

# Mouse models and the interpretation of human GWAS in type 2 diabetes and obesity

Roger D. Cox<sup>1,\*</sup> and Christopher D. Church<sup>1</sup>

Within the last 3 years, genome-wide association studies (GWAS) have had unprecedented success in identifying loci that are involved in common diseases. For example, more than 35 susceptibility loci have been identified for type 2 diabetes and 32 for obesity thus far. However, the causal gene and variant at a specific linkage disequilibrium block is often unclear. Using a combination of different mouse alleles, we can greatly facilitate the understanding of which candidate gene at a particular disease locus is associated with the disease in humans, and also provide functional analysis of variants through an allelic series, including analysis of hypomorph and hypermorph point mutations, and knockout and overexpression alleles. The phenotyping of these alleles for specific traits of interest, in combination with the functional analysis of the genetic variants, may reveal the molecular and cellular mechanism of action of these disease variants, and ultimately lead to the identification of novel therapeutic strategies for common human diseases. In this Commentary, we discuss the progress of GWAS in identifying common disease loci for metabolic disease, and the use of the mouse as a model to confirm candidate genes and provide mechanistic insights.

## The power of GWAS for studying human disease

Genome-wide association studies (GWAS) in human populations have been carried out for a broad range of common diseases and disease subtypes, including bipolar disorder, coronary heart disease, Crohn's disease, hypertension, rheumatoid arthritis, obesity, and type 1 and type 2 diabetes (Wellcome Trust Case Control Consortium, 2007). These studies have been huge undertakings, with large numbers of individuals in case-control groups and the application of very-high-density single-nucleotide polymorphism (SNP) genotyping. In many cases, this has required

a high degree of international cooperation, including meta-analysis in order to maximise the statistical power for the discovery and replication of newly identified loci that are associated with disease traits (Zeggini et al., 2007). This approach has been very successful, identifying numerous loci and novel candidate genes that are directly implicated in human disease; for type 2 diabetes, the current confirmed total stands at approximately 35 loci (Voight et al., 2010). However, each of the identified genes has a relatively small effect on disease risk, explaining only a very small part of the observed familiarity in type 2 diabetes (Saxena et al., 2007; Sladek et al., 2007; Zeggini et al., 2007; Dupuis et al., 2010; Voight et al., 2010). Thus, the predictive value of these loci for assessing an individual's disease risk is disappointing. There might be additional low-frequency or rare alleles of the identified genes in the population that further contribute to heritability, as is being found for hypertriglyceridemia (Gloyn and McCarthy, 2010; Johansen et al., 2010); however, the true value of identifying these disease-associated loci is that they provide insight into the biology of the disease and uncover potential new targets and pathways for future therapeutic intervention. Although the magnitude of the effect of common disease variants identified at the population level is small, there is still the possibility that large therapeutic (or detrimental) effects can be generated through pharmacological manipulation of a potential target or, indeed, through more extreme alleles.

The ultimate translation of disease-associated loci identified through GWAS into benefit for patients presents a number of challenges. First, GWAS identify novel loci that are defined by linkage disequilibrium (LD) blocks (also known as haplotype blocks; sequences that are inherited together in a non-random fashion), rather than identifying the causal genetic variant(s) [although small genomic regions are identified owing to the precision of high-density single-nucleotide polymorphism (SNP) panels]. Second, the SNPs that are associated with disease are not necessarily functional (in that they might have an effect on another gene that gives rise to a phenotypic consequence). Functional identification of the causal gene(s) within a disease-associated locus is the first essential step towards gaining new biological insight and identifying potential therapeutic targets.

## GWAS of type 2 diabetes and obesity

Some of the greatest success in identifying previously unknown susceptibility loci for common diseases has been in the identification of loci for type 2 diabetes and obesity. In 2006, the *TCF7L2* gene [a key Wnt signalling transcription factor encoded within a previously mapped chromosomal region that was suggestively linked with type 2 diabetes (Reynisdottir et al., 2003)] was directly implicated in the disease, initially through strong association in an Icelandic population that was then

<sup>1</sup>Metabolism and Inflammation, MRC Harwell Mammalian Genetics Unit, Harwell Science and Innovation Campus, Oxfordshire, OX11 0RD, UK  
\*Author for correspondence (r.cox@har.mrc.ac.uk)

© 2011. Published by The Company of Biologists Ltd  
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial Share Alike License (<http://creativecommons.org/licenses/by-nc-sa/3.0>), which permits unrestricted non-commercial use, distribution and reproduction in any medium provided that the original work is properly cited and all further distributions of the work or adaptation are subject to the same Creative Commons License terms.

### Glossary

- **Dominant negative:** a mutation that adversely alters the gene product by interfering with the function of the normal gene product (in heterozygotes).
- **Effect size** (also known as odds ratio): a measure of relative risk for a disease association compared with controls.
- **GWAS:** an unbiased method, using a dense array of single-nucleotide polymorphisms as genetic markers, to detect associations between genotype frequency and a trait of interest in large populations of patients and healthy controls.
- **Hypermorph:** a mutation that causes an increase in normal gene function; also known as gain of function.
- **Hypomorph:** a mutation that causes a partial loss of gene function.
- **Knockout:** complete ablation of gene function (also referred to as null), often involving the removal or replacement of large genomic regions including coding and intronic sequence.
- **Linkage disequilibrium:** the tendency of alleles located close to each other to be inherited together due to a recent mutation, genetic drift or selection; often observed as a correlation between genotypes and closely linked markers.
- **Overexpression:** the increased expression of a gene of interest or mutant variant controlled by the endogenous regulatory sequence or a defined promoter sequence. Can be either expressed from a defined genomic locus or randomly integrated into the genome.
- **Power:** the probability of detecting a statistically significant variant assumed to be causal.
- **Replication:** the evaluation of specific association signals in additional independent samples and populations.
- **SNPs:** single-nucleotide polymorphisms are defined as variations of a single base pair that occur at a significant frequency in a population.

replicated in two other populations of European ancestry (Grant et al., 2006). This result was later widely replicated in other populations. Mechanistic work indicates that this gene is involved in insulin secretion, probably at the level of regulating genes that are required for secretory granule exocytosis (see Xavier et al., 2009). The following year, 2007, was a landmark for type 2 diabetes and obesity GWAS. One of the first associations to be published was of SNPs within the first intron of a gene called *FTO* (Dina et al., 2007; Frayling et al., 2007; Scuteri et al., 2007): these SNPs were strongly linked to type 2 diabetes, but this association was lost when an adjustment for body mass index (BMI) was made (Frayling et al., 2007). The *FTO* locus predisposes both children and adults to obesity, and consequently increases the risk of type 2 diabetes (Dina et al., 2007; Frayling et al., 2007). Most importantly, *FTO* was a gene of unknown function in an unknown pathway (Frayling et al., 2007) and, therefore, there was much interest in determining whether *FTO* was the causal gene underlying the obesity association. This initial set of studies was rapidly followed by the publication of additional loci that were directly associated with type 2 diabetes (Scott et al., 2007; Sladek et al., 2007; Wellcome Trust Case Control Consortium, 2007; Zeggini et al., 2008; Voight et al., 2010). Of the loci identified to date, and the putative candidate genes contained within them, many are likely to affect insulin secretion, whereas few are potentially involved in insulin action (Voight et al., 2010) (Table 1). Interestingly, some loci are associated with different pathologies: for example, the *KLF14* locus has been identified in studies of both type 2 diabetes and basal cell carcinoma, although the functional significance of these associations remains to be determined (Stacey et al., 2009; Voight et al., 2010).

Similar success has been achieved in GWAS of obesity: following a flurry of papers in 2009, the total number of loci associated with obesity is at least 32 (Table 2) (Meyre et al., 2009; Thorleifsson et al., 2009; Willer et al., 2009; Heid et al., 2010; Scherag et al., 2010; Speliotes et al., 2010). Many of these loci contain genes that point towards a role for neurons and the brain in the disease, particularly the hypothalamus [which has a major role in obesity (for a review, see Belgardt et al., 2009)], whereas fewer loci might influence adipocyte biology and still others are completely novel.

### Approaches for identifying the underlying disease-associated gene

There are several complementary approaches being applied to identify the underlying causative gene within each disease-associated locus. The availability of the human genome sequence allows a list of genes within a locus to be compiled, and the haplotype map (HapMap) data on linkage disequilibrium between SNPs allows the size of a locus to be defined. Available information about each of the listed genes – for example, about SNP proximity, gene expression, expression quantitative trait loci (eQTL; see below), known animal models and the phenotype of interest – can then be collated, and the plausibility of each gene as a candidate involved in disease can be evaluated.

In some cases there might be very strong candidates; for example, SNPs 188 kb downstream of the *MC4R* gene were found to be associated with elevated BMI at the population level (Loos et al., 2008). The *MC4R* gene is well characterised and is known to be involved in the control of food intake by the hypothalamus. Furthermore, there are known monogenic mutations in *MC4R* that cause severe childhood obesity (Vaisse et al., 1998; Yeo et al., 1998). Although an obvious candidate, the functional evidence showing how the DNA marked by the associated SNPs influences *MC4R* expression or *MC4R* protein function is missing. Understanding how the SNPs, or the true functional variants that they mark, influence *MC4R* and BMI is an important scientific question. The fact that this gene is involved in obesity indicates that *MC4R* might be a potential therapeutic target, although an *MC4R* receptor agonist has been shown to increase blood pressure in humans (Greenfield et al., 2009).

When there is no obvious candidate gene within an identified locus, the search for a causative gene is more problematic. One approach being used to address this problem is to exploit eQTL data (e.g. Nica et al., 2010; Nicolae et al., 2010). These datasets are constructed by correlating expression levels of genes, usually measured using a highly parallel expression profiling technology, with the presence of different SNPs. Using natural genetic polymorphisms between individuals, differences in expression of specific genes between individuals can be mapped to SNPs. Many of these SNPs are in *cis* with the gene that is being regulated, although this is not always the case. Using these datasets, it is possible to ask: are the associated SNPs, or flanking DNA, linked with the expression of another gene in the region? If there is a gene being regulated according to the genotype of a GWAS-associated SNP (or one in LD with it), then that gene is a candidate to test for its potential involvement in the disease. One potential limitation of this approach is that much of the available highly parallel expression data is from human lymphoblastoid cell lines, which might not faithfully replicate endogenous patterns of expression in

**Table 1. Type 2 diabetes susceptibility loci identified through GWAS**

Candidate gene symbol	Gene name	Location	Effect size <sup>a</sup>	References
<i>NOTCH2</i>	Notch homologue 2	1p12	1.13	Zeggini et al., 2008
<i>PROX1</i>	Prospero homeobox 1	1q41	1.07	Dupuis et al., 2010
<i>BCL11A</i>	B-cell lymphoma/leukemia 11A	2p16.1	1.08	Voight et al., 2010
<i>THADA</i>	Thyroid adenoma associated	2p21	1.15	Zeggini et al., 2008
<i>GCKR</i>	Glucokinase (hexokinase 4) regulator	2p23	1.06	Dupuis et al., 2010
<i>G6PC2</i>	Glucose-6-phosphatase, catalytic, 2	2q24.3	0.97	Zeggini et al., 2007; Bouatia-Naji et al., 2008; Chen et al., 2008; Dupuis et al., 2010
<i>IRS1</i>	Insulin receptor substrate 1	2q36	1.11	Voight et al., 2010
<i>ADAMTS9</i>	ADAM metallopeptidase with thrombospondin type 1 motif 9	3p14	1.09	Zeggini et al., 2008
<i>ADCY5</i>	Adenylate cyclase 5	3q13.2	1.12	Dupuis et al., 2010
<i>PPARG</i>	Peroxisome proliferator-activated receptor $\gamma$	3q25	0.78	Altshuler et al., 2000
<i>SLC2A2</i>	Solute carrier family 2 (facilitated glucose transporter), member 2	3q26.2	1.01	Dupuis et al., 2010
<i>IGF2BP2</i>	Insulin-like growth factor 2 mRNA-binding protein 2	3q28	1.14	Saxena et al., 2007; Scott et al., 2007; Zeggini et al., 2008
<i>WFS1</i>	Wolfram syndrome 1	4p16.1	1.12	Young et al., 2007
<i>ZBED3</i>	Zinc finger, BED-type containing 3	5q13.2	1.08	Voight et al., 2010
<i>CDKAL1</i>	CDK5 regulatory subunit-associated protein 1-like 1	6p22.2	1.14	Saxena et al., 2007; Scott et al., 2007; Steinthorsdottir et al., 2007; Zeggini et al., 2008
<i>JAZF1</i>	Juxtaposed with another zinc finger gene 1	7p15	1.10	Zeggini et al., 2008
<i>GCK</i>	Glucokinase (hexokinase 4)	7p15.3	1.07	Dupuis et al., 2010
<i>DGKB</i> and <i>TMEM195</i>	Diacylglycerol kinase $\beta$ ; transmembrane protein 195 (alkylglycerol monooxygenase)	7p21.1	1.06	Dupuis et al., 2010
<i>KLF14</i>	Kruppel-like factor 14	7q32.3	1.07	Voight et al., 2010
<i>TP53INP1</i>	Kumor protein p53 inducible nuclear protein 1	8q22	1.06	Voight et al., 2010
<i>SLC30A8</i>	Solute carrier family 30 member 8 (zinc transporter 8)	8q24.11	1.15	Sladek et al., 2007; Dupuis et al., 2010
<i>CDKN2A</i> and <i>CDKN2B</i>	Cyclin-dependent kinase 4 inhibitor 2A and B	9p21	1.20	Saxena et al., 2007; Scott et al., 2007; Zeggini et al., 2008
<i>GLIS3</i>	GLIS family zinc finger 3	9p24.2	1.03	Dupuis et al., 2010
<i>CHCHD9</i>	Coiled-coil-helix-coiled-coil-helix domain containing 9	9q21.3	1.11	Voight et al., 2010
<i>CDC123</i> and <i>CAMK1D</i>	Cell division cycle protein 123 homologue ( <i>Saccharomyces cerevisiae</i> ); calcium/calmodulin-dependent protein kinase type 1D	10p13	1.11	Zeggini et al., 2008
<i>HHEX</i> and <i>IDE</i>	Haematopoietically-expressed homeobox; insulin-degrading enzyme	10q23	1.15	Sladek et al., 2007
<i>ADRA2A</i>	Alpha-2A adrenergic receptor	10q24	1.04	Dupuis et al., 2010
<i>TCF7L2</i>	Transcription factor 7-like 2	10q25.3	1.40	Grant et al., 2006; Sladek et al., 2007; Dupuis et al., 2010
<i>MADD</i>	MAP-kinase activating death domain	11p11.2	1.01	Dupuis et al., 2010
<i>CRY2</i>	Cryptochrome 2	11p11.2	1.04	Dupuis et al., 2010
<i>KCNJ11</i>	Potassium channel, inwardly rectifying, subfamily J, member 11	11p15.1	1.23	Gloyn et al., 2003
<i>KCNQ1</i>	Potassium voltage-gated channel, KQT-like subfamily, member 1	11p15.5	1.29	Yasuda et al., 2008; Maeda et al., 2010
<i>FADS1</i>	Fatty acid desaturase 1	11q12	1.04	Dupuis et al., 2010
<i>ARAP1</i>	ArfGAP with RhoGAP domain, ankyrin repeat and PH domain 1	11q13.3	1.14	Voight et al., 2010
<i>MTNR1B</i>	Melatonin receptor 1B	11q21	1.09	Prokopenko et al., 2009; Dupuis et al., 2010; Voight et al., 2010
<i>HMGA2</i>	High mobility group AT-hook 2	12q15	1.10	Voight et al., 2010
<i>TSPAN8</i> and <i>LGR5</i>	Tetraspanin-8; leucine-rich repeat-containing G-protein coupled receptor 5	12q21	1.09	Zeggini et al., 2008
<i>IGF1</i>	Insulin-like growth factor 1	12q23.2	1.04	Dupuis et al., 2010
<i>HNF1A</i>	Hepatocyte nuclear factor-1 $\alpha$	12q24.3	1.07	Voight et al., 2010
<i>C2CD4B</i>	C2 calcium-dependent domain containing 4B	15q22.2	1.03	Dupuis et al., 2010
<i>ZFAND6</i>	Zinc finger, AN1-type domain 6	15q24.3	1.06	Voight et al., 2010
<i>PRC1</i>	Protein regulator of cytokinesis 1	15q26.1	1.07	Voight et al., 2010
<i>FTO</i>	Fat mass and obesity associated	16q12.2	1.17	Dina et al., 2007; Frayling et al., 2007; Scuteri et al., 2007
<i>HNF1B</i>	Hepatocyte nuclear factor 1 homeobox B	17q12	1.13	Gudmundsson et al., 2007; Winckler et al., 2007
<i>DUSP9</i>	Dual specificity phosphatase 9	Xq28	1.27	Voight et al., 2010

<sup>a</sup>The effect size (also known as odds ratio) is a measure of relative risk for a disease association compared with controls.

**Table 2. Obesity-associated loci identified through GWAS**

Candidate gene symbol	Gene name	Location	Effect size <sup>a</sup>	References
<i>NEGR1</i>	Neuronal growth regulator 1	1p31	1.05	Thorleifsson et al., 2009; Willer et al., 2009
<i>SEC16B</i> and <i>RASAL2</i>	SEC16 homologue B ( <i>S. cerevisiae</i> ); RAS protein activator like 2	1q25	1.11	Thorleifsson et al., 2009
<i>SDCCAG8</i>	Serologically defined colon cancer antigen 8 gene	1q43	1.19	Scherag et al., 2010
<i>TMEM18</i>	Transmembrane protein 18	2p25	1.41	Willer et al., 2009
<i>ETV5</i>	Ets variant gene	3q27	1.11	Thorleifsson et al., 2009
<i>GNPDA2</i>	Glucosamine-6-phosphate deaminase	4p13	1.12	Willer et al., 2009
<i>PRL</i>	Prolactin	6p22.2	0.31	Meyre et al., 2009
<i>NCR3</i> , <i>AIF1</i> and <i>BAT2</i>	Natural cytotoxicity triggering receptor 3 precursor prolactin; allograft inflammatory factor 1; HLA-B associated transcript 2	6p21	1.07	Thorleifsson et al., 2009
<i>TNKS</i> and <i>MSRA</i>	Tankyrase-1 (TRF1-interacting ankyrin-related ADP-ribose polymerase); methionine sulfoxide reductase A gene	8p23	1.00	Scherag et al., 2010
<i>PTER</i>	Phosphotriesterase-related	10p12	0.68	Meyre et al., 2009
<i>BDNF</i>	Brain-derived neurotrophic factor	11p14	1.11	Thorleifsson et al., 2009
<i>MTCH2</i>	Mitochondrial carrier homologue 2	11p11.2	1.03	Willer et al., 2009
<i>FAIM</i> and <i>BCDIN3D</i>	Fas apoptotic inhibitory molecule; BCDIN3 domain containing	12q13	1.14	Willer et al., 2009
<i>SH2B1</i>	SH2B adaptor protein 1	16p11.2	1.11	Thorleifsson et al., 2009; Willer et al., 2009
<i>MAF</i>	V-maf musculoaponeurotic fibrosarcoma oncogene homologue	16q22	1.39	Meyre et al., 2009
<i>FTO</i>	Fat mass and obesity associated	16q22.2	1.67	Frayling et al., 2007; Thorleifsson et al., 2009; Willer et al., 2009
<i>NPC1</i>	Niemann-Pick disease type C1	18q11.2	0.75	Meyre et al., 2009
<i>MC4R</i>	Melanocortin 4 receptor	18q22	1.12	Loos et al., 2008; Meyre et al., 2009; Thorleifsson et al., 2009; Willer et al., 2009
<i>KCTD15</i>	Potassium channel tetramerization domain containing 15	19q13.11	1.04	Thorleifsson et al., 2009; Willer et al., 2009

<sup>a</sup>The effect size (also known as odds ratio) is a measure of relative risk for a disease association compared with controls. Eighteen additional obesity loci have now been identified (Speliotes et al., 2010) for BMI, as well as 13 for body fat (Heid et al., 2010).

vivo – as in, for example, a neuron in the hypothalamus of the brain that controls food intake (Nicolae et al., 2010). It has also been suggested that many genes have an eQTL associated with them; as these datasets grow in size, the correlations between trait-associated SNPs and eQTLs might occur by chance (Nica et al., 2010). Approaches are being developed to address some of these issues, such as taking account of LD in the analysis (Nica et al., 2010). For example, in a study of fasting diabetes-related traits, eQTL analysis indicated that *FAQDS1* is a candidate gene, because the most closely associated SNP is in LD with an SNP that is thought to be involved in regulating the levels of *FAQDS1* mRNA (Dupuis et al., 2010). Furthermore, a recent type 2 diabetes GWAS reported a strong *IRS-1 cis* eQTL (Voight et al., 2010).

As we alluded to above, disease-associated SNPs might tag functional variants that are in LD with the SNPs used in mapping. For a small number of genes, the coding variants are known and are in LD with associated SNPs; for example, an SNP that is linked to type 2 diabetes and is located near the glucokinase regulatory protein (*GCKR*) gene is in LD with a P446L variant that was previously associated with glucose phenotypes, such as elevated fasting glucose and insulin levels, and for which there is some functional evidence that it influences glucokinase (*GCK*) activity

(Beer et al., 2009; Dupuis et al., 2010). Consequently, another potentially valuable approach for identifying causal variants is the sequencing of individuals with and without the disease to investigate a locus for other potentially causal variants in genes. Furthermore, there might be other variants in different population subsets or extreme cases. Such an approach was described above for hypertriglyceridemia (Johansen et al., 2010).

A novel approach using zebrafish to identify causal genes has been described by Ragvin et al. (Ragvin et al., 2010). They took the sequence within LD blocks for three loci from type 2 diabetes associations and searched across species for highly conserved noncoding elements that potentially marked genomic regulatory blocks. They then tested these blocks using a transgenic reporter in zebrafish embryos and correlated expression patterns with other genes in the syteny region (where the gene order is conserved between species). For example, the region containing *CDKALI* contains several additional genes, including *SOX4*, and the associated SNPs seem to mark a highly conserved element within the fifth intron of *CDKALI*. Ragvin et al. took this element and tested it for enhancer activity in zebrafish embryos; they found that the pattern of expression directed by this element closely resembled that of *sox4* and not *cdk11*. They concluded that *SOX4*, for which

there is biological evidence supporting a role in insulin secretion, was a more likely candidate for involvement in type 2 diabetes than was *CDKALI*, which was more probably a bystander (Ragvin et al., 2010). This is a very interesting approach, although it might be limited owing to the requirement to use embryos. In addition, the restricted patterns of expression of some of the other genes within the tested regions might result in some genes being rejected as the causal gene because they are not expressed in zebrafish either at that time or location.

The approaches discussed above only identify candidate genes – thus, it is still essential to functionally validate each gene. Ideally this would include a definition of the causal genetic variant linked to the mapped SNPs within a gene, or its functional regulatory elements, and how these changes affect the function of the target gene(s). Furthermore, this validation would address the function of those genes in an intact organism and how they give rise to increased disease risk. As discussed above, another complementary approach would be to test the candidate genes for a functional role in disease without necessarily understanding the details of the functional polymorphisms marked by the human SNPs. In this latter approach, the mouse, as well as other model organisms, can play a key role.

### The value of mouse models in diabetes and obesity research

Mouse models have great potential to help take GWAS data to the next level by investigating the function of associated genes in vivo. This section will focus on the use of the mouse to model human metabolic disease. The mouse is a good mammalian model organism for several reasons: the availability of its complete genomic sequence, genetically defined strains and extensive genetic manipulation tools; the ease of breeding; the ability to control breeding and the phenotyping environment; and the availability of a wide array of phenotyping tests, many of which are standardised and easily applied. But how good a model is the mouse for human obesity and type 2 diabetes? As discussed below, many aspects of the physiology of these disorders appear to be the same in both species, although there are some differences. Notably, the control of blood glucose is similar in many respects: both humans and mice exhibit impaired glucose tolerance and overt diabetes that can be diagnosed with essentially the same symptoms and tests.

### Mouse models of type 2 diabetes and dysregulated glucose metabolism

An interesting difference between mice and humans is that mice have high levels of high-density lipoprotein (HDL) cholesterol that might protect them against atherosclerosis. Dyslipidaemia is an important component of metabolic syndrome (diabetes, obesity and hypertension) in patients. The basis of this difference might be explained by two recent papers showing that *miR-33a* and *miR-33b*, which are located within the two genes encoding sterol regulatory element-binding proteins (SREBPs), are involved in regulating cholesterol biosynthesis (Najafi-Shoushtari et al., 2010; Rayner et al., 2010) (for a review, see Brown et al., 2010). These microRNAs target several mRNAs for destruction, including that of the *ABCA1* gene, which is required for cholesterol transport and, in the liver, for HDL production. Unlike humans, mice lack *miR-33b*, which is encoded in an intron of *SREBP1*, and so might have

higher levels of *ABCA1* and thus higher levels of HDL. It will be interesting to see whether this is indeed the case and, more importantly, what roles these microRNAs have in human metabolic syndrome (Brown et al., 2010; Najafi-Shoushtari et al., 2010).

Mutations in the mouse can give rise to similar phenotypes to those observed in humans carrying the corresponding mutations. For example, point mutations in the *GCK* gene give rise to maturity-onset diabetes of the young 2 (*GCK-MODY2*) in humans and to a similar phenotype in the mouse (e.g. Toye et al., 2004). In the case of at least one *GCK* mutation in the mouse, in addition to inducing a dominant phenotype, the mutation is homozygous viable, and there is a more severe glucose intolerance phenotype in homozygous mice compared with heterozygous mice (Toye et al., 2004). By contrast, mouse homozygous *GCK*-knockout mutations are lethal within a week of birth due to severe diabetes (Bali et al., 1995; Grupe et al., 1995). Rare homozygous viable human mutations in *GCK* have also been described, although disease in these patients is managed through insulin treatment (Njolstad et al., 2001). Therefore, the effect of the expression of a disease-associated gene between human and mouse can vary depending on the type of allele.

Not all mutations in mice clearly recapitulate the effects of disease-causing mutations in humans. An example of this is hyperinsulinism of infancy (HI) caused by inactivating mutations in the  $K_{ATP}$  channel subunit *KCNJ11*. Patients with HI exhibit life-threatening hypoglycaemia; however, mice carrying the disease-causing mutations show a more varied phenotype (Seino et al., 2000). Some knockouts show transient neonatal hypoglycaemia, whereas some dominant-negative transgenic mice go on to show overt diabetes due to loss of  $\beta$ -cell mass following initial hypersecretion (Miki et al., 1997; Seino et al., 2000). We recently described a homozygous *Kcnj11* point mutation in the mouse that was identical to that found in a young HI patient (Hugill et al., 2010). Unlike the patients carrying this mutation, our mouse did not exhibit hypersecretion of insulin or hypoglycaemia, but we did observe impaired glucose tolerance and reduced insulin secretion (Hugill et al., 2010). By contrast, mice carrying gain-of-function mutations in *Kcnj11* (i.e. channel-open mutations), which give rise to neonatal diabetes in humans, have a phenotype that closely resembles the human disease (Koster et al., 2000). This is illustrated in the study of the  $\beta$ -cell-specific expression of transgenic mice carrying *Kcnj11* with the V59M mutation (corresponding to a well-characterised human mutation); this mutation gives rise to severe neonatal diabetes and the *Kcnj11* mutant mice are providing a valuable tool for investigating the pathology and physiology of this disease (Girard et al., 2009). Furthermore, neuronal- or muscle-specific expression of this mutation shows that the muscle weakness observed in some patients with neonatal diabetes is of neurological origin. Overall, these studies show the importance of tissue-specific mouse models in revealing novel therapeutic strategies and drug-tissue specificity. In the case of neonatal diabetes, the treatment of muscle problems will require drugs that can pass the blood-brain barrier (Clark et al., 2010).

The morbidity and mortality associated with type 2 diabetes is the result of complications in multiple tissues, such as heart, kidney, retina and peripheral vasculature [for information on the prevalence of diabetes and complications in the UK see, for example, the International Diabetes Federation, IDF Diabetes Atlas 4th Edition (2009) (<http://www.diabetesatlas.org>)]. Whether the mouse is a

good model for diabetic complications is an open question. These complications occur in patients over a long period of time, well beyond the lifespan of a mouse. Understandably, mouse models are usually not aged, as many experiments are done in the first few weeks or months of life. Ideally, mouse models of type 2 diabetes would be aged to a year or 18 months to monitor complications. There are a number of potential mouse models of diabetic nephropathy (Brosius, 2010); for example, susceptible mouse strains, such as FVB, db/db or diabetic DBA/2 mice, have been successfully used to identify the novel nephropathy *Dbnph1* locus (Brosius, 2010; Chua et al., 2010).

### Mouse models of obesity

The mouse has also been a pivotal model in obesity research. The identification, by the Friedman laboratory, of the genes underlying the diabetes mouse (*db/db; LepR/LepR*) and obesity mouse (*ob/ob; Lep/Lep*) as the leptin receptor and leptin itself, respectively, has been essential to the identification of hypothalamic pathways controlling food intake that are highly conserved across species (for a recent review, see Konner et al., 2009).

In summary, the mouse is a reasonably faithful model for type 2 diabetes and obesity, and offers great potential for investigating the underlying molecular and cellular mechanisms of disease development. Clearly, there are differences between mouse and human (although some of these differences might even help to provide insights into disease mechanisms) but, as one of the approaches available alongside work in humans and in other species and systems including in vitro experiments, the mouse is set to be a key contributor to the challenges that we face.

### Taking on the functional challenge of GWAS using mouse models

Given that mouse models of type 2 diabetes and obesity are reasonable models of the human disorders, mice can be used to functionally test the genes within a GWAS locus. This can be achieved by examining the phenotypic consequences of genetic differences within candidate genes, including natural polymorphisms, knockout alleles, overexpression alleles, dominant-negative alleles and point mutations. An advantage of using, at least in the initial analysis, alleles that are globally expressed is that no assumptions are made about the tissues involved in a complex multiorgan disease such as type 2 diabetes; this is especially important for genes with poorly characterised or unknown function.

Knockout alleles have been used, for example, to assess whether the *FTO* gene, in which the associated obesity and/or diabetes SNPs are within intron 1, is in fact causal for these diseases. Homozygous *Fto*-knockout mice exhibit a lean phenotype characterised by reduced lean mass, reduced adipose tissue, reduced spontaneous locomotor activity, increased energy expenditure and relative hyperphagia, as well as increased postnatal lethality and postnatal growth retardation (Fischer et al., 2009). This suggests that the *FTO* gene is involved in determining adiposity, although this is complicated by the additional phenotypes of lethality and growth retardation observed when the gene is knocked out. Mice with a missense point mutation in *Fto* causing a loss of the protein's demethylase function exhibited a milder phenotype of reduced fat mass without the additional complications, further supporting a

role for this gene in determining BMI (Church et al., 2009). In addition, we have observed that overexpression of *Fto* in mice results in obesity essentially due to increased food intake (Church et al., 2010). Increased food intake, in individuals with the at-risk allele, has been observed in the majority of human *FTO* studies (Speakman et al., 2008; Timpson et al., 2008). Whether the human at-risk genotype affects expression levels of *FTO* in humans is unclear; however, there is one report that shows that, in fibroblasts and blood samples from heterozygous individuals, the expression levels of primary mRNA transcripts encoding the at-risk genotype are higher than those encoding the low-risk genotype (Berulava and Horsthemke, 2010). It would be interesting to know what the relative *FTO* expression levels are in, for example, neurons that are involved in food intake. The next issue that can be addressed using mouse models to investigate the function of *FTO* is which tissue(s) are involved in causing the phenotype. This can be studied using conditional alleles (knockout or overexpression floxed alleles in combination with tissue-specific Cre-recombinase lines) that allow knockout of the *Fto* gene only in specific tissues. An alternative to generating tissue-specific Cre mouse lines is to use adenovirus technology for targeted delivery of the recombinase to specific tissues or regions, as has been carried out for *Fto* in the brain (Tung et al., 2010).

Type 2 diabetes GWAS also identified the islet-restricted *SLC30A8* gene (which encodes the zinc transporter ZNT8) (Sladek et al., 2007). Inactivation of *SLC30A8* in the mouse supports the role of *SLC30A8* as the causative gene underlying the susceptibility to type 2 diabetes conferred by the *rs13266634* locus: the mice exhibited a defect in insulin secretion as well as glucose intolerance (Lemaire et al., 2009; Nicolson et al., 2009; Pound et al., 2009). Furthermore,  $\beta$ -cell-specific inactivation of *SLC30A8* in the mouse also results in glucose intolerance (Wijesekara et al., 2010). *SLC30A8* might therefore be a promising therapeutic target for the treatment of type 2 diabetes.

These models that functionally test disease-associated genes illustrate how the mouse can help to identify causative genes among GWAS data by providing functional evidence to support the involvement of a candidate gene in a disease phenotype. These models also provide tools for mechanistic studies that would be difficult to perform in humans, because mice enable access to tissues and the opportunity to carry out detailed in vivo physiological analysis.

There is an increasing availability of different mouse models in which to study disease-associated genes. The aim of the International Knockout Mouse Consortium is to make mutations in all mouse protein-coding genes, and is systematically generating gene-targeting and gene-trap constructs in C57BL/6N embryonic stem cells (see <http://www.knockoutmouse.org/about>). Other programmes, such as EUMODIC and the International Mouse Phenotyping Consortium, are carrying out or proposing systematic high-throughput phenotyping of these mouse lines, which might provide evidence that disease-associated mutations identified in GWAS directly influence BMI or alter blood glucose (see <http://www.eumodic.org/>). In addition, phenotype data on mice is being collected in a number of databases, including EUROPHENOME (<http://www.europhenome.org/>) and the Mouse Phenome Database (<http://phenome.jax.org/>). Together, these efforts to genotype and phenotype mutant mice should provide

additional data that could aid in validating candidate disease genes identified in GWAS.

**Approaches to the translational challenges**

Even if we succeed in discovering the underlying disease-causing genes, how can we translate this into benefits for patients? Although an animal model is useful for testing potential therapeutic strategies, it is alone insufficient to bridge this gap. A clear understanding of the mechanisms by which a gene influences a disease, and of its function in the organism, might help to promote it as a pharmacological target to industry. Mouse models can facilitate mechanistic studies by providing tissue for experiments. They can also be used for the testing of other pathway components, such as upstream or downstream molecules and post-translational modifications, by perturbing them pharmacologically or genetically. Through clinical collaboration, the phenotype observed in mouse models can be compared with the human disease phenotype and, similarly, observations in patients can be examined and investigated further in the models.

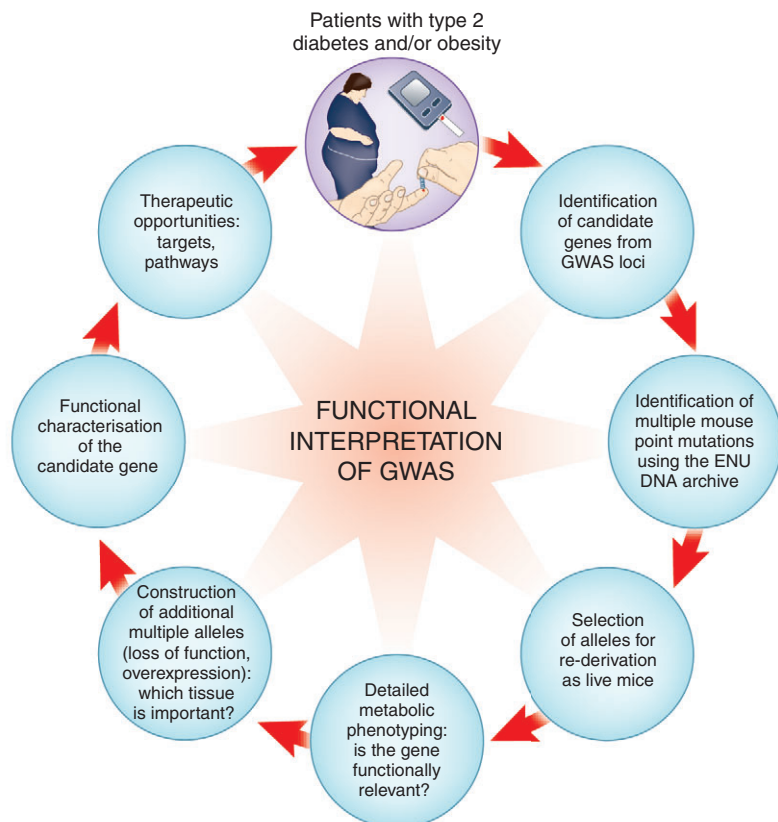
The mouse models investigating the role of the *FTO* gene suggest that pharmacological inhibition of the encoded protein's enzymatic activity might lead to a reduction in adiposity (see above), which is the basis of an anti-obesity drug. However, the high level of postnatal lethality in the *Fto* knockout and the severe multiorgan and developmental problems observed in a human carrying a catalytic inactive mutation in *FTO* suggests that inhibition of *FTO* as an anti-obesity therapy might have detrimental side effects (Boissel et al., 2009). Therefore, further research is needed on the function of this gene before it can be

considered a viable therapeutic target that can be translated to the clinic. Additional experiments involving genetic manipulations in the mouse might aid our understanding of the tissue specificity and function of *FTO* in vivo.

**A future strategy using the mouse**

A broad-ranging strategy is needed to exploit the success of GWAS and translate the labour-intensive data they have generated into benefit for patients. Although this might take years, such a strategy may ultimately identify molecular pathways for common human diseases, some of which might overlap, that are potential pharmacological targets. The mouse can contribute to this goal by providing a platform for the functional analysis of candidate genes. Once genes have been selected for testing, there are a number of options for allele selection, including making conditional knockout alleles, or obtaining gene-trap or targeted allele embryonic stem cells from distribution centres. Another approach that we have exploited is to search for point mutations in DNA archives of male mice mutagenised with the powerful point mutagen ENU (Coghill et al., 2002; Quwailid et al., 2004; Church et al., 2009). This is advantageous because it can reveal a series of point mutations in the coding sequence of a candidate gene that give a variety of allele types, including nulls, hypomorphs, hypermorphs and dominant-negative alleles. Investigating the effects of the different allele types could be very informative, as described above for *FTO*. Moreover, a single base pair change in the mouse genome induced by ENU mutagenesis is unlikely to affect the expression or function of nearby genes. Alleles can then be selected for generation as live mice that are

Disease Models & Mechanisms • DMM



**Fig. 1. A strategy for functional interpretation of GWAS.** GWAS candidate genes are selected on the basis of SNP proximity, expression and the trait of interest. An allelic series in the mouse can be identified using ENU mutagenesis archive screens and mutations selected on the basis of potential functional consequences and evolutionary conservation. Next generation sequencing (NGS) can allow a catalogue of mutations to be made. Each variant is characterised for the trait of interest – for example, glucose tolerance or body weight. Conditional mouse models can also be generated to investigate the tissue specificity and temporal role of candidate genes in disease using mice expressing Cre recombinase. Further functional analysis using expression analysis and proteomics can reveal novel therapeutic targets for common diseases. This scheme is based on proof-of-principle studies that defined a function for *FTO*, a gene of previously unknown function in an unknown pathway (Frayling et al., 2007), in obesity.

### Online resources

- ENSEMBL; genomic sequence and annotation: <http://www.ensembl.org/index.html>
- HapMap; linkage disequilibrium information: <http://hapmap.ncbi.nlm.nih.gov/whatis/hapmap.html>
- GWAS; searchable lists of GWAS loci: <http://www.genome.gov/26525384>
- EMPRESS; standardised mouse phenotyping protocols: <http://empress.har.mrc.ac.uk/>
- Diabetes information and reports: <http://www.diabetes.org.uk/>  
<http://www.cdc.gov/diabetes/>  
<http://www.diabetes.org/>  
[www.idf.org](http://www.idf.org)

then subjected to detailed metabolic phenotyping (weight, body composition, glucose tolerance, insulin sensitivity, food intake, metabolic rate, etc.) using assays that are well established in the mouse. If phenotypes or subphenotypes relevant to the human disease are detected, then further alleles could be generated (such as overexpression alleles and conditional alleles) to investigate tissue specificity. These refined models then provide a resource for further functional analysis of candidate genes and pathways, which can include the application of proteomics, expression profiling and partner-protein analysis. A schematic of this strategy is illustrated in Fig. 1. The combination of multiple disease alleles in separate mouse lines, originally identified through GWAS in human populations, should provide excellent models of human disease for functional and therapeutic analysis.

This strategy is potentially very powerful for uncovering genes that are associated with human disease, but, if applied systematically to all GWAS candidates, it is also resource hungry. Several genes per locus must be analysed in some detail in order to confirm or reject them as causal candidates. However, the potential gains are also great, because new functionally validated targets for development by pharmaceutical and biotech industries will probably be identified. A single approach will not be enough, however, because even “The best laid schemes of mice and men, Go often askew...” (Robert Burns’ poem ‘To A Mouse, On Turning Her Up In Her Nest With The Plough’, 1785). So, success will rely on a flexible and targeted approach, using all available tools as appropriate, as well as close collaboration between basic and clinical researchers.

### COMPETING INTERESTS

Research in the authors’ lab was funded by the UK Medical Research Council, Diabetes UK and the Wellcome Trust. Christopher D. Church holds an MRC doctoral training studentship. The funders had no role in the preparation of this article, and authors declare that there are no competing financial interests.

### REFERENCES

- Altshuler, D., Hirschhorn, J. N., Klannemark, M., Lindgren, C. M., Vohl, M. C., Nemes, J., Lane, C. R., Schaffner, S. F., Bolk, S., Brewer, C. et al. (2000). The common PPAR $\gamma$  Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat. Genet.* **26**, 76–80.
- Bali, D., Svetlanov, A., Lee, H. W., Fusco-DeMane, D., Leiser, M., Li, B., Barzilai, N., Surana, M., Hou, H., Fleischer, N. et al. (1995). Animal model for maturity-onset diabetes of the young generated by disruption of the mouse glucokinase gene. *J. Biol. Chem.* **270**, 21464–21467.
- Beer, N. L., Tribble, N. D., McCulloch, L. J., Roos, C., Johnson, P. R., Orho-Melander, M. and Gloyn, A. L. (2009). The P446L variant in GCKR associated with fasting plasma glucose and triglyceride levels exerts its effect through increased glucokinase activity in liver. *Hum. Mol. Genet.* **18**, 4081–4088.

- Belgardt, B. F., Okamura, T. and Brüning, J. C. (2009). Hormone and glucose signalling in POMC and AgRP neurons. *J. Physiol.* **587**, 5305–5314.
- Berulava, T. and Horsthemke, B. (2010). The obesity-associated SNPs in intron 1 of the FTO gene affect primary transcript levels. *Eur. J. Hum. Genet.* **18**, 1054–1056.
- Boissel, S., Reish, O., Proulx, K., Kawagoe-Takaki, H., Sedgwick, B., Yeo, G. S., Meyre, D., Golzio, C., Molinari, F., Kadhon, N. et al. (2009). Loss-of-function mutation in the dioxygenase-encoding FTO gene causes severe growth retardation and multiple malformations. *Am. J. Hum. Genet.* **85**, 106–111.
- Boutia-Naji, N., Rocheleau, G., Van Lommel, L., Lemaire, K., Schuit, F., Cavalcanti-Proenca, C., Marchand, M., Hartikainen, A. L., Sovio, U., De Graeve, F. et al. (2008). A polymorphism within the G6PC2 gene is associated with fasting plasma glucose levels. *Science* **320**, 1085–1088.
- Brosius, F. C. (2010). Susceptible mice: identifying a diabetic nephropathy disease locus using a murine model. *Kidney Int.* **78**, 431–432.
- Brown, M. S., Ye, J. and Goldstein, J. L. (2010). Medicine. HDL miR-ed down by SREBP introns. *Science* **328**, 1495–1496.
- Chen, W. M., Erdos, M. R., Jackson, A. U., Saxena, R., Sanna, S., Silver, K. D., Timpson, N. J., Hansen, T., Orru, M., Grazia Piras, M. et al. (2008). Variations in the G6PC2/ABCB11 genomic region are associated with fasting glucose levels. *J. Clin. Invest.* **118**, 2620–2628.
- Chua, S., Jr, Li, Y., Liu, S. M., Liu, R., Chan, K. T., Martino, J., Zheng, Z., Susztak, K., D’Agati, V. D. and Gharavi, A. G. (2010). A susceptibility gene for kidney disease in an obese mouse model of type II diabetes maps to chromosome 8. *Kidney Int.* **78**, 453–462.
- Church, C., Lee, S., Bagg, E. A., McTaggart, J. S., Deacon, R., Gerken, T., Lee, A., Moir, L., Mecinovic, J., Quwailid, M. M. et al. (2009). A mouse model for the metabolic effects of the human fat mass and obesity associated FTO gene. *PLoS Genet.* **5**, e1000599.
- Church, C., Moir, L., McMurray, F., Girard, C., Banks, G. T., Teboul, L., Wells, S., Bruning, J. C., Nolan, P. M., Ashcroft, F. M. et al. (2010). Overexpression of Fto leads to increased food intake and results in obesity. *Nat. Genet.* **42**, 1086–1092.
- Clark, R. H., McTaggart, J. S., Webster, R., Mannikko, R., Iberl, M., Sim, X. L., Rorsman, P., Glitsch, M., Beeson, D. and Ashcroft, F. M. (2010). Muscle dysfunction caused by a KATP channel mutation in neonatal diabetes is neuronal in origin. *Science* **329**, 458–461.
- Coghill, E. L., Hugill, A., Parkinson, N., Davison, C., Glenister, P., Clements, S., Hunter, J., Cox, R. D. and Brown, S. D. (2002). A gene-driven approach to the identification of ENU mutants in the mouse. *Nat. Genet.* **30**, 255–256.
- Dina, C., Meyre, D., Gallina, S., Durand, E., Korner, A., Jacobson, P., Carlsson, L. M., Kiess, W., Vatin, V., Lecoeur, C. et al. (2007). Variation in FTO contributes to childhood obesity and severe adult obesity. *Nat. Genet.* **39**, 724–726.
- Dupuis, J., Langenberg, C., Prokopenko, I., Saxena, R., Soranzo, N., Jackson, A. U., Wheeler, E., Glazer, N. L., Bouatia-Naji, N., Gloyn, A. L. et al. (2010). New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat. Genet.* **42**, 105–116.
- Fischer, J., Koch, L., Emmerling, C., Vierkotten, J., Peters, T., Bruning, J. C. and Ruther, U. (2009). Inactivation of the Fto gene protects from obesity. *Nature* **458**, 894–898.
- Frayling, T. M., Timpson, N. J., Weedon, M. N., Zeggini, E., Freathy, R. M., Lindgren, C. M., Perry, J. R., Elliott, K. S., Lango, H., Rayner, N. W. et al. (2007). A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* **316**, 889–894.
- Girard, C. A., Wunderlich, F. T., Shimomura, K., Collins, S., Kaizik, S., Proks, P., Abdulkader, F., Clark, A., Ball, V., Zubcevic, L. et al. (2009). Expression of an activating mutation in the gene encoding the KATP channel subunit Kir6.2 in mouse pancreatic beta cells recapitulates neonatal diabetes. *J. Clin. Invest.* **119**, 80–90.
- Gloyn, A. L. and McCarthy, M. I. (2010). Variation across the allele frequency spectrum. *Nat. Genet.* **42**, 648–650.
- Gloyn, A. L., Weedon, M. N., Owen, K. R., Turner, M. J., Knight, B. A., Hitman, G., Walker, M., Levy, J. C., Sampson, M., Halford, S. et al. (2003). Large-scale association studies of variants in genes encoding the pancreatic beta-cell KATP channel subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) confirm that the KCNJ11 E23K variant is associated with type 2 diabetes. *Diabetes* **52**, 568–572.
- Grant, S. F., Thorleifsson, G., Reynisdottir, I., Benediktsson, R., Manolescu, A., Sainz, J., Helgason, A., Stefansson, H., Emilsson, V., Helgadóttir, A. et al. (2006). Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat. Genet.* **38**, 320–323.
- Greenfield, J. R., Miller, J. W., Keogh, J. M., Henning, E., Satterwhite, J. H., Cameron, G. S., Astruc, B., Mayer, J. P., Brage, S., See, T. C. et al. (2009). Modulation of blood pressure by central melanocortinergic pathways. *N. Engl. J. Med.* **360**, 44–52.
- Grube, A., Hultgren, B., Ryan, A., Ma, Y. H., Bauer, M. and Stewart, T. A. (1995). Transgenic knockouts reveal a critical requirement for pancreatic beta cell glucokinase in maintaining glucose homeostasis. *Cell* **83**, 69–78.



Gudmundsson, J., Sulem, P., Steinthorsdottir, V., Bergthorsson, J. T., Thorleifsson, G., Manolescu, A., Rafnar, T., Gudbjartsson, D., Agnarsson, B. A., Baker, A. et al. (2007). Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. *Nat. Genet.* **39**, 977-983.

Heid, I. M., Jackson, A. U., Randall, J. C., Winkler, T. W., Qi, L., Steinthorsdottir, V., Thorleifsson, G., Zillikens, M. C., Speliotes, E. K., Magi, R. et al. (2010). 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.

Hugill, A., Shimomura, K., Ashcroft, F. M. and Cox, R. D. (2010). A mutation in KCNJ11 causing human hyperinsulinism (Y12X) results in a glucose-intolerant phenotype in the mouse. *Diabetologia* **53**, 2352-2356.

Johansen, C. T., Wang, J., Lanktree, M. B., Cao, H., McIntyre, A. D., Ban, M. R., Martins, R. A., Kennedy, B. A., Hassell, R. G., Visser, M. E. et al. (2010). Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. *Nat. Genet.* **42**, 684-687.

Konner, A. C., Klockener, T. and Bruning, J. C. (2009). Control of energy homeostasis by insulin and leptin: targeting the arcuate nucleus and beyond. *Physiol. Behav.* **97**, 632-638.

Koster, J. C., Marshall, B. A., Ensor, N., Corbett, J. A. and Nichols, C. G. (2000). Targeted overactivity of beta cell K(ATP) channels induces profound neonatal diabetes. *Cell* **100**, 645-654.

Lemaire, K., Ravier, M. A., Schraenen, A., Creemers, J. W., Van de Plas, R., Granvik, M., Van Lommel, L., Waelkens, E., Chimienti, F., Rutter, G. A. et al. (2009). Insulin crystallization depends on zinc transporter Znt8 expression, but is not required for normal glucose homeostasis in mice. *Proc. Natl. Acad. Sci. USA* **106**, 14872-14877.

Loos, R. J., Lindgren, C. M., Li, S., Wheeler, E., Zhao, J. H., Prokopenko, I., Inouye, M., Freathy, R. M., Attwood, A. P., Beckmann, J. S. et al. (2008). Common variants near MC4R are associated with fat mass, weight and risk of obesity. *Nat. Genet.* **40**, 768-775.

Maeda, S., Kobayashi, M. A., Araki, S., Babazono, T., Freedman, B. I., Bostrom, M. A., Cooke, J. N., Toyoda, M., Umezono, T., Tarnow, L. et al. (2010). A single nucleotide polymorphism within the acetyl-coenzyme A carboxylase beta gene is associated with proteinuria in patients with type 2 diabetes. *PLoS Genet.* **6**, e1000842.

Meyre, D., Delplanque, J., Chevre, J. C., Lecoq, C., Lobbens, S., Gallina, S., Durand, E., Vatin, V., Degraeve, F., Proenca, C. et al. (2009). Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. *Nat. Genet.* **41**, 157-159.

Miki, T., Tashiro, F., Iwanaga, T., Nagashima, K., Yoshitomi, H., Aihara, H., Nitta, Y., Gono, T., Inagaki, N., Miyazaki, J. et al. (1997). Abnormalities of pancreatic islets by targeted expression of a dominant-negative KATP channel. *Proc. Natl. Acad. Sci. USA* **94**, 11969-11973.

Najafi-Shoushtari, S. H., Kristo, F., Li, Y., Shioda, T., Cohen, D. E., Gerszten, R. E. and Naar, A. M. (2010). MicroRNA-33 and the SREBP host genes cooperate to control cholesterol homeostasis. *Science* **328**, 1566-1569.

Nica, A. C., Montgomery, S. B., Dimas, A. S., Stranger, B. E., Beazley, C., Barroso, I. and Dermitzakis, E. T. (2010). Candidate causal regulatory effects by integration of expression QTLs with complex trait genetic associations. *PLoS Genet.* **6**, e1000895.

Nicolae, D. L., Gamazon, E., Zhang, W., Duan, S., Dolan, M. E. and Cox, N. J. (2010). Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet.* **6**, e1000888.

Nicolson, T. J., Bellomo, E. A., Wijesekara, N., Loder, M. K., Baldwin, J. M., Gyulkhandanyan, A. V., Koshkin, V., Tarasov, A. I., Carzaniga, R., Kronenberger, K. et al. (2009). Insulin storage and glucose homeostasis in mice null for the granule zinc transporter Znt8 and studies of the type 2 diabetes-associated variants. *Diabetes* **58**, 2070-2083.

Njolstad, P. R., Sovik, O., Cuesta-Munoz, A., Bjorkhaug, L., Massa, O., Barbetti, F., Undlien, D. E., Shiota, C., Magnuson, M. A., Molven, A. et al. (2001). Neonatal diabetes mellitus due to complete glucokinase deficiency. *N. Engl. J. Med.* **344**, 1588-1592.

Pound, L. D., Sarkar, S. A., Benninger, R. K., Wang, Y., Suwanichkul, A., Shadoan, M. K., Printz, R. L., Oeser, J. K., Lee, C. E., Piston, D. W. et al. (2009). Deletion of the mouse Slc30a8 gene encoding zinc transporter-8 results in impaired insulin secretion. *Biochem. J.* **421**, 371-376.

Prokopenko, I., Langenberg, C., Florez, J. C., Saxena, R., Soranzo, N., Thorleifsson, G., Loos, R. J., Manning, A. K., Jackson, A. U., Aulchenko, Y. et al. (2009). Variants in MTNR1B influence fasting glucose levels. *Nat. Genet.* **41**, 77-81.

Quwallid, M. M., Hugill, A., Dear, N., Vizor, L., Wells, S., Horner, E., Fuller, S., Weedon, J., McMath, H., Woodman, P. et al. (2004). A gene-driven ENU-based approach to generating an allelic series in any gene. *Mamm. Genome* **15**, 585-591.

Ragvin, A., Moro, E., Fredman, D., Navratilova, P., Drivenes, O., Engstrom, P. G., Alonso, M. E., Mustienes Ede, L., Skarmeta, J. L., Tavares, M. J. et al. (2010). Long-range gene regulation links genomic type 2 diabetes and obesity risk regions to HHEX, SOX4, and IRX3. *Proc. Natl. Acad. Sci. USA* **107**, 775-780.

Rayner, K. J., Suarez, Y., Davalos, A., Parathath, S., Fitzgerald, M. L., Tamehiro, N., Fisher, E. A., Moore, K. J. and Fernandez-Hernando, C. (2010). MIR-33 contributes to the regulation of cholesterol homeostasis. *Science* **328**, 1570-1573.

Reynisdottir, I., Thorleifsson, G., Benediktsson, R., Sigurdsson, G., Emilsson, V., Einarsson, A. S., Hjorleifsdottir, E. E., Orlygsdottir, G. T., Bjornsdottir, G. T., Saemundsdottir, J. et al. (2003). Localization of a susceptibility gene for type 2 diabetes to chromosome 5q34-q35.2. *Am. J. Hum. Genet.* **73**, 323-335.

Saxena, R., Voight, B. F., Lyssenko, V., Burt, N. P., de Bakker, P. I., Chen, H., Roix, J. J., Kathiresan, S., Hirschhorn, J. N., Daly, M. J. et al. (2007). Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* **316**, 1331-1336.

Scherag, A., Dina, C., Hinney, A., Vatin, V., Scherag, S., Vogel, C. I., Muller, T. D., Grallert, H., Wichmann, H. E., Balkau, B. et al. (2010). Two new Loci for body-weight regulation identified in a joint analysis of genome-wide association studies for early-onset extreme obesity in French and German study groups. *PLoS Genet.* **6**, e1000916.

Scott, L. J., Mohlke, K. L., Bonnycastle, L. L., Willer, C. J., Li, Y., Duren, W. L., Erdos, M. R., Stringham, H. M., Chines, P. S., Jackson, A. U. et al. (2007). A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* **316**, 1341-1345.

Scuteri, A., Sanna, S., Chen, W. M., Uda, M., Albai, G., Strait, J., Najjar, S., Nagaraja, R., Orru, M., Usala, G. et al. (2007). Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet.* **3**, e115.

Seino, S., Iwanaga, T., Nagashima, K. and Miki, T. (2000). Diverse roles of K(ATP) channels learned from Kir6.2 genetically engineered mice. *Diabetes* **49**, 311-318.

Sladek, R., Rocheleau, G., Rung, J., Dina, C., Shen, L., Serre, D., Boutin, P., Vincent, D., Belisle, A., Hadjadj, S. et al. (2007). A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* **445**, 881-885.

Speakman, J. R., Rance, K. A. and Johnstone, A. M. (2008). Polymorphisms of the FTO gene are associated with variation in energy intake, but not energy expenditure. *Obesity* **16**, 1961-1965.

Speliotes, E. K., Willer, C. J., Berndt, S. I., Monda, K. L., Thorleifsson, G., Jackson, A. U., Allen, H. L., Lindgren, C. M., Luan, J., Magi, R. et al. (2010). Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.* **42**, 937-948.

Stacey, S. N., Sulem, P., Masson, G., Gudjonsson, S. A., Thorleifsson, G., Jakobsdottir, M., Sigurdsson, A., Gudbjartsson, D. F., Sigurgeirsson, B., Benediktsson, K. R. et al. (2009). New common variants affecting susceptibility to basal cell carcinoma. *Nat. Genet.* **41**, 909-914.

Steinthorsdottir, V., Thorleifsson, G., Reynisdottir, I., Benediktsson, R., Jonsdottir, T., Walters, G. B., Styrkarsdottir, U., Gretarsdottir, S., Emilsson, V., Ghosh, S. et al. (2007). A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat. Genet.* **39**, 770-775.

Thorleifsson, G., Walters, G. B., Gudbjartsson, D. F., Steinthorsdottir, V., Sulem, P., Helgadóttir, A., Styrkarsdottir, U., Gretarsdottir, S., Thorlacius, S., Jonsdottir, I. et al. (2009). Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. *Nat. Genet.* **41**, 18-24.

Timpson, N. J., Emmett, P. M., Frayling, T. M., Rogers, I., Hattersley, A. T., McCarthy, M. I. and Davey Smith, G. (2008). The fat mass- and obesity-associated locus and dietary intake in children. *Am. J. Clin. Nutr.* **88**, 971-978.

Toye, A. M., Moir, L., Hugill, A., Bentley, L., Quarterman, J., Mijat, V., Hough, T., Goldsworthy, M., Haynes, A., Hunter, A. J. et al. (2004). A new mouse model of type 2 diabetes, produced by N-ethyl-nitrosourea mutagenesis, is the result of a missense mutation in the glucokinase gene. *Diabetes* **53**, 1577-1583.

Tung, Y. C., Ayuso, E., Shan, X., Bosch, F., O'Rahilly, S., Coll, A. P. and Yeo, G. S. (2010). Hypothalamic-specific manipulation of Fto, the ortholog of the human obesity gene FTO, affects food intake in rats. *PLoS One* **5**, e8771.

Vaisse, C., Clement, K., Guy-Grand, B. and Froguel, P. (1998). A frameshift mutation in human MC4R is associated with a dominant form of obesity. *Nat. Genet.* **20**, 113-114.

Voight, B. F., Scott, L. J., Steinthorsdottir, V., Morris, A. P., Dina, C., Welch, R. P., Zeggini, E., Huth, C., Aulchenko, Y. S., Thorleifsson, G. et al. (2010). Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat. Genet.* **42**, 579-589.

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

Wijesekara, N., Dai, F. F., Hardy, A. B., Giglou, P. R., Bhattacharjee, A., Koshkin, V., Chimienti, F., Gaisano, H. Y., Rutter, G. A. and Wheeler, M. B. (2010). Beta cell-specific Znt8 deletion in mice causes marked defects in insulin processing, crystallisation and secretion. *Diabetologia* **53**, 1656-1668.

Willer, C. J., Speliotes, E. K., Loos, R. J., Li, S., Lindgren, C. M., Heid, I. M., Berndt, S. I., Elliott, A. L., Jackson, A. U., Lamina, C. et al. (2009). Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat. Genet.* **41**, 25-34.

- Winckler, W., Weedon, M. N., Graham, R. R., McCarroll, S. A., Purcell, S., Almgren, P., Tuomi, T., Gaudet, D., Bostrom, K. B., Walker, M. et al. (2007). Evaluation of common variants in the six known maturity-onset diabetes of the young (MODY) genes for association with type 2 diabetes. *Diabetes* **56**, 685-693.
- Xavier, G. D. S., Loder, M. K., McDonald, A., Tarasov, A. I., Carzaniga, R., Kronenberger, K., Barg, S. and Rutter, G. A. (2009). TCF7L2 regulates late events in insulin secretion from pancreatic islet  $\beta$ -cells. *Diabetes* **58**, 894-905.
- Yasuda, K., Miyake, K., Horikawa, Y., Hara, K., Osawa, H., Furuta, H., Hirota, Y., Mori, H., Jonsson, A., Sato, Y. et al. (2008). Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus. *Nat. Genet.* **40**, 1092-1097.
- Yeo, G. S., Farooqi, I. S., Aminian, S., Halsall, D. J., Stanhope, R. G. and O'Rahilly, S. (1998). A frameshift mutation in MC4R associated with dominantly inherited human obesity. *Nat. Genet.* **20**, 111-112.
- Young, E. H., Wareham, N. J., Farooqi, S., Hinney, A., Hebebrand, J., Scherag, A., O'Rahilly, S., Barroso, I. and Sandhu, M. S. (2007). The V103I polymorphism of the MC4R gene and obesity: population based studies and meta-analysis of 29 563 individuals. *Int. J. Obes. (Lond.)* **31**, 1437-1441.
- Zeggini, E., Weedon, M. N., Lindgren, C. M., Frayling, T. M., Elliott, K. S., Lango, H., Timpson, N. J., Perry, J. R., Rayner, N. W., Freathy, R. M. et al. (2007). Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* **316**, 1336-1341.
- Zeggini, E., Scott, L. J., Saxena, R., Voight, B. F., Marchini, J. L., Hu, T., de Bakker, P. I., Abecasis, G. R., Almgren, P., Andersen, G. et al. (2008). Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat. Genet.* **40**, 638-645.

## Forthcoming DMM articles

For articles published online ahead of print see <http://dmm.biologists.org/content/early/recent>

Sign up for email alerts (eToC) at <http://dmm.biologists.org/cgi/alerts/etoc>

### Review Articles

- **PRIMER: *C. elegans*: model host and tool for antimicrobial drug discovery.** Olivier Zugasti and Jonathan J. Ewbank
- **AT A GLANCE: Technical approaches for mouse models of human disease.** Monica J. Justice, Linda D. Siracusa and A. Francis Stewart
- **COMMENTARY: Worms, flies and four-legged friends: the applicability of biological models to the understanding of intestinal inflammatory diseases.** Joyce Lin and David J. Hackam

### Research Articles

- **Zebrafish deficient for Muscblind-like 2 exhibit features of myotonic dystrophy.** Laura E. Machuca-Tzili, Sarah Buxton, Aaran Thorpe, Cathy M. Timson, Peter Wigmore, Pradeep K. Luther and J. David Brook
- **Physiological homology between *Drosophila melanogaster* and vertebrate cardiovascular systems.** Michael A. Choma, Melissa J. Suter, Benjamin J. Vakoc, Brett E. Bouma and Guillermo J. Tearney
- **Boc modifies the holoprosencephaly spectrum of *Cdo* mutant mice.** Wei Zhang, Mingi Hong, Gyu-un Bae, Jong-Sun Kang and Robert S. Krauss
- **Defects in hepatic Notch signaling result in disruption of the communicating intrahepatic bile duct network in mice.** Erin E. Sparks, Daniel S. Perrien, Kari A. Huppert, Todd E. Peterson and Stacey S. Huppert
- **Human cardiomyopathy mutations induce myocyte hyperplasia and activate hypertrophic pathways during cardiogenesis in zebrafish.** Jason R. Becker, Rahul C. Deo, Andreas A. Werdich, Daniela Panàková, Shannon Coy and Calum A. MacRae
- **Mammary-specific inactivation of E-cadherin and p53 impairs functional gland development and leads to pleomorphic invasive lobular carcinoma in mice.** Patrick W. B. Derksen, Tanya M. Braumuller, Eline van der Burg, Marten Hornsveld, Elly Mesman, Jelle Wesseling, Paul Krimpenfort and Jos Jonkers
- **A *Drosophila* model for fetal alcohol syndrome disorders: role for the insulin pathway.** Kimberly D. McClure, Rachael L. French and Ulrike Heberlein
- **Noonan syndrome gain-of-function mutations in *NRAS* cause zebrafish gastrulation defects.** Vincent Runtuwene, Mark van Eekelen, John Overvoorde, Holger Rehmann, Helger G. Yntema, Willy M. Nillesen, Arie van Haeringen, Ineke van der Burgt, Boudewijn Burgering and Jeroen den Hertog