. 9a–d**).

For each trait, evaluation of the 66,000 Metabochip follow-up SNPs revealed a significant excess of SNPs showing directionally consistent

© 2012 Nature America, Inc. All rights reserved.

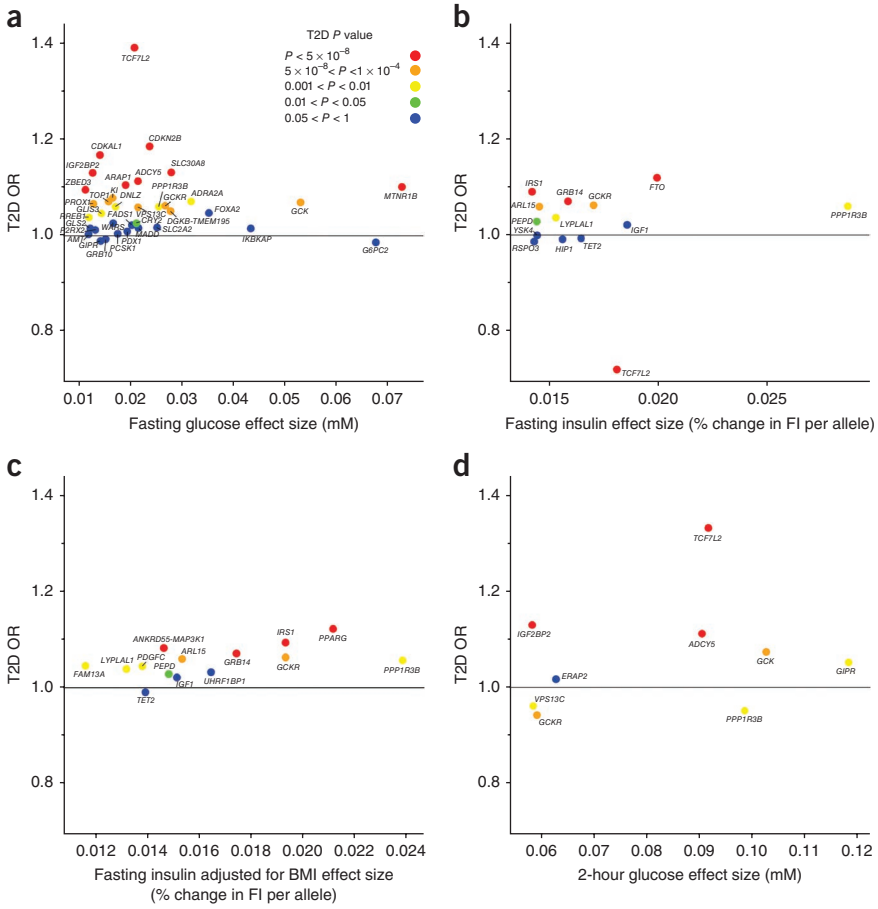

Figure 2 Per-allele β coefficients for glucose and insulin concentrations versus ORs for T2D. (**a**) Fasting glucose concentration versus T2D. (**b**) Fasting insulin (FI) concentration versus T2D. (**c**) Fasting insulin concentration adjusted for BMI versus T2D. (**d**) 2-hour glucose versus T2D.

associations ($P < 0.05$) compared to that expected by chance (fasting glucose, $P_{binomial} = 5.01 \times 10^{-12}$; fasting insulin, $P_{binomial} = 7.58 \times 10^{-13}$; fasting insulin (adjusted for BMI), $P_{binomial} = 9.76 \times 10^{-9}$; 2hGlu, $P_{binomial} = 2.37 \times 10^{-6}$; **Supplementary Fig. 9a–d** and **Supplementary Table 7**). FDR analyses suggested that a number of these nominal associations in the follow-up studies are true positives for fasting glucose and fasting insulin in particular (fasting glucose, 23%; fasting insulin, 24%; **Supplementary Table 7**). Notably, when we evaluated consistency of association with fasting insulin (between discovery and follow-up stages) among SNPs submitted to the Metabochip by other consortia, SNPs submitted by the Genetic Investigation of ANthropometric Traits (GIANT) Consortium (anthropometric traits) ($P_{binomial} = 1.52 \times 10^{-8}$) and the Global Lipids Genetics Consortium (GLGC) (lipid traits) ($P_{binomial} = 1.15 \times 10^{-6}$) and for BMI and triglycerides in particular also showed a marked excess of directional consistency (**Supplementary Fig. 9b** and **Supplementary Table 7**). When we performed the same test for fasting insulin concentration adjusted for BMI, the observed enrichment among SNPs submitted by GIANT and GLGC was attenuated (**Supplementary Fig. 9c** and **Supplementary Table 7**), although SNPs nominated to follow up on triglyceride associations remained the most significant ($P = 3.18 \times 10^{-7}$; **Supplementary Fig. 9c** and **Supplementary Table 7**). Of the 3,353 SNPs submitted for follow-up study of triglyceride associations, 158 SNPs showed nominal significance ($P < 0.05$) in follow-up studies and consistent direction of association with fasting insulin (adjusted for BMI) in both discovery and follow-up stages (**Supplementary Table 7**). In 139 (88%) of these SNPs, the insulin-raising alleles were associated with higher levels of triglycerides, consistent with the positive correlations between fasting insulin and triglyceride associations observed among the genome-wide significant loci for fasting insulin concentration (**Fig. 1**).

DISCUSSION

In the current meta-analysis of ~66,000 Metabochip follow-up SNPs in up to 133,010 individuals, we identified a large number of loci that associated with glycemic traits, explaining 4.8%, 1.2% and 1.7% of the variance in fasting glucose, fasting insulin and 2hGlu, respectively. Of the 53 glycemic loci, 33 are also associated with increased T2D risk ($q < 0.05$), extending the overlap between glycemic and T2D loci. Given the current DIAGRAM effective sample size of 106,953 individuals, we can exclude an effect on T2D of 1.04 with 80% power to detect alleles more frequent than 5%, effectively confirming that the overlap is incomplete and that many loci associated with glycemic traits have no discernible effect on T2D (**Figs. 1** and **2**).

Previously, we had detected only two loci associated with fasting insulin concentration and had hypothesized that this might be due to a different genetic architecture for this trait compared to fasting glucose, with potentially smaller effect sizes, lower frequency alleles or greater environmental influence on fasting insulin[1]. In the current meta-analysis including up to 108,557 individuals (compared to 62,264 individuals previously), we expanded the number of loci associated with this trait to 19. Of note was the effect of BMI adjustment on our ability to detect additional loci (five nonoverlapping with unadjusted results)[19]. We also noted that some of the loci influencing fasting insulin that were uncovered after BMI adjustment are likely to have been negatively confounded in previous efforts: at some loci, the insulin-raising allele was nominally associated with lower BMI (potentially via insulin resistance, attenuating the anabolic effects of insulin). Given the positive correlation between BMI and fasting insulin, it is likely that this association previously masked their effect on fasting insulin. Fasting insulin loci showed directionally consistent association with lipid levels (HDL and triglycerides); that is, the insulin-raising allele was associated with lower HDL and higher triglyceride levels, a hallmark combination in insulin-resistant individuals. We also observed some overlap between fasting insulin loci and those associated with abdominal obesity (**Fig. 1**). Jointly, these data suggest links of these fasting insulin loci to insulin resistance–related phenotypes. Indeed, some of the fasting insulin loci identified, such as *IRS1* and *PPARG*, are classically known to exert effects on insulin action or sensitivity[50,51].

There are now 36 established fasting glucose loci, many of which contain compelling biological candidate genes with plausible causality, including those encoding transcription factors with known roles in pancreas development (for example, *PDX1*, *FOXA2*, *PROX1* and *GLIS3*) and genes involved in β-cell function and insulin secretion

© 2012 Nature America, Inc. All rights reserved.

pathways (*SLC2A2*, *GCK* and *PCSK1*). For 2hGlu, only nine associated loci have been established to date, which likely reflects the smaller sample size available and the consequent reduction in power.

Comparing the consistency of the direction of associations for glycemic traits between discovery and follow-up studies suggests that we are observing more directionally consistent associations than expected by chance among Metabochip follow-up SNPs (**Supplementary Fig. 9a–d**). This finding, combined with the excess of biologically plausible genes among the loci on the edge of being significant (**Supplementary Table 6a–c**), suggests that, beyond the genome-wide significant loci, there is a more extensive list of loci still likely to contain true associations. Indeed, some of these loci are implicated by gene-based analyses, which identify genes with compelling biological credentials. For fasting insulin, these analyses revealed additional loci with previously suggested links to insulin resistance (*PPARD* and *PTEN*). These results lend further support to the proposal that a long tail of common variants of small effect size is likely to account for a substantial proportion of the variance of complex traits[7,8].

Of note is the number of glycemic loci associated with other metabolic traits ($q < 0.05$; 34 of 53) and also at genome-wide levels of significance ($P < 5 \times 10^{-8}$; 14 of 53) (**Fig. 1**), potentially implicating pleiotropic effects. Further support for this notion comes from the analysis of loci nominated for the Metabochip by other consortia and their associations with glycemic traits (**Supplementary Fig. 9a–d**). Indeed, some of the loci associated with glycemic traits at genome-wide significance levels were not originally nominated for the Metabochip for follow-up study by MAGIC (**Table 1**). Metabochip data available across all contributing consortia will facilitate systematic exploration of these correlated phenotypes with more sophisticated statistical methods for joint analysis[52–54], yielding greater insight into the underlying pathways and genetic networks they represent. As data from human genetic networks accrue, we will be better placed to test whether there is support for the notion of 'hub' genes—that is, genes highly connected with others in the network, proposed by experiments in *Caenorhabditis elegans* to act as buffers for genetic variation, that could act as modifier genes for many different disorders[55].

In summary, we present a large number of genome-wide significant loci influencing glycemic traits, many with a compelling biological basis for their association, as well as a number of loci not previously implicated in glycemic regulation, for which fine mapping and functional follow-up study will expand and improve understanding. Use of the Metabochip for deep follow up has identified additional loci involved in glycemic regulation that, due to insufficient sample size and power, did not reach genome-wide significance. Consideration of such loci in future studies will better exploit data from GWAS and complimentary approaches and further improve our biological understanding of glycemic control and the etiology of diabetes.

URLs. Cardiovascular Health Study (CHS), http://www.chs-nhlbi.org/pi.htm; Health2000, http://www.terveys2000.fi/; SNP&SEQ Technology Platform, www.genotyping.se/; Genetic Cluster Computer, http://www.geneticcluster.org/; UCSC Lift Genomes tool, http://genome.ucsc.edu/cgi-bin/hgLiftOver; GLGC data, http://www.sph.umich.edu/csg/abecasis/public/lipids2010/; GIANT Consortium, http://www.broadinstitute.org/collaboration/giant/index.php/Main_Page; MAGENTA, http://www.broadinstitute.org/mpg/magenta/; Human Protein Atlas, http://www.proteinatlas.org/.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

AGES: The AGES-Reykjavik study was supported by a contract from the National Institutes of Health (N01-AG-1-2100), National Institute on Aging Intramural Research Program, Hjartavernd (the Icelandic Heart Association) and the Althingi (the Icelandic Parliament).

ALSPAC: We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council (grant 74882), the Wellcome Trust (grants 076467 and 092731) and the University of Bristol provide core support for ALSPAC.

AMC-PAS: AMC-PAS is grateful to M.D. Trip and S. Sivapalaratnam for their input in collecting the data.

Amish: We gratefully thank our Amish community and research volunteers for their long-standing partnership in research and acknowledge the dedication of our Amish liaisons, field workers and the Amish Research Clinic staff, without whom these studies would not have been possible. The Amish studies are supported by grants and contracts from the US NIH, including R01 AG18728, R01 HL088119, U01 GM074518, U01 HL072515-06, U01 HL84756, R01 DK54261, the University of Maryland General Clinical Research Center grant M01 RR 16500, the Mid-Atlantic Nutrition Obesity Research Center grant P30 DK72488, the Baltimore Diabetes Research and Training Center grant P60DK79637 and the T32 training grant AG000219 (M.E.M.). In addition, this project was supported by National Research Initiative Competitive Grant 2007-35205-17883 from the US Department of Agriculture (USDA) National Institute of Food and Agriculture.

ARIC: We thank the staff and participants of the ARIC study for their important contributions. The Atherosclerosis Risk in Communities (ARIC) Study is carried out as a collaborative study that is supported by National Heart, Lung, and Blood Institute (NHLBI) contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, N01-HC-55022, R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and US NIH contract HHSN268200625226C. Infrastructure was partly supported by grant UL1RR025005, a component of the NIH and NIH Roadmap for Medical Research. We are grateful for resources provided by the University of Minnesota Supercomputing Institute.

ASAP: The ASAP study was funded by a donation from F. Lundberg.

ASCOT: We thank all ASCOT trial participants, physicians, nurses and practices in the participating countries for their important contributions to the study. In particular, we thank C. Muckian and D. Toomey for their help in DNA extraction, storage and handling. This work was supported by Pfizer for the ASCOT study and the collection of the ASCOT DNA repository, by Servier Research Group and by Leo Laboratories. Genotyping was funded by a Wellcome Trust Strategic Award (083948).

BLSA: The BLSA was supported in part by the Intramural Research Program of the National Institute on Aging of the US NIH. A portion of that support was through an R&D contract with the MedStar Research Institute.

Busselton Health Study (BSN): The Busselton Health Study acknowledges the generous support for the 1994/5 follow-up study from Healthway, Western Australia, the numerous Busselton community volunteers who assisted with data collection and the study participants from the Shire of Busselton. The BHS is supported by The Great Wine Estates of the Margaret River region of Western Australia.

CHS: This research at CHS was supported by NHLBI contracts N01-HC-85239, N01-HC-85079 through N01-HC-85086, N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150 and N01-HC-45133 and NHLBI grants HL080295, HL075366, HL087652 and HL105756, with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through AG-023629, AG-15928, AG-20098 and AG-027058 from the National Institute on Aging (see URLs). DNA handling and genotyping were supported in part by National Center for Research Resources grant M01-RR00425 to the Cedars-Sinai General Clinical Research Center Genotyping Core and National Institute of Diabetes and Digestive and Kidney Diseases grant DK063491 to the Southern California Diabetes Endocrinology Research Center. B. Psaty serves on a Data and Safety Monitoring Board (DSMB) for a clinical trial of a device funded by the manufacturer (Zoll-Lifecor).

CLHNS: We thank the Office of Population Studies Foundation research and data collection teams and the study participants who generously provided their time for this study. This work was supported by US NIH grants DK078150, TW05596, HL085144, RR20649, ES10126 and DK56350.

© 2012 Nature America, Inc. All rights reserved.

CoLaus: The authors also express their gratitude to the participants in the Lausanne CoLaus study and to the investigators who have contributed to the recruitment, in particular, Y. Barreau, A.-L. Bastian, B. Ramic, M. Moranville, M. Baumer, M. Sagette, J. Ecoffey and S. Mermoud for data collection. The CoLaus study was supported by research grants from the Swiss National Science Foundation (grant 33CSCO-122661), from GlaxoSmithKline and from the Faculty of Biology and Medicine of Lausanne, Switzerland. P.V. and G.W. received an unrestricted grant from GlaxoSmithKline to build the CoLaus study.

CROATIA-Korcula: The authors collectively thank a large number of individuals for their help in organizing, planning and carrying out the field work related to the project: S. Jankovic and staff at the University of Split Medical School; B. Salzer from the biochemistry laboratory 'Salzer'; local general practitioners and nurses; and the employees of several other Croatian institutions who participated in the field work, including but not limited to the University of Rijeka; the Croatian Institute of Public Health; and the Institutes of Public Health in Split and Dubrovnik. SNP genotyping of the Korcula samples was carried out by Helmholtz Zentrum München. The work is supported by the European Union framework program 6 EUROSPAN project (contract LSHG-CT-2006-018947) and grant ~216-1080315-0302 (to I.R.) from the Croatian Ministry of Science, Education and Sport. Studies carried out on the Croatian island of Korcula were supported by Medical Research Council, UK.

CROATIA-Split: The authors collectively thank a large number of individuals for their help in organizing, planning and carrying out the field work related to the project: S. Jankovic and staff at the University of Split Medical School; B. Salzer from the biochemistry laboratory 'Salzer'; local general practitioners and nurses; and the employees of several other Croatian institutions who participated in the field work. including but not limited to the University of Rijeka; the Croatian Institute of Public Health; and the Institutes of Public Health in Split and Dubrovnik. SNP genotyping of the Split samples was carried out by AROS Applied Biotechnology. The work is supported by grant ~216-1080315-0302 (to I.R.) from the Croatian Ministry of Science, Education and Sport. Studies carried out in the Croatian city of Split were supported by Medical Research Council, UK.

CROATIA-Vis: The authors collectively thank a large number of individuals for their help in organizing, planning and carrying out the field work related to the project: P. Rudan and staff of the Institute for Anthropological Research in Zagreb; S. Jankovic and staff at the University of Split Medical School; A. Vorko-Jovic and staff and medical students of the Andrija Stampar School of Public Health of the Faculty of Medicine at the University of Zagreb; B. Salzer from the biochemistry laboratory 'Salzer'; local general practitioners and nurses; and the employees of several other Croatian institutions who participated in the field work. including but not limited to the University of Rijeka; the Croatian Institute of Public Health; and the Institutes of Public Health in Split and Dubrovnik. SNP genotyping of the Vis samples was carried out by the Genetics Core Laboratory at the Wellcome Trust Clinical Research Facility at Western General Hospital. The work is supported by the European Union framework program 6 EUROSPAN project (contract LSHG-CT-2006-018947) and grant ~216-1080315-0302 (to I.R.) from the Croatian Ministry of Science, Education and Sport. Studies carried out on the Croatian island of Vis were supported by Medical Research Council, UK.

DESIR: We thank all the participants of the D.E.S.I.R study, E. Eury and S. Lobbens for technical support for the genotyping and O. Lantieri and M. Marre from the D.E.S.I.R study. Genotyping was supported by the Conseil Régional Nord-Pas-de-Calais Fonds Européen de Développement Economique et Regional CPER axe Cartdiodiabète 2010–2011 grant to N.B.-N.

deCODE: We thank the individuals who participated in the study and whose contributions made this work possible. The research performed at deCODE genetics was in part funded through the European Union's Seventh Framework Programme (FP7/2007-2013) ENGAGE project, grant HEALTH-F4-2007-201413.

DIAGEN: We are grateful to all of the patients who cooperated in this study and to their referring physicians and diabetologists in Saxony. The presented study was supported by the Commission of the European Communities, Directorate C–Public Health and Risk Assessment, Health & Consumer Protection, grant agreement 2004310, and by the Dresden University of Technology Funding Grant Med Drive.

DPS: The DPS has been financially supported by grants from the Academy of Finland (117844, 40758, 211497 and 118590), the EVO funding of the Kuopio University Hospital from the Ministry of Health and Social Affairs (5254), the Finnish Funding Agency for Technology and Innovation (40058/07), the Nordic Centre of Excellence on Systems biology in controlled dietary interventions and cohort studies (SYSDIET; 070014), the Finnish Diabetes Research Foundation, the Yrjö Jahnsson Foundation (56358), the Sigrid Juselius Foundation, the Juho Vainio Foundation and TEKES grants 70103/06 and 40058/07.

DR's EXTRA: The DR's EXTRA Study was supported by grants from the Ministry of Education and Culture of Finland (627;2004-2011), the Academy of Finland (102318 and 123885), Kuopio University Hospital, the Finnish Diabetes Association, the Finnish Heart Association, the Päivikki and Sakari Sohlberg Foundation, a European Union FP6 Integrated Project (EXGENESIS, LSHM-CT-2004-005272), the city of Kuopio and the Social Insurance Institution of Finland (4/26/2010).

EAS (Metabochip): EAS was supported by the British Heart Foundation. Genotyping was supported by a grant from the Chief Scientist Office, Scotland, and was performed at the Wellcome Trust Clinical Research Facility in Edinburgh.

EGCUT: EGCUT was financed by FP7 grants (201413 and 245536), a grant from the Estonian government (SF0180142s08), the European Union through the European Regional Development Fund, in the framework of the Centre of Excellence in Genomics and a grant from the University of Tartu (SP1GVARENG).

Ely: We are grateful to all the volunteers and to the staff of St. Mary's Street Surgery, Ely and the study team. The Ely Study was funded by the MRC (MC_U106179471) and Diabetes UK. Genotyping in the Ely and Fenland studies was supported in part by an MRC-GlaxoSmithKline pilot programme grant (G0701863).

ERF: We thank the participants from Genetic Research in Isolated Populations, Erasmus Rucphen Family and their general practitioners, who made this work possible. This study is financially supported by the Netherlands Organization for Scientific Research (NWO), the European Union framework program 6 EUROSPAN project (contract LSHG-CT-2006-018947), the ENGAGE project (grant HEALTH-F4-2007-201413), the International Stichting Alzheimer Onderzoek (ISAO), Hersenstichning Nederland (HSN) and the Centre for Medical Systems Biology (CMSB 1&2) in the framework of the Netherlands Genomics Initiative (NGI).

FamHS: The Family Heart Study (FamHS) was supported by US NIH grants RO1-HL-087700 and RO1-HL-088215 (to M.A.P.) from HNLBI and RO1-DK-8925601 and RO1-DK-075681 (to I.B.B.) from NIDDK.

Fenland: The Fenland Study is funded by the Wellcome Trust and the Medical Research Council (MC_U106179471). We are grateful to all the volunteers for their time and help, and to the General Practitioners and practice staff for assistance with recruitment. We thank the Fenland Study Investigators, Fenland Study Co-ordination team and the Epidemiology Field, Data and Laboratory teams. Biochemical assays were performed by the National Institute for Health Research, Cambridge Biomedical Research Centre, Core Biochemistry Assay Laboratory, and the Cambridge University Hospitals NHS Foundation Trust, Department of Clinical Biochemistry.

FIN-D2D 2007: The FIN-D2D study has been financially supported by the hospital districts of Pirkanmaa, South Ostrobothnia and Central Finland, the Finnish National Public Health Institute (current National Institute for Health and Welfare), the Finnish Diabetes Association, the Ministry of Social Affairs and Health in Finland, the Academy of Finland (grant 129293), the Commission of the European Communities, Directorate C–Public Health (grant agreement 2004310) and Finland's Slottery Machine Association.

FINRISK/DILGOM: The DILGOM study received support from the Etelä-Pohjanmaa Hospital District, the Pohjois-Pohjanmaa Hospital District, the Keski-Suomi Hospital District, the Pirkanmaa Hospital District and the Pohjois-Savo Hospital District. The DILGOM survey was funded by the Academy of Finland (grant 118065). V.S. was supported by the Academy of Finland (grant 139635 and 129494).

Framingham Heart Study: This research was conducted in part using data and resources from the Framingham Heart Study of the NHLBI of the US NIH and the Boston University School of Medicine. The analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. This work was partially supported by the NHLBI's Framingham Heart Study (contract N01-HC-25195) and its contract with Affymetrix for genotyping services (contract N02-HL-6-4278). A portion of this research used the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. The study is also supported by NIDDK R01 DK078616 to J.B.M., J.D. and J.C.F. and NIDDK K24 DK080140 to J.B.M.

FUSION: Support for FUSION was provided by NIH grants R01-DK062370 (to M. Boehnke), R01-DK072193 (to K.L.M.) and intramural project number 1Z01-HG000024 (to F.S.C.). Genome-wide genotyping was conducted by the Johns Hopkins

© 2012 Nature America, Inc. All rights reserved.

University Genetic Resources Core Facility SNP Center at the Center for Inherited Disease Research (CIDR), with support from CIDR NIH contract N01-HG-65403.

GEMS: We thank the investigators S. Grundy, P. Barter, R. McPherson, R. Mahley, T. Bersot and A. Kesaniemi for collection of the samples. The GEMS study was sponsored in part by GlaxoSmithKline.

GENOA: We thank E. Boerwinkle and J. Cunningham for their help with genotyping. The Genetic Epidemiology Network of Arteriopathy (GENOA) study is supported by US NIH grants HL087660 and HL100245 from the NHLBI.

GenomEUtwin: We acknowledge support from the European Commission under Quality of Life and Management of the Living Resources of the Fifth Framework Program (GenomEUtwin QLG2-CT-2002-01254). The study is also supported by ENGAGE–European Network for Genetic and Genomic Epidemiology, FP7-HEALTH-F4-2007, grant agreement 201413.

GLACIER: The GLACIER Study is nested within the Northern Sweden Health and Disease Study, and phenotyping was conducted as part of the Västerbotten Intervention Project. We thank the participants and the investigators from these studies for their valuable contributions, with specific thanks to L. Weinehall, Å. Agren, K. Enquist and T. Johansson. The GLACIER Study and part the salary of P.W.F. were funded by grants from the Swedish Research Council, the Swedish Heart-Lung Foundation, Novo Nordisk, the Umeå Medical Research Foundation and the Swedish Diabetes Association (to P.W.F.). Genotyping for this specific project was funded by the Wellcome Trust Sanger Institute. I.B. acknowledges funding from Wellcome Trust grant 098051, the UK NIHR Cambridge Biomedical Research Centre and the MRC Centre for Obesity and Related Metabolic Diseases. We thank E. Gray, D. Simpkin, S. Hunt and the staff of the Wellcome Trust Sanger Institute Sample Logistics, Genotyping and Variation Informatics Facilities.

GoDARTS: The study was funded by the Wellcome Trust, Tenovus Tayside and the Medical Research Council, UK.

Health ABC: This study was supported by National Institute on Aging contracts N01AG62101, N01AG62103 and N01AG62106 and in part by the Intramural Research Program of the US NIH, the National Institute on Aging. The GWAS was funded by National Institute on Aging grant 1R01AG032098-01A1 to Wake Forest University Health Sciences, and genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the NIH to Johns Hopkins University, contract HHSN268200782096C.

Health2000: The Health 2000 Study is funded by the National Institute for Health and Welfare (THL), the Finnish Centre for Pensions (ETK), the Social Insurance Institution of Finland (KELA), the Local Government Pensions Institution (KEVA) and other organizations listed on the website of the survey. GWAS genotyping was supported by the Wellcome Trust Sanger Institute.

InChianti: The InCHIANTI study baseline (1998–2000) was supported as a targeted project (ICS110.1/RF97.71) by the Italian Ministry of Health and in part by the US National Institute on Aging (contracts 263 MD 9164 and 263 MD 821336).

KORA F4: We thank all members of field staff who were involved in the planning and conduct of the MONICA/KORA Augsburg studies. The MONICA/KORA Augsburg studies were financed by the Helmholtz Zentrum München–Research Center for Environment and Health and supported by grants from the German Federal Ministry of Education and Research (BMBF), the Federal Ministry of Health, the Ministry of Innovation, Science, Research and Technology of North Rhine-Westphalia, the German National Genome Research Network (NGFN) and the Munich Center of Health Sciences (MC Health) as part of LMUinnovativ.

LEIPZIG_ADULT_IFB: We thank all those who participated in the studies. This work was supported by grants from Integrated Research and Treatment Centre (IFB) Adiposity Diseases (K7-36 to A. Körner) and from the Clinical Research Group Atherobesity KFO 152 (projects BL 833/1-1 and Stu192/6-1).

LEIPZIG_CHILHOOD_IFB: We are grateful to all the patients and families for contributing to the study. We greatly appreciate the support of the Obesity Team and the Auxo Team of the Leipzig University Children's Hospital for management of the patients and to the Pediatric Research Center Lab Team for support with DNA banking. This work was supported by grants from Integrated Research and Treatment Centre (IFB) Adiposity Diseases (K7-36 to A. Körner) and from the Clinical Research Group Atherobesity KFO 152 (projects KO3512/1-2 to A.K.).

LURIC: The authors extend appreciation to the participants in the LURIC study; without their collaboration, this article would not have been written. We thank the LURIC study team members who are either temporarily or permanently involved in patient recruitment, sample collection and data handling and the laboratory staff at the Ludwigshafen General Hospital and the Universities of Freiburg and Ulm. LURIC received funding through the Sixth Framework Programme (integrated project Bloodomics, grant LSHM-CT-2004-503485) and Seventh Framework Programme (integrated project AtheroRemo, grant agreement 201668) of the European Union.

METSIM: The METSIM study was funded by the Academy of Finland (grants 77299 and 124243).

MICROS: For the MICROS study, we thank the primary care practitioners R. Stocker, S. Waldner, T. Pizzecco, J. Plangger and U. Marcadent and the personnel of the Hospital of Silandro (Department of Laboratory Medicine) for their participation and collaboration in the research project. In South Tyrol, the study was supported by the Ministry of Health and Department of Educational Assistance, University and Research of the Autonomous Province of Bolzano and the South Tyrolean Sparkasse Foundation.

NFBC66: We thank P. Rantakallio (launch of NFBC1966 and 1986) and O. Tornwall and M. Jussila (DNA biobanking). The authors would like to acknowledge the contribution of the late Academian of Science Leena Peltonen. NFBC1986(1966) received financial support from the Academy of Finland (project grants 104781, 120315, 129269, 1114194, Center of Excellence in Complex Disease Genetics and SALVE), University Hospital Oulu, Biocenter, University of Oulu, Finland (75617), the European Commission (EURO-BLCS, Framework 5 award QLG1-CT-2000-01643), NHLBI grant 5R01HL087679-02 through the STAMPEED program (1RL1MH083268-01), US NIH/National Institute of Mental Health (NIMH) (5R01MH63706:02), ENGAGE project and grant agreement HEALTH-F4-2007-201413, the Medical Research Council, UK (G0500539, G0600705 and PrevMetSyn/SALVE) and the Wellcome Trust (project grant GR069224). DNA extraction, sample quality control, biobank upkeeping and aliquotting were performed at the National Public Health Institute at Biomedicum Helsinki and were supported financially by the Academy of Finland and Biocentrum Helsinki.

NFBC86: The research of V.L. is funded in part through the European Community's Seventh Framework Programme (FP7/2007-2013), ENGAGE project, grant agreement HEALTH-F4-2007- 201413. The research of I.P. is funded in part through the European Community's Seventh Framework Programme (FP7/2007-2013), ENGAGE project, grant agreement HEALTH-F4-2007-201413.

NTRNESDA: Funding was obtained from the Netherlands Organization for Scientific Research (NWO; MagW/ZonMW, 904-61-090, 985-10-002, 904-61-193, 480-04-004, 400-05-717, Addiction-31160008 and Middelgroot- 911-09-032); Spinozapremie (SPI 56-464-14192); the Center for Medical Systems Biology (CMSB) (NWO Genomics); NBIC/BioAssist/RK/2008.024; Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL; 184.021.007); the VU University Institute for Health and Care Research (EMGO+) and Neuroscience Campus Amsterdam (NCA); the European Science Foundation (ESF): Genomewide Analyses of European Twin and Population Cohorts (EU/QLRT-2001-01254); the European Community's Seventh Framework Program (FP7/2007-2013): ENGAGE (HEALTH-F4-2007-201413); the European Science Council (ERC) Genetics of Mental Illness (230374); Rutgers University Cell and DNA Repository cooperative agreement (NIMH U24 MH068457-06); Collaborative Study of the Genetics of DZ Twinning (US NIH R01D0042157-01A); and the Genetic Association Information Network, a public-private partnership between the NIH and Pfizer, Affymetrix and Abbott Laboratories.

ORCADES: We would like to acknowledge the invaluable contributions of L. Anderson, the research nurses in Orkney and the administrative team in Edinburgh. ORCADES was supported by the Scottish Executive Health Department, the Royal Society and the European Union Framework Programme 6 EUROSPAN project (contract LSHG-CT-2006-018947). DNA extractions were performed at the Wellcome Trust Clinical Research Facility in Edinburgh.

PIVUS: Genotyping was performed by the SNP&SEQ Technology Platform in Uppsala (see URLs). We thank T. Axelsson, A.-C. Wiman and C. Pöntinen for their excellent assistance with genotyping. The SNP Technology Platform is supported by Uppsala University, Uppsala University Hospital and the Swedish Research Council for Infrastructures. E.I. is supported by grants from the Swedish Research Council, the Swedish Heart-Lung Foundation, the Swedish Foundation for Strategic Research and the Royal Swedish Academy of Science.

PREVEND: PREVEND genetics is supported by the Dutch Kidney Foundation (grant E033), the US National Institutes of Health (grant LM010098), The Netherlands Organization for Health Research and Development (NWO VENI grant 916.761.70) and the Dutch Interuniversity Cardiology Institute Netherlands (ICIN).

© 2012 Nature America, Inc. All rights reserved.

PROCARDIS: The PROCARDIS study was supported by the European Community Sixth Framework Programme (LSHM-CT- 2007-037273), AstraZeneca, the British Heart Foundation, the Oxford BHF Centre of Research Excellence, the Wellcome Trust (075491/Z/04), the Swedish Research Council, the Knut and Alice Wallenberg Foundation, the Swedish Heart-Lung Foundation, the Torsten and Ragnar Söderberg Foundation, the Strategic Cardiovascular Program of Karolinska Institutet and Stockholm County Council, the Foundation for Strategic Research and the Stockholm County Council (560283).

PROSPER: The PROSPER/PHASE study has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement HEALTH-F2-2009-223004. PROSPER/PHASE is supported by grants from the Interuniversity Cardiology Institute of the Netherlands (ICIN) and the Durrer Center for Cardiogenetic Research, both of which are Institutes of the Netherlands Royal Academy of Arts and Sciences (KNAW); the Center for Medical Systems Biology (CMSB), a center of excellence approved by the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research (NWO); and the Netherlands Consortium for Healthy Ageing (NCHA). The research leading to the PROSPER study was sponsored by Bristol Myers Squibb.

Rotterdam Study: The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. We thank P. Arp, M. Jhamai, M. Verkerk, L. Herrera and M. Peters for their help in creating the GWAS database and K. Estrada and M.V. Struchalin for their support in the creation and analysis of imputed data. The generation and management of GWAS genotype data for the Rotterdam Study are supported by the Netherlands Organization of Scientific Research NWO Investments (175.010.2005.011 and 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), Netherlands Genomics Initiative (NGI)/Netherlands Organization for Scientific Research (NWO) project 050-060-810, the European Community's Seventh Framework Programme (FP7/2007-2013) and the ENGAGE project, grant agreement HEALTH-F4-2007-201413. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, the Netherlands Organization for Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII) and the Municipality of Rotterdam.

SardiNIA: The authors thank all the volunteers and the mayors of the four towns involved. This work was supported in part by the Intramural Research Program of the National Institute on Aging through the US NIH and by contract NO1-AG-1-2109 from the NIA to the SardiNIA (ProgeNIA) team.

SCARFSHEEP: Funding for the study was provided by the European Commission (LSHM-CT-2007-037273), the Swedish Heart-Lung Foundation, the Swedish Research Council (2669, 8691), the Knut and Alice Wallenberg Foundation, the Foundation for Strategic Research, the Torsten and Ragnar Söderberg Foundation, the Strategic Cardiovascular Programme of Karolinska Institutet and the Stockholm County Council (560183).

SORBS: We thank all those who participated in the studies. This work was supported by grants from Integrated Research and Treatment Centre (IFB) Adiposity Diseases (K7-36 to M.S. and A. Körner), from the Clinical Research Group Atherobesity KFO 152 (projects BL 833/1-1 to M.B. and Stu192/6-1 to M.S.). R.M. acknowledges financial support from the European Commission under a Marie Curie Intra-European Fellowship. P.K. acknowledges financial support from the Boehringer Ingelheim Foundation.

SUVIMAX : This work was supported by the Institut National de la Santé et de la Recherche Médicale, the Institut National de la Recherche Agronomique, the Université Paris 13, the Centre National de Génotypage and the Commissariat à L'Energie Atomique.

Swedish Twin Registry: We thank T. Axelsson, A.-C. Wiman and C. Pöntinen for their excellent assistance with genotyping. This work was supported by grants from the US NIH (AG028555, AG08724, AG04563, AG10175 and AG08861), the Swedish Research Council, the Swedish Heart-Lung Foundation, the Swedish Foundation for Strategic Research, the Royal Swedish Academy of Science and ENGAGE (within the European Union Seventh Framework Programme, HEALTH-F4-2007-201413). Genotyping was performed by the SNP&SEQ Technology Platform (see URLs). The SNP Technology Platform is supported by Uppsala University, Uppsala University Hospital and the Swedish Research Council for Infrastructures.

THISEAS: We thank all the dieticians and clinicians for their contribution to the project. Recruitment for THISEAS was partially funded by a research grant (PENED 2003) from the Greek General Secretary of Research and Technology.

TRAILS: TRAILS (TRacking Adolescents' Individual Lives Survey) is a collaborative project involving various departments of the University Medical Center and University of Groningen, the Erasmus University Medical Center Rotterdam, the University of Utrecht, the Radboud Medical Center Nijmegen and the Parnassia Bavo group, all in The Netherlands. We are grateful to all adolescents and their parents and teachers who participated in this research and to everyone who worked on this project and made it possible. TRAILS has been financially supported by grants from the Netherlands Organization for Scientific Research NWO (Medical Research Council program grant GB-MW 940-38-011; ZonMW Brainpower grant 100-001-004; ZonMw Risk Behavior and Dependence grants 60-60600-98-018 and 60-60600-97-118; ZonMw Culture and Health grant 261-98-710; Social Sciences Council medium-sized investment grants GB-MaGW 480-01-006 and GB-MaGW 480-07-001; Social Sciences Council project grants GB-MaGW 457-03-018, GB-MaGW 452-04-314 and GB-MaGW 452-06-004; NWO large-sized investment grant 175.010.2003.005; and NWO Longitudinal Survey and Panel Funding 481-08-013); the Sophia Foundation for Medical Research (projects 301 and 393), the Dutch Ministry of Justice (WODC), the European Science Foundation (EuroSTRESS project FP-006) and the participating universities. Statistical analyses were carried out on the Genetic Cluster Computer (see URLs), which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003) along with a supplement from the Dutch Brain Foundation.

TwinsUK: The study was funded by the Wellcome Trust; European Community's Seventh Framework Programme (FP7/2007-2013), grant agreement HEALTH-F2-2008-201865-GEFOS, (FP7/2007-2013), ENGAGE project grant agreement HEALTH-F4-2007-201413, and the Fifth Framework Programme GenomEUtwin Project (QLG2-CT-2002-01254). The study also receives support from the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St. Thomas' National Health Service (NHS) Foundation Trust in partnership with King"s College London. T.D.S. is an NIHR senior Investigator. The project also received support from a Biotechnology and Biological Sciences Research Council (BBSRC) project grant (G20234). The authors acknowledge the funding and support of the National Eye Institute via an NIH/CIDR genotyping project (to T. Young).

ULSAM: Genotyping was performed by the SNP&SEQ Technology Platform in Uppsala (see URLs). We thank T. Axelsson, A.-C. Wiman and C. Pöntinen for their excellent assistance with genotyping. The SNP Technology Platform is supported by Uppsala University, Uppsala University Hospital and the Swedish Research Council for Infrastructures. E.I. is supported by grants from the Swedish Research Council, the Swedish Heart-Lung Foundation, the Swedish Foundation for Strategic Research, and the Royal Swedish Academy of Science.

Whitehall II: The WHII study has been supported by grants from the Medical Research Council; the Economic and Social Research Council; British Heart Foundation (BHF); the Health and Safety Executive; the Department of Health; the NHLBI (HL36310), the US, NIH: National Institute on Aging (AG13196), US, NIH; the Agency for Health Care Policy Research (HS06516); and the John D. and Catherine T. MacArthur Foundation Research Networks on Successful Midlife Development and Socioeconomic Status and Health. Genotyping in WHII was supported by BHF grant PG/07/133/24260 and by an MRC-GlaxoSmithKline pilot programme grant (85374). S. Raychaudhuri is supported by the US NIH (K08AR055688).

AUTHOR CONTRIBUTIONS

Writing group: R.A.S., V.L., R.P.W., E.W., M.M., R.J.S., N.B.-N., M.I.M., P.W.F., J.B.M., T.M.T., J.C.F., C. Langenberg, E.I., I.P. and I.B. wrote the manuscript. All authors reviewed the manuscript.

Project design, management and coordination: (AGES) V.G., T.B.H.; (ALSPAC) G.D.S.; (AMC-PAS) G.K.H.; (Amish) A.R.S.; (ASAP) L. Folkersen; (ASCOT) M.C., P.B.M.; B.F.P.; (BLSA) L. Ferrucci.; (BSN) J.P.B.; (CHS) B.M.P.; (CLHNS) L.S.A., K.L.M.; (CoLaus) P.V.; (CROATIA-Vis) N.D.H., I.R.; (deCODE) K. Stefansson; (DESIR) B.B., P.F.; (DIAGEN) S.R.B., P.E.H.S.; (DPS) M.U.; (DR's EXTRA) T.A.L., R.R.; (EAS) J.F. Price; (EGCUT) A. Metspalu; (Ely/Fenland) N.J.W., N.G.F., C. Langenberg, R.J.F.L.; (ERF) B.A.O.; (FIN-D2D 2007) S.M.K.-K., T.E.S.; (FINRISK/DILGOM) S. Ripatti, V.S.; (FamHS) M.A.P., I.B.B.; (Framingham Heart Study) J.B.M.; (FUSION) R.N.B., M. Boehnke, F.S.C., K.L.M., J.T., R.M.W.; (GEMs) D.M.W.; (GENOA) P.A.P.; (GenomEUtwin) D.I.B., J. Kaprio, K.O.K., P.K.E.M., N.L.P.; (GLACIER) I.B., P.W.F., T.V.V.; (GRAIL) S. Raychaudhuri; (Health ABC) T.B.H.; (InChianti) L. Ferrucci; (KORA F3) H.G., A. Peters; (KORA F4) H.G., T.I.; (LEIPZIG_ADULT_IFB) A.T.; (LEIPZIG_CHILDHOOD_IFB) A. Körner, W.K.; (LURIC) W.M., B.O.B.; (METSIM) J. Kuusisto, M.L.; (MICROS) P.P.P.; (NFBC66) M.-R.J., M.I.M.; (NFBC86) M.-R.J., M.I.M., A. Pouta; (NTRNESDA) D.I.B.,

© 2012 Nature America, Inc. All rights reserved.

E.J.C.d.G., B.W.P.; (ORCADES-Discovery) H. Campbell; (ORCADES-Metabochip) J.F.W., A.F.W.; (PIVUS) E.I., L.L.; (PREVEND) P.v.d.H.; (PROCARDIS) M.F., H.W.; (PROSPER) N.S.; (Rotterdam) A. Hofman, C.M.v.D.; (SardiNIA) F.C.; (SCARFSHEEP) U.d.F., A. Hamsten; (SORBS) I.P., M.S., A.T.; (SUVIMAX) P.M.; (Swedish Twin Registry) E.I.; (THISEAS) G.V.D.; (THISEAS-PROMIS-AMCPAS) P.D.; (TRAILS) A.J.O.; (TwinsUK) T.D.S.; (GoDARTS) A.D.M.; (ULSAM) E.I.; (Whitehall II) A. Hingorani, M. Kivimaki.

Sample collection and phenotyping: (AGES) V.G., T.B.H.; (AMC-PAS) H.V.B.; (Amish) A.R.S.; (ARIC) J.S.P.; (BSN) J.P.B., J.H., A.L.J.; (CHS) B.M.P.; (CLHNS) L.S.A.; (CoLaus) P.V.; (CROATIA-Korcula) O.P.; (DESIR) B.B.; (DPS) J. Lindström; (EAS) G.R.F., J.F. Price; (EGCUT) A. Metspalu; (Ely/Fenland) N.J.W., N.G.F.; (ERF) B.A.O.; (FINRISK/DILGOM) H.O., V.S.; (Framingham Heart Study) J.C.F., J.B.M.; (GEMs) D.M.W.; (GenomEUtwin) D.I.B., J. Kaprio, K.O.K., P.K.E.M., N.L.P.; (GLACIER) G.H.; (Health ABC) I.M.; (Health 2000) J.A., M.P.; (IMPROVE) B.S.; (InChianti) S.B.; (KORA F3) W.R.; (KORA F4) C. Herder; (LEIPZIG_ADULT_IFB) P.K., A.T.; (LEIPZIG_CHILDHOOD_IFB) W.K., A. Körner; (METSIM) J. Kuusisto, M.L.; (MICROS) P.P.P.; (NTRNESDA) J.H.S., G.W.; (ORCADES-Discovery) S.H.W.; (ORCADES-Metabochip) J.F.W.; (PIVUS) E.I., L.L.; (PREVEND) S.J.L.B.; (PROSPER) J.W.J.; (Rotterdam) A. Hofman, C.M.v.D.; (SORBS) A.T.; (SUVIMAX) S.H.; (Swedish Twin Registry) E.I., N.L.P.; (THISEAS) M.D.; (TwinsUK) T.D.S.; (GoDARTS) A.S.F.D., C.N.A.P.; (ULSAM) E.I.

Genotyping: (ARIC) K.E.N.; (BSN) J.P.B., J.H.; (CoLaus) J.S.B.; (CROATIA-Korcula) S.C.; (CROATIA-Split) C. Hayward, T.Z.; (CROATIA-Vis) N.D.H.; (DESIR) N.B.-N., P.F.; (DR's EXTRA) M. Atalay; (EAS) J.L.B.; (EGCUT) T.E.; (Ely/Fenland) C. Langenberg, R.J.F.L., N.J.W.; (ERF) A.I., B.A.O.; (FINRISK/DILGOM) S. Ripatti; (Framingham Heart Study) J.D., J.C.F., J.B.M.; (FUSION) L.L.B.; (GEMs) D.M.W.; (GenomEUtwin) D.I.B., J.-J.H., J. Kaprio, K.O.K., P.K.E.M., N.L.P.; (GLACIER) S.E.; (Health ABC) Y.L.; (KORA F3) H.G., M.M.-N.; (KORA F4) H.G., T.I.; (LEIPZIG_CHILDHOOD_IFB) A. Körner; (MICROS) A.A.H.; (NTRNESDA) J.-J.H., B.W.P., G.W.; (ORCADES-Discovery) H. Campbell; (ORCADES-Metabochip) J.F.W., A.F.W.; (PIVUS) E.I.; (PREVEND) P.v.d.H.; (PROSPER) S.T., J.W.J.; (Rotterdam) F.R., A.G.U.; (SardiNIA) M. Antonella, R.N.; (SCARFSHEEP) B.G.; (Swedish Twin Registry) E.I., N.L.P.; (THISEAS-PROMIS-AMCPAS) K. Stirrups; (TRAILS) M. Bruinenberg; (TwinsUK) M.M., K. Small; (GoDARTS) A.S.F.D., C.N.A.P.; (ULSAM) E.I.; (Whitehall II) M. Kumari, C. Langenberg, N.J.W.

Statistical analysis: (AGES) A.V.S.; (ALSPAC) D.M.E., B.S.P., N.J.T.; (Amish) M.E.M., J.R.O.; (ARIC) L.J.R.-T.; (ASAP) A.F.-C.; (ASCOT) T.J.; B.F.P.; (BLSA) T.T.; (BSN) J.P.B., J.H.; (CLHNS) Y.W.; (CoLaus) J.S.B.; (CROATIA-Korcula) P.N.; (CROATIOA-Split) C. Hayward, V.V.; (deCODE) A. Kong, G.T.; (DESIR) N.B.-N., C. Lecoeur, L.Y.; (EGCUT) T.E., E.M.; (Ely/Fenland) J. Luan, R.A.S.; (ERF) A.I.; (FINRISK/DILGOM) S. Ripatti; (FamHS) P.A.; (Framingham Heart Study) H. Chen, J.D., C.-T.L., D.R.; (FUSION) P.S.C., A.U.J., H.M.K., H.M.S., T.M.T., R.P.W.; (GEMs) K. Song; (GENOA) L.F.B., W.Z.; (GenomEUtwin) J.-J.H.; (GLACIER) D.S., E.W.; (Health ABC) J.S.A.; (Health 2000) M.P., P.S.; (InChianti) T.T.; (KORA F3) H.G., M.M.-N.; (KORA F4) C.G., H.G.; (LEIPZIG_ADULT_IFB) A. Mahajan; (LEIPZIG_CHILDHOOD_IFB) A. Mahajan; L.M.; (LURIC) M.E.K.; (MICROS) A.A.H.; (NFBC66) C.M.L., A.P.M., I.P.; (NFBC86) V.L., I.P., N.W.R., N.R.; (NTRNESDA) J.-J.H.; (ORCADES-Metabochip) R.M.F.; (PIVUS) E.I., E.R.; (PREVEND) P.v.d.H., N.V.; (PROCARDIS) M.F., A.G., J.F. Peden; (PROSPER) I.F., P.C.D.J., J.W.J.; (Rotterdam) S.M.W.; (SardiNIA) S. Sanna, C.S.; (SCARFSHEEP) R.J.S.; (SORBS) R.M.; (SUVIMAX) T.J.; (Swedish Twin Reg.) T.F., E.I.; (THISEAS-PROMIS-AMCPAS) S.K., K. Stirrups; (TRAILS) H.J.; (GoDARTS) A.S.F.D., M.H., C.N.A.P.; (ULSAM) S.G., E.I., E.R.; (Whitehall II) J. Luan, S. Shah.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

Published online at http://www.nature.com/doifinder/10.1038/ng.2385.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Dupuis, J. *et al.* New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat. Genet.* **42**, 105–116 (2010).
2. Saxena, R. *et al.* Genetic variation in *GIPR* influences the glucose and insulin responses to an oral glucose challenge. *Nat. Genet.* **42**, 142–148 (2010).
3. Prokopenko, I. *et al.* Variants in *MTNR1B* influence fasting glucose levels. *Nat. Genet.* **41**, 77–81 (2009).
4. Barker, A. *et al.* Association of genetic loci with glucose levels in childhood and adolescence: a meta-analysis of over 6,000 children. *Diabetes* **60**, 1805–1812 (2011).
5. Ingelsson, E. *et al.* Detailed physiologic characterization reveals diverse mechanisms for novel genetic loci regulating glucose and insulin metabolism in humans. *Diabetes* **59**, 1266–1275 (2010).
6. Voight, B.F. *et al.* Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat. Genet.* **42**, 579–589 (2010).
7. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565–569 (2010).
8. Park, J.H. *et al.* Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nat. Genet.* **42**, 570–575 (2010).
9. Voight, B.F. *et al.* The Metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. *PLoS Genet.* **8**, e1002793 (2012).
10. Liu, J.Z. *et al.* A versatile gene-based test for genome-wide association studies. *Am. J. Hum. Genet.* **87**, 139–145 (2010).
11. Kuro-o, M. *et al.* Mutation of the mouse *klotho* gene leads to a syndrome resembling ageing. *Nature* **390**, 45–51 (1997).
12. Ohnishi, M., Kato, S., Akiyoshi, J., Atfi, A. & Razzaque, M.S. Dietary and genetic evidence for enhancing glucose metabolism and reducing obesity by inhibiting klotho functions. *FASEB J.* **25**, 2031–2039 (2011).
13. Utsugi, T. *et al.* Decreased insulin production and increased insulin sensitivity in the *klotho* mutant mouse, a novel animal model for human aging. *Metabolism* **49**, 1118–1123 (2000).
14. Rhee, E.J. *et al.* Relationship between polymorphisms G395A in promoter and C1818T in exon 4 of the *KLOTHO* gene with glucose metabolism and cardiovascular risk factors in Korean women. *J. Endocrinol. Invest.* **29**, 613–618 (2006).
15. Paroni, G. *et al.* Klotho locus, metabolic traits, and serum hemoglobin in hospitalized older patients: a genetic association analysis. *Age (Omaha)* published online (22 June 2011); doi:10.1007/s11357–011–9273-x.
16. Rampersaud, E. *et al.* Identification of novel candidate genes for type 2 diabetes from a genome-wide association scan in the Old Order Amish: evidence for replication from diabetes-related quantitative traits and from independent populations. *Diabetes* **56**, 3053–3062 (2007).
17. Tabassum, R., Chavali, S., Dwivedi, O.P., Tandon, N. & Bharadwaj, D. Genetic variants of *FOXA2*: risk of type 2 diabetes and effect on metabolic traits in North Indians. *J. Hum. Genet.* **53**, 957–965 (2008).
18. Xing, C., Cohen, J.C. & Boerwinkle, E. A weighted false discovery rate control procedure reveals alleles at *FOXA2* that influence fasting glucose levels. *Am. J. Hum. Genet.* **86**, 440–446 (2010).
19. Manning, A.K. *et al.* A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat. Genet.* **44**, 659–669 (2012).
20. Jonsson, J., Carlsson, L., Edlund, T. & Edlund, H. Insulin-promoter-factor 1 is required for pancreas development in mice. *Nature* **371**, 606–609 (1994).
21. Stoffers, D.A., Ferrer, J., Clarke, W.L. & Habener, J.F. Early-onset type-II diabetes mellitus (MODY4) linked to *IPF1*. *Nat. Genet.* **17**, 138–139 (1997).
22. Stoffers, D.A., Zinkin, N.T., Stanojevic, V., Clarke, W.L. & Habener, J.F. Pancreatic agenesis attributable to a single nucleotide deletion in the human *IPF1* gene coding sequence. *Nat. Genet.* **15**, 106–110 (1997).
23. Nicolino, M. *et al.* A novel hypomorphic *PDX1* mutation responsible for permanent neonatal diabetes with subclinical exocrine deficiency. *Diabetes* **59**, 733–740 (2010).
24. Morris, A.P. *et al.* Large-scale association analysis of the Metabochip provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat. Genet.* published online (12 August 2012); doi:10.1038/ng.2383.
25. 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).
26. 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).
27. Teslovich, T.M. *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707–713 (2010).
28. Irwin, D.M. & Tan, H. Molecular evolution of the vertebrate hexokinase gene family: identification of a conserved fifth vertebrate hexokinase gene. *Comp. Biochem. Physiol. Part D Genomics Proteomics* **3**, 96–107 (2008).
29. 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).
30. Richards, J.B. *et al.* A genome-wide association study reveals variants in *ARL15* that influence adiponectin levels. *PLoS Genet.* **5**, e1000768 (2009).
31. Yujiri, T. *et al.* MEK kinase 1 interacts with focal adhesion kinase and regulates insulin receptor substrate-1 expression. *J. Biol. Chem.* **278**, 3846–3851 (2003).
32. Meyer, C.F., Wang, X., Chang, C., Templeton, D. & Tan, T.H. Interaction between c-Rel and the mitogen-activated protein kinase kinase kinase 1 signaling cascade in mediating κB enhancer activation. *J. Biol. Chem.* **271**, 8971–8976 (1996).
33. Lee, F.S., Hagler, J., Chen, Z.J. & Maniatis, T. Activation of the IκBα kinase complex by MEKK1, a kinase of the JNK pathway. *Cell* **88**, 213–222 (1997).
34. Yujiri, T., Sather, S., Fanger, G.R. & Johnson, G.L. Role of MEKK1 in cell survival and activation of JNK and ERK pathways defined by targeted gene disruption. *Science* **282**, 1911–1914 (1998).
35. Hirosumi, J. *et al.* A central role for JNK in obesity and insulin resistance. *Nature* **420**, 333–336 (2002).
36. Cai, D. *et al.* Local and systemic insulin resistance resulting from hepatic activation of IKK-β and NF-κB. *Nat. Med.* **11**, 183–190 (2005).
37. Barish, G.D., Narkar, V.A. & Evans, R.M. PPARδ: a dagger in the heart of the metabolic syndrome. *J. Clin. Invest.* **116**, 590–597 (2006).
38. Tanaka, T. *et al.* Activation of peroxisome proliferator–activated receptor δ induces fatty acid β-oxidation in skeletal muscle and attenuates metabolic syndrome. *Proc. Natl. Acad. Sci. USA* **100**, 15924–15929 (2003).

© 2012 Nature America, Inc. All rights reserved.

39. Butler, M. *et al.* Specific inhibition of PTEN expression reverses hyperglycemia in diabetic mice. *Diabetes* **51**, 1028–1034 (2002).
40. Wijesekara, N. *et al.* Muscle-specific Pten deletion protects against insulin resistance and diabetes. *Mol. Cell. Biol.* **25**, 1135–1145 (2005).
41. Fajans, S.S., Bell, G.I. & Polonsky, K.S. Molecular mechanisms and clinical pathophysiology of maturity-onset diabetes of the young. *N. Engl. J. Med.* **345**, 971–980 (2001).
42. Schmid, J. *et al.* Modulation of pancreatic islets–stress axis by hypothalamic releasing hormones and 11β-hydroxysteroid dehydrogenase. *Proc. Natl. Acad. Sci. USA* **108**, 13722–13727 (2011).
43. Strawbridge, R.J. *et al.* Genome-wide association identifies nine common variants associated with fasting proinsulin levels and provides new insights into the pathophysiology of type 2 diabetes. *Diabetes* **60**, 2624–2634 (2011).
44. Cui, B. *et al.* A genome-wide association study confirms previously reported loci for type 2 diabetes in Han Chinese. *PLoS ONE* **6**, e22353 (2011).
45. Raychaudhuri, S. *et al.* Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. *PLoS Genet.* **5**, e1000534 (2009).
46. Withers, D.J. *et al.* Disruption of IRS-2 causes type 2 diabetes in mice. *Nature* **391**, 900–904 (1998).
47. Withers, D.J. *et al.* Irs-2 coordinates Igf-1 receptor–mediated β-cell development and peripheral insulin signalling. *Nat. Genet.* **23**, 32–40 (1999).
48. Aitman, T.J. *et al.* Identification of Cd36 (Fat) as an insulin-resistance gene causing defective fatty acid and glucose metabolism in hypertensive rats. *Nat. Genet.* **21**, 76–83 (1999).
49. Yamagata, K. *et al.* Mutations in the hepatocyte nuclear factor-1α gene in maturity-onset diabetes of the young (MODY3). *Nature* **384**, 455–458 (1996).
50. White, M.F. The IRS-signalling system: a network of docking proteins that mediate insulin action. *Mol. Cell. Biochem.* **182**, 3–11 (1998).
51. Spiegelman, B.M. PPAR-γ: adipogenic regulator and thiazolidinedione receptor. *Diabetes* **47**, 507–514 (1998).
52. Curtis, R.E., Yin, J., Kinnaird, P. & Xing, E.P. Finding genome-transcriptome-phenome association with structured association mapping and visualization in genamap. *Pac. Symp. Biocomput.* 327–338 (2012).
53. Kim, S. & Xing, E.P. Statistical estimation of correlated genome associations to a quantitative trait network. *PLoS Genet.* **5**, e1000587 (2009).
54. Kim, S., Sohn, K.A. & Xing, E.P. A multivariate regression approach to association analysis of a quantitative trait network. *Bioinformatics* **25**, i204–i212 (2009).
55. Lehner, B., Crombie, C., Tischler, J., Fortunato, A. & Fraser, A.G. Systematic mapping of genetic interactions in *Caenorhabditis elegans* identifies common modifiers of diverse signaling pathways. *Nat. Genet.* **38**, 896–903 (2006).
56. Higgins, J.P. & Thompson, S.G. Quantifying heterogeneity in a meta-analysis. *Stat. Med.* **21**, 1539–1558 (2002).
57. Anderson, S.L. *et al.* Familial dysautonomia is caused by mutations of the *IKAP* gene. *Am. J. Hum. Genet.* **68**, 753–758 (2001).
58. Hiromura, M. *et al.* YY1 is regulated by O-linked N-acetylglucosaminylation (O-glcNAcylation). *J. Biol. Chem.* **278**, 14046–14052 (2003).
59. Shimoyama, Y., Nishio, K., Hamajima, N. & Niwa, T. *KLOTHO* gene polymorphisms G-395A and C1818T are associated with lipid and glucose metabolism, bone mineral density and systolic blood pressure in Japanese healthy subjects. *Clin. Chim. Acta* **406**, 134–138 (2009).
60. Oguro, R. *et al.* Association of carotid atherosclerosis with genetic polymorphisms of the klotho gene in patients with hypertension. *Geriatr. Gerontol. Int.* **10**, 311–318 (2010).
61. Freathy, R.M. *et al.* The functional "KL-VS" variant of *KLOTHO* is not associated with type 2 diabetes in 5028 UK Caucasians. *BMC Med. Genet.* **7**, 51 (2006).
62. Mullin, B.H. *et al.* *Klotho* gene polymorphisms are associated with osteocalcin levels but not bone density of aged postmenopausal women. *Calcif. Tissue Int.* **77**, 145–151 (2005).
63. Zarrabeitia, M.T. *et al.* *Klotho* gene polymorphism and male bone mass. *Calcif. Tissue Int.* **80**, 10–14 (2007).
64. Reue, K., Xu, P., Wang, X.P. & Slavin, B.G. Adipose tissue deficiency, glucose intolerance, and increased atherosclerosis result from mutation in the mouse fatty liver dystrophy (fld) gene. *J. Lipid Res.* **41**, 1067–1076 (2000).
65. Aulchenko, Y.S. *et al.* *LPIN2* is associated with type 2 diabetes, glucose metabolism, and body composition. *Diabetes* **56**, 3020–3026 (2007).
66. Anderson, C.A. *et al.* Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat. Genet.* **43**, 246–252 (2011).
67. Franke, A. *et al.* Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat. Genet.* **42**, 1118–1125 (2010).
68. McGovern, D.P. *et al.* Genome-wide association identifies multiple ulcerative colitis susceptibility loci. *Nat. Genet.* **42**, 332–337 (2010).
69. Tang, A.T., Campbell, W.B. & Nithipatikom, K. ROCK1 feedback regulation of the upstream small GTPase RhoA. *Cell. Signal.* **24**, 1375–1380 (2012).
70. Nakamura, Y. *et al.* Marked increase of insulin gene transcription by suppression of the Rho/Rho-kinase pathway. *Biochem. Biophys. Res. Commun.* **350**, 68–73 (2006).
71. Furukawa, N. *et al.* Role of Rho-kinase in regulation of insulin action and glucose homeostasis. *Cell Metab.* **2**, 119–129 (2005).
72. Chun, K.H. *et al.* *In vivo* activation of ROCK1 by insulin is impaired in skeletal muscle of humans with type 2 diabetes. *Am. J. Physiol. Endocrinol. Metab.* **300**, E536–E542 (2011).
73. Yuneva, M.O. *et al.* The metabolic profile of tumors depends on both the responsible genetic lesion and tissue type. *Cell Metab.* **15**, 157–170 (2012).
74. Yang, Q. *et al.* Multiple genetic loci influence serum urate levels and their relationship with gout and cardiovascular disease risk factors. *Circ. Cardiovasc. Genet.* **3**, 523–530 (2010).
75. Hurlbert, M.S. *et al.* Mice transgenic for an expanded CAG repeat in the Huntington's disease gene develop diabetes. *Diabetes* **48**, 649–651 (1999).
76. Björkqvist, M. *et al.* The R6/2 transgenic mouse model of Huntington's disease develops diabetes due to deficient β-cell mass and exocytosis. *Hum. Mol. Genet.* **14**, 565–574 (2005).
77. Bradley, S.V. *et al.* Degenerative phenotypes caused by the combined deficiency of murine HIP1 and HIP1r are rescued by human HIP1. *Hum. Mol. Genet.* **16**, 1279–1292 (2007).
78. Hancock, D.B. *et al.* Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. *Nat. Genet.* **42**, 45–52 (2010).
79. Cho, M.H. *et al.* Variants in *FAM13A* are associated with chronic obstructive pulmonary disease. *Nat. Genet.* **42**, 200–202 (2010).
80. Nomiyama, T. *et al.* Osteopontin mediates obesity-induced adipose tissue macrophage infiltration and insulin resistance in mice. *J. Clin. Invest.* **117**, 2877–2888 (2007).
81. Pietiläinen, K.H. *et al.* Global transcript profiles of fat in monozygotic twins discordant for BMI: pathways behind acquired obesity. *PLoS Med.* **5**, e51 (2008).
82. Lyssenko, V. *et al.* Pleiotropic effects of GIP on islet function involve osteopontin. *Diabetes* **60**, 2424–2433 (2011).
83. Olofsson, L.E. *et al.* CCAAT/enhancer binding protein α (C/EBPα) in adipose tissue regulates genes in lipid and glucose metabolism and a genetic variation in C/EBPα is associated with serum levels of triglycerides. *J. Clin. Endocrinol. Metab.* **93**, 4880–4886 (2008).
84. Wu, Z. *et al.* Cross-regulation of C/EBPα and PPARγ controls the transcriptional pathway of adipogenesis and insulin sensitivity. *Mol. Cell* **3**, 151–158 (1999).
85. Hollenberg, A.N. *et al.* Functional antagonism between CCAAT/enhancer binding protein-α and peroxisome proliferator–activated receptor-γ on the leptin promoter. *J. Biol. Chem.* **272**, 5283–5290 (1997).
86. Keller, S.R. Role of the insulin-regulated aminopeptidase IRAP in insulin action and diabetes. *Biol. Pharm. Bull.* **27**, 761–764 (2004).

Robert A Scott[1,166], Vasiliki Lagou[2,3,166], Ryan P Welch[4–6,166], Eleanor Wheeler[7], May E Montasser[8], Jian'an Luan[1], Reedik Mägi[2,9], Rona J Strawbridge[10,11], Emil Rehnberg[12], Stefan Gustafsson[12], Stavroula Kanoni[7], Laura J Rasmussen-Torvik[13], Loïc Yengo[14,15], Cecile Lecoeur[14,15], Dmitry Shungin[16–18], Serena Sanna[19], Carlo Sidore[5,6,19,20], Paul C D Johnson[21], J Wouter Jukema[22,23], Toby Johnson[24,25], Anubha Mahajan[2], Niek Verweij[26], Gudmar Thorleifsson[27], Jouke-Jan Hottenga[28], Sonia Shah[29], Albert V Smith[30,31], Bengt Sennblad[10], Christian Gieger[32], Perttu Salo[33], Markus Perola[9,33,34], Nicholas J Timpson[35], David M Evans[35], Beate St Pourcain[36], Ying Wu[37], Jeanette S Andrews[38], Jennie Hui[39–42], Lawrence F Bielak[43], Wei Zhao[43], Momoko Horikoshi[2,3], Pau Navarro[44], Aaron Isaacs[45,46], Jeffrey R O'Connell[8], Kathleen Stirrups[7], Veronique Vitart[44], Caroline Hayward[44], Tõnu Esko[9,47], Evelin Mihailov[47], Ross M Fraser[48], Tove Fall[12], Benjamin F Voight[49,50], Soumya Raychaudhuri[51,52], Han Chen[53], Cecilia M Lindgren[2], Andrew P Morris[2], Nigel W Rayner[2,3], Neil Robertson[2,3], Denis Rybin[54], Ching-Ti Liu[53], Jacques S Beckmann[55,56], Sara M Willems[46], Peter S Chines[57], Anne U Jackson[5,6], Hyun Min Kang[5,6], Heather M Stringham[5,6], Kijoung Song[58], Toshiko Tanaka[59], John F Peden[2,60], Anuj Goel[2,61], Andrew A Hicks[62], Ping An[63],

© 2012 Nature America, Inc. All rights reserved.

Martina Müller-Nurasyid[32,64,65], Anders Franco-Cereceda[66], Lasse Folkersen[10,11], Letizia Marullo[2,67], Hanneke Jansen[68], Albertine J Oldehinkel[69], Marcel Bruinenberg[70], James S Pankow[71], Kari E North[72,73], Nita G Forouhi[1], Ruth J F Loos[1], Sarah Edkins[7], Tibor V Varga[16], Göran Hallmans[74], Heikki Oksa[75], Mulas Antonella[19], Ramaiah Nagaraja[76], Stella Trompet[22,23], Ian Ford[21], Stephan J L Bakker[77], Augustine Kong[27], Meena Kumari[78], Bruna Gigante[79], Christian Herder[80], Patricia B Munroe[24,25], Mark Caulfield[24,25], Jula Antti[33], Massimo Mangino[81], Kerrin Small[81], Iva Miljkovic[82], Yongmei Liu[83], Mustafa Atalay[84], Wieland Kiess[85,86], Alan L James[39,87,88], Fernando Rivadeneira[45,89–91], Andre G Uitterlinden[45,89–91], Colin N A Palmer[92], Alex S F Doney[92], Gonneke Willemsen[28], Johannes H Smit[93], Susan Campbell[44], Ozren Polasek[94], Lori L Bonnycastle[57], Serge Hercberg[95], Maria Dimitriou[96], Jennifer L Bolton[48], Gerard R Fowkes[48], Peter Kovacs[97], Jaana Lindström[98], Tatijana Zemunik[94], Stefania Bandinelli[99], Sarah H Wild[48], Hanneke V Basart[100], Wolfgang Rathmann[101], Harald Grallert[102], DIAbetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium[103], Winfried Maerz[104,105], Marcus E Kleber[105,106], Bernhard O Boehm[107], Annette Peters[108], Peter P Pramstaller[61,109,110], Michael A Province[63], Ingrid B Borecki[63], Nicholas D Hastie[44], Igor Rudan[48], Harry Campbell[48], Hugh Watkins[2,61], Martin Farrall[2,61], Michael Stumvoll[85,111], Luigi Ferrucci[59], Dawn M Waterworth[58], Richard N Bergman[112], Francis S Collins[57], Jaakko Tuomilehto[98,113–115], Richard M Watanabe[116,117], Eco J C de Geus[28], Brenda W Penninx[93], Albert Hofman[91], Ben A Oostra[45,46,90], Bruce M Psaty[118–121], Peter Vollenweider[122], James F Wilson[48], Alan F Wright[44], G Kees Hovingh[100], Andres Metspalu[9,47], Matti Uusitupa[123,124], Patrik K E Magnusson[12], Kirsten O Kyvik[125,126], Jaakko Kaprio[34,127,128], Jackie F Price[48], George V Dedoussis[96], Panos Deloukas[7], Pierre Meneton[129], Lars Lind[130], Michael Boehnke[5,6], Alan R Shuldiner[8,131], Cornelia M van Duijn[45,46,90,91], Andrew D Morris[92], Anke Toenjes[85,110], Patricia A Peyser[43], John P Beilby[39,41,42], Antje Körner[85,86], Johanna Kuusisto[132], Markku Laakso[132], Stefan R Bornstein[133], Peter E H Schwarz[133], Timo A Lakka[84,134], Rainer Rauramaa[134,135], Linda S Adair[136], George Davey Smith[35], Tim D Spector[81], Thomas Illig[102,137], Ulf de Faire[79], Anders Hamsten[10,11,138], Vilmundur Gudnason[30,31], Mika Kivimaki[78], Aroon Hingorani[78], Sirkka M Keinanen-Kiukaanniemi[139,140], Timo E Saaristo[75,141], Dorret I Boomsma[28], Kari Stefansson[27,31], Pim van der Harst[26], Josée Dupuis[53,142], Nancy L Pedersen[12], Naveed Sattar[143], Tamara B Harris[144], Francesco Cucca[19,20], Samuli Ripatti[34,145,146], Veikko Salomaa[147], Karen L Mohlke[37], Beverley Balkau[148,149], Philippe Froguel[14,15,150], Anneli Pouta[151,152], Marjo-Riitta Jarvelin[152–155], Nicholas J Wareham[1], Nabila Bouatia-Naji[14,15,156], Mark I McCarthy[2,3,157], Paul W Franks[16,17,158], James B Meigs[159,160], Tanya M Teslovich[5,6], Jose C Florez[160–163,167], Claudia Langenberg[1,78,167], Erik Ingelsson[12,167], Inga Prokopenko[2,3,167] & Inês Barroso[7,164,165,167]

[1]Medical Research Council (MRC) Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK. [2]Wellcome Trust Center for Human Genetics, University of Oxford, Oxford, UK. [3]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK. [4]Bioinformatics Graduate Program, University of Michigan Medical School, Ann Arbor, Michigan, USA. [5]Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan, USA. [6]Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, USA. [7]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK. [8]Division of Endocrinology, Diabetes and Nutrition, University of Maryland School of Medicine, Baltimore, Maryland, USA. [9]Estonian Genome Center, University of Tartu, Tartu, Estonia. [10]Atherosclerosis Research Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden. [11]Center for Molecular Medicine, Karolinska University Hospital, Stockholm, Sweden. [12]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. [13]Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA. [14]Universite Lille Nord de France, Lille, France. [15]Centre National de la Recherche Scientifique (CNRS) Unité Mixte de Recherche (UMR) 8199, Institut Pasteur de Lille, Lille, France. [16]Genetic and Molecular Epidemiology Unit, Department of Clinical Sciences, Lund University, Skåne University Hospital Malmö, Malmö, Sweden. [17]Department of Public Health & Clinical Medicine, Umeå University, Umeå, Sweden. [18]Department of Odontology, Umeå University, Umeå, Sweden. [19]Istituto di Ricerca Genetica e Biomedica, Consiglio Nazionale delle Ricerche (CNR), Monserrato, Italy. [20]Dipartimento di Scienze Biomediche, Università di Sassari, Sassari, Italy. [21]Robertson Centre for Biostatistics, University of Glasgow, Glasgow, UK. [22]Interuniversity Cardiology Institute of the Netherlands (ICIN), Durrer Center for Cardiogenetic Research, Utrecht, The Netherlands. [23]Departmentt of Cardiology, Leiden University Medical Center, Leiden, The Netherlands. [24]Department of Clinical Pharmacology, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK. [25]The Genome Centre, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK. [26]Department of Cardiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. [27]deCODE genetics, Rekjavik, Iceland. [28]Department of Biological Psychology, VU University & EMGO+ Institute, Amsterdam, The Netherlands. [29]University College London Genetics Institute (UGI), University College London, London, UK. [30]Icelandic Heart Association, Kopavogur, Iceland. [31]Faculty of Medicine, University of Iceland, Reykjavík, Iceland. [32]Institute of Genetic Epidemiology, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany. [33]Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland. [34]Institute of Molecular Medicine, Finland (FIMM), University of Helsinki, Helsinki, Finland. [35]MRC Council Centre for Causal Analyses in Translational Epidemiology (CAiTE) Centre, School of Social and Community Medicine, University of Bristol, Bristol, UK. [36]School of Social and Community Medicine, University of Bristol, Bristol, UK. [37]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. [38]Department of Biostatistical Sciences, Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA. [39]Busselton Population Medical Research Institute, Sir Charles Gairdner Hospital, Nedlands, Western Australia, Australia. [40]School of Population Health, The University of Western Australia, Nedlands, Western Australia, Australia. [41]School of Pathology and Laboratory Medicine, The University of Western Australia, Nedlands, Western Australia, Australia. [42]PathWest Laboratory Medicine WA, Nedlands, Western Australia, Australia. [43]Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, Michigan, USA. [44]MRC Human Genetics Unit at the Medical Research Council Institute of Genetics and Molecular Medicine, University of Edinburgh, Western

© 2012 Nature America, Inc. All rights reserved.

General Hospital, Edinburgh, UK. [45]Centre for Medical Systems Biology (CMSB), Leiden, The Netherlands. [46]Genetic Epidemiology Unit, Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands. [47]Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia. [48]Centre for Population Health Sciences, University of Edinburgh, Edinburgh, UK. [49]The Broad Institute of Harvard and MIT, Boston, Massachusetts, USA. [50]Department of Pharmacology, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA. [51]Division of Genetics, Brigham and Women's Hospital, Boston, Massachusetts, USA. [52]Division of Rheumatology, Brigham and Women's Hospital, Boston, Massachusetts, USA. [53]Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, USA. [54]Boston University Data Coordinating Center, Boston, Massachusetts, USA. [55]Department of Medical Genetics, University of Lausanne, Lausanne, Switzerland. [56]The Service of Medical Genetics, Centre Hospitalier Universitaire Vaudois (CHUV), University Hospital, Lausanne, Switzerland. [57]Genome Technology Branch, National Human Genome Research Institute, US National Institutes of Health (NIH), Bethesda, Maryland, USA. [58]Genetics, GlaxoSmithKline, Upper Merion, Pennsylvania, USA. [59]Clinical Research Branch, National Institute on Aging, NIH, Baltimore, Maryland, USA. [60]Illumina, Inc., Essex, UK. [61]Department of Cardiovascular Medicine, University of Oxford, Oxford, UK. [62]Centre for Biomedicine, European Academy Bozen/Bolzano (EURAC), Bolzano, Italy (affiliated Institute of the University of Lübeck). [63]Division of Statistical Genomics, Washington University School of Medicine, St. Louis, Missouri, USA. [64]Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany. [65]Department of Medicine I, University Hospital Grosshadern, Ludwig-Maximilians-Universität, Munich, Germany. [66]Cardiothoracic Surgery Unit, Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden. [67]Department of Evolutionary Biology, Genetic Section, University of Ferrara, Ferrara, Italy. [68]Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [69]Interdisciplinary Center for Pathology of Emotions, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. [70]University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [71]Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, Minnesota, USA. [72]Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. [73]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. [74]Section for Nutritional Research, Department of Public Health and Clinical Medicine, Umeå University Hospital, Umeå, Sweden. [75]Pirkanmaa Hospital District, Tampere, Finland. [76]Laboratory of Genetics, National Institute on Aging, NIH, Baltimore, Maryland, USA. [77]Department of Internal Medicine, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. [78]Department of Epidemiology and Public Health, University College London, London, UK. [79]Division of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. [80]Institute for Clinical Diabetology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf, Düsseldorf, Germany. [81]Department of Twin Research and Genetic Epidemiology, King's College London, London, UK. [82]Department of Epidemiology, Center for Aging and Population Health, University of Pittsburgh, Pittsburgh, Pennsylvania, USA. [83]Department of Epidemiology and Prevention, Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA. [84]Institute of Biomedicine, Physiology, University of Eastern Finland, Kuopio Campus, Kuopio, Finland. [85]Integrated Research and Treatment Center (IFB) Adiposity Diseases, University of Leipzig, Leipzig, Germany. [86]Pediatric Research Center, Department of Women's & Child Health, University of Leipzig, Leipzig, Germany. [87]School of Medicine and Pharmacology, The University of Western Australia, Nedlands, Western Australia, Australia. [88]Pulmonary Physiology, Sir Charles Gairdner Hospital, Nedlands, Western Australia, Australia. [89]Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, The Netherlands. [90]Netherlands Consortium for Healthy Ageing (NCHAH) of the Genomics Initiative (NGI), Leiden, The Netherlands. [91]Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands. [92]Medical Research Institute, University of Dundee, Dundee, UK. [93]Department of Psychiatry, VU University Medical Centre, Amsterdam, The Netherlands. [94]Faculty of Medicine, University of Split, Split, Croatia. [95]U557 Institut National de la Santé et de la Recherche Médicale (INSERM), U1125 Institut National de la Recherche Agronomique (INRA), Université Paris 13, Bobigny, France. [96]Department of Dietetics and Nutrition, Harokopio University, Athens, Greece. [97]University of Leipzig, Interdisciplinary Center for Clinical Research, Leipzig, Germany. [98]Diabetes Prevention Unit, National Institute for Health and Welfare, Helsinki, Finland. [99]Geriatric Department, Azienda Sanitaria Firenze, Florence, Italy. [100]Department Vascular Medicine, Academic Medical Center, Amsterdam, The Netherlands. [101]Institute of Biometrics and Epidemiology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf, Düsseldorf, Germany. [102]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany. [103]A full list of members and affiliations is given in the **Supplementary Note**. [104]Synlab Academy, Mannheim, Germany. [105]Mannheim Institute of Public Health, Social and Preventive Medicine, Medical Faculty of Mannheim, University of Heidelberg, Mannheim, Germany. [106]Ludwigshafen Risk and Cardiovascular Health (LURIC) Study, Freiburg, Germany. [107]Division of Endocrinology and Diabetes, Department of Medicine, University Hospital, Ulm, Germany. [108]Institute of Epidemiology II, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany. [109]Department of Neurology, General Central Hospital, Bolzano, Italy. [110]Department of Neurology, University of Lübeck, Lübeck, Germany. [111]Department of Medicine, University of Leipzig, Leipzig, Germany. [112]Diabetes and Obesity Research Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA. [113]South Ostrobothnia Central Hospital, Seinäjoki, Finland. [114]Red RECAVA (La Red Temática de Investigación Cooperativa en Enfermedades Cardiovasculares) Grupo RD06/0014/0015, Hospital Universitario La Paz, Madrid, Spain. [115]Centre for Vascular Prevention, Danube–University Krems, Krems, Austria. [116]Department of Preventive Medicine, Keck School of Medicine of the University of Southern California, Los Angeles, California, USA. [117]Department of Physiology & Biophysics, Keck School of Medicine of the University of Souterhn California, Los Angeles, California, USA. [118]Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, Washington, USA. [119]Group Health Research Institute, Group Health Cooperative, Seattle, Washington, USA. [120]Department of Epidemiology, University of Washington, Seattle, Washington, USA. [121]Department of Health Services, University of Washington, Seattle, Washington, USA. [122]Department of Internal Medicine, University Hospital and University of Lausanne, Lausanne, Switzerland. [123]Department of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland. [124]Research Unit, Kuopio University Hospital, Kuopio, Finland. [125]Odense Patient data Explorative Network (OPEN), Odense, Denmark. [126]Institute of Regional Health Services Research, Odense, Denmark. [127]Department of Public Health, Hjelt Institute, University of Helsinki, Helsinki, Finland. [128]Department of Mental Health and Substance Abuse Services, National Institute for Health and Welfare, Helsinki, Finland. [129]U872 INSTERM, Centre de Recherche des Cordeliers, Paris, France. [130]Department of Medical Sciences, Uppsala University, Uppsala, Sweden. [131]Geriatric Research and Education Clinical Center, Veterans Administration Medical Center, Baltimore, Maryland, USA. [132]Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland. [133]Department of Medicine III, University of Dresden, Medical Faculty Carl Gustav Carus, Dresden, Germany. [134]Kuopio Research Institute of Exercise Medicine, Kuopio, Finland. [135]Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland. [136]Department of Nutrition, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. [137]Hannover Unified Biobank, Hannover Medical School, Hannover, Germany. [138]Department of Cardiology, Karolinska University Hospital, Stockholm, Sweden. [139]Faculty of Medicine, Institute of Health Sciences, University of Oulu, Oulu, Finland. [140]Unit of General Practice, Oulu University Hospital, Oulu, Finland. [141]Finnish Diabetes Association, Tampere, Finland. [142]National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, Massachusetts, USA. [143]British Heart Foundation (BHF) Building, Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK. [144]Laboratory of Epidemiology, Demography, and Biometry, National Institute on Ageing, NIH, Bethesda, Maryland, USA. [145]Public Health Genomics Unit, National Institute for Health and Welfare, Helsinki, Finland. [146]Wellcome Trust Sanger Institute, Hinxton, UK. [147]Unit of Chronic Disease Epidemiology and Prevention, National Institute for Health and Welfare, Helsinki, Finland. [148]INSERM, Centre de Recherche en Epidémiologie et Santé des Populations (CESP) Center for Research in Epidemiology and Public Health, U1018, Epidemiology of Diabetes, Obesity and Chronic Kidney Disease over the Lifecourse, Villejuif, France. [149]University Paris Sud 11, UMRS 1018, Villejuif, France. [150]Department of Genomics of Common Disease, School of Public Health, Imperial College London, Hammersmith Hospital, London, UK. [151]Department of Clinical Sciences/Obstetrics and Gynecology, University of Oulu, Oulu, Finland. [152]Department of Lifecourse and Services, National Institute for Health and Welfare, Oulu, Finland. [153]Biocenter Oulu, University of Oulu, Oulu, Finland. [154]Department of Epidemiology and Biostatistics, School of Public Health, MRC–Health Protection Agency (HPA) Centre for Environment and Health, Faculty of Medicine, Imperial College London, London, UK. [155]Institute of Health Sciences, University of Oulu, Oulu, Finland. [156]INSERM U970, Paris Cardiovascular Research Center (PARCC), Paris, France. [157]Oxford National Institute for Health Research (NIHR) Biomedical Research Centre, Churchill Hospital, Oxford, UK. [158]Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts, USA. [159]General Medicine Division, Massachusetts General Hospital, Boston, Massachusetts, USA. [160]Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA. [161]Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA. [162]Diabetes Research Center, Diabetes Unit, Massachusetts General Hospital, Boston, Massachusetts, USA. [163]Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA. [164]NIHR Cambridge Biomedical Research Centre, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK. [165]University of Cambridge Metabolic Research Laboratories, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK. [166]These authors contributed equally to this work. [167]These authors jointly directed this work. Correspondence should be addressed to J.C.F. (jcflorez@partners.org), C. L. (claudia.langenberg@mrc-epid.cam.ac.uk), E.I. (erik.ingelsson@ki.se), I.P. (inga@well.ox.ac.uk) or I.B. (ib1@sanger.ac.uk).

© 2012 Nature America, Inc. All rights reserved.

npg

ONLINE METHODS

Study design. The Illumina CardioMetabochip (Metabochip) is a custom Illumina iSELECT array of 196,725 SNPs. It has been designed to support efficient large-scale follow-up analysis of putative associations for glycemic (including fasting glucose, fasting insulin and post-challenge glucose (2hGlu) concentrations and other metabolic and cardiovascular traits (**Supplementary Fig. 1**)[9] and to enable the fine mapping of established loci. Overall, there were 65,435 SNPs genotyped on the Metabochip for follow up of previous associations, including a total of 23 cardio-metabolic traits. Traits contributing SNPs to the Metabochip were prioritized into primary (including fasting glucose) and secondary (including fasting insulin and 2hGlu), contributing ~5,000 and ~1,000 SNPs, respectively, from the most significantly associated variants for each phenotype in the discovery meta-analyses from each contributing consortium. This included 5,055 SNPs for follow up of fasting glucose, 1,046 for fasting insulin and 1,038 for 2hGlu associations. In the present analysis, we focused our analysis on this set of follow-up SNPs available on the Metabochip to establish variants among these SNP associated with glycemic traits. Although we also included newly available studies genotyped on genome-wide platforms, we limited our primary analyses to only these ~66,000 SNPs.

Studies. In the present effort, collaborating studies within the Meta-Analysis of Glucose and Insulin related traits Consortium (MAGIC) provided results for the 66,000 follow-up SNPs genotyped on Metabochip on a maximum total of 133,010 (fasting glucose), 108,557 (fasting insulin) and 42,854 (2hGlu) individuals. In addition to those newly genotyped on the Metabochip platform, in our overall meta-analysis, we were able to include further studies that had genotyped or imputed the same SNPs on other platforms. The largest proportion of our entire sample was directly genotyped on the Metabochip and comprised 53,622 (fasting glucose), 42,384 (fasting insulin) and 27,602 (2hGlu) individuals from 26, 21 and 12 studies, respectively. We were also able to recruit 11,690 (fasting glucose) and 8,813 (fasting insulin) individuals from up to 4 additional GWAS (Prevend, Ascot (fasting glucose only), Prosper and TRAILS) (**Supplementary Table 1**) not included in the original meta-analysis[1]. From another MAGIC study of sex-specific associations with glycemic traits (I.P. on behalf of the MAGIC authors, personal communication), we were able to recruit another 15 and 13 independent studies comprising up to 25,618 and 23,130 individuals for fasting glucose and fasting insulin, respectively. The above studies were combined in a single fixed-effects meta-analysis with those studies included in the original GWAS[1,2], including 20 (fasting glucose), 19 (fasting insulin) and 9 (2hGlu) studies and 42,080 (fasting glucose), 34,230 (fasting insulin) and 15,252 (2hGlu) individuals, as described previously[1,2]. The study and individual counts from the original GWAS excluded the family-based SardiNIA study, where, initially, a large number of the individuals had imputed genotype data only. The entire sample was directly genotyped on Metabochip, and the resulting data were included in place of the original GWAS data. Some studies had genotyping data available from both Metabochip and genome-wide arrays but from entirely independent samples within the study (**Supplementary Table 1**). Full study characteristics of all Metabochip studies are shown in **Supplementary Table 1**, and data from discovery genome-wide studies, and those from the sex-specific analyses are reported elsewhere (refs. 1,2 and I.P. on behalf of the MAGIC authors, personal communication). All participants of the main analysis were of European descent and were mostly adults, although data from a total of 7,872 and 7,164 adolescents were also included in the fasting glucose and fasting insulin meta-analyses, respectively (NFBC86, Leipzig-childhood_IFB, TRAILS and ALSPAC studies). All studies were approved by local research ethic committees, and all participants gave informed consent. Results from the CLHNS study of Filipino women (n = 1,682 and 1,635 for fasting glucose and fasting insulin, respectively) genotyped on Metabochip were also available and were included in supplementary analyses to compare effect directions with studies of individuals of European descent alone.

Phenotypes. Analyses were undertaken for fasting glucose and fasting insulin measured in mM and pM, respectively. 2hGlu was measured in mM. As in the previous MAGIC discovery analysis[1,2], individuals were excluded from the analysis if they had a physician diagnosis of diabetes, were on diabetes treatment (oral or insulin) or had a fasting plasma glucose concentration equal to or greater than 7 mM. Individual studies applied further sample exclusions, including for pregnancy, non-fasting individuals and type 1 diabetes, as detailed in **Supplementary Table 1**. Individuals from case-control studies (**Supplementary Table 1**) were excluded if they had undergone hospitalization or blood transfusion in the 2–3 months before phenotyping took place. 2hGlu measures were carried out 2 h after a glucose challenge during an oral glucose tolerance test (OGTT). Measures of fasting glucose and 2hGlu made in whole blood were corrected to plasma level using the correction factor of 1.13 (ref. 87). Fasting insulin was measured in serum. Detailed descriptions of study-specific glycemic measurements are given in **Supplementary Table 1**.

Trait transformations and adjustment. Analyses were performed for untransformed levels of fasting glucose, natural logarithm–transformed fasting insulin and untransformed 2hGlu using a linear regression model. All analyses were adjusted (if applicable) for age, study site and geographic covariates to evaluate the association using an additive genetic model at each genetic SNP variant.

BMI-adjusted analysis. In the Fenland study (**Supplementary Table 1**), we investigated the correlation between BMI and natural logarithm–transformed fasting insulin, fasting glucose and 2hGlu to establish the variation in each trait explained by BMI. Meta-analyses for each trait were also adjusted for BMI. Metabochip studies and new GWAS performed study-level analyses adjusted for BMI. Most studies in the original GWAS (except deCode, GEMs, KORAF4 and TwinsUK) as well as from the studies analyzed in a sex-specific manner were included in BMI-adjusted meta-analysis. The original discovery 2hGlu meta-analysis adjusted for BMI[2] was also included in these analyses. We also performed an analysis for 2hGlu adjusted for fasting glucose to investigate whether additional variants would be identified with an effect on 2hGlu independent of fasting glucose and also to establish whether identified 2hGlu associations were driven by fasting glucose.

Genotyping and quality control. The Metabochip or other commercial genome-wide arrays were used by individual studies for genotyping. Details are presented in **Supplementary Table 1** or are reported elsewhere[1,2]. The quality control criteria for both Metabochip and genome-wide arrays for filtering of poorly genotyped individuals or low-quality SNPs before imputation included the occurrence of (i) a call rate of <0.95; (ii) sex discrepancies; (iii) ancestry outliers; (iv) heterozygosity (**Supplementary Table 1**); (v) SNP minor allele frequency of <0.01; (vi) SNP Hardy-Weinberg equilibrium $P < 1 \times 10^{-4}$; (vii) SNP effect estimate standard error (SE) ≥10; and (viii) SNP minor allele count (MAC) of <10 (calculated as the total number of observed alleles at each SNP multiplied by the MAF).

Studies with genome-wide arrays undertook imputation using the HapMap CEU reference panel using MACH[88,89] and IMPUTE[90,91] software (**Supplementary Table 1**). Parameters used in imputation and filters applied to imputed genotypes are described in **Supplementary Table 1** or were reported previously[1,2]. From a total of ~2.5 million directly genotyped or imputed autosomal SNPs across the genome, study-specific results for the ~66,000 Metabochip follow-up SNPs were considered for the present meta-analyses. SNPs at which there were meta-analysis results in more than 10,000 individuals were included in the analysis.

Statistical analysis. Analyses of previous discovery studies are reported elsewhere[1,2], and those studies genotyped on the Metabochip are described in **Supplementary Table 1**. SNP effect estimates and their standard errors (for an additive genetic model) were combined by inverse variance–weighted fixed-effects meta-analysis using METAL[92] and GWAMA[93]. Two parallel meta-analyses for each trait by different analysts were compared for consistency. Individual cohort results were corrected for residual inflation of the test statistics using genomic control (λ) estimates. The genomic control values were estimated for each study, using either test statistics from all SNPs for the GWAS, whereas, for those studies genotyped on the Metabochip, genomic control λ estimates were derived from test statistics for 5,041 SNPs selected for follow-up analysis of QT-interval associations, as we perceived these to have the lowest likelihood of common architecture of associations with glycemic traits. Individual study–level λ genomic control estimates are shown in

© 2012 Nature America, Inc. All rights reserved.

doi:10.1038/ng.2385

Supplementary Table 1. Overall quantile-quantile plots for the QT follow-up SNPs are shown in **Supplementary Figure 10**.

Trait-associated signal selection strategy. Meta-analysis results for each trait were considered to have reached genome-wide significance if they had $P \leq 5 \times 10^{-8}$ and were not in LD ($r^2 < 0.05$) or within 500 kb of an established signal. The most significantly associated SNP (lowest *P* value) in each region (of 500 kb) was selected as the lead SNP. Associated loci are referred to by the name of the nearest gene, unless a more biologically plausible gene was nearby or a nearby gene was previously associated with another glycemic trait. In such cases, we maintained consistency with the previous naming but list the nearest genes in **Supplementary Table 2a–d**. To establish the variance in each trait explained by these SNPs, in the Framingham Heart Study, we included all SNPs in a model adjusted for age, sex, BMI and cohort.

Fine mapping of known glycemic trait loci. To undertake preliminary fine-mapping analyses, we investigated the patterns of association at 17 known fasting glucose and fasting insulin loci[1] and at 5 known 2hGlu loci[2] using meta-analysis results from 13,644, 1,309 and 1,249 SNPs genotyped on the Metabochip in 53,622, 42,384 and 27,602 individuals for fasting glucose, fasting insulin and 2hGlu, respectively. Only studies genotyped directly on the Metabochip were used for fine-mapping purposes to have equal sample size and availability of all SNPs. Regional plots for each locus were created using the previous lead SNP[1] or a suitable proxy ($r^2 > 0.8$) as the index SNP if that marker was not present on Metabochip. The plots were generated on LocusZoom web-based plotting software[94] using LD information from the 1000 Genomes Project (hg19; November 2010 European (EUR) data). Before generating the plots, all SNP names and positions from the Metabochip-only meta-analysis files were aligned to Build 37 using the Lift Genome annotation tool on the UCSC website to be compatible with the 1000 Genomes SNP naming format (chr: position) and allow more thorough assessment of the pairwise LD patterns around the established SNPs.

Associations of glycemic trait variants with related traits. For those SNPs that we identified that associated with genome-wide significance, we also investigated their association with other metabolic and disease traits. We exchanged reciprocal data for such SNPs with the latest DIAGRAM Metabochip analyses[24] and examined associations of these SNPs in publicly available data from previous studies of lipid traits from the GLGC[27] (triglycerides, HDL-cholesterol and low-density lipoprotein (LDL)-cholesterol) as well as BMI and WHR from the GIANT Consortium[25,26]. From these data, we were able to establish the presence of any association and the direction of effect for these other traits aligned to our trait-raising alleles. We highlighted associations with other traits at $P < 0.05$ and also performed FDR analyses. We performed FDR analyses for each trait separately (removing duplicate loci that were associated with more than one glycemic trait) and identified those with $q < 0.05$.

Expression quantitative trait locus (eQTL) analyses. Liver gene expression data from the Advanced Study of Aortic Pathology (ASAP) has been described previously[95]. In brief, liver biopsies were collected from individuals at the Karolinska University Hospital who were undergoing aortic valve surgery, alone or combined with surgery for aortic aneurysm, starting from 13 February 2007. All subjects gave their informed consent, and the study was approved by the ethics committee of Karolinska Institutet. After hybridization of extracted RNA to Affymetrix ST 1.0 Exon arrays, data were robust multi-array average (RMA) normalized and log transformed. DNA was extracted from whole blood, and genotyping was carried out using the Illumina 610w-Quad bead array platform. Imputation was carried out on SNPs with a call rate exceeding 95%, using the MACH algorithm. Imputation quality scores of RSQ < 0.3 were excluded from analysis. An additive genetic model was used to test for association between SNPs and gene expression.

VEGAS. To identify genes with multiple associated SNPs, we performed gene-based analysis using VEGAS, described in detail previously[10]. Briefly, on all available samples and among the ~66,000 follow-up SNPs, VEGAS pooled the information for all SNPs within each gene (±50 kb) to identify genes with higher evidence of association than expected by chance, while adjusting for gene size and the LD structure of the SNPs, by simulation (the maximum number of simulations used was 1,000,000). We identified genomic regions (separated by >1 Mb) showing evidence of association and described the genes contained within those regions. Although we often identified multiple genes within an associated region, it is probable that some of these are significant via LD. Bonferroni correction was used to adjust for multiple testing on the basis of the number of independent tests (number of genes tested) (~9,300), and *P* values of $<5.0 \times 10^{-6}$ were considered significant. Although the number of genes represented was constrained by those SNPs submitted for inclusion on the Metabochip, our analyses asked the question: of the genes represented on the Metabochip, all with a slightly raised previous likelihood of association, which show the most evidence for association with glycemic traits?

GRAIL. We used GRAIL[45] to evaluate whether loci across the genome associated with glycemic traits were enriched for connectivity between genes representing particular pathways or molecular processes. As described in detail previously[45], to define the genes near each SNP, GRAIL finds the furthest neighboring SNPs in the 3′ and 5′ direction that are in LD (HapMap CEU $r^2 > 0.5$) and proceeds outward in each direction to the nearest recombination hotspot[96]. All genes that overlap that interval are considered implicated by the SNP. If there are no genes in that region, the interval is extended by 250 kb in either direction. The method performs a text-based analysis, looking at abstracts in PubMed before December 2006 (to avoid confounding from GWAS results arising after that date). We performed two analyses for each trait. First, we took all genome-wide signals for each trait as a seed and queried loci to investigate biological connectivity among those loci (fasting glucose = 35, fasting insulin = 16, 2hGlu = 9). For fasting insulin, we did not include *FTO*, as the association with fasting insulin was entirely mediated by BMI. Second, we investigated connectivity between established signals (as seed regions) and those that did not reach genome-wide significance but were suggestively associated with each trait ($P < 0.0005$) (as query regions), as described previously[97]. For fasting insulin, we used BMI-adjusted results to define the query regions. Query regions were defined by taking all SNPs more significant than $P < 0.0005$, removing those associated at genome-wide levels of significance and pruning SNPs of $r^2 > 0.05$ in each region using PLINK[98]. As GRAIL tests connectivity of regions, we also removed any duplicates where a region was represented by more than one SNP. For those SNPs not found by the software, we submitted the region as a 500-kb window centered at the location of the SNP. This approach identified 218, 155 and 100 query regions (representing 715, 639 and 298 genes) for fasting glucose, fasting insulin (adjusted for BMI) and 2hGlu, respectively. The number of loci reaching $P_{GRAIL} < 0.01$ was determined from these analyses, and, to establish the level of enrichment, we randomly sampled 1,000 random sets of matched numbers of SNPs and calculated the proportion with as many or more reaching $P_{GRAIL} < 0.01$ to derive a permutation-based *P* value ($P_{permutation}$).

Pathway analyses. Pathway analysis was carried out for fasting glucose, fasting insulin and 2hGlu (uniform or adjusted for fasting glucose or adjusted for BMI) using data from previous discovery GWAS only[1] to avoid bias toward pathways represented on the Metabochip (Build 36; $n > 10{,}000$ and MAF ≥ 1% cutoff used). The software used for this analysis was MAGENTA 2.4 (July 2011; see URLs). SNPs from the meta-analysis file were assigned to a gene if they mapped within 110 kb upstream and 40 kb downstream of transcript boundaries. The smallest *P* value for the set of SNPs assigned to the gene was adjusted for confounders, such as gene length, marker density and LD, in a linear regression, creating a gene association score. If a top SNP was assigned to multiple genes, only the gene with the lowest score was kept to avoid positional clustering. The human leukocyte antigen (HLA) region was removed due to high LD and gene density. Pathway terms from multiple databases (GO, PANTHER, Ingenuity and KEGG) were attached to each gene. The genes were ranked on their association score, and a GSEA test was performed that tests all pathway terms using 5% and 75% cutoffs. Initially, 10,000 gene set permutations were performed for GSEA *P*-value estimation. This number was then increased with GSEA $P < 1 \times 10^{-4}$, and up to 1,000,000 permutations were performed. Results were sorted on the basis of FDR (5% cutoff), and FDR < 0.05 was considered to indicate significance.

© 2012 Nature America, Inc. All rights reserved.
npg

Analyses of directional consistency of associations between discovery and follow-up studies. We investigated whether the Metabochip follow-up SNPs were likely to contain further true associations in addition to those SNPs that reached genome-wide significance. To do so, we performed meta-analysis of those studies involved in the original discovery analyses[1,2], comprising 42,078 individuals for fasting glucose, 34,230 for fasting insulin and 15,252 for 2hGlu, and we then separately performed meta-analysis of all studies newly available to follow up, comprising 85,710 individuals for fasting glucose, 69,240 for fasting insulin and 27,602 for 2hGlu. For each trait (fasting glucose, fasting insulin, FI-BMIadj and 2hGlu), we identified all SNPs that had a nominally significant association ($P < 0.05$) in the follow-up studies alone and, for these SNPs, performed a two-sided binomial test of whether more SNPs than expected by chance (50%) had a consistent direction of effect with that observed in the discovery analyses. Before performing these analyses, SNPs were filtered by LD ($r^2 < 0.01$) to identify independent variants, and all SNPs (and those in LD, $r^2 \geq 0.01$) associated with glycemic traits (fasting glucose, fasting insulin, 2hGlu, HbA1c and proinsulin) at genome-wide levels of significance (including those SNPs identified in the present study) were excluded. These analyses were initially performed for all 66,000 SNPs, but we were then able to compare across SNPs submitted to the Metabochip by different consortia and for SNPs submitted to follow up on particular traits among these consortia. The results of each of these tests were plotted overall within SNPs from each consortium and within SNPs submitted for follow up of each trait (**Supplementary Fig. 9**). The numbers of SNPs meeting these criteria are shown in **Supplementary Table 7**. We supplemented these results with FDR analyses and noted the q value at $P = 0.05$ in the follow-up studies to identify the likelihood of true positives among these nominally significant SNPs (**Supplementary Table 7**).

87. D'Orazio, P. *et al.* Approved IFCC recommendation on reporting results for blood glucose (abbreviated). *Clin. Chem.* **51**, 1573–1576 (2005).
88. Li, Y., Willer, C., Sanna, S. & Abecasis, G.R. Genotype imputation. *Annu. Rev. Genomics Hum. Genet.* **10**, 387–406 (2009).
89. Li, Y., Willer, C.J., Ding, J., Scheet, P. & Abecasis, G.R. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet. Epidemiol.* **34**, 816–834 (2010).
90. Howie, B.N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
91. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P.R. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.* **39**, 906–913 (2009).
92. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
93. Mägi, R. & Morris, A.P. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics* **11**, 288 (2010).
94. Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336–2337 (2010).
95. Folkersen, L. *et al.* Association of genetic risk variants with expression of proximal genes identifies novel susceptibility genes for cardiovascular disease. *Circ. Cardiovasc. Genet.* **3**, 365–373 (2010).
96. Myers, S., Bottolo, L., Freeman, C., McVean, G. & Donnelly, P. A fine-scale map of recombination rates and hotspots across the human genome. *Science* **310**, 321–324 (2005).
97. Raychaudhuri, S. *et al.* Genetic variants at *CD28*, *PRDM1* and *CD2/CD58* are associated with rheumatoid arthritis risk. *Nat. Genet.* **41**, 1313–1318 (2009).
98. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).

© 2012 Nature America, Inc. All rights reserved.