

Canine genome-wide association study identifies *DENND1B* as an obesity gene in dogs and humans.

Authors: Natalie J. Wallis¹, Alyce McClellan¹, Alexander Mörseburg^{2,3}, Katherine A. Kentistou⁴, Aqfan Jamaluddin⁵, Georgina K.C. Dowsett², Ellen Schofield^{1,6}, Anna Morros-Nuevo¹, Sadia Saeed⁷, Brian Y.H. Lam^{2,3}, Natasha T. Sumanasekera¹, Justine Chan¹, Sambhavi S.Kumar¹, Rey M Zhang^{1†}, Jodie F. Wainwright^{1‡}, Marie Dittmann¹, Gabriella Lakatos^{2, §}, Kara Rainbow², David Withers², Rebecca Bounds^{2, 8}, Marcella Ma^{2, 3}, Alexander J. German⁹, Jane Ladlow⁶, David Sargan⁶, Philippe Froguel⁷, I. Sadaf Farooqi^{2, 8}, Ken K. Ong⁴, Giles S.H. Yeo^{2, 3}, John A. Tadross^{2, 10}, John R.B. Perry^{2, 4}, Caroline M. Gorvin⁵, Eleanor Raffan^{1, 2*}.

Affiliations:

¹ Department of Physiology, Development and Neuroscience, University of Cambridge; Cambridge, CB2 3EG, UK.

² Metabolic Research Laboratories, Institute of Metabolic Science, University of Cambridge; Cambridge, CB2 0QQ, UK.

³ MRC Metabolic Diseases Unit, Institute of Metabolic Science, University of Cambridge; Cambridge, CB2 0QQ, UK.

⁴ MRC Epidemiology Unit, Institute of Metabolic Science, University of Cambridge; Cambridge, CB2 0QQ, UK.

⁵ Institute of Metabolism and Systems Research, University of Birmingham; Birmingham, B15 2TT, UK and Centre of Membrane Proteins and Receptors (COMPARE), Universities of Birmingham & Nottingham, UK

⁶ Department of Veterinary Medicine, University of Cambridge; Cambridge, CB3 0ES, UK.

⁷ INSERM UMR 1283, CNRS UMR 8199, European Genomic Institute for Diabetes, Institut Pasteur de Lille, Lille, France; University of Lille, Lille University Hospital, Lille, France; and Department of Metabolism, Digestion and Reproduction, Imperial College London, London, London, SW7 2AZ, UK.

⁸ National Institute for Health and Care Research (NIHR) Cambridge Biomedical Research Centre, Hills Road, Cambridge, UK.

⁹ Institute of Life Course & Medical Sciences & School of Veterinary Science, University of Liverpool; Neston, CH64 7TE, UK.

¹⁰ Department of Histopathology and Cambridge Genomics Laboratory, Cambridge University Hospitals NHS Foundation Trust, Cambridge, CB2 0QQ, UK.

[†] Present address: Ontario Veterinary College, University of Guelph; Guelph, Canada.

[‡] Present address: Animal Health & Welfare Directorate, Ministry for Primary Industries; Wellington 6140, New Zealand.

[§] Present address: Robotics Research Group, School of Physics, Engineering, and Computer Science, University of Hertfordshire; Hatfield, AL10 9AB, UK.

*Corresponding author. Email: er311@cam.ac.uk

Abstract:

Obesity is a heritable disease, but its genetic basis is incompletely understood. Canine population history facilitates trait mapping. We performed a canine genome-wide association study for body condition score, a measure of obesity, in 241 Labrador retrievers. Using a cross-species approach, we showed canine obesity genes are also associated with rare and common forms of obesity in humans. The lead canine association was within the gene DENN domain containing 1B (*DENND1B*). Each copy of the alternate allele was associated with ~7% greater body fat. We demonstrate a role for this gene in regulating signaling and trafficking of melanocortin 4 receptor, a critical controller of energy homeostasis. Thus, canine genetics identified obesity genes and mechanisms relevant to both dogs and humans.

Main Text:

Obesity is a complex disease resulting from a multitude of biological and environmental factors and is a major threat to both human and animal health worldwide. Studies in human patients with severe, early onset obesity highlighted the critical role of hypothalamic leptin-melanocortin signaling in the central control of energy balance. This homeostatic pathway integrates peripheral signals of energy status, translating them into alterations in energy expenditure and eating behavior (1). Large-scale population genomic studies in humans have identified >1000 BMI-associated loci, but moving from genetic association to mechanistic insight has been challenging, in part because it is hard to know which small-effect and non-coding loci justify resource-intensive follow-up (2).

Dogs are a compelling model of human obesity because they develop obesity subject to similar environmental influences and, notably, offer the opportunity for genetic discovery due to their distinctive genetic structure. Ancestral dog populations were diverse and genetically heterogeneous but narrow population bottlenecks at breed formation mean modern breeds are genetically homogeneous with a relatively long-range linkage disequilibrium (LD) structure which renders even complex trait mapping remarkably tractable (3–7). This population history also resulted in a high frequency of genetic disease in some dog breeds (8, 9), including obesity (10–12). Additionally, there is greater homology between the dog and human genomes than between those of human and mouse (4). Despite these compelling reasons to use dogs for scientific discovery relevant to both canine and human biology, dogs have been underused as a model organism to date.

About 40-60% of pet dogs are overweight or obese (13, 14), predisposing them to a range of health problems (15, 16). Dogs are exposed to similar environmental risk factors for obesity as humans; most have limited exercise and easy access to food (17). Owner management of diet and exercise is important in determining obesity outcomes, as are other risk factors such as sex, gonadectomy status and age (15). Inheritance of obesity in dogs is complex and its genetic basis is poorly understood.

Only one small genome wide association study (GWAS) for obesity has been performed in dogs and it found no significant associations (18). However, a few causative mutations have been identified in candidate gene studies (16). One is a large effect mutation in the pro-opiomelanocortin gene (*POMC* p.P187fs), which is found in a quarter of pet Labrador retrievers and is associated with increased weight, adiposity and hunger, and lower energy expenditure (19, 20). Those canine studies corroborated the role of different *POMC* derived neuropeptides for activation of melanocortin 4 receptor in the hypothalamus, well recognized as a critical nexus of energy homeostasis in humans and other species (21–23). We hypothesized that other large effect genetic variants would influence obesity in the breed.

Labrador retrievers are particularly obesity-prone and tend to be highly food motivated (10, 11, 24). We studied a population of British Labrador retrievers and performed a GWAS which revealed multiple obesity-associated loci. We developed polygenic risk scores which explain previously observed obesity variation in the breed and quantify gene-environment interaction. Comparative genomics identified that canine obesity genes were also associated with human obesity. The gene most strongly associated in dogs was *DENND1B* which we studied in vitro to reveal it has a role in regulating MC4R signaling.

Results

Phenotypic characteristics and genetic data

We studied pet and working Labrador retrievers. Only adult dogs (age 1-10 years, mean 6 years) were included, free of known or suspected systemic illness and not being treated with medications likely to affect obesity status. Body fat mass was assessed using a well validated measure of adiposity, Body Condition Score (BCS) which uses a combination of haptic and visual cues to assign dogs to BCS categories 1-9 according to standardized descriptors (Fig. S1). On this scale BCS 4-5 are considered to represent optimal body fat mass and each point increase equates to ~8% increase in body fat mass with BCS 8-9 generally considered obese (25–27). To measure food motivation and owner control of diet and exercise, we used the validated, owner-reported Dog Obesity Risk Assessment questionnaire (DORA) which scores responses to a series of statements about dog behavior related to food in the home environment and owners' management of diet and exercise (Table S1 and Supplementary Methods) (24).

For the discovery GWAS, we studied 241 pet dogs, all of which lived with their owners and some of which were also used as working dogs (for example, gun-dogs). Since the amount of time 'working' was highly variable, we quantified activity levels using the DORA questionnaire, rather than by owner-reported role. Most of the dogs had undergone gonadectomy (female = 82, male = 75) but some were sexually intact (female = 26, male = 58). BCS ranged from 3 (slightly underweight) to 9 (severely overweight) with mean BCS 5.7 (SD = 1.3). Weight ranged from 17 to 59 kg (mean = 32.7kg, SD = 7.0). Further information, including summary statistics in different sub-groups of Labradors, is presented in Table S2.

Direct genotyping was performed on the CanineHD Genotyping BeadChip (Illumina) array and data were then imputed to 9.4 million single nucleotide polymorphisms (SNPs) against a reference panel of genomes from 676 dogs of 91 breeds, including 31 Labrador retrievers. For the GWAS, we retained SNPs called with 70% confidence and which were called in >95% of dogs with an allele frequency > 5% and a Hardy Weinberg equilibrium test $p > 0.001\%$. There were 4.5 million SNPs included in the GWAS.

Canine GWAS for body condition score

We performed a GWAS for BCS in 241 Labrador retriever dogs applying a linear mixed effects model (GCTA MLMA-LOCO). Regression modelling was used to identify factors significantly affecting BCS in the population which were then included as covariates for the GWAS. These included sex, neuter status, and sex:neuter status interaction term (Fig. 1) (28). Our stringent Bonferroni corrected significance threshold ($p = 8.31 \times 10^{-7}$) was surpassed by one variant within the gene *DENND1B*, rs24430444. A more lenient nominal significance threshold was determined by the point at which the observed versus expected p value diverged outside the 95% confidence interval on a quantile-quantile (QQ) plot, an approach previously applied in canine GWAS studies (29, 30) (Fig. S2). This threshold of $p = 1.54 \times 10^{-5}$ was surpassed by a further 109 SNPs (Fig. 1A). Heritability of BCS in this canine cohort, measured using GCTA-LDMS GREML analysis from GCTA, was estimated at 70% (+/-22%).

Conditional analysis identified seven independent, non-overlapping, signals surpassing the suggestive significance threshold. Haplotype mapping and LD structure was used to define regions of interest ranging from 4.5 kb – 2.2 Mb long (mean = 549 kb, median = 65 kb, Table S3), of which five contained protein coding genes (Fig. 1B-K). Further information about lead SNP at each locus is detailed in the Supplementary Text. Three regions contained just one

protein coding gene (*CSNK1A1*, *SEMA3D* and *CDH8*). At the chromosome 6 locus there were two genes (*SDK1*, *CARD11*) with the lead SNP positioned within an intron of *CARD11*. At the chromosome 7 locus there were seven genes (*NR5A2*, *PTPRC*, *ATP6V1G3*, *NEK7*, *LHX9*, *DENND1B*, *CRB1*) with the lead SNP positioned within an intron of *DENND1B*. We interrogated Labrador whole genome sequences across each locus in dogs carrying both risk and non-risk alleles to search for genetic variants which might be considered candidates for causation. Multiple non-coding variants were identified which are listed in Table S4; no protein coding mutations predicted to have a deleterious consequence were identified and the canine LD structure meant no single causative variant could be defined. None of the genes within these loci were previously well characterized as having roles in obesity, although some had epidemiological or functional data to suggest a plausible role in energy homeostasis (detailed in Table S5).

Canine obesity genes are implicated in human obesity

To test if regions and genes identified on the canine GWAS were also relevant to human obesity, we identified regions of the human genome that were syntenic to the regions of interest defined in dogs, and examined for BMI association with all annotated genes within those human regions. We hypothesized that if canine candidate genes also regulate human BMI there would be a statistically significant association implicating the gene in one or more of the following analyses: a GWAS for BMI on 806,834 participants from the GIANT study (31); an exome-wide association study (ExWAS) of rare (MAF < 0.1%), deleterious exome variants from 454,787 individuals from the UK Biobank study (UKB) (32–34); and rare variant enrichment tests in the Severe Childhood Onset Obesity Project (SCOOP-UK) (35) (n = 982), specifically testing for enrichment of very rare (MAF < 0.0026%), predicted deleterious (CADD ≥ 25) variants compared to reference exomes of similar ancestry (gnomAD v2.1.1, n = 56,885) (36). Furthermore, we investigated the Severe Obesity in Pakistani Population (SOPP) cohort which includes patients who presented with severe, early onset obesity and in whom no monogenic causes of obesity were identified with exome sequencing. Since SOPP patients have normal weight parents and come from a highly consanguineous population, they are likely enriched for homozygous carriers of as yet unknown genetic causes of obesity.

Using this approach, we identified evidence of a genetic association with human obesity for all of the five top canine loci (Fig. 2, Table S6, S7). Full details, including clinical descriptions of the patients identified, are included in the Supplementary Text but, in brief, *CARD11* was associated in the GIANT GWAS for BMI; *CSNK1A1* was enriched for rare, deleterious variants in SCOOP with variants segregating with obesity in two families; *CDH8* was enriched for rare, deleterious variants in both SCOOP and UKB; and a proband with a rare, predicted deleterious homozygous *SEMA3D* mutation was identified in SOPP. Multiple approaches showed a human *DENND1B* association, as expanded below.

DENND1B is associated with canine and human obesity

The top canine association was within the *DENND1B* gene which encodes DENN Domain Containing 1B, a guanine nucleotide exchange factor for Rab35 that binds to the adaptor protein 2 (AP2) complex, and has a critical role in clathrin-mediated endocytosis of membrane proteins (37). Each allele of the intronic 7:5004016:T>C variant in dogs conferred a 0.94 increase in BCS (Fig. 1J). This association was replicated in regression modelling of its effect in a large population of golden retrievers (n = 1793; n = 2229) for BCS (p = 0.029) and body weight (p = 0.0022).

Canine *DENND1B* has high homology with human (89.4%) and mouse (82.7%) orthologues, particularly in functionally important domains (Fig. S3). In humans, GWAS on 806,834 participants from the GIANT study (31) showed significant association with BMI within the human region syntenic to the canine association signal (Table S7). The lead signal rs6702421 (0.011 kg/m² increase per copy of the T allele, 24% frequency, $p = 9.42 \times 10^{-9}$) is intronic to *DENND1B* (Fig. 2F), while the secondary signal rs1009188 (0.012 kg/m² increase per copy of T allele, 71% frequency, $p = 7.15 \times 10^{-11}$) is further upstream. We used activity-by-contact (ABC) enhancer maps (38) to identify whether these GWAS SNP or their proxies fell within regulatory elements for any of their proximal genes (promoters/enhancer/etc.), restricting our investigations to tissues where our candidate genes were actively expressed.

For the *DENND1B* signal, rs6702421, we found that SNPs in high LD with the signal ($r^2 > 0.8$) lay within two regulatory elements identified by the ABC enhancer maps. One was the *DENND1B* promoter itself as identified using the HAP1 human cell line from ENCODE. The other was an enhancer element active in bipolar iPSC neurons from ENCODE. Colocalization analyses using expression QTL data showed alleles for decreased *DENND1B* expression in blood also associated with decreased BMI (Table S7). This corroborates the hypothesis that the BMI GWAS signal and its closely correlated SNPs alter the sequence of established enhancers of *DENND1B* and consequently the expression of *DENND1B*.

All of those data were integrated as part of the GWAS 2 Gene (G2G) pipeline (39), and further information in supplementary methods) which predicted *DENND1B* as the most likely causal gene at this locus and in the 96th centile of likely causal prioritized genes in the BMI GWAS (Table S7). Furthermore, rare damaging variants in *DENND1B* are nominally associated with BMI in UKB ($p = 0.0087$, $\beta = 0.35$ kg/m², Fig. 2B, Fig. S4).

DENND1B is co-expressed with hypothalamic receptors involved in energy homeostasis

Since *DENND1B* has a role in clathrin mediated endocytosis of signaling receptors (40), we hypothesized that variation in *DENND1B* activity would affect the internalization, cell surface expression and/or recycling of receptors involved in energy homeostasis. We focused initially on *MC4R*, mutations in which cause human obesity (23). Canine RNAseq data from BarkBase (41) confirmed *DENND1B* is expressed in the canine brain (cortex 0.26, cerebellum 0.27, and pituitary 1.01 fragments per kilobase of transcript per million read pairs, other brain regions not available) (Table S8).

To look at co-expression of *DENND1B* and *MC4R* in the hypothalamus, we interrogated HypoMap: a unified single cell gene expression atlas of the mouse hypothalamus (42), and HYPOMAP: A comprehensive spatio-cellular map of the human hypothalamus (43). We found high *DENND1B/Dennd1b* expression in hypothalamic neuronal clusters (with lower expression levels in non-neuronal cell types) in all regions of the hypothalamus, including in the paraventricular nucleus of the hypothalamus (PVH) (Fig. 3A-C, Fig. S5, S6). In mouse, *Dennd1b* was expressed in 22.2% of all *Mc4r*-expressing cells (Fig. 3B, C, Table S9). In humans, *DENND1B* was expressed in 79.9% of *MC4R*-expressing neurons and expressed in 63.3-87.5% of cells in the 5 clusters with the highest percentages of *MC4R* expression (Fig. S6, Table S9, S10). Duplex RNAscope in situ hybridization in human hypothalamic tissue sections confirmed *DENND1B/MC4R* co-expression in neurons within the PVH. (Fig. 3D).

In addition to *MC4R* we examined multiple other hypothalamic receptors with known roles in energy homeostasis. In both the murine (Fig. S5) and human (Fig. S6) hypothalamus,

there was co-expression of *DENND1B/Dennd1b* with growth hormone secretagogue receptor (*GHSR*), melanocortin 3 receptor (*MC3R/Mc3r*), Neuropeptide Y Receptors Y1 and Y5 (*NPY1R/Npy1r*, *NPY5R/Npy5r*), leptin receptor (*LEPR/Lepr*), insulin receptor (*INSR/Insr*), 5-Hydroxytryptamine Receptors 1B and 2C (*HTR1B/Htr1b*, *HTR2C/Htr2c*), and glucagon like peptide 1 receptor (*GLP1R/Glp1r*). We included *GHSR* in functional studies to examine *DENND1B* activity as its orexigenic effect contrasts with the anorexigenic effect of *MC4R*.

DENND1B* expression affects internalization and signaling of *MC4R

To test whether *DENND1B* expression affects signaling or receptor internalization of *MC4R* and *GHSR*, receptors were overexpressed in HEK293 cells and ligand-induced cAMP generation and internalization were assessed under conditions of *DENND1B* overexpression or *DENND1B* knockdown and compared to control conditions (empty vector or scrambled siRNA, respectively). Neither condition affected basal cell surface expression of *MC4R* (Fig. S7). However, overexpression of *DENND1B* increased *MC4R* internalization after ligand activation and reduced cAMP signaling (Fig. 3E, F, G). Conversely, knockdown of *DENND1B* reduced *MC4R* internalization and increased cAMP signaling although only at maximal ligand concentrations (Fig. 3E, H, Fig. S7).

In contrast, altering *DENDD1B* expression had no effect on *GHSR* cell surface expression, cAMP signaling or internalization (Fig. S8). However, *DENDD1B* overexpression did increase signaling (reduced pEC₅₀) by the canonical IP-1 pathway downstream of *GHSR* (Fig. 3I, Fig. S7C).

A human DENND1B missense variant in a morbidly obese patient affects MC4R expression

A patient with severe childhood obesity was identified in the SOPP cohort as homozygous for a *DENND1B* p.R501C (Fig. S4B). This variant is extremely rare, with only a single heterozygous carrier found in gnomAD (MAF = 6.7×10^{-6} , Table S11). No alternative genetic diagnosis for variants in established candidate obesity genes was identified by exome sequencing (44, 45). The proband presented at 2.4 years of age with body weight of 32 kg (BMI 32, BMI standard deviation score, SDS, 7.01) accompanied by hyperphagia. At 7 years of age, she weighed 63 kg (BMI 34.5, BMI SDS 4.72). Neurodevelopmental milestones were normal. Both parents were heterozygous and did not have obesity. This variant has a CADD Score of 23.9 and is predicted to affect a binding motif that interacts with AP2 (Fig. 3J), a key binding partner of *DENND1B* at the initiation of endocytosis (40).

We tested the functional effect of *DENND1B* p.R501C in vitro as above. It caused a reduction in *MC4R* protein abundance at the cell surface, compared with both the empty vector ($p = 0.006$) and wildtype *DENND1B* ($p = 0.013$, Fig. 3K). Additionally, it caused a reduction in cAMP response to ligand activation of *MC4R* compared to the empty vector although to a lesser extent than wildtype *DENND1B* ($p \leq 0.0001$, Fig. S7E).

Polygenic risk score to quantify obesity risk in dogs

Polygenic risk scores (PRS) have not previously been applied in dogs. We constructed a PRS comprising 16 SNPs weighted for GWAS effect size on BCS using the ‘clumping and thresholding’ technique (Fig. S9) (46). The PRS improved prediction of BCS and body weight in an independent set of Labrador retrievers (Fig. 4A). When we included PRS, the model predicted 11% of the variability in BCS compared to just 4.5% when PRS is not included (Table S12).

We applied the PRS to determine its utility in other breeds, to test whether it explained known obesity risk factors in the breed and to examine how genetic risk interacted with dogs' environmental exposure to food and exercise (Fig. 4B-F and Supplementary Text). The Labrador PRS retained a small but significant predictive value for BCS and body weight in a closely related breed, the golden retriever ($p = 0.0078$, $\beta = 0.0041$, $n = 1765$), in which adding PRS to the model predicted 7.4% of the variability in BCS compared to 7.1% when it was not included. PRS was not predictive in more distantly related breeds (Supplementary Text, Fig. 4A, Table S13). We also observed that in dogs with high polygenic risk, stricter owner control of diet and exercise significantly reduced BCS ($p = 0.0077$) but that it had no statistically significant impact on BCS in dogs with low polygenic risk (Fig. 4E).

Discussion

A canine GWAS for body condition score in Labrador retrievers identified multiple genes associated with human obesity. The genes have previously not been well studied for their effect on energy homeostasis because the association has not been reported or their effect size in humans is small. In dogs, large effect sizes provide orthogonal evidence these genes can strongly influence energy homeostasis and are worthy of more in-depth study.

The lead canine GWAS signal was at *DENND1B* for which we identified a role in the regulation of hypothalamic melanocortin signaling. Human genomics revealed significant associations between *DENND1B* and BMI using both common (GWAS) and rare variant (ExWAS) approaches. Furthermore, we studied the molecular consequences of a mutation implicated in causing severe, early onset obesity in a single homozygous proband. In dogs, we generated a common variant PRS which provided multiple insights into known within-breed differences in obesity susceptibility, as well as evidence of gene-environment interaction in the regulation of body fat mass.

DENND1B variants were associated with obesity in both dogs and humans. Previously, this gene has been implicated in the pathogenesis of childhood asthma and other immune disorders by modifying T cell receptor function (47). Based on its previously characterized role in clathrin-mediated endocytosis (48), we hypothesized that *DENND1B* may regulate the trafficking and consequently the signaling of MC4R and GHSR.

Our data show *DENND1B* promotes MC4R internalization and reduces cAMP mediated anorexigenic signaling downstream of the receptor, suggesting that *DENND1B* can regulate MC4R trafficking and signaling, with the proposed mechanism summarized in Fig. S10. This finding is consistent with the human genetic evidence that the protective allele at the *DENND1B* locus is associated with reduced expression of the gene. Since even minor alterations in MC4R activation have been shown to have a clinically observable effect, this would be consistent with altering obesity risk (49, 50). Functionally deleterious mutations in two other regulators of MC4R signaling cause human obesity, the chaperone protein MRAP2 and transcription factor SIM1 (51, 52).

DENND1B also caused increased IP-1 signaling by the orexigenic receptor GHSR. This remains consistent with the human genetic findings. The finding is reminiscent of how other regulatory proteins, notably MRAP2, regulate the signaling and trafficking of multiple G protein-coupled receptors (GPCRs) (53) and may suggest a role for *DENND1B* in orchestrating a wider repertoire of responses in energy homeostasis.

In a single morbidly obese human patient, we identified a deleterious homozygous *DENND1B* missense variant. Overexpression of the variant reduced cell surface expression of MC4R as compared with wild type *DENND1B*, consistent with previous findings that most obesity-associated *MC4R* mutations reduce cell surface expression (49). The variant caused a lesser reduction in cAMP accumulation after ligand activation of MC4R compared to wild type *DENND1B*. This suggests that, in common with ~25% of obesity-associated *MC4R* mutations, it may not cause obesity by impairing the canonical Gs-cAMP pathway. Instead, such *MC4R* mutations can impact receptor homodimerization, recycling or alternative signaling pathways (for example, ERK1/2 phosphorylation). This *DENND1B* variant may have similar complex effects, or effects on other GPCRs which warrant further investigation.

Our work advances the understanding of the genetics of obesity in dogs. The PRS provided a meaningful increase in predictive value of BCS in Labradors over conventional risk factors (6-7%), which was comparable to the predictive value of polygenic scores developed specifically for human BMI (54–56). Its utility was shown to be restricted to the discovery breed, which is expected given that LD structure varies across breeds (57). Even so, it is important to highlight this at a time when canine disease prediction is increasingly desired and discussed in veterinary medicine.

Notably, the *POMC* p.P187fs variant was not statistically significantly associated with BCS in the GWAS. This may be due to variant stratification within the population (it is more common in assistance dogs which were not included in the discovery GWAS but were included in the original research reports of this mutation), low allele frequency (MAF 0.14), modest effect size, and variable penetrance in the study dogs, illustrating the complex genetic architecture of canine obesity.

We showed stratification of genetic risk exists even within the breed, with previously recognized risk factors - chocolate coat color and being purpose bred for assistance work – being associated with higher polygenic risk which was reflected by varying degrees of genetic stratification. The high polygenic risk in the genetically distinct assistance dog population of Labradors is reminiscent of the high frequency of the *POMC* p.P187fs variant in the same cohort (19). This may be due to genetic drift but raises the possibility of inadvertent selection for obesity-promoting genetic variants in this population, perhaps because dogs with a high food drive are easier to train using food to positively reinforce desirable behavior, meaning they are more likely to be selected for breeding future generations of assistance dogs.

Polygenic risk was shown to be mediated in part via eating behavior in dogs, as in other species, measured as food motivation score using a validated questionnaire (24). This means dogs with higher polygenic risk were more likely to seek out food in the home environment, to ‘beg’ for food, and to eat any food on offer. Labradors with low polygenic risk tended to remain normal weight irrespective of owner control of diet and exercise, but high-risk dogs were prone to developing obesity if dog activity was limited and owners were permissive with food (for instance, by offering human food or not restraining their dogs’ intake by limiting the food available). These canine data provide a compelling illustration of gene-environment interaction and supports data from human populations that show individuals with high appetite are particularly vulnerable to developing obesity in an permissive environment and so need to exercise greater cognitive restraint to maintain a healthy body weight (58, 59).

We have identified obesity-related genes in humans by studying the canine model, with findings relevant to preventative and therapeutic interventions in both species. The discovery of

DENND1B as a regulator of MC4R activity informs our understanding of melanocortin signaling, a critical pathway in hypothalamic regulation of energy homeostasis. Importantly, our findings show that even high polygenic risk can be mitigated. These findings demonstrate the benefits of studying complex disease in non-traditional animal models such as the dog and have practical implications for improved management of canine obesity.

Materials and methods summary

The materials and methods are summarized here, and further detail is found in the supplementary materials document.

Canine and human research was approved by the relevant local ethical review committees and the appropriate consent obtained. We studied pet and working adult Labrador retriever dogs (age 1-10 years), free of systemic illness, not being treated with medications likely to affect obesity status, and which lived with their owners. Body fat mass was assessed using a well validated measure of adiposity, Body Condition Score (BCS) which uses visual and haptic descriptors to score dogs from 1-9 where 4-5 represents optimal body fat mass and 8-9 is considered obese (Fig. S1) (25–27). Food motivation and owner control of diet and exercise was determined using the validated, owner-reported Dog Obesity Risk Assessment questionnaire (DORA) which scores responses to a series of statements about dog behavior related to food, and owners' management of their dog's diet, and dogs' activity levels (Table S1 and Supplementary Methods) (24). Canine DNA samples were extracted from saliva collected using oral swabs (Performagene, DNA Genotek) or from residual EDTA blood samples left over after veterinary investigation (Qiagen, UK). Direct genotyping was performed on the CanineHD Genotyping BeadChip (Illumina) array and data were imputed to 9.4 million single nucleotide polymorphisms (SNPs) against a reference panel of genomes from 676 dogs of 91 breeds, including 31 Labrador retrievers.

We performed a GWAS for BCS in 241 Labrador retriever dogs using the 4.5 million SNPs retained after data quality control. To identify factors significantly affecting BCS in the study population we performed regression modelling, using Akaike's Information Criterion to identify the minimal model. Significant factors in the regression (sex, neuter status and a sex:neuter interaction term) were included as covariates in the GWAS which applied a linear mixed effects model (GCTA MLMA-LOCO) to identify variants associated with BCS (60). A stringent, conservative significance threshold ($p = 8.31 \times 10^{-7}$) was determined by Bonferroni correction, using the number of independent SNP in the analysis (determined by LD pruning of the data set using a cut-off of $r^2 < 0.7$ in PLINK v.1.9) (61). A more lenient nominal significance threshold was determined ($p = 1.54 \times 10^{-4.81}$) by the point at which the observed versus expected p value diverged outside the 95% confidence interval on a quantile-quantile (QQ) plot, an approach previously applied in canine GWAS studies (29, 30). Heritability of BCS was estimated using GCTA-LDMS GREML (60, 62, 63). Stepwise conditional analysis was performed to identify independent signals followed by haplotype mapping and LD structure ($r^2 \geq 0.8$ with the lead SNP) analysis to define regions of interest PLINK v.1.9 (61). Each locus was interrogated in whole genome sequences from Labrador dogs carrying both risk and non-risk alleles to search for genetic variants which might be candidates for causation.

Canine polygenic risk scores (PRS) were constructed using GWAS SNP weighted by effect size on BCS, using the 'clumping and thresholding' technique to include only independent variants from loci most strongly associated with BCS (Fig. S9) (46). A secondary test set of Labrador retrievers was genotyped by low-pass sequencing with imputation using skimSEEK™ technology (Neogen Europe Ltd). We tested whether PRS was a predictor of BCS, weight or food motivation in the test set of Labradors and dogs of other breeds (flat-coated

retrievers, pugs and golden retrievers), and if it was associated with known obesity risk factors (coat color, assistance dog status). To examine how genetic risk interacted with environmental exposure to food and exercise, we modelled the predictive effect of owner control of diet and exercise (measured using the DORA questionnaire) on BCS for dogs of contrasting PRS.

5 To determine whether loci and genes identified on the canine GWAS were also relevant to human obesity, we identified human genome loci syntenic to the regions of interest defined in dogs. To test whether genes in those regions were associated with human obesity, we examined whether there was a statistically significant association with BMI in both large population based studies and in cohorts of patients with severe, early onset obesity. We interrogated data from a human GWAS for BMI in 806,834 participants from the GIANT study (31) and used the GWAS 2 Gene (G2G) pipeline (39) to identify independent GWAS signals and predict causal genes for human GWAS associations at each locus. Additionally, in 454,787 individuals from the UK Biobank study (UKB) (32–34), we performed an ExWAS study, implementing BOLT-LMM v2.3.551(64) and using a set of dummy genotypes representing the per-gene carrier status for rare (MAF < 0.1%), deleterious exome variants.

10 Focusing on patients with severe, early onset obesity, we first analyzed data from the Severe Childhood Onset Obesity Project (SCOOP) (n = 982), a subset of the Genetics of Obesity Study (GOOS) consisting of patients who presented with severe obesity in childhood, all of UK British origin (35). Specifically, we tested for enrichment of very rare (MAF < 0.0026%), predicted deleterious (CADD ≥ 25) variants compared to reference exomes of similar ancestry (gnomAD v2.1.1, n = 56,885 (36)). We also investigated the Severe Obesity in Pakistani Population (SOPP) which is comprised of individuals for which selection criteria include having a BMI >35 or BMI SDS (standard deviation score/Z score compared to WHO global reference data (65, 66) for age) >3.5; onset of obesity prior to 5 years of age; pronounced hyperphagia; and having parents with either first- or second-degree consanguinity who are of normal weight or overweight (explicitly excluding parental obesity). We hypothesized that affected probands would be homozygous for deleterious variants in the canine genes of interest. Where variants were identified, we examined their frequency in ancestry diverse public comparator populations including gnomAD v.2.1.1, NCBI (67), TopMED (68), and NIH ClinVar (69).

20 We examined the expression of genes of interest by analyzing canine RNAseq data from BarkBase (41). To test their expression in brain regions important in energy homeostasis and to find out if they were co-expressed receptors involved in neuroendocrine control of body weight, we interrogated data from HypoMap: a unified single cell gene expression atlas of the mouse hypothalamus (42), and HYPOMAP: A comprehensive spatio-cellular map of the human hypothalamus (43). RNAscope in situ hybridization in human hypothalamic tissue sections was performed as previously described (70) to confirm co-expression of *DENND1B*, with *MC4R*.

25 To test the effect of *DENND1B* on the function of hypothalamic receptors, we performed molecular experiments in HEK293 cells, cultured as previously described (71). Specifically we tested *MC4R*, mutations in which cause obesity, and *GHSR*, a contrasting orexigenic receptor (23). Briefly, SNAP-tagged receptors (*MC4R* and *GHSR*) were transiently transfected using Lipofectamine 2000 (LifeTechnologies) in combination with *DENND1B* overexpression (WT or *DENND1B* p.R501C) or knockdown (siRNA). After forty-eight hours, endogenous surface expression of SNAP-647 (NEB) labelled receptor was determined by co-localization (JACoP) (72) with co-expressed Venus-Kras using live HILO microscopy. Subsequently, receptor translocation away from the membrane was measured after ligand activation using the same technique. Receptor surface expression was compared using a one-way ANOVA and receptor internalization using a Mann-Whitney test. Ligand-induced cAMP generation was assessed using the cAMP GloSensor assay (Promega) after co-expression with pGloSensor-22F. Ligand-

induced GHSR canonical signaling was measured using the Cisbio IP-One Gq HTRF kit (Revvity, Codolet, France). For MC4R assays, α -MSH (Bachem AG) was added (100 μ M – 10 nM). For GHSR assays, high affinity agonist MK-0677 (Tocris, Abingdon, UK) was added (10 μ M - 10nM). For all concentration-response curve assays, pEC50 values from independent experiments were grouped, normalized and compared by one-way ANOVA.

References and Notes

1. A. A. van der Klaauw, I. S. Farooqi, The hunger genes: Pathways to obesity. *Cell* **161**, 119–132 (2015). [doi:10.1016/j.cell.2015.03.008](https://doi.org/10.1016/j.cell.2015.03.008) [Medline](#)
2. R. J. F. Loos, G. S. H. Yeo, The genetics of obesity: From discovery to biology. *Nat. Rev. Genet.* **23**, 120–133 (2022). [doi:10.1038/s41576-021-00414-z](https://doi.org/10.1038/s41576-021-00414-z) [Medline](#)
3. N. B. Sutter, M. A. Eberle, H. G. Parker, B. J. Pullar, E. F. Kirkness, L. Kruglyak, E. A. Ostrander, Extensive and breed-specific linkage disequilibrium in *Canis familiaris*. *Genome Res.* **14**, 2388–2396 (2004). [doi:10.1101/gr.3147604](https://doi.org/10.1101/gr.3147604) [Medline](#)
4. K. Lindblad-Toh, C. M. Wade, T. S. Mikkelsen, E. K. Karlsson, D. B. Jaffe, M. Kamal, M. Clamp, J. L. Chang, E. J. Kulbokas 3rd, M. C. Zody, E. Mauceli, X. Xie, M. Breen, R. K. Wayne, E. A. Ostrander, C. P. Ponting, F. Galibert, D. R. Smith, P. J. DeJong, E. Kirkness, P. Alvarez, T. Biagi, W. Brockman, J. Butler, C.-W. Chin, A. Cook, J. Cuff, M. J. Daly, D. DeCaprio, S. Gnerre, M. Grabherr, M. Kellis, M. Kleber, C. Bardeleben, L. Goodstadt, A. Heger, C. Hitte, L. Kim, K.-P. Koepfli, H. G. Parker, J. P. Pollinger, S. M. J. Searle, N. B. Sutter, R. Thomas, C. Webber, Broad Sequencing Platform Members, E. S. Lander, Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* **438**, 803–819 (2005). [doi:10.1038/nature04338](https://doi.org/10.1038/nature04338) [Medline](#)
5. J. V. Halo, A. L. Pendleton, F. Shen, A. J. Doucet, T. Derrien, C. Hitte, L. E. Kirby, B. Myers, E. Sliwerska, S. Emery, J. V. Moran, A. R. Boyko, J. M. Kidd, Long-read assembly of a Great Dane genome highlights the contribution of GC-rich sequence and mobile elements to canine genomes. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2016274118 (2021). [doi:10.1073/pnas.2016274118](https://doi.org/10.1073/pnas.2016274118) [Medline](#)
6. C. Wang, O. Wallerman, M.-L. Arendt, E. Sundström, Å. Karlsson, J. Nordin, S. Mäkeläinen, G. R. Pielberg, J. Hanson, Å. Ohlsson, S. Saellström, H. Rönnerberg, I. Ljungvall, J. Häggström, T. F. Bergström, Å. Hedhammar, J. R. S. Meadows, K. Lindblad-Toh, A novel canine reference genome resolves genomic architecture and uncovers transcript complexity. *Commun. Biol.* **4**, 185 (2021). [doi:10.1038/s42003-021-01698-x](https://doi.org/10.1038/s42003-021-01698-x) [Medline](#)
7. J. R. S. Meadows, J. M. Kidd, G.-D. Wang, H. G. Parker, P. Z. Schall, M. Bianchi, M. J. Christmas, K. Bougiouri, R. M. Buckley, C. Hitte, A. K. Nguyen, C. Wang, V. Jagannathan, J. E. Niskanen, L. A. F. Frantz, M. Arumilli, S. Hundi, K. Lindblad-Toh, C. Ginja, K. K. Agustina, C. André, A. R. Boyko, B. W. Davis, M. Drögemüller, X.-Y. Feng, K. Gkagkavouzis, G. Iliopoulos, A. C. Harris, M. K. Hytönen, D. C. Kalthoff, Y.-H. Liu, P. Lymberakis, N. Poulakakis, A. E. Pires, F. Racimo, F. Ramos-Almodovar, P. Savolainen, S. Venetsani, I. Tammen, A. Triantafyllidis, B. vonHoldt, R. K. Wayne, G. Larson, F. W. Nicholas, H. Lohi, T. Leeb, Y.-P. Zhang, E. A. Ostrander, Genome sequencing of 2000 canids by the Dog10K consortium advances the understanding of demography, genome function and architecture. *Genome Biol.* **24**, 187 (2023). [doi:10.1186/s13059-023-03023-7](https://doi.org/10.1186/s13059-023-03023-7) [Medline](#)

8. E. V. Dutrow, J. A. Serpell, E. A. Ostrander, Domestic dog lineages reveal genetic drivers of behavioral diversification. *Cell* **185**, 4737–4755.e18 (2022). [doi:10.1016/j.cell.2022.11.003](https://doi.org/10.1016/j.cell.2022.11.003) [Medline](#)
- 5 9. J. Donner, J. Freyer, S. Davison, H. Anderson, M. Blades, L. Honkanen, L. Inman, C. A. Brookhart-Knox, A. Louviere, O. P. Forman, R. Chodroff Foran, Genetic prevalence and clinical relevance of canine Mendelian disease variants in over one million dogs. *PLOS Genet.* **19**, e1010651 (2023). [doi:10.1371/journal.pgen.1010651](https://doi.org/10.1371/journal.pgen.1010651) [Medline](#)
- 10 10. E. M. Lund, P. J. Armstrong, C. A. Kirk, J. S. Klausner, Prevalence and risk factors for obesity in adult dogs from private US veterinary practices. *Intern. J. Appl. Res. Vet. Med.* **4**, 177–186 (2006).
- 11 11. C. Pegram, E. Raffan, E. White, A. H. Ashworth, D. C. Brodbelt, D. B. Church, D. G. O’Neill, Frequency, breed predisposition and demographic risk factors for overweight status in dogs in the UK. *J. Small Anim. Pract.* **62**, 521–530 (2021). [doi:10.1111/jsap.13325](https://doi.org/10.1111/jsap.13325) [Medline](#)
- 15 12. J. F. Summers, D. G. O’Neill, D. Church, L. Collins, D. Sargan, D. C. Brodbelt, Health-related welfare prioritisation of canine disorders using electronic health records in primary care practice in the UK. *BMC Vet. Res.* **15**, 163 (2019). [doi:10.1186/s12917-019-1902-0](https://doi.org/10.1186/s12917-019-1902-0) [Medline](#)
- 20 13. A. J. German, G. R. T. Woods, S. L. Holden, L. Brennan, C. Burke, Dangerous trends in pet obesity. *Vet. Rec.* **182**, 25 (2018). [doi:10.1136/vr.k2](https://doi.org/10.1136/vr.k2) [Medline](#)
- 14 14. E. A. Courcier, R. M. Thomson, D. J. Mellor, P. S. Yam, An epidemiological study of environmental factors associated with canine obesity. *J. Small Anim. Pract.* **51**, 362–367 (2010). [doi:10.1111/j.1748-5827.2010.00933.x](https://doi.org/10.1111/j.1748-5827.2010.00933.x) [Medline](#)
- 25 15. A. J. German, The growing problem of obesity in dogs and cats. *J. Nutr.* **136**, 1940S–1946S (2006). [doi:10.1093/jn/136.7.1940S](https://doi.org/10.1093/jn/136.7.1940S) [Medline](#)
- 16 16. N. Wallis, E. Raffan, The Genetic Basis of Obesity and Related Metabolic Diseases in Humans and Companion Animals. *Genes* **11**, 1378 (2020). [doi:10.3390/genes11111378](https://doi.org/10.3390/genes11111378) [Medline](#)
- 30 17. E. Raffan, The big problem: Battling companion animal obesity. *Vet. Rec.* **173**, 287–291 (2013). [doi:10.1136/vr.f5815](https://doi.org/10.1136/vr.f5815) [Medline](#)
- 18 18. S. Sheet, S. Krishnamoorthy, J. Cha, S. Choi, B.-H. Choi, Identification of Candidate Genes and Pathways Associated with Obesity-Related Traits in Canines via Gene-Set Enrichment and Pathway-Based GWAS Analysis. *Animals* **10**, 2071 (2020). [doi:10.3390/ani10112071](https://doi.org/10.3390/ani10112071) [Medline](#)
- 35 19. E. Raffan, R. J. Dennis, C. J. O’Donovan, J. M. Becker, R. A. Scott, S. P. Smith, D. J. Withers, C. J. Wood, E. Conci, D. N. Clements, K. M. Summers, A. J. German, C. S. Mellersh, M. L. Arendt, V. P. Iyemere, E. Withers, J. Söder, S. Wernersson, G. Andersson, K. Lindblad-Toh, G. S. H. Yeo, S. O’Rahilly, A Deletion in the Canine *POMC* Gene Is Associated with Weight and Appetite in Obesity-Prone Labrador Retriever Dogs. *Cell Metab.* **23**, 893–900 (2016). [doi:10.1016/j.cmet.2016.04.012](https://doi.org/10.1016/j.cmet.2016.04.012) [Medline](#)
- 40

20. M. T. Dittmann, G. Lakatos, J. F. Wainwright, J. Mokrosinski, E. Cross, I. S. Farooqi, N. J. Wallis, L. G. Halsey, R. Wilson, S. O’Rahilly, G. S. H. Yeo, E. Raffan, Low resting metabolic rate and increased hunger due to β -MSH and β -endorphin deletion in a canine model. *Sci. Adv.* **10**, eadj3823 (2024). [doi:10.1126/sciadv.adj3823](https://doi.org/10.1126/sciadv.adj3823) [Medline](#)
- 5 21. H. Biebermann, T. R. Castañeda, F. van Landeghem, A. von Deimling, F. Escher, G. Brabant, J. Hebebrand, A. Hinney, M. H. Tschöp, A. Grüters, H. Krude, A role for β -melanocyte-stimulating hormone in human body-weight regulation. *Cell Metab.* **3**, 141–146 (2006). [doi:10.1016/j.cmet.2006.01.007](https://doi.org/10.1016/j.cmet.2006.01.007) [Medline](#)
- 10 22. Y. S. Lee, B. G. Challis, D. A. Thompson, G. S. H. Yeo, J. M. Keogh, M. E. Madonna, V. Wraight, M. Sims, V. Vatin, D. Meyre, J. Shield, C. Burren, Z. Ibrahim, T. Cheetham, P. Swift, A. Blackwood, C.-C. C. Hung, N. J. Wareham, P. Froguel, G. L. Millhauser, S. O’Rahilly, I. S. Farooqi, A POMC variant implicates β -melanocyte-stimulating hormone in the control of human energy balance. *Cell Metab.* **3**, 135–140 (2006). [doi:10.1016/j.cmet.2006.01.006](https://doi.org/10.1016/j.cmet.2006.01.006) [Medline](#)
- 15 23. G. S. H. Yeo, D. H. M. Chao, A.-M. Siegert, Z. M. Koerperich, M. D. Ericson, S. E. Simonds, C. M. Larson, S. Luquet, I. Clarke, S. Sharma, K. Clément, M. A. Cowley, C. Haskell-Luevano, L. Van Der Ploeg, R. A. H. Adan, The melanocortin pathway and energy homeostasis: From discovery to obesity therapy. *Mol. Metab.* **48**, 101206 (2021). [doi:10.1016/j.molmet.2021.101206](https://doi.org/10.1016/j.molmet.2021.101206) [Medline](#)
- 20 24. E. Raffan, S. P. Smith, S. O’Rahilly, J. Wardle, Development, factor structure and application of the Dog Obesity Risk and Appetite (DORA) questionnaire. *PeerJ* **3**, e1278 (2015). [doi:10.7717/peerj.1278](https://doi.org/10.7717/peerj.1278) [Medline](#)
- 25 25. A. J. German, S. L. Holden, G. L. Moxham, K. L. Holmes, R. M. Hackett, J. M. Rawlings, A simple, reliable tool for owners to assess the body condition of their dog or cat. *J. Nutr.* **136**, 2031S–2033S (2006). [doi:10.1093/jn/136.7.2031S](https://doi.org/10.1093/jn/136.7.2031S) [Medline](#)
26. D. P. Laflamme, Development and validation of a body condition score system for dogs. *Canine Pract.* **22**, 10–15 (1997).
- 30 27. D. I. Mawby, J. W. Bartges, A. d’Avignon, D. P. Laflamme, T. D. Moyers, T. Cottrell, Comparison of various methods for estimating body fat in dogs. *J. Am. Anim. Hosp. Assoc.* **40**, 109–114 (2004). [doi:10.5326/0400109](https://doi.org/10.5326/0400109) [Medline](#)
28. N. J. Wallis, N. T. Sumanasekera, E. Raffan, Obesity risk factors in British Labrador retrievers: Effect of sex, neuter status, age, chocolate coat colour and food motivation. *Vet. Rec.* **194**, e3410 (2024). [doi:10.1002/vetr.3410](https://doi.org/10.1002/vetr.3410) [Medline](#)
- 35 29. I. N. Kieler, S. M. Persson, R. Hagman, V. D. Marinescu, Å. Hedhammar, E. Strandberg, K. Lindblad-Toh, M. L. Arendt, Genome wide association study in Swedish Labrador retrievers identifies genetic loci associated with hip dysplasia and body weight. *Sci. Rep.* **14**, 6090 (2024). [doi:10.1038/s41598-024-56060-y](https://doi.org/10.1038/s41598-024-56060-y) [Medline](#)
- 40 30. M. L. Arendt, M. Melin, N. Tonomura, M. Koltookian, C. Courtay-Cahen, N. Flindall, J. Bass, K. Boerkamp, K. Megquir, L. Youell, S. Murphy, C. McCarthy, C. London, G. R. Rutteman, M. Starkey, K. Lindblad-Toh, Genome-Wide Association Study of Golden Retrievers Identifies Germ-Line Risk Factors Predisposing to Mast Cell Tumours. *PLOS Genet.* **11**, e1005647 (2015). [doi:10.1371/journal.pgen.1005647](https://doi.org/10.1371/journal.pgen.1005647) [Medline](#)

31. L. Yengo, J. Sidorenko, K. E. Kemper, Z. Zheng, A. R. Wood, M. N. Weedon, T. M. Frayling, J. Hirschhorn, J. Yang, P. M. Visscher, GIANT Consortium, Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Hum. Mol. Genet.* **27**, 3641–3649 (2018). [doi:10.1093/hmg/ddy271](https://doi.org/10.1093/hmg/ddy271) [Medline](#)
32. C. Sudlow, J. Gallacher, N. Allen, V. Beral, P. Burton, J. Danesh, P. Downey, P. Elliott, J. Green, M. Landray, B. Liu, P. Matthews, G. Ong, J. Pell, A. Silman, A. Young, T. Sprosen, T. Peakman, R. Collins, UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLOS Med.* **12**, e1001779 (2015). [doi:10.1371/journal.pmed.1001779](https://doi.org/10.1371/journal.pmed.1001779) [Medline](#)
33. L. R. Kaisinger, K. A. Kentistou, S. Stankovic, E. J. Gardner, F. R. Day, Y. Zhao, A. Mörseburg, C. J. Carnie, G. Zagnoli-Vieira, F. Puddu, S. P. Jackson, S. O’Rahilly, I. S. Farooqi, L. Dearden, L. C. Pantaleão, S. E. Ozanne, K. K. Ong, J. R. B. Perry, Large-scale exome sequence analysis identifies sex- and age-specific determinants of obesity. *Cell Genom.* **3**, 100362 (2023). [doi:10.1016/j.xgen.2023.100362](https://doi.org/10.1016/j.xgen.2023.100362) [Medline](#)
34. Y. Zhao, M. Chukanova, K. A. Kentistou, Z. Fairhurst-Hunter, A. M. Siegert, R. Y. Jia, G. K. C. Dowsett, E. J. Gardner, K. Lawler, F. R. Day, L. R. Kaisinger, Y. L. Tung, B. Y. H. Lam, H. C. Chen, Q. Wang, J. Berumen-Campos, P. Kuri-Morales, R. Tapia-Conyer, J. Alegre-Diaz, I. Barroso, J. Emberson, J. M. Torres, R. Collins, D. Saleheen, K. R. Smith, D. S. Paul, F. Merkle, I. S. Farooqi, N. J. Wareham, S. Petrovski, S. O’Rahilly, K. K. Ong, G. S. H. Yeo, J. R. B. Perry, Protein-truncating variants in *BSN* are associated with severe adult-onset obesity, type 2 diabetes and fatty liver disease. *Nat. Genet.* **56**, 579–584 (2024). [doi:10.1038/s41588-024-01694-x](https://doi.org/10.1038/s41588-024-01694-x) [Medline](#)
35. E. Wheeler, N. Huang, E. G. Bochukova, J. M. Keogh, S. Lindsay, S. Garg, E. Henning, H. Blackburn, R. J. F. Loos, N. J. Wareham, S. O’Rahilly, M. E. Hurles, I. Barroso, I. S. Farooqi, Genome-wide SNP and CNV analysis identifies common and low-frequency variants associated with severe early-onset obesity. *Nat. Genet.* **45**, 513–517 (2013). [doi:10.1038/ng.2607](https://doi.org/10.1038/ng.2607) [Medline](#)
36. K. J. Karczewski, L. C. Francioli, G. Tiao, B. B. Cummings, J. Alföldi, Q. Wang, R. L. Collins, K. M. Laricchia, A. Ganna, D. P. Birnbaum, L. D. Gauthier, H. Brand, M. Solomonson, N. A. Watts, D. Rhodes, M. Singer-Berk, E. M. England, E. G. Seaby, J. A. Kosmicki, R. K. Walters, K. Tashman, Y. Farjoun, E. Banks, T. Poterba, A. Wang, C. Seed, N. Whiffin, J. X. Chong, K. E. Samocha, E. Pierce-Hoffman, Z. Zappala, A. H. O’Donnell-Luria, E. V. Minikel, B. Weisburd, M. Lek, J. S. Ware, C. Vittal, I. M. Armean, L. Bergelson, K. Cibulskis, K. M. Connolly, M. Covarrubias, S. Donnelly, S. Ferreira, S. Gabriel, J. Gentry, N. Gupta, T. Jeandet, D. Kaplan, C. Llanwarne, R. Munshi, S. Novod, N. Petrillo, D. Roazen, V. Ruano-Rubio, A. Saltzman, M. Schleicher, J. Soto, K. Tibbetts, C. Tolonen, G. Wade, M. E. Talkowski, Genome Aggregation Database Consortium, B. M. Neale, M. J. Daly, D. G. MacArthur, The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020). [doi:10.1038/s41586-020-2308-7](https://doi.org/10.1038/s41586-020-2308-7) [Medline](#)
37. S. Yoshimura, A. Gerondopoulos, A. Linford, D. J. Rigden, F. A. Barr, Family-wide characterization of the DENN domain Rab GDP-GTP exchange factors. *J. Cell Biol.* **191**, 367–381 (2010). [doi:10.1083/jcb.201008051](https://doi.org/10.1083/jcb.201008051) [Medline](#)

38. J. Nasser, D. T. Bergman, C. P. Fulco, P. Guckelberger, B. R. Doughty, T. A. Patwardhan, T. R. Jones, T. H. Nguyen, J. C. Ulirsch, F. Leek, K. Mualim, H. M. Natri, E. M. Weeks, G. Munson, M. Kane, H. Y. Kang, A. Cui, J. P. Ray, T. M. Eisenhaure, R. L. Collins, K. Dey, H. Pfister, A. L. Price, C. B. Epstein, A. Kundaje, R. J. Xavier, M. J. Daly, H. Huang, H. K. Finucane, N. Hacohen, E. S. Lander, J. M. Engreitz, Genome-wide enhancer maps link risk variants to disease genes. *Nature* **593**, 238–243 (2021). [doi:10.1038/s41586-021-03446-x](https://doi.org/10.1038/s41586-021-03446-x) [Medline](#)
39. K. A. Kentistou, L. R. Kaisinger, S. Stankovic, M. Vaudel, E. Mendes de Oliveira, A. Messina, R. G. Walters, X. Liu, A. S. Busch, H. Helgason, D. J. Thompson, F. Santoni, K. M. Petricek, Y. Zouaghi, I. Huang-Doran, D. F. Gudbjartsson, E. Bratland, K. Lin, E. J. Gardner, Y. Zhao, R. Y. Jia, C. Terao, M. J. Riggan, M. K. Bolla, M. Yazdanpanah, N. Yazdanpanah, J. P. Bradfield, L. Broer, A. Campbell, D. I. Chasman, D. L. Cousminer, N. Franceschini, L. H. Franke, G. Girotto, C. He, M.-R. Jarvelin, P. K. Joshi, Y. Kamatani, R. Karlsson, J. Luan, K. L. Lunetta, R. Mägi, M. Mangino, S. E. Medland, C. Meisinger, R. Noordam, T. Nutile, M. P. Concas, O. Polašek, E. Porcu, S. M. Ring, C. Sala, A. V. Smith, T. Tanaka, P. J. van der Most, V. Vitart, C. A. Wang, G. Willemsen, M. Zygmont, T. U. Ahearn, I. L. Andrulis, H. Anton-Culver, A. C. Antoniou, P. L. Auer, C. L. K. Barnes, M. W. Beckmann, A. Berrington de Gonzalez, N. V. Bogdanova, S. E. Bojesen, H. Brenner, J. E. Buring, F. Canzian, J. Chang-Claude, F. J. Couch, A. Cox, L. Crisponi, K. Czene, M. B. Daly, E. W. Demerath, J. Dennis, P. Devilee, I. De Vivo, T. Dörk, A. M. Dunning, M. Dwek, J. G. Eriksson, P. A. Fasching, L. Fernandez-Rhodes, L. Ferrel, O. Fletcher, M. Gago-Dominguez, M. García-Closas, J. A. García-Sáenz, A. González-Neira, H. Grallert, P. Guénel, C. A. Haiman, P. Hall, U. Hamann, H. Hakonarson, R. J. Hart, M. Hickey, M. J. Hooning, R. Hoppe, J. L. Hopper, J.-J. Hottenga, F. B. Hu, H. Huebner, D. J. Hunter, ABCTB Investigators, H. Jernström, E. M. John, D. Karasik, E. K. Khusnutdinova, V. N. Kristensen, J. V. Lacey, D. Lambrechts, L. J. Launer, P. A. Lind, A. Lindblom, P. K. E. Magnusson, A. Mannermaa, M. I. McCarthy, T. Meitinger, C. Menni, K. Michailidou, I. Y. Millwood, R. L. Milne, G. W. Montgomery, H. Nevanlinna, I. M. Nolte, D. R. Nyholt, N. Obi, K. M. O'Brien, K. Offit, A. J. Oldehinkel, S. R. Ostrowski, A. Palotie, O. B. Pedersen, A. Peters, G. Pianigiani, D. Plaseska-Karanfilska, A. Pouta, A. Pozarickij, P. Radice, G. Rennert, F. R. Rosendaal, D. Ruggiero, E. Saloustros, D. P. Sandler, S. Schipf, C. O. Schmidt, M. K. Schmidt, K. Small, B. Spedicati, M. Stampfer, J. Stone, R. M. Tamimi, L. R. Teras, E. Tikkanen, C. Turman, C. M. Vachon, Q. Wang, R. Winqvist, A. Wolk, B. S. Zemel, W. Zheng, K. W. van Dijk, B. Z. Alizadeh, S. Bandinelli, E. Boerwinkle, D. I. Boomsma, M. Ciullo, G. Chenevix-Trench, F. Cucca, T. Esko, C. Gieger, S. F. A. Grant, V. Gudnason, C. Hayward, I. Kolčić, P. Kraft, D. A. Lawlor, N. G. Martin, E. A. Nøhr, N. L. Pedersen, C. E. Pennell, P. M. Ridker, A. Robino, H. Snieder, U. Sovio, T. D. Spector, D. Stöckl, C. Sudlow, N. J. Timpson, D. Toniolo, A. Uitterlinden, S. Ulivi, H. Völzke, N. J. Wareham, E. Widen, J. F. Wilson, Lifelines Cohort Study, Danish Blood Donor Study, Ovarian Cancer Association Consortium, Breast Cancer Association Consortium, Biobank Japan Project, China Kadoorie Biobank Collaborative Group, P. D. P. Pharoah, L. Li, D. F. Easton, P. R. Njølstad, P. Sulem, J. M. Murabito, A. Murray, D. Manousaki, A. Juul, C. Erikstrup, K. Stefansson, M. Horikoshi, Z. Chen, I. S. Farooqi, N. Pitteloud, S. Johansson, F. R. Day, J. R. B. Perry, K. K. Ong, Understanding the genetic complexity of puberty timing across the allele frequency spectrum. *Nat. Genet.* **56**, 1397–1411 (2024). [doi:10.1038/s41588-024-01798-4](https://doi.org/10.1038/s41588-024-01798-4) [Medline](#)

40. A. L. Marat, P. S. McPherson, The connecdenn family, Rab35 guanine nucleotide exchange factors interfacing with the clathrin machinery. *J. Biol. Chem.* **285**, 10627–10637 (2010). [doi:10.1074/jbc.M109.050930](https://doi.org/10.1074/jbc.M109.050930) [Medline](#)
- 5 41. K. Megquier, D. P. Genereux, J. Hekman, R. Swofford, J. Turner-Maier, J. Johnson, J. Alonso, X. Li, K. Morrill, L. J. Anguish, M. Koltookian, B. Logan, C. R. Sharp, L. Ferrer, K. Lindblad-Toh, V. N. Meyers-Wallen, A. Hoffman, E. K. Karlsson, BarkBase: Epigenomic Annotation of Canine Genomes. *Genes* **10**, 433 (2019). [doi:10.3390/genes10060433](https://doi.org/10.3390/genes10060433) [Medline](#)
- 10 42. L. Steuernagel, B. Y. H. Lam, P. Klemm, G. K. C. Dowsett, C. A. Bauder, J. A. Tadross, T. S. Hitschfeld, A. Del Rio Martin, W. Chen, A. J. de Solis, H. Fenselau, P. Davidsen, I. Cimino, S. N. Kohnke, D. Rimmington, A. P. Coll, A. Beyer, G. S. H. Yeo, J. C. Brüning, HypoMap—A unified single-cell gene expression atlas of the murine hypothalamus. *Nat. Metab.* **4**, 1402–1419 (2022). [doi:10.1038/s42255-022-00657-y](https://doi.org/10.1038/s42255-022-00657-y) [Medline](#)
- 15 43. J. A. Tadross, L. Steuernagel, G. K. C. Dowsett, K. A. Kentistou, S. Lundh, M. Porniece-Kumar, P. Klemm, K. Rainbow, H. Hvid, K. Kania, J. Poley-Wolf, L. Bjerre-Knudsen, C. Pyke, J. R. B. Perry, B. Y. H. Lam, J. C. Brüning, G. S. H. Yeo, Human HYPOMAP: A comprehensive spatio-cellular map of the human hypothalamus. *bioRxiv* 2023.09.15.557967 [Preprint] (2023). [doi:10.1101/2023.09.15.557967](https://doi.org/10.1101/2023.09.15.557967)
- 20 44. S. Saeed, Q. M. Janjua, A. Haseeb, R. Khanam, E. Durand, E. Vaillant, L. Ning, A. Badreddine, L. Berberian, M. Boissel, S. Amanzougarene, M. Canouil, M. Derhourhi, A. Bonnefond, M. Arslan, P. Froguel, Rare Variant Analysis of Obesity-Associated Genes in Young Adults With Severe Obesity From a Consanguineous Population of Pakistan. *Diabetes* **71**, 694–705 (2022). [doi:10.2337/db21-0373](https://doi.org/10.2337/db21-0373) [Medline](#)
- 25 45. S. Saeed, M. Arslan, J. Manzoor, S. M. Din, Q. M. Janjua, H. Ayesha, Q.-T. Ain, L. Inam, S. Lobbens, E. Vaillant, E. Durand, M. Derhourhi, S. Amanzougarene, A. Badreddine, L. Berberian, S. Gaget, W. I. Khan, T. A. Butt, A. Bonnefond, P. Froguel, Genetic Causes of Severe Childhood Obesity: A Remarkably High Prevalence in an Inbred Population of Pakistan. *Diabetes* **69**, 1424–1438 (2020). [doi:10.2337/db19-1238](https://doi.org/10.2337/db19-1238) [Medline](#)
- 30 46. S. W. Choi, T. S.-H. Mak, P. F. O'Reilly, Tutorial: A guide to performing polygenic risk score analyses. *Nat. Protoc.* **15**, 2759–2772 (2020). [doi:10.1038/s41596-020-0353-1](https://doi.org/10.1038/s41596-020-0353-1) [Medline](#)
- 35 47. C.-W. Yang, C. D. Hojer, M. Zhou, X. Wu, A. Wuster, W. P. Lee, B. L. Yaspan, A. C. Chan, Regulation of T Cell Receptor Signaling by DENND1B in T_H2 Cells and Allergic Disease. *Cell* **164**, 141–155 (2016). [doi:10.1016/j.cell.2015.11.052](https://doi.org/10.1016/j.cell.2015.11.052) [Medline](#)
48. P. D. Allaire, A. L. Marat, C. Dall'Armi, G. Di Paolo, P. S. McPherson, B. Ritter, The Connecdenn DENN domain: A GEF for Rab35 mediating cargo-specific exit from early endosomes. *Mol. Cell* **37**, 370–382 (2010). [doi:10.1016/j.molcel.2009.12.037](https://doi.org/10.1016/j.molcel.2009.12.037) [Medline](#)
- 40 49. B. Brouwers, E. M. de Oliveira, M. Marti-Solano, F. B. F. Monteiro, S.-A. Laurin, J. M. Keogh, E. Henning, R. Bounds, C. A. Daly, S. Houston, V. Ayinampudi, N. Wasiluk, D. Clarke, B. Plouffe, M. Bouvier, M. M. Babu, I. S. Farooqi, J. Mokrosiński, Human MC4R variants affect endocytosis, trafficking and dimerization revealing multiple cellular mechanisms involved in weight regulation. *Cell Rep.* **34**, 108862 (2021). [doi:10.1016/j.celrep.2021.108862](https://doi.org/10.1016/j.celrep.2021.108862) [Medline](#)

50. G. S. H. Yeo, E. J. Lank, I. S. Farooqi, J. Keogh, B. G. Challis, S. O’Rahilly, Mutations in the human melanocortin-4 receptor gene associated with severe familial obesity disrupts receptor function through multiple molecular mechanisms. *Hum. Mol. Genet.* **12**, 561–574 (2003). [doi:10.1093/hmg/ddg057](https://doi.org/10.1093/hmg/ddg057) [Medline](#)
- 5 51. S. Ramachandrapa, A. Raimondo, A. M. G. Cali, J. M. Keogh, E. Henning, S. Saeed, A. Thompson, S. Garg, E. G. Bochukova, S. Brage, V. Trowse, E. Wheeler, A. E. Sullivan, M. Dattani, P. E. Clayton, V. Datta, J. B. Bruning, N. J. Wareham, S. O’Rahilly, D. J. Peet, I. Barroso, M. L. Whitelaw, I. S. Farooqi, Rare variants in single-minded 1 (*SIMI*) are associated with severe obesity. *J. Clin. Invest.* **123**, 3042–3050 (2013).
10 [doi:10.1172/JCI68016](https://doi.org/10.1172/JCI68016) [Medline](#)
52. M. Asai, S. Ramachandrapa, M. Joachim, Y. Shen, R. Zhang, N. Nuthalapati, V. Ramanathan, D. E. Stochlic, P. Ferket, K. Linhart, C. Ho, T. V. Novoselova, S. Garg, M. Ridderstråle, C. Marcus, J. N. Hirschhorn, J. M. Keogh, S. O’Rahilly, L. F. Chan, A. J. Clark, I. S. Farooqi, J. A. Majzoub, Loss of function of the melanocortin 2 receptor accessory protein 2 is associated with mammalian obesity. *Science* **341**, 275–278 (2013).
15 [doi:10.1126/science.1233000](https://doi.org/10.1126/science.1233000) [Medline](#)
53. M. Wang, X. Wang, B. Jiang, Y. Zhai, J. Zheng, L. Yang, X. Tai, Y. Li, S. Fu, J. Xu, X. Lei, Z. Kuang, C. Zhang, X. Bai, M. Li, T. Zan, S. Qu, Q. Li, C. Zhang, Identification of MRAP protein family as broad-spectrum GPCR modulators. *Clin. Transl. Med.* **12**, e1091 (2022). [doi:10.1002/ctm2.1091](https://doi.org/10.1002/ctm2.1091) [Medline](#)
- 20 54. A. E. Locke, B. Kahali, S. I. Berndt, A. E. Justice, T. H. Pers, F. R. Day, C. Powell, S. Vedantam, M. L. Buchkovich, J. Yang, D. C. Croteau-Chonka, T. Esko, T. Fall, T. Ferreira, S. Gustafsson, Z. Kutalik, J. Luan, R. Mägi, J. C. Randall, T. W. Winkler, A. R. Wood, T. Workalemahu, J. D. Faul, J. A. Smith, J. H. Zhao, W. Zhao, J. Chen, R. Fehrmann, Å. K. Hedman, J. Karjalainen, E. M. Schmidt, D. Absher, N. Amin, D. Anderson, M. Beekman, J. L. Bolton, J. L. Bragg-Gresham, S. Buyske, A. Demirkan, G. Deng, G. B. Ehret, B. Feenstra, M. F. Feitosa, K. Fischer, A. Goel, J. Gong, A. U. Jackson, S. Kanoni, M. E. Kleber, K. Kristiansson, U. Lim, V. Lotay, M. Mangino, I. M. Leach, C. Medina-Gomez, S. E. Medland, M. A. Nalls, C. D. Palmer, D. Pasko, S. Pechlivanis, M. J. Peters, I. Prokopenko, D. Shungin, A. Stančáková, R. J. Strawbridge, Y. J. Sung, T. Tanaka, A. Teumer, S. Trompet, S. W. van der Laan, J. van Setten, J. V. Van Vliet-Ostaptchouk, Z. Wang, L. Yengo, W. Zhang, A. Isaacs, E. Albrecht, J. Ärnlöv, G. M. Arscott, A. P. Attwood, S. Bandinelli, A. Barrett, I. N. Bas, C. Bellis, A. J. Bennett, C. Berne, R. Blagieva, M. Blüher, S. Böhringer, L. L. Bonnycastle, Y. Böttcher, H. A. Boyd, M. Bruinenberg, I. H. Caspersen, Y. I. Chen, R. Clarke, E. W. Daw, A. J. M. de Craen, G. Delgado, M. Dimitriou, A. S. F. Doney, N. Eklund, K. Estrada, E. Eury, L. Folkersen, R. M. Fraser, M. E. Garcia, F. Geller, V. Giedraitis, B. Gigante, A. S. Go, A. Golay, A. H. Goodall, S. D. Gordon, M. Gorski, H.-J. Grabe, H. Grallert, T. B. Grammer, J. Gräßler, H. Grönberg, C. J. Groves, G. Gusto, J. Haessler, P. Hall, T. Haller, G. Hallmans, C. A. Hartman, M. Hassinen, C. Hayward, N. L. Heard-Costa, Q. Helmer, C. Hengstenberg, O. Holmen, J.-J. Hottenga, A. L. James, J. M. Jeff, Å. Johansson, J. Jolley, T. Juliusdottir, L. Kinnunen, W. Koenig, M. Koskenvuo, W. Kratzer, J. Laitinen, C. Lamina, K. Leander, N. R. Lee, P. Lichtner, L. Lind, J. Lindström, K. S. Lo, S. Lobbens, R. Lorbeer, Y. Lu, F. Mach, P. K. E. Magnusson, A. Mahajan, W. L. McArdle, S. McLachlan, C. Menni, S. Merger, E. Mihailov, L. Milani, A. Moayyeri, K. L. Monda, M. A. Morcken, A. Mulas, G. Müller, M. Müller-Nurasyid, A. W. Musk, R. Nagaraja, M. M.
- 25
30
35
40
45

Nöthen, I. M. Nolte, S. Pilz, N. W. Rayner, F. Renstrom, R. Rettig, J. S. Ried, S. Ripke, N. R. Robertson, L. M. Rose, S. Sanna, H. Scharnagl, S. Scholtens, F. R. Schumacher, W. R. Scott, T. Seufferlein, J. Shi, A. V. Smith, J. Smolonska, A. V. Stanton, V. Steinthorsdottir, K. Stirrups, H. M. Stringham, J. Sundström, M. A. Swertz, A. J. Swift, A.-C. Syvänen, S.-T. Tan, B. O. Tayo, B. Thorand, G. Thorleifsson, J. P. Tyrer, H.-W. Uh, L. Vandenput, F. C. Verhulst, S. H. Vermeulen, N. Verweij, J. M. Vonk, L. L. Waite, H. R. Warren, D. Waterworth, M. N. Weedon, L. R. Wilkens, C. Willenborg, T. Wilsgaard, M. K. Wojczynski, A. Wong, A. F. Wright, Q. Zhang, LifeLines Cohort Study, E. P. Brennan, M. Choi, Z. Dastani, A. W. Drong, P. Eriksson, A. Franco-Cereceda, J. R. Gådin, A. G. Gharavi, M. E. Goddard, R. E. Handsaker, J. Huang, F. Karpe, S. Kathiresan, S. Keildson, K. Kiryluk, M. Kubo, J.-Y. Lee, L. Liang, R. P. Lifton, B. Ma, S. A. McCarroll, A. J. McKnight, J. L. Min, M. F. Moffatt, G. W. Montgomery, J. M. Murabito, G. Nicholson, D. R. Nyholt, Y. Okada, J. R. B. Perry, R. Dorajoo, E. Reinmaa, R. M. Salem, N. Sandholm, R. A. Scott, L. Stolk, A. Takahashi, T. Tanaka, F. M. van 't Hooft, A. A. E. Vinkhuyzen, H.-J. Westra, W. Zheng, K. T. Zondervan, ADIPOGen Consortium, AGEN-BMI Working Group, CARDIOGRAMplusC4D Consortium, CKDGen Consortium, GLGC, ICBP, MAGIC Investigators, MuTHER Consortium, MIGen Consortium, PAGE Consortium, ReproGen Consortium, GENIE Consortium, International Endogene Consortium, A. C. Heath, D. Arveiler, S. J. L. Bakker, J. Beilby, R. N. Bergman, J. Blangero, P. Bovet, H. Campbell, M. J. Caulfield, G. Cesana, A. Chakravarti, D. I. Chasman, P. S. Chines, F. S. Collins, D. C. Crawford, L. A. Cupples, D. Cusi, J. Danesh, U. de Faire, H. M. den Ruijter, A. F. Dominiczak, R. Erbel, J. Erdmann, J. G. Eriksson, M. Farrall, S. B. Felix, E. Ferrannini, J. Ferrières, I. Ford, N. G. Forouhi, T. Forrester, O. H. Franco, R. T. Gansevoort, P. V. Gejman, C. Gieger, O. Gottesman, V. Gudnason, U. Gyllensten, A. S. Hall, T. B. Harris, A. T. Hattersley, A. A. Hicks, L. A. Hindorf, A. D. Hingorani, A. Hofman, G. Homuth, G. K. Hovingh, S. E. Humphries, S. C. Hunt, E. Hyppönen, T. Illig, K. B. Jacobs, M.-R. Jarvelin, K.-H. Jöckel, B. Johansen, P. Jousilahti, J. W. Jukema, A. M. Jula, J. Kaprio, J. J. P. Kastelein, S. M. Keinanen-Kiukaanniemi, L. A. Kiemeny, P. Knekt, J. S. Kooner, C. Kooperberg, P. Kovacs, A. T. Kraja, M. Kumari, J. Kuusisto, T. A. Lakka, C. Langenberg, L. L. Marchand, T. Lehtimäki, V. Lyssenko, S. Männistö, A. Marette, T. C. Matisse, C. A. McKenzie, B. McKnight, F. L. Moll, A. D. Morris, A. P. Morris, J. C. Murray, M. Nelis, C. Ohlsson, A. J. Oldehinkel, K. K. Ong, P. A. F. Madden, G. Pasterkamp, J. F. Peden, A. Peters, D. S. Postma, P. P. Pramstaller, J. F. Price, L. Qi, O. T. Raitakari, T. Rankinen, D. C. Rao, T. K. Rice, P. M. Ridker, J. D. Rioux, M. D. Ritchie, I. Rudan, V. Salomaa, N. J. Samani, J. Saramies, M. A. Sarzynski, H. Schunkert, P. E. H. Schwarz, P. Sever, A. R. Shuldiner, J. Sinisalo, R. P. Stolk, K. Strauch, A. Tönjes, D.-A. Trégouët, A. Tremblay, E. Tremoli, J. Virtamo, M.-C. Vohl, U. Völker, G. Waeber, G. Willemssen, J. C. Witteman, M. C. Zillikens, L. S. Adair, P. Amouyel, F. W. Asselbergs, T. L. Assimes, M. Bochud, B. O. Boehm, E. Boerwinkle, S. R. Bornstein, E. P. Bottinger, C. Bouchard, S. Cauchi, J. C. Chambers, S. J. Chanock, R. S. Cooper, P. I. W. de Bakker, G. Dedoussis, L. Ferrucci, P. W. Franks, P. Froguel, L. C. Groop, C. A. Haiman, A. Hamsten, J. Hui, D. J. Hunter, K. Hveem, R. C. Kaplan, M. Kivimaki, D. Kuh, M. Laakso, Y. Liu, N. G. Martin, W. März, M. Melbye, A. Metspalu, S. Moebus, P. B. Munroe, I. Njølstad, B. A. Oostra, C. N. A. Palmer, N. L. Pedersen, M. Perola, L. Pérusse, U. Peters, C. Power, T. Quertermous, R. Rauramaa, F. Rivadeneira, T. E. Saaristo, D. Saleheen, N. Sattar, E. E. Schadt, D. Schlessinger, P. E. Slagboom, H. Snieder, T. D. Spector, U. Thorsteinsdottir, M. Stumvoll, J. Tuomilehto, A. G.

- Uitterlinden, M. Uusitupa, P. van der Harst, M. Walker, H. Wallaschofski, N. J. Wareham, H. Watkins, D. R. Weir, H.-E. Wichmann, J. F. Wilson, P. Zanen, I. B. Borecki, P. Deloukas, C. S. Fox, I. M. Heid, J. R. O'Connell, D. P. Strachan, K. Stefansson, C. M. van Duijn, G. R. Abecasis, L. Franke, T. M. Frayling, M. I. McCarthy, P. M. Visscher, A. Scherag, C. J. Willer, M. Boehnke, K. L. Mohlke, C. M. Lindgren, J. S. Beckmann, I. Barroso, K. E. North, E. Ingelsson, J. N. Hirschhorn, R. J. F. Loos, E. K. Speliotes, Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518**, 197–206 (2015). [doi:10.1038/nature14177](https://doi.org/10.1038/nature14177) [Medline](#)
55. F. Seyednasrollah, J. Mäkelä, N. Pitkänen, M. Juonala, N. Hutri-Kähönen, T. Lehtimäki, J. Viikari, T. Kelly, C. Li, L. Bazzano, L. L. Elo, O. T. Raitakari, Prediction of Adulthood Obesity Using Genetic and Childhood Clinical Risk Factors in the Cardiovascular Risk in Young Finns Study. *Circ. Cardiovasc. Genet.* **10**, e001554 (2017). [doi:10.1161/CIRCGENETICS.116.001554](https://doi.org/10.1161/CIRCGENETICS.116.001554) [Medline](#)
56. P. R. Jansen, N. Vos, J. van Uhm, I. A. Dekkers, R. van der Meer, M. M. A. M. Mannens, M. M. van Haelst, The utility of obesity polygenic risk scores from research to clinical practice: A review. *Obes. Rev.* **25**, e13810 (2024). [doi:10.1111/obr.13810](https://doi.org/10.1111/obr.13810) [Medline](#)
57. E. K. Karlsson, K. Lindblad-Toh, Leader of the pack: Gene mapping in dogs and other model organisms. *Nat. Rev. Genet.* **9**, 713–725 (2008). [doi:10.1038/nrg2382](https://doi.org/10.1038/nrg2382) [Medline](#)
58. B. de Lauzon-Guillain, E. A. Clifton, F. R. Day, K. Clément, S. Brage, N. G. Forouhi, S. J. Griffin, Y. A. Koudou, V. Pelloux, N. J. Wareham, M.-A. Charles, B. Heude, K. K. Ong, Mediation and modification of genetic susceptibility to obesity by eating behaviors. *Am. J. Clin. Nutr.* **106**, 996–1004 (2017). [doi:10.3945/ajcn.117.157396](https://doi.org/10.3945/ajcn.117.157396) [Medline](#)
59. C. H. Llewellyn, A. R. Kininmonth, M. Herle, Z. Nas, A. D. Smith, S. Carnell, A. Fildes, Behavioural susceptibility theory: The role of appetite in genetic susceptibility to obesity in early life. *Phil. Trans. R. Soc. B* **378**, 20220223 (2023). [doi:10.1098/rstb.2022.0223](https://doi.org/10.1098/rstb.2022.0223) [Medline](#)
60. J. Yang, S. H. Lee, M. E. Goddard, P. M. Visscher, GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011). [doi:10.1016/j.ajhg.2010.11.011](https://doi.org/10.1016/j.ajhg.2010.11.011) [Medline](#)
61. S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, P. C. Sham, PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007). [doi:10.1086/519795](https://doi.org/10.1086/519795) [Medline](#)
62. R Core Team, R: A Language and Environment for Statistical Computing, version 4.2.2 (R Foundation for Statistical Computing, 2021).
63. J. Yang, A. Bakshi, Z. Zhu, G. Hemani, A. A. E. Vinkhuyzen, S. H. Lee, M. R. Robinson, J. R. B. Perry, I. M. Nolte, J. V. van Vliet-Ostaptchouk, H. Snieder, LifeLines Cohort Study, T. Esko, L. Milani, R. Mägi, A. Metspalu, A. Hamsten, P. K. E. Magnusson, N. L. Pedersen, E. Ingelsson, N. Soranzo, M. C. Keller, N. R. Wray, M. E. Goddard, P. M. Visscher, Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat. Genet.* **47**, 1114–1120 (2015). [doi:10.1038/ng.3390](https://doi.org/10.1038/ng.3390) [Medline](#)

64. P.-R. Loh, G. Tucker, B. K. Bulik-Sullivan, B. J. Vilhjálmsson, H. K. Finucane, R. M. Salem, D. I. Chasman, P. M. Ridker, B. M. Neale, B. Berger, N. Patterson, A. L. Price, Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat. Genet.* **47**, 284–290 (2015). [doi:10.1038/ng.3190](https://doi.org/10.1038/ng.3190) [Medline](#)
- 5 65. H. Pan, T. Cole, LMSgrowth, a Microsoft Excel add-in to access growth references based on the LMS method, version 2.77 (2012); <http://www.healthforallchildren.co.uk/>.
66. WHO Multicentre Growth Reference Study Group, *WHO Child Growth Standards: Length/Height-for-Age, Weight-for-Age, Weight-for-Length, Weight-for-Height and Body Mass Index-for-Age: Methods and Development* (World Health Organization, 2006).
- 10 67. E. W. Sayers, E. E. Bolton, J. R. Brister, K. Canese, J. Chan, D. C. Comeau, R. Connor, K. Funk, C. Kelly, S. Kim, T. Madej, A. Marchler-Bauer, C. Lanczycki, S. Lathrop, Z. Lu, F. Thibaud-Nissen, T. Murphy, L. Phan, Y. Skripchenko, T. Tse, J. Wang, R. Williams, B. W. Trawick, K. D. Pruitt, S. T. Sherry, Database resources of the national center for biotechnology information. *Nucleic Acids Res.* **50**, D20–D26 (2022).
15 [doi:10.1093/nar/gkab1112](https://doi.org/10.1093/nar/gkab1112) [Medline](#)
68. D. Taliun, D. N. Harris, M. D. Kessler, J. Carlson, Z. A. Szpiech, R. Torres, S. A. G. Taliun, A. Corvelo, S. M. Gogarten, H. M. Kang, A. N. Pitsillides, J. LeFaive, S.-B. Lee, X. Tian, B. L. Browning, S. Das, A.-K. Emde, W. E. Clarke, D. P. Loesch, A. C. Shetty, T. W. Blackwell, A. V. Smith, Q. Wong, X. Liu, M. P. Conomos, D. M. Bobo, F. Aguet, C. Albert, A. Alonso, K. G. Ardlie, D. E. Arking, S. Aslibekyan, P. L. Auer, J. Barnard, R. G. Barr, L. Barwick, L. C. Becker, R. L. Beer, E. J. Benjamin, L. F. Bielak, J. Blangero, M. Boehnke, D. W. Bowden, J. A. Brody, E. G. Burchard, B. E. Cade, J. F. Casella, B. Chalazan, D. I. Chasman, Y.-D. I. Chen, M. H. Cho, S. H. Choi, M. K. Chung, C. B. Clish, A. Correa, J. E. Curran, B. Custer, D. Darbar, M. Daya, M. de Andrade, D. L. DeMeo, S. K. Dutcher, P. T. Ellinor, L. S. Emery, C. Eng, D. Fatkin, T. Fingerlin, L. Forer, M. Fornage, N. Franceschini, C. Fuchsberger, S. M. Fullerton, S. Germer, M. T. Gladwin, D. J. Gottlieb, X. Guo, M. E. Hall, J. He, N. L. Heard-Costa, S. R. Heckbert, M. R. Irvin, J. M. Johnsen, A. D. Johnson, R. Kaplan, S. L. R. Kardia, T. Kelly, S. Kelly, E. E. Kenny, D. P. Kiel, R. Klemmer, B. A. Konkle, C. Kooperberg, A. Köttgen, L. A. Lange, J. Lasky-Su, D. Levy, X. Lin, K.-H. Lin, C. Liu, R. J. F. Loos, L. Garman, R. Gerszten, S. A. Lubitz, K. L. Lunetta, A. C. Y. Mak, A. Manichaikul, A. K. Manning, R. A. Mathias, D. D. McManus, S. T. McGarvey, J. B. Meigs, D. A. Meyers, J. L. Mikulla, M. A. Minear, B. D. Mitchell, S. Mohanty, M. E. Montasser, C. Montgomery, A. C. Morrison, J. M. Murabito, A. Natale, P. Natarajan, S. C. Nelson, K. E. North, J. R. O’Connell, N. D. Palmer, N. Pankratz, G. M. Peloso, P. A. Peyser, J. Pleiness, W. S. Post, B. M. Psaty, D. C. Rao, S. Redline, A. P. Reiner, D. Roden, J. I. Rotter, I. Ruczinski, C. Sarnowski, S. Schoenherr, D. A. Schwartz, J.-S. Seo, S. Seshadri, V. A. Sheehan, W. H. Sheu, M. B. Shoemaker, N. L. Smith, J. A. Smith, N. Sotoodehnia, A. M. Stilp, W. Tang, K. D. Taylor, M. Telen, T. A. Thornton, R. P. Tracy, D. J. Van Den Berg, 40 R. S. Vasan, K. A. Viaud-Martinez, S. Vrieze, D. E. Weeks, B. S. Weir, S. T. Weiss, L.-C. Weng, C. J. Willer, Y. Zhang, X. Zhao, D. K. Arnett, A. E. Ashley-Koch, K. C. Barnes, E. Boerwinkle, S. Gabriel, R. Gibbs, K. M. Rice, S. S. Rich, E. K. Silverman, P. Qasba, W. Gan, NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium, G. J. Papanicolaou, D. A. Nickerson, S. R. Browning, M. C. Zody, S. Zöllner, J. G. Wilson, 45 L. A. Cupples, C. C. Laurie, C. E. Jaquish, R. D. Hernandez, T. D. O’Connor, G. R.

Abecasis, Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature* **590**, 290–299 (2021). [doi:10.1038/s41586-021-03205-y](https://doi.org/10.1038/s41586-021-03205-y) [Medline](#)

69. M. J. Landrum, J. M. Lee, M. Benson, G. R. Brown, C. Chao, S. Chitipiralla, B. Gu, J. Hart, D. Hoffman, W. Jang, K. Karapetyan, K. Katz, C. Liu, Z. Maddipatla, A. Malheiro, K. McDaniel, M. Ovetsky, G. Riley, G. Zhou, J. B. Holmes, B. L. Kattman, D. R. Maglott, ClinVar: Improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* **46**, D1062–D1067 (2018). [doi:10.1093/nar/gkx1153](https://doi.org/10.1093/nar/gkx1153) [Medline](#)
70. A. E. Adriaenssens, E. K. Biggs, T. Darwish, J. Tadross, T. Sukthankar, M. Girish, J. Poley-Wolf, B. Y. Lam, I. Zvetkova, W. Pan, D. Chiarugi, G. S. H. Yeo, C. Blouet, F. M. Gribble, F. Reimann, Glucose-Dependent Insulinotropic Polypeptide Receptor-Expressing Cells in the Hypothalamus Regulate Food Intake. *Cell Metab.* **30**, 987–996.e6 (2019). [doi:10.1016/j.cmet.2019.07.013](https://doi.org/10.1016/j.cmet.2019.07.013) [Medline](#)
71. H. A. Abid, A. Inoue, C. M. Gorvin, Heterogeneity of G protein activation by the calcium-sensing receptor. *J. Mol. Endocrinol.* **67**, 41–53 (2021). [doi:10.1530/JME-21-0058](https://doi.org/10.1530/JME-21-0058) [Medline](#)
72. S. Bolte, F. P. Cordelières, A guided tour into subcellular colocalization analysis in light microscopy. *J. Microsc.* **224**, 213–232 (2006). [doi:10.1111/j.1365-2818.2006.01706.x](https://doi.org/10.1111/j.1365-2818.2006.01706.x) [Medline](#)
73. N. J. Wallis, A. McClellan, A. Mörseburg, K. A. Kentistou, A. Jamaluddin, G. K. C. Dowsett, E. Schofield, A. Morros-Nuevo, S. Saeed, B. Y. H. Lam, N. T. Sumanasekera, J. Chan, S. S. Kumar, R. M. Zhang, J. F. Wainwright, M. Dittmann, G. Lakatos, K. Rainbow, D. Withers, R. Bounds, M. Ma, A. J. German, J. Ladlow, D. Sargan, P. Froguel, I. S. Farooqi, K. K. Ong, G. S. H. Yeo, J. A. Tadross, J. R. B. Perry, C. M. Gorvin, E. Raffan, Canine genome-wide association study identifies DENND1B as an obesity gene in dogs and humans, Dryad (2024). [doi:10.5061/dryad.0vt4b8h85](https://doi.org/10.5061/dryad.0vt4b8h85)
74. E. Schofield, GOdogs-Project/Imputation: Version 0.1.1, version-0.1.1, Zenodo (2024). [doi:10.5281/zenodo.14541383](https://doi.org/10.5281/zenodo.14541383)
75. J. O’Connell, D. Gurdasani, O. Delaneau, N. Pirastu, S. Ulivi, M. Cocca, M. Traglia, J. Huang, J. E. Huffman, I. Rudan, R. McQuillan, R. M. Fraser, H. Campbell, O. Polasek, G. Asiki, K. Ekoru, C. Hayward, A. F. Wright, V. Vitart, P. Navarro, J.-F. Zagury, J. F. Wilson, D. Toniolo, P. Gasparini, N. Soranzo, M. S. Sandhu, J. Marchini, A general approach for haplotype phasing across the full spectrum of relatedness. *PLOS Genet.* **10**, e1004234 (2014). [doi:10.1371/journal.pgen.1004234](https://doi.org/10.1371/journal.pgen.1004234) [Medline](#)
76. R. M. Buckley, A. C. Harris, G.-D. Wang, D. T. Whitaker, Y.-P. Zhang, E. A. Ostrander, Best practices for analyzing imputed genotypes from low-pass sequencing in dogs. *Mamm. Genome* **33**, 213–229 (2022). [doi:10.1007/s00335-021-09914-z](https://doi.org/10.1007/s00335-021-09914-z) [Medline](#)
77. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010). [doi:10.1093/bioinformatics/btq033](https://doi.org/10.1093/bioinformatics/btq033) [Medline](#)
78. P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, H. Li, Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021). [doi:10.1093/gigascience/giab008](https://doi.org/10.1093/gigascience/giab008) [Medline](#)

79. B. N. Howie, P. Donnelly, J. Marchini, A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLOS Genet.* **5**, e1000529 (2009). [doi:10.1371/journal.pgen.1000529](https://doi.org/10.1371/journal.pgen.1000529) [Medline](#)
- 5 80. J. H. Li, C. A. Mazur, T. Berisa, J. K. Pickrell, Low-pass sequencing increases the power of GWAS and decreases measurement error of polygenic risk scores compared to genotyping arrays. *Genome Res.* **31**, 529–537 (2021). [doi:10.1101/gr.266486.120](https://doi.org/10.1101/gr.266486.120) [Medline](#)
- 10 81. K. Morrill, J. Hekman, X. Li, J. McClure, B. Logan, L. Goodman, M. Gao, Y. Dong, M. Alonso, E. Carmichael, N. Snyder-Mackler, J. Alonso, H. J. Noh, J. Johnson, M. Koltookian, C. Lieu, K. Megquier, R. Swofford, J. Turner-Maier, M. E. White, Z. Weng, A. Colubri, D. P. Genereux, K. A. Lord, E. K. Karlsson, Ancestry-inclusive dog genomics challenges popular breed stereotypes. *Science* **376**, eabk0639 (2022). [doi:10.1126/science.abk0639](https://doi.org/10.1126/science.abk0639) [Medline](#)
- 15 82. J. Labadie, B. Swafford, M. DePena, K. Tietje, R. Page, J. Patterson-Kane, Cohort profile: The Golden Retriever Lifetime Study (GRLS). *PLOS ONE* **17**, e0269425 (2022). [doi:10.1371/journal.pone.0269425](https://doi.org/10.1371/journal.pone.0269425) [Medline](#)
83. X. Zhou, M. Stephens, Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **44**, 821–824 (2012). [doi:10.1038/ng.2310](https://doi.org/10.1038/ng.2310) [Medline](#)
- 20 84. F. J. Martin, M. R. Amode, A. Aneja, O. Austine-Orimoloye, A. G. Azov, I. Barnes, A. Becker, R. Bennett, A. Berry, J. Bhai, S. K. Bhurji, A. Bignell, S. Boddu, P. R. Branco Lins, L. Brooks, S. B. Ramaraju, M. Charkhchi, A. Cockburn, L. Da Rin Fiorretto, C. Davidson, K. Dodiya, S. Donaldson, B. El Houdaigui, T. El Naboulsi, R. Fatima, C. G. Giron, T. Genez, G. S. Ghattaoraya, J. G. Martinez, C. Guijarro, M. Hardy, Z. Hollis, T. Hourlier, T. Hunt, M. Kay, V. Kaykala, T. Le, D. Lemos, D. Marques-Coelho, J. C. Marugán, G. A. Merino, L. P. Mirabueno, A. Mushtaq, S. N. Hossain, D. N. Ogeh, M. P. Sakthivel, A. Parker, M. Perry, I. Piližota, I. Prosovetskaia, J. G. Pérez-Silva, A. I. A. Salam, N. Saraiva-Agostinho, H. Schuilenburg, D. Sheppard, S. Sinha, B. Sipos, W. Stark, E. Steed, R. Sukumaran, D. Sumathipala, M.-M. Suner, L. Surapaneni, K. Sutinen, M. Szpak, F. F. Tricomi, D. Urbina-Gómez, A. Veidenberg, T. A. Walsh, B. Walts, E. 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 900 905 910 915 920 925 930 935 940 945 950 955 960 965 970 975 980 985 990 995 1000 1005 1010 1015 1020 1025 1030 1035 1040 1045 1050 1055 1060 1065 1070 1075 1080 1085 1090 1095 1100 1105 1110 1115 1120 1125 1130 1135 1140 1145 1150 1155 1160 1165 1170 1175 1180 1185 1190 1195 1200 1205 1210 1215 1220 1225 1230 1235 1240 1245 1250 1255 1260 1265 1270 1275 1280 1285 1290 1295 1300 1305 1310 1315 1320 1325 1330 1335 1340 1345 1350 1355 1360 1365 1370 1375 1380 1385 1390 1395 1400 1405 1410 1415 1420 1425 1430 1435 1440 1445 1450 1455 1460 1465 1470 1475 1480 1485 1490 1495 1500 1505 1510 1515 1520 1525 1530 1535 1540 1545 1550 1555 1560 1565 1570 1575 1580 1585 1590 1595 1600 1605 1610 1615 1620 1625 1630 1635 1640 1645 1650 1655 1660 1665 1670 1675 1680 1685 1690 1695 1700 1705 1710 1715 1720 1725 1730 1735 1740 1745 1750 1755 1760 1765 1770 1775 1780 1785 1790 1795 1800 1805 1810 1815 1820 1825 1830 1835 1840 1845 1850 1855 1860 1865 1870 1875 1880 1885 1890 1895 1900 1905 1910 1915 1920 1925 1930 1935 1940 1945 1950 1955 1960 1965 1970 1975 1980 1985 1990 1995 2000 2005 2010 2015 2020 2025 2030 2035 2040 2045 2050 2055 2060 2065 2070 2075 2080 2085 2090 2095 2100 2105 2110 2115 2120 2125 2130 2135 2140 2145 2150 2155 2160 2165 2170 2175 2180 2185 2190 2195 2200 2205 2210 2215 2220 2225 2230 2235 2240 2245 2250 2255 2260 2265 2270 2275 2280 2285 2290 2295 2300 2305 2310 2315 2320 2325 2330 2335 2340 2345 2350 2355 2360 2365 2370 2375 2380 2385 2390 2395 2400 2405 2410 2415 2420 2425 2430 2435 2440 2445 2450 2455 2460 2465 2470 2475 2480 2485 2490 2495 2500 2505 2510 2515 2520 2525 2530 2535 2540 2545 2550 2555 2560 2565 2570 2575 2580 2585 2590 2595 2600 2605 2610 2615 2620 2625 2630 2635 2640 2645 2650 2655 2660 2665 2670 2675 2680 2685 2690 2695 2700 2705 2710 2715 2720 2725 2730 2735 2740 2745 2750 2755 2760 2765 2770 2775 2780 2785 2790 2795 2800 2805 2810 2815 2820 2825 2830 2835 2840 2845 2850 2855 2860 2865 2870 2875 2880 2885 2890 2895 2900 2905 2910 2915 2920 2925 2930 2935 2940 2945 2950 2955 2960 2965 2970 2975 2980 2985 2990 2995 3000 3005 3010 3015 3020 3025 3030 3035 3040 3045 3050 3055 3060 3065 3070 3075 3080 3085 3090 3095 3100 3105 3110 3115 3120 3125 3130 3135 3140 3145 3150 3155 3160 3165 3170 3175 3180 3185 3190 3195 3200 3205 3210 3215 3220 3225 3230 3235 3240 3245 3250 3255 3260 3265 3270 3275 3280 3285 3290 3295 3300 3305 3310 3315 3320 3325 3330 3335 3340 3345 3350 3355 3360 3365 3370 3375 3380 3385 3390 3395 3400 3405 3410 3415 3420 3425 3430 3435 3440 3445 3450 3455 3460 3465 3470 3475 3480 3485 3490 3495 3500 3505 3510 3515 3520 3525 3530 3535 3540 3545 3550 3555 3560 3565 3570 3575 3580 3585 3590 3595 3600 3605 3610 3615 3620 3625 3630 3635 3640 3645 3650 3655 3660 3665 3670 3675 3680 3685 3690 3695 3700 3705 3710 3715 3720 3725 3730 3735 3740 3745 3750 3755 3760 3765 3770 3775 3780 3785 3790 3795 3800 3805 3810 3815 3820 3825 3830 3835 3840 3845 3850 3855 3860 3865 3870 3875 3880 3885 3890 3895 3900 3905 3910 3915 3920 3925 3930 3935 3940 3945 3950 3955 3960 3965 3970 3975 3980 3985 3990 3995 4000 4005 4010 4015 4020 4025 4030 4035 4040 4045 4050 4055 4060 4065 4070 4075 4080 4085 4090 4095 4100 4105 4110 4115 4120 4125 4130 4135 4140 4145 4150 4155 4160 4165 4170 4175 4180 4185 4190 4195 4200 4205 4210 4215 4220 4225 4230 4235 4240 4245 4250 4255 4260 4265 4270 4275 4280 4285 4290 4295 4300 4305 4310 4315 4320 4325 4330 4335 4340 4345 4350 4355 4360 4365 4370 4375 4380 4385 4390 4395 4400 4405 4410 4415 4420 4425 4430 4435 4440 4445 4450 4455 4460 4465 4470 4475 4480 4485 4490 4495 4500 4505 4510 4515 4520 4525 4530 4535 4540 4545 4550 4555 4560 4565 4570 4575 4580 4585 4590 4595 4600 4605 4610 4615 4620 4625 4630 4635 4640 4645 4650 4655 4660 4665 4670 4675 4680 4685 4690 4695 4700 4705 4710 4715 4720 4725 4730 4735 4740 4745 4750 4755 4760 4765 4770 4775 4780 4785 4790 4795 4800 4805 4810 4815 4820 4825 4830 4835 4840 4845 4850 4855 4860 4865 4870 4875 4880 4885 4890 4895 4900 4905 4910 4915 4920 4925 4930 4935 4940 4945 4950 4955 4960 4965 4970 4975 4980 4985 4990 4995 5000 5005 5010 5015 5020 5025 5030 5035 5040 5045 5050 5055 5060 5065 5070 5075 5080 5085 5090 5095 5100 5105 5110 5115 5120 5125 5130 5135 5140 5145 5150 5155 5160 5165 5170 5175 5180 5185 5190 5195 5200 5205 5210 5215 5220 5225 5230 5235 5240 5245 5250 5255 5260 5265 5270 5275 5280 5285 5290 5295 5300 5305 5310 5315 5320 5325 5330 5335 5340 5345 5350 5355 5360 5365 5370 5375 5380 5385 5390 5395 5400 5405 5410 5415 5420 5425 5430 5435 5440 5445 5450 5455 5460 5465 5470 5475 5480 5485 5490 5495 5500 5505 5510 5515 5520 5525 5530 5535 5540 5545 5550 5555 5560 5565 5570 5575 5580 5585 5590 5595 5600 5605 5610 5615 5620 5625 5630 5635 5640 5645 5650 5655 5660 5665 5670 5675 5680 5685 5690 5695 5700 5705 5710 5715 5720 5725 5730 5735 5740 5745 5750 5755 5760 5765 5770 5775 5780 5785 5790 5795 5800 5805 5810 5815 5820 5825 5830 5835 5840 5845 5850 5855 5860 5865 5870 5875 5880 5885 5890 5895 5900 5905 5910 5915 5920 5925 5930 5935 5940 5945 5950 5955 5960 5965 5970 5975 5980 5985 5990 5995 6000 6005 6010 6015 6020 6025 6030 6035 6040 6045 6050 6055 6060 6065 6070 6075 6080 6085 6090 6095 6100 6105 6110 6115 6120 6125 6130 6135 6140 6145 6150 6155 6160 6165 6170 6175 6180 6185 6190 6195 6200 6205 6210 6215 6220 6225 6230 6235 6240 6245 6250 6255 6260 6265 6270 6275 6280 6285 6290 6295 6300 6305 6310 6315 6320 6325 6330 6335 6340 6345 6350 6355 6360 6365 6370 6375 6380 6385 6390 6395 6400 6405 6410 6415 6420 6425 6430 6435 6440 6445 6450 6455 6460 6465 6470 6475 6480 6485 6490 6495 6500 6505 6510 6515 6520 6525 6530 6535 6540 6545 6550 6555 6560 6565 6570 6575 6580 6585 6590 6595 6600 6605 6610 6615 6620 6625 6630 6635 6640 6645 6650 6655 6660 6665 6670 6675 6680 6685 6690 6695 6700 6705 6710 6715 6720 6725 6730 6735 6740 6745 6750 6755 6760 6765 6770 6775 6780 6785 6790 6795 6800 6805 6810 6815 6820 6825 6830 6835 6840 6845 6850 6855 6860 6865 6870 6875 6880 6885 6890 6895 6900 6905 6910 6915 6920 6925 6930 6935 6940 6945 6950 6955 6960 6965 6970 6975 6980 6985 6990 6995 7000 7005 7010 7015 7020 7025 7030 7035 7040 7045 7050 7055 7060 7065 7070 7075 7080 7085 7090 7095 7100 7105 7110 7115 7120 7125 7130 7135 7140 7145 7150 7155 7160 7165 7170 7175 7180 7185 7190 7195 7200 7205 7210 7215 7220 7225 7230 7235 7240 7245 7250 7255 7260 7265 7270 7275 7280 7285 7290 7295 7300 7305 7310 7315 7320 7325 7330 7335 7340 7345 7350 7355 7360 7365 7370 7375 7380 7385 7390 7395 7400 7405 7410 7415 7420 7425 7430 7435 7440 7445 7450 7455 7460 7465 7470 7475 7480 7485 7490 7495 7500 7505 7510 7515 7520 7525 7530 7535 7540 7545 7550 7555 7560 7565 7570 7575 7580 7585 7590 7595 7600 7605 7610 7615 7620 7625 7630 7635 7640 7645 7650 7655 7660 7665 7670 7675 7680 7685 7690 7695 7700 7705 7710 7715 7720 7725 7730 7735 7740 7745 7750 7755 7760 7765 7770 7775 7780 7785 7790 7795 7800 7805 7810 7815 7820 7825 7830 7835 7840 7845 7850 7855 7860 7865 7870 7875 7880 7885 7890 7895 7900 7905 7910 7915 7920 7925 7930 7935 7940 7945 7950 7955 7960 7965 7970 7975 7980 7985 7990 7995 8000 8005 8010 8015 8020 8025 8030 8035 8040 8045 8050 8055 8060 8065 8070 8075 8080 8085 8090 8095 8100 8105 8110 8115 8120 8125 8130 8135 8140 8145 8150 8155 8160 8165 8170 8175 8180 8185 8190 8195 8200 8205 8210 8215 8220 8225 8230 8235 8240 8245 8250 8255 8260 8265 8270 8275 8280 8285 8290 8295 8300 8305 8310 8315 8320 8325 8330 8335 8340 8345 8350 8355 8360 8365 8370 8375 8380 8385 8390 8395 8400 8405 8410 8415 8420 8425 8430 8435 8440 8445 8450 8455 8460 8465 8470 8475 8480 8485 8490 8495 8500 8505 8510 8515 8520 8525 8530 8535 8540 8545 8550 8555 8560 8565 8570 8575 8580 8585 8590 8595 8600 8605 8610 8615 8620 8625 8630 8635 8640 8645 8650 8655 8660 8665 8670 8675 8680 8685 8690 8695 8700 8705 8710 8715 8720 8725 8730 8735 8740 8745 8750 8755 8760 8765 8770 8775 8780 8785 8790 8795 8800 8805 8810 8815 8820 8825 8830 8835 8840 8845 8850 8855 8860 8865 8870 8875 8880 8885 8890 8895 8900 8905 8910 8915 8920 8925 8930 8935 8940 8945 8950 8955 8960 8965 8970 8975 8980 8985 8990 8995 9000 9005 9010 9015 9020 9025 9030 9035 9040 9045 9050 9055 9060 9065 9070 9075 9080 9085 9090 9095 9100 9105 9110 9115 9120 9125 9130 9135 9140 9145 9150 9155 9160 9165 9170 9175 9180 9185 9190 9195 9200 9205 9210 9215 9220 9225 9230 9235 9240 9245 9250 9255 9260 9265 9270 9275 9280 9285 9290 9295 9300 9305 9310 9315 9320 9325 9330 9335 9340 9345 9350 9355 9360 9365 9370 9375 9380 9385 9390 9395 9400 9405 9410 9415 9420 9425 9430 9435 9440 9445 9450 9455 9460 9465 9470 9475 9480 9485 9490 9495 9500 9505 9510 9515 9520 9525 9530 9535 9540 9545 9550 9555 9560 9565 9570 9575 9580 9585 9590 9595 9600 9605 9610 9615 9620 9625 9630 9635 9640 9645 9650 9655 9660 9665 9670 9675 9680 9685 9690 9695 9700 9705 9710 9715 9720 9725 9730 9735 9740 9745 9750 9755 9760 9765 9770 9775 9780 9785 9790 9795 9800 9805 9810 9815 9820 9825 9830 9835 9840 9845 9850 9855 9860 9865 9870 9875 9880 9885 9890 9895 9900 9905 9910 9915 9920 9925 9930 9935 9940 9945 9950 9955 9960 9965 9970 9975 9980 9985 9990 9995 10000 10005 10010 10015 10020 10025 10030 10035 10040 10045 10050 10055 10060 10065 10070 10075 10080 10085 10090 10095 10100 10105 10110 10115 10120 10125 10130 10135 10140 10145 10150 10155 10160 10165 10170 10175 10180 10185 10190 10195 10200 10205 10210 10215 10220 10225 10230 10235 10240 10245 10250 10255 10260 10265 10270 10275 10280 10285 10290 10295 10300 10305 10310 10315 10320 10325 10330 10335 10340 10345 10350 10355 10360 10365 10370 10375 10380 10385 10390 10395 10400 10405 10410 10415 10420 10425 10430 10435 10440 10445 10450 10455 10460 10465 10470 10475 10480 10485 10490 10495 10500 10505 10510 10515 10520 10525 10530 10535 10540 10545 10550 10555 10560 10565 10570 10575 10580 10585 10590 10595 10600 10605 10610 10615 10620 10625 10630 10635 10640 10645 10650 10655 10660 10665 10670 10675 10680 10685 10690 10695 10700 10705 10710 10715 10720 10725 10730 10735 10740 10745 10750 10755 10760 10765 10770 10775 10780 10785 10790 10795 10800 10805 10810 10815 10820 10825 10830 10835 10840 10845 10850 10855 10860 10865 10870 10875 10880 10885 10890 10895 10900 10905 10910 10915 10920 10925 10930 10935 10940 10945 10950 10955 10960 10965 10970 10975 10980 10985 10990 10995 11000 11005 11010 11015 11020 11025 11030 11035 11040 11045 11050 11055 11060 11065 11070 11075 11080 11085 11090 1109

88. W. McLaren, L. Gil, S. E. Hunt, H. S. Riat, G. R. S. Ritchie, A. Thormann, P. Flicek, F. Cunningham, The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).
[doi:10.1186/s13059-016-0974-4](https://doi.org/10.1186/s13059-016-0974-4) [Medline](#)
89. S. B. Gabriel, S. F. Schaffner, H. Nguyen, J. M. Moore, J. Roy, B. Blumenstiel, J. Higgins, M. DeFelice, A. Lochner, M. Faggart, S. N. Liu-Cordero, C. Rotimi, A. Adeyemo, R. Cooper, R. Ward, E. S. Lander, M. J. Daly, D. Altshuler, The structure of haplotype blocks in the human genome. *Science* **296**, 2225–2229 (2002).
[doi:10.1126/science.1069424](https://doi.org/10.1126/science.1069424) [Medline](#)
90. D. Szklarczyk, R. Kirsch, M. Koutrouli, K. Nastou, F. Mehryary, R. Hachilif, A. L. Gable, T. Fang, N. T. Doncheva, S. Pyysalo, P. Bork, L. J. Jensen, C. von Mering, The STRING database in 2023: Protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* **51**, D638–D646 (2023). [doi:10.1093/nar/gkac1000](https://doi.org/10.1093/nar/gkac1000) [Medline](#)
91. P. D. Thomas, D. Ebert, A. Muruganujan, T. Mushayahama, L.-P. Albou, H. Mi, PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Sci.* **31**, 8–22 (2022).
[doi:10.1002/pro.4218](https://doi.org/10.1002/pro.4218) [Medline](#)
92. T. Groza, F. L. Gomez, H. H. Mashhadi, V. Muñoz-Fuentes, O. Gunes, R. Wilson, P. Cacheiro, A. Frost, P. Keskiivali-Bond, B. Vardal, A. McCoy, T. K. Cheng, L. Santos, S. Wells, D. Smedley, A.-M. Mallon, H. Parkinson, The International Mouse Phenotyping Consortium: Comprehensive knockout phenotyping underpinning the study of human disease. *Nucleic Acids Res.* **51**, D1038–D1045 (2023). [doi:10.1093/nar/gkac972](https://doi.org/10.1093/nar/gkac972) [Medline](#)
93. J. A. Blake, R. Baldarelli, J. A. Kadin, J. E. Richardson, C. L. Smith, C. J. Bult, Mouse Genome Database Group, Mouse Genome Database (MGD): Knowledgebase for mouse-human comparative biology. *Nucleic Acids Res.* **49**, D981–D987 (2021).
[doi:10.1093/nar/gkaa1083](https://doi.org/10.1093/nar/gkaa1083) [Medline](#)
94. J. MacArthur, E. Bowler, M. Cerezo, L. Gil, P. Hall, E. Hastings, H. Junkins, A. McMahon, A. Milano, J. Morales, Z. M. Pendlington, D. Welter, T. Burdett, L. Hindorff, P. Flicek, F. Cunningham, H. Parkinson, The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.* **45**, D896–D901 (2017).
[doi:10.1093/nar/gkw1133](https://doi.org/10.1093/nar/gkw1133) [Medline](#)
95. K. Watanabe, S. Stringer, O. Frei, M. Umićević Mirkov, C. de Leeuw, T. J. C. Polderman, S. van der Sluis, O. A. Andreassen, B. M. Neale, D. Posthuma, A global overview of pleiotropy and genetic architecture in complex traits. *Nat. Genet.* **51**, 1339–1348 (2019).
[doi:10.1038/s41588-019-0481-0](https://doi.org/10.1038/s41588-019-0481-0) [Medline](#)
96. K. J. Karczewski, M. Solomonson, K. R. Chao, J. K. Goodrich, G. Tiao, W. Lu, B. M. Riley-Gillis, E. A. Tsai, H. I. Kim, X. Zheng, F. Rahimov, S. Esmaeeli, A. J. Grundstad, M. Reppell, J. Waring, H. Jacob, D. Sexton, P. G. Bronson, X. Chen, X. Hu, J. I. Goldstein, D. King, C. Vittal, T. Poterba, D. S. Palmer, C. Churchhouse, D. P. Howrigan, W. Zhou, N. A. Watts, K. Nguyen, H. Nguyen, C. Mason, C. Farnham, C. Tolonen, L. D. Gauthier, N. Gupta, D. G. MacArthur, H. L. Rehm, C. Seed, A. A. Philippakis, M. J. Daly, J. W. Davis, H. Runz, M. R. Miller, B. M. Neale, Systematic single-variant and gene-based association testing of thousands of phenotypes in 394,841 UK Biobank exomes. *Cell Genom.* **2**, 100168 (2022). [doi:10.1016/j.xgen.2022.100168](https://doi.org/10.1016/j.xgen.2022.100168) [Medline](#)

97. A. Thormann, M. Halachev, W. McLaren, D. J. Moore, V. Svinti, A. Campbell, S. M. Kerr, M. Tischkowitz, S. E. Hunt, M. G. Dunlop, M. E. Hurles, C. F. Wright, H. V. Firth, F. Cunningham, D. R. FitzPatrick, Flexible and scalable diagnostic filtering of genomic variants using G2P with Ensembl VEP. *Nat. Commun.* **10**, 2373 (2019).
5 [doi:10.1038/s41467-019-10016-3](https://doi.org/10.1038/s41467-019-10016-3) [Medline](#)
98. M. Uhlén, L. Fagerberg, B. M. Hallström, C. Lindskog, P. Oksvold, A. Mardinoglu, Å. Sivertsson, C. Kampf, E. Sjöstedt, A. Asplund, I. Olsson, K. Edlund, E. Lundberg, S. Navani, C. A.-K. Szigyaró, J. Odeberg, D. Djureinovic, J. O. Takanen, S. Hober, T. Alm, P.-H. Edqvist, H. Berling, H. Tegel, J. Mulder, J. Rockberg, P. Nilsson, J. M. Schwenk, 10 M. Hamsten, K. von Feilitzen, M. Forsberg, L. Persson, F. Johansson, M. Zwahlen, G. von Heijne, J. Nielsen, F. Pontén, Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015). [doi:10.1126/science.1260419](https://doi.org/10.1126/science.1260419) [Medline](#)
99. J. Lonsdale, J. Thomas, M. Salvatore, R. Phillips, E. Lo, S. Shad, R. Hasz, G. Walters, F. Garcia, N. Young, B. Foster, M. Moser, E. Karasik, B. Gillard, K. Ramsey, S. Sullivan, J. 15 Bridge, H. Magazine, J. Syron, J. Fleming, L. Siminoff, H. Traino, M. Mosavel, L. Barker, S. Jewell, D. Rohrer, D. Maxim, D. Filkins, P. Harbach, E. Cortadillo, B. Berghuis, L. Turner, E. Hudson, K. Feenstra, L. Sobin, J. Robb, P. Branton, G. Korzeniewski, C. Shive, D. Tabor, L. Qi, K. Groch, S. Nampally, S. Buia, A. Zimmerman, A. Smith, R. Burges, K. Robinson, K. Valentino, D. Bradbury, M. 20 Cosentino, N. Diaz-Mayoral, M. Kennedy, T. Engel, P. Williams, K. Erickson, K. Ardlie, W. Winckler, G. Getz, D. DeLuca, D. MacArthur, M. Kellis, A. Thomson, T. Young, E. Gelfand, M. Donovan, Y. Meng, G. Grant, D. Mash, Y. Marcus, M. Basile, J. Liu, J. Zhu, Z. Tu, N. J. Cox, D. L. Nicolae, E. R. Gamazon, H. K. Im, A. Konkashbaev, J. Pritchard, M. Stevens, T. Flutre, X. Wen, E. T. Dermitzakis, T. Lappalainen, R. Guigo, J. Monlong, 25 M. Sammeth, D. Koller, A. Battle, S. Mostafavi, M. McCarthy, M. Rivas, J. Maller, I. Rusyn, A. Nobel, F. Wright, A. Shabalín, M. Feolo, N. Sharopova, A. Sturcke, J. Paschal, J. M. Anderson, E. L. Wilder, L. K. Derr, E. D. Green, J. P. Struewing, G. Temple, S. Volpi, J. T. Boyer, E. J. Thomson, M. S. Guyer, C. Ng, A. Abdallah, D. Colantuoni, T. R. Insel, S. E. Koester, A. R. Little, P. K. Bender, T. Lehner, Y. Yao, C. C. Compton, J. B. 30 Vaught, S. Sawyer, N. C. Lockhart, J. Demchok, H. F. Moore, The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013). [doi:10.1038/ng.2653](https://doi.org/10.1038/ng.2653) [Medline](#)
100. R. M. Baldarelli, C. M. Smith, J. H. Finger, T. F. Hayamizu, I. J. McCright, J. Xu, D. R. Shaw, J. S. Beal, O. Blodgett, J. Campbell, L. E. Corbani, P. J. Frost, S. C. Giannatto, D. 35 B. Miers, J. A. Kadin, J. E. Richardson, M. Ringwald, The mouse Gene Expression Database (GXD): 2021 update. *Nucleic Acids Res.* **49**, D924–D931 (2021). [doi:10.1093/nar/gkaa914](https://doi.org/10.1093/nar/gkaa914) [Medline](#)
101. Y. Hao, S. Hao, E. Andersen-Nissen, W. M. Mauck 3rd, S. Zheng, A. Butler, M. J. Lee, A. J. Wilk, C. Darby, M. Zager, P. Hoffman, M. Stoeckius, E. Papalexi, E. P. Mimitou, J. 40 Jain, A. Srivastava, T. Stuart, L. M. Fleming, B. Yeung, A. J. Rogers, J. M. McElrath, C. A. Blish, R. Gottardo, P. Smibert, R. Satija, Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021). [doi:10.1016/j.cell.2021.04.048](https://doi.org/10.1016/j.cell.2021.04.048) [Medline](#)
102. W. K. Kroeze, M. F. Sassano, X.-P. Huang, K. Lansu, J. D. McCorvy, P. M. Giguère, N. Sciaky, B. L. Roth, PRESTO-Tango as an open-source resource for interrogation of the

druggable human GPCRome. *Nat. Struct. Mol. Biol.* **22**, 362–369 (2015).
[doi:10.1038/nsmb.3014](https://doi.org/10.1038/nsmb.3014) [Medline](#)

5 103. K. J. Livak, T. D. Schmittgen, Analysis of relative gene expression data using real-time
quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* **25**, 402–408 (2001).
[doi:10.1006/meth.2001.1262](https://doi.org/10.1006/meth.2001.1262) [Medline](#)

10 104. S. L. Pulit, C. Stoneman, A. P. Morris, A. R. Wood, C. A. Glastonbury, J. Tyrrell, L.
Yengo, T. Ferreira, E. Marouli, Y. Ji, J. Yang, S. Jones, R. Beaumont, D. C. Croteau-
Chonka, T. W. Winkler, GIANT Consortium, A. T. Hattersley, R. J. F. Loos, J. N.
Hirschhorn, P. M. Visscher, T. M. Frayling, H. Yaghoobkar, C. M. Lindgren, Meta-
analysis of genome-wide association studies for body fat distribution in 694 649
individuals of European ancestry. *Hum. Mol. Genet.* **28**, 166–174 (2019).
[doi:10.1093/hmg/ddy327](https://doi.org/10.1093/hmg/ddy327) [Medline](#)

15 105. S. E. Graham, S. L. Clarke, K.-H. H. Wu, S. Kanoni, G. J. M. Zajac, S. Ramdas, I. Surakka,
I. Ntalla, S. Vedantam, T. W. Winkler, A. E. Locke, E. Marouli, M. Y. Hwang, S. Han,
A. Narita, A. Choudhury, A. R. Bentley, K. Ekoru, A. Verma, B. Trivedi, H. C. Martin,
K. A. Hunt, Q. Hui, D. Klarin, X. Zhu, G. Thorleifsson, A. Helgadóttir, D. F.
Gudbjartsson, H. Holm, I. Olafsson, M. Akiyama, S. Sakaue, C. Terao, M. Kanai, W.
Zhou, B. M. Brumpton, H. Rasheed, S. E. Ruotsalainen, A. S. Havulinna, Y. Veturri, Q.
Feng, E. A. Rosenthal, T. Lingren, J. A. Pacheco, S. A. Pendergrass, J. Haessler, F.
20 Giulianini, Y. Bradford, J. E. Miller, A. Campbell, K. Lin, I. Y. Millwood, G. Hindy, A.
Rasheed, J. D. Faul, W. Zhao, D. R. Weir, C. Turman, H. Huang, M. Graff, A. Mahajan,
M. R. Brown, W. Zhang, K. Yu, E. M. Schmidt, A. Pandit, S. Gustafsson, X. Yin, J.
Luan, J.-H. Zhao, F. Matsuda, H.-M. Jang, K. Yoon, C. Medina-Gomez, A. Pitsillides, J.
J. Hottenga, G. Willemsen, A. R. Wood, Y. Ji, Z. Gao, S. Haworth, R. E. Mitchell, J. F.
25 Chai, M. Aadahl, J. Yao, A. Manichaikul, H. R. Warren, J. Ramirez, J. Bork-Jensen, L.
L. Kårhus, A. Goel, M. Sabater-Lleal, R. Noordam, C. Sidore, E. Fiorillo, A. F. McDaid,
P. Marques-Vidal, M. Wielscher, S. Trompet, N. Sattar, L. T. Møllehave, B. H. Thuesen,
M. Munz, L. Zeng, J. Huang, B. Yang, A. Poveda, A. Kurbasic, C. Lamina, L. Forer, M.
Scholz, T. E. Galesloot, J. P. Bradfield, E. W. Daw, J. M. Zmuda, J. S. Mitchell, C.
30 Fuchsberger, H. Christensen, J. A. Brody, M. F. Feitosa, M. K. Wojczynski, M. Preuss,
M. Mangino, P. Christofidou, N. Verweij, J. W. Benjamins, J. Engmann, R. L. Kember,
R. C. Slieker, K. S. Lo, N. R. Zilhao, P. Le, M. E. Kleber, G. E. Delgado, S. Huo, D. D.
Ikeda, H. Iha, J. Yang, J. Liu, H. L. Leonard, J. Marten, B. Schmidt, M. Arendt, L. J.
Smyth, M. Cañadas-Garre, C. Wang, M. Nakatochi, A. Wong, N. Hutri-Kähönen, X.
35 Sim, R. Xia, A. Huerta-Chagoya, J. C. Fernandez-Lopez, V. Lyssenko, M. Ahmed, A. U.
Jackson, N. A. Yousri, M. R. Irvin, C. Oldmeadow, H.-N. Kim, S. Ryu, P. R. H. J.
Timmers, L. Arbeeve, R. Dorajoo, L. A. Lange, X. Chai, G. Prasad, L. Lorés-Motta, M.
Pauper, J. Long, X. Li, E. Theusch, F. Takeuchi, C. N. Spracklen, A. Loukola, S.
Bollepalli, S. C. Warner, Y. X. Wang, W. B. Wei, T. Nutile, D. Ruggiero, Y. J. Sung, Y.-
40 J. Hung, S. Chen, F. Liu, J. Yang, K. A. Kentistou, M. Gorski, M. Brumat, K. Meidtner,
L. F. Bielak, J. A. Smith, P. Hebbar, A.-E. Farmaki, E. Hofer, M. Lin, C. Xue, J. Zhang,
M. P. Concas, S. Vaccargiu, P. J. van der Most, N. Pitkänen, B. E. Cade, J. Lee, S. W.
van der Laan, K. N. Chitrala, S. Weiss, M. E. Zimmermann, J. Y. Lee, H. S. Choi, M.
Nethander, S. Freitag-Wolf, L. Southam, N. W. Rayner, C. A. Wang, S.-Y. Lin, J.-S.
45 Wang, C. Couture, L.-P. Lyytikäinen, K. Nikus, G. Cuellar-Partida, H. Vestergaard, B.
Hidalgo, O. Giannakopoulou, Q. Cai, M. O. Obura, J. van Setten, X. Li, K. Schwander,

5 N. Terzikhan, J. H. Shin, R. D. Jackson, A. P. Reiner, L. W. Martin, Z. Chen, L. Li, H.
M. Highland, K. L. Young, T. Kawaguchi, J. Thiery, J. C. Bis, G. N. Nadkarni, L. J.
Launer, H. Li, M. A. Nalls, O. T. Raitakari, S. Ichihara, S. H. Wild, C. P. Nelson, H.
Campbell, S. Jäger, T. Nabika, F. Al-Mulla, H. Niinikoski, P. S. Braund, I. Kolcic, P.
10 Kovacs, T. Giardoglou, T. Katsuya, K. F. Bhatti, D. de Kleijn, G. J. de Borst, E. K. Kim,
H. H. H. Adams, M. A. Ikram, X. Zhu, F. W. Asselbergs, A. O. Kraaijeveld, J. W. J.
Beulens, X.-O. Shu, L. S. Rallidis, O. Pedersen, T. Hansen, P. Mitchell, A. W. Hewitt, M.
Kähönen, L. Pérusse, C. Bouchard, A. Tönjes, Y. I. Chen, C. E. Pennell, T. A. Mori, W.
Lieb, A. Franke, C. Ohlsson, D. Mellström, Y. S. Cho, H. Lee, J.-M. Yuan, W.-P. Koh, S.
15 Y. Rhee, J.-T. Woo, I. M. Heid, K. J. Stark, H. Völzke, G. Homuth, M. K. Evans, A. B.
Zonderman, O. Polasek, G. Pasterkamp, I. E. Hofer, S. Redline, K. Pahlala, A. J.
Oldehinkel, H. Snieder, G. Biino, R. Schmidt, H. Schmidt, Y. E. Chen, S. Bandinelli, G.
Dedoussis, T. A. Thanaraj, S. L. R. Kardina, N. Kato, M. B. Schulze, G. Girotto, B. Jung,
C. A. Böger, P. K. Joshi, D. A. Bennett, P. L. De Jager, X. Lu, V. Mamakou, M. Brown,
20 M. J. Caulfield, P. B. Munroe, X. Guo, M. Ciullo, J. B. Jonas, N. J. Samani, J. Kaprio, P.
Pajukanta, L. S. Adair, S. A. Bechayda, H. J. de Silva, A. R. Wickremasinghe, R. M.
Krauss, J.-Y. Wu, W. Zheng, A. I. den Hollander, D. Bharadwaj, A. Correa, J. G. Wilson,
L. Lind, C.-K. Heng, A. E. Nelson, Y. M. Golightly, J. F. Wilson, B. Penninx, H.-L. Kim,
J. Attia, R. J. Scott, D. C. Rao, D. K. Arnett, S. C. Hunt, M. Walker, H. A. Koistinen, G.
25 R. Chandak, C. S. Yajnik, J. M. Mercader, T. Tusié-Luna, C. A. Aguilar-Salinas, C. G.
Villalpando, L. Orozco, M. Fornage, E. S. Tai, R. M. van Dam, T. Lehtimäki, N.
Chaturvedi, M. Yokota, J. Liu, D. F. Reilly, A. J. McKnight, F. Kee, K.-H. Jöckel, M. I.
McCarthy, C. N. A. Palmer, V. Vitart, C. Hayward, E. Simonsick, C. M. van Duijn, F.
Lu, J. Qu, H. Hishigaki, X. Lin, W. März, E. J. Parra, M. Cruz, V. Gudnason, J.-C.
30 Tardif, G. Lettre, L. M. 't Hart, P. J. M. Elders, S. M. Damrauer, M. Kumari, M.
Kivimaki, P. van der Harst, T. D. Spector, R. J. F. Loos, M. A. Province, B. M. Psaty, I.
Brandslund, P. P. Pramstaller, K. Christensen, S. Ripatti, E. Widén, H. Hakonarson, S. F.
A. Grant, L. A. L. M. Kiemeny, J. de Graaf, M. Loeffler, F. Kronenberg, D. Gu, J.
Erdmann, H. Schunkert, P. W. Franks, A. Linneberg, J. W. Jukema, A. V. Khera, M.
35 Männikkö, M.-R. Jarvelin, Z. Kutalik, F. Cucca, D. O. Mook-Kanamori, K. W. van Dijk,
H. Watkins, D. P. Strachan, N. Grarup, P. Sever, N. Poulter, J. I. Rotter, T. M. Dantoft, F.
Karpe, M. J. Neville, N. J. Timpson, C.-Y. Cheng, T.-Y. Wong, C. C. Khor, C.
Sabanayagam, A. Peters, C. Gieger, A. T. Hattersley, N. L. Pedersen, P. K. E.
Magnusson, D. I. Boomsma, E. J. C. de Geus, L. A. Cupples, J. B. J. van Meurs, M.
40 Ghanbari, P. Gordon-Larsen, W. Huang, Y. J. Kim, Y. Tabara, N. J. Wareham, C.
Langenberg, E. Zeggini, J. Kuusisto, M. Laakso, E. Ingelsson, G. Abecasis, J. C.
Chambers, J. S. Kooner, P. S. de Vries, A. C. Morrison, K. E. North, M. Daviglus, P.
Kraft, N. G. Martin, J. B. Whitfield, S. Abbas, D. Saleheen, R. G. Walters, M. V.
Holmes, C. Black, B. H. Smith, A. E. Justice, A. Baras, J. E. Buring, P. M. Ridker, D. I.
45 Chasman, C. Kooperberg, W.-Q. Wei, G. P. Jarvik, B. Namjou, M. G. Hayes, M. D.
Ritchie, P. Jousilahti, V. Salomaa, K. Hveem, B. O. Åsvold, M. Kubo, Y. Kamatani, Y.
Okada, Y. Murakami, U. Thorsteinsdottir, K. Stefansson, Y.-L. Ho, J. A. Lynch, D. J.
Rader, P. S. Tsao, K.-M. Chang, K. Cho, C. J. O'Donnell, J. M. Gaziano, P. Wilson, C.
N. Rotimi, S. Hazelhurst, M. Ramsay, R. C. Trembath, D. A. van Heel, G. Tamiya, M.
50 Yamamoto, B.-J. Kim, K. L. Mohlke, T. M. Frayling, J. N. Hirschhorn, S. Kathiresan,
VA Million Veteran Program, Global Lipids Genetics Consortium, M. Boehnke, P.
Natarajan, G. M. Peloso, C. D. Brown, A. P. Morris, T. L. Assimes, P. Deloukas, Y. V.

Sun, C. J. Willer, The power of genetic diversity in genome-wide association studies of lipids. *Nature* **600**, 675–679 (2021). [doi:10.1038/s41586-021-04064-3](https://doi.org/10.1038/s41586-021-04064-3) [Medline](#)

106. C. A. de Leeuw, J. M. Mooij, T. Heskes, D. Posthuma, MAGMA: Generalized gene-set analysis of GWAS data. *PLOS Comput. Biol.* **11**, e1004219 (2015).
5 [doi:10.1371/journal.pcbi.1004219](https://doi.org/10.1371/journal.pcbi.1004219) [Medline](#)

107. Z. Zhu, F. Zhang, H. Hu, A. Bakshi, M. R. Robinson, J. E. Powell, G. W. Montgomery, M. E. Goddard, N. R. Wray, P. M. Visscher, J. Yang, Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016). [doi:10.1038/ng.3538](https://doi.org/10.1038/ng.3538) [Medline](#)

108. C. Giambartolomei, D. Vukcevic, E. E. Schadt, L. Franke, A. D. Hingorani, C. Wallace, V. Plagnol, Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLOS Genet.* **10**, e1004383 (2014).
10 [doi:10.1371/journal.pgen.1004383](https://doi.org/10.1371/journal.pgen.1004383) [Medline](#)

109. H. K. Finucane, Y. A. Reshef, V. Anttila, K. Slowikowski, A. Gusev, A. Byrnes, S. Gazal, P.-R. Loh, C. Lareau, N. Shores, G. Genovese, A. Saunders, E. Macosko, S. Pollack, Brainstorm Consortium, J. R. B. Perry, J. D. Buenrostro, B. E. Bernstein, S. Raychaudhuri, S. McCarroll, B. M. Neale, A. L. Price, Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat. Genet.* **50**, 621–629 (2018). [doi:10.1038/s41588-018-0081-4](https://doi.org/10.1038/s41588-018-0081-4) [Medline](#)

110. GTEx Consortium, K. G. Ardlie, D. S. Deluca, A. V. Segrè, T. J. Sullivan, T. R. Young, E. T. Gelfand, C. A. Trowbridge, J. B. Maller, T. Tukiainen, M. Lek, L. D. Ward, P. Kheradpour, B. Iriarte, Y. Meng, C. D. Palmer, T. Esko, W. Winckler, J. N. Hirschhorn, M. Kellis, D. G. MacArthur, G. Getz, A. A. Shabalín, G. Li, Y.-H. Zhou, A. B. Nobel, I. Rusyn, F. A. Wright, T. Lappalainen, P. G. Ferreira, H. Ongen, M. A. Rivas, A. Battle, S. Mostafavi, J. Monlong, M. Sammeth, M. Mele, F. Reverter, J. M. Goldmann, D. Koller, R. Guigó, M. I. McCarthy, E. T. Dermitzakis, E. R. Gamazon, H. K. Im, A. Konkashbaev, D. L. Nicolae, N. J. Cox, T. Flutre, X. Wen, M. Stephens, J. K. Pritchard, Z. Tu, B. Zhang, T. Huang, Q. Long, L. Lin, J. Yang, J. Zhu, J. Liu, A. Brown, B. Mestichelli, D. Tidwell, E. Lo, M. Salvatore, S. Shad, J. A. Thomas, J. T. Lonsdale, M. T. Moser, B. M. Gillard, E. Karasik, K. Ramsey, C. Choi, B. A. Foster, J. Syron, J. Fleming, H. Magazine, R. Hasz, G. D. Walters, J. P. Bridge, M. Miklos, S. Sullivan, L. K. Barker, H. M. Traino, M. Mosavel, L. A. Siminoff, D. R. Valley, D. C. Rohrer, S. D. Jewell, P. A. Branton, L. H. Sobin, M. Barcus, L. Qi, J. McLean, P. Hariharan, K. S. Um, S. Wu, D. Tabor, C. Shive, A. M. Smith, S. A. Buia, A. H. Undale, K. L. Robinson, N. Roche, K. M. Valentino, A. Britton, R. Burges, D. Bradbury, K. W. Hambright, J. Seleski, G. E. Korzeniewski, K. Erickson, Y. Marcus, J. Tejada, M. Taherian, C. Lu, M. Basile, D. C. Mash, S. Volpi, J. P. Struewing, G. F. Temple, J. Boyer, D. Colantuoni, R. Little, S. Koester, L. J. Carithers, H. M. Moore, P. Guan, C. Compton, S. J. Sawyer, J. P. Demchok, J. B. Vaught, C. A. Rabiner, N. C. Lockhart, K. G. Ardlie, G. Getz, F. A. Wright, M. Kellis, S. Volpi, E. T. Dermitzakis, The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science* **348**, 648–660 (2015).
20
25
30
35
40 [doi:10.1126/science.1262110](https://doi.org/10.1126/science.1262110) [Medline](#)

111. U. Võsa, A. Claringbould, H.-J. Westra, M. J. Bonder, P. Deelen, B. Zeng, H. Kirsten, A. Saha, R. Kreuzhuber, S. Yazar, H. Brugge, R. Oelen, D. H. de Vries, M. G. P. van der Wijst, S. Kasela, N. Pervjakova, I. Alves, M.-J. Favé, M. Agbessi, M. W. Christiansen, R.

- Jansen, I. Seppälä, L. Tong, A. Teumer, K. Schramm, G. Hemani, J. Verlouw, H. Yaghootkar, R. Sönmez Flitman, A. Brown, V. Kukulshkina, A. Kalnapenkis, S. Rüeger, E. Porcu, J. Kronberg, J. Kettunen, B. Lee, F. Zhang, T. Qi, J. A. Hernandez, W. Arindrarto, F. Beutner, BIOS Consortium, i2QTL Consortium, J. Dmitrieva, M. Elansary, B. P. Fairfax, M. Georges, B. T. Heijmans, A. W. Hewitt, M. Kähönen, Y. Kim, J. C. Knight, P. Kovacs, K. Krohn, S. Li, M. Loeffler, U. M. Marigorta, H. Mei, Y. Momozawa, M. Müller-Nurasyid, M. Nauck, M. G. Nivard, B. W. J. H. Penninx, J. K. Pritchard, O. T. Raitakari, O. Rotzschke, E. P. Slagboom, C. D. A. Stehouwer, M. Stumvoll, P. Sullivan, P. A. C. 't Hoen, J. Thiery, A. Tönjes, J. van Dongen, M. van Iterson, J. H. Veldink, U. Völker, R. Warmerdam, C. Wijmenga, M. Swertz, A. Andiappan, G. W. Montgomery, S. Ripatti, M. Perola, Z. Kutalik, E. Dermitzakis, S. Bergmann, T. Frayling, J. van Meurs, H. Prokisch, H. Ahsan, B. L. Pierce, T. Lehtimäki, D. I. Boomsma, B. M. Psaty, S. A. Gharib, P. Awadalla, L. Milani, W. H. Ouwehand, K. Downes, O. Stegle, A. Battle, P. M. Visscher, J. Yang, M. Scholz, J. Powell, G. Gibson, T. Esko, L. Franke, Large-scale *cis*- and *trans*-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat. Genet.* **53**, 1300–1310 (2021). [doi:10.1038/s41588-021-00913-z](https://doi.org/10.1038/s41588-021-00913-z) [Medline](#)
112. T. Qi, Y. Wu, J. Zeng, F. Zhang, A. Xue, L. Jiang, Z. Zhu, K. Kemper, L. Yengo, Z. Zheng, eQTLGen Consortium, R. E. Marioni, G. W. Montgomery, I. J. Deary, N. R. Wray, P. M. Visscher, A. F. McRae, J. Yang, Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. *Nat. Commun.* **9**, 2282 (2018). [doi:10.1038/s41467-018-04558-1](https://doi.org/10.1038/s41467-018-04558-1) [Medline](#)
113. E. M. Weeks, J. C. Ulirsch, N. Y. Cheng, B. L. Trippe, R. S. Fine, J. Miao, T. A. Patwardhan, M. Kanai, J. Nasser, C. P. Fulco, K. C. Tashman, F. Aguet, T. Li, J. Ordovas-Montanes, C. S. Smillie, M. Biton, A. K. Shalek, A. N. Ananthakrishnan, R. J. Xavier, A. Regev, R. M. Gupta, K. Lage, K. G. Ardlie, J. N. Hirschhorn, E. S. Lander, J. M. Engreitz, H. K. Finucane, Leveraging polygenic enrichments of gene features to predict genes underlying complex traits and diseases. *Nat. Genet.* **55**, 1267–1276 (2023). [doi:10.1038/s41588-023-01443-6](https://doi.org/10.1038/s41588-023-01443-6) [Medline](#)
114. E. J. Gardner, K. A. Kentistou, S. Stankovic, S. Lockhart, E. Wheeler, F. R. Day, N. D. Kerrison, N. J. Wareham, C. Langenberg, S. O’Rahilly, K. K. Ong, J. R. B. Perry, Damaging missense variants in *IGF1R* implicate a role for IGF-1 resistance in the etiology of type 2 diabetes. *Cell Genom.* **2**, 100208 (2022). [doi:10.1016/j.xgen.2022.100208](https://doi.org/10.1016/j.xgen.2022.100208) [Medline](#)
115. C. Bycroft, C. Freeman, D. Petkova, G. Band, L. T. Elliott, K. Sharp, A. Motyer, D. Vukcevic, O. Delaneau, J. O’Connell, A. Cortes, S. Welsh, A. Young, M. Effingham, G. McVean, S. Leslie, N. Allen, P. Donnelly, J. Marchini, The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018). [doi:10.1038/s41586-018-0579-z](https://doi.org/10.1038/s41586-018-0579-z) [Medline](#)
116. P. Rentzsch, D. Witten, G. M. Cooper, J. Shendure, M. Kircher, CADD: Predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* **47**, D886–D894 (2019). [doi:10.1093/nar/gky1016](https://doi.org/10.1093/nar/gky1016) [Medline](#)
117. G. Marenne, A. E. Hendricks, A. Perdikari, R. Bounds, F. Payne, J. M. Keogh, C. J. Lelliott, E. Henning, S. Pathan, S. Ashford, E. G. Bochukova, V. Mistry, A. Daly, C. Hayward, INTERVAL, UK10K Consortium, N. J. Wareham, S. O’Rahilly, C.

Langenberg, E. Wheeler, E. Zeggini, I. S. Farooqi, I. Barroso, Exome Sequencing Identifies Genes and Gene Sets Contributing to Severe Childhood Obesity, Linking *PHIP* Variants to Repressed *POMC* Transcription. *Cell Metab.* **31**, 1107–1119.e12 (2020). [doi:10.1016/j.cmet.2020.05.007](https://doi.org/10.1016/j.cmet.2020.05.007) [Medline](#)

- 5 118. A. E. Hendricks, E. G. Bochukova, G. Marenne, J. M. Keogh, N. Atanassova, R. Bounds, E. Wheeler, V. Mistry, E. Henning, A. Körner, D. Muddyman, S. McCarthy, A. Hinney, J. Hebebrand, R. A. Scott, C. Langenberg, N. J. Wareham, P. Surendran, J. M. Howson, A. S. Butterworth, J. Danesh, B. G. Nordestgaard, S. F. Nielsen, S. Afzal, S. Papadia, S. Ashford, S. Garg, G. L. Millhauser, R. I. Palomino, A. Kwasniewska, I. Tachmazidou, S. O’Rahilly, E. Zeggini, I. Barroso, I. S. Farooqi, Understanding Society Scientific Group, EPIC-CVD Consortium, UK10K Consortium, Rare Variant Analysis of Human and
10 Rodent Obesity Genes in Individuals with Severe Childhood Obesity. *Sci. Rep.* **7**, 4394 (2017). [doi:10.1038/s41598-017-03054-8](https://doi.org/10.1038/s41598-017-03054-8) [Medline](#)
119. UniProt Consortium, UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* **51**, D523–D531 (2023). [doi:10.1093/nar/gkac1052](https://doi.org/10.1093/nar/gkac1052) [Medline](#)
120. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Židek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T.
20 Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021). [doi:10.1038/s41586-021-03819-2](https://doi.org/10.1038/s41586-021-03819-2) [Medline](#)
121. M. Lek, K. J. Karczewski, E. V. Minikel, K. E. Samocha, E. Banks, T. Fennell, A. H. O’Donnell-Luria, J. S. Ware, A. J. Hill, B. B. Cummings, T. Tukiainen, D. P. Birnbaum, J. A. Kosmicki, L. E. Duncan, K. Estrada, F. Zhao, J. Zou, E. Pierce-Hoffman, J. Berghout, D. N. Cooper, N. DeFlaux, M. DePristo, R. Do, J. Flannick, M. Fromer, L. Gauthier, J. Goldstein, N. Gupta, D. Howrigan, A. Kiezun, M. I. Kurki, A. L. Moonshine, P. Natarajan, L. Orozco, G. M. Peloso, R. Poplin, M. A. Rivas, V. Ruano-Rubio, S. A. Rose, D. M. Ruderfer, K. Shakir, P. D. Stenson, C. Stevens, B. P. Thomas, G. Tiao, M. T. Tusie-Luna, B. Weisburd, H.-H. Won, D. Yu, D. M. Altshuler, D. Ardissino, M.
30 Boehnke, J. Danesh, S. Donnelly, R. Elosua, J. C. Florez, S. B. Gabriel, G. Getz, S. J. Glatt, C. M. Hultman, S. Kathiresan, M. Laakso, S. McCarroll, M. I. McCarthy, D. McGovern, R. McPherson, B. M. Neale, A. Palotie, S. M. Purcell, D. Saleheen, J. M. Scharf, P. Sklar, P. F. Sullivan, J. Tuomilehto, M. T. Tsuang, H. C. Watkins, J. G. Wilson, M. J. Daly, D. G. MacArthur, Exome Aggregation Consortium, Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016). [doi:10.1038/nature19057](https://doi.org/10.1038/nature19057) [Medline](#)
122. H. Wickham, M. Averick, J. Bryan, W. Chang, L. D. McGowan, R. François, G. Grolemond, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. L. Pedersen, E. Miller, S. M. Bache, K. Müller, J. Ooms, D. Robinson, D. P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, H. Yutani, Welcome to the Tidyverse. *J. Open Source Softw.* **4**, 1686 (2019). [doi:10.21105/joss.01686](https://doi.org/10.21105/joss.01686)
- 40 123. J. R. Bedsaul, N. M. Carter, K. E. Deibel, S. M. Hutcherson, T. A. Jones, Z. Wang, C. Yang, Y.-K. Yang, J. L. Pomerantz, Mechanisms of Regulated and Dysregulated

CARD11 Signaling in Adaptive Immunity and Disease. *Front. Immunol.* **9**, 2105 (2018).
[doi:10.3389/fimmu.2018.02105](https://doi.org/10.3389/fimmu.2018.02105) [Medline](#)

124. C. Lu, J. Zhang, B. Wang, Q. Gao, K. Ma, S. Pei, J. Li, S. Cui, Casein kinase 1 α is required to maintain murine hypothalamic pro-opiomelanocortin expression. *iScience* **26**, 106670 (2023). [doi:10.1016/j.isci.2023.106670](https://doi.org/10.1016/j.isci.2023.106670) [Medline](#)

125. P. Xu, P. Fischer-Posovszky, J. Bischof, P. Radermacher, M. Wabitsch, D. Henne-Bruns, A.-M. Wolf, A. Hillenbrand, U. Knippschild, Gene expression levels of Casein kinase 1 (CK1) isoforms are correlated to adiponectin levels in adipose tissue of morbid obese patients and site-specific phosphorylation mediated by CK1 influences multimerization of adiponectin. *Mol. Cell. Endocrinol.* **406**, 87–101 (2015).
[doi:10.1016/j.mce.2015.02.010](https://doi.org/10.1016/j.mce.2015.02.010) [Medline](#)

126. R. E. Mesías, Y. Zaki, C. A. Guevara, L. G. Friedman, A. Hussein, K. Therrien, A. R. Magee, N. Tzavaras, P. Del Valle, M. G. Baxter, G. W. Huntley, D. L. Benson, Development and cadherin-mediated control of prefrontal corticostriatal projections in mice. *iScience* **26**, 108002 (2023). [doi:10.1016/j.isci.2023.108002](https://doi.org/10.1016/j.isci.2023.108002) [Medline](#)

127. A. A. van der Klaauw, S. Croizier, E. Mendes de Oliveira, L. K. J. Stadler, S. Park, Y. Kong, M. C. Banton, P. Tandon, A. E. Hendricks, J. M. Keogh, S. E. Riley, S. Papadia, E. Henning, R. Bounds, E. G. Bochukova, V. Mistry, S. O’Rahilly, R. B. Simerly, INTERVAL, UK10K Consortium, J. E. N. Minchin, I. Barroso, E. Y. Jones, S. G. Bouret, I. S. Farooqi, Human Semaphorin 3 Variants Link Melanocortin Circuit Development and Energy Balance. *Cell* **176**, 729–742.e18 (2019).
[doi:10.1016/j.cell.2018.12.009](https://doi.org/10.1016/j.cell.2018.12.009) [Medline](#)

128. Catalogue of Somatic Mutations in Cancer (Cosmic), Cancer Gene Census;
<https://cancer.sanger.ac.uk/census>.

129. B. Schitteck, T. Sinnberg, Biological functions of casein kinase 1 isoforms and putative roles in tumorigenesis. *Mol. Cancer* **13**, 231 (2014). [doi:10.1186/1476-4598-13-231](https://doi.org/10.1186/1476-4598-13-231) [Medline](#)

130. J. Shi, C. Cheng, J. Ma, C.-C. Liew, X. Geng, Gene expression signature for detection of gastric cancer in peripheral blood. *Oncol. Lett.* **15**, 9802–9810 (2018).
[doi:10.3892/ol.2018.8577](https://doi.org/10.3892/ol.2018.8577) [Medline](#)

131. D. Chen, Y. Zhang, L. Meng, L. Lu, G. Meng, circRNA DENND1B inhibits tumorigenicity of clear cell renal cell carcinoma via miR-122-5p/TIMP2 axis. *Open Med.* **17**, 2085–2097 (2022). [doi:10.1515/med-2022-0536](https://doi.org/10.1515/med-2022-0536) [Medline](#)

132. Z. Wang, M. Ding, N. Qian, B. Song, J. Yu, J. Tang, J. Wang, Decreased expression of semaphorin 3D is associated with genesis and development in colorectal cancer. *World J. Surg. Oncol.* **15**, 67 (2017). [doi:10.1186/s12957-017-1128-1](https://doi.org/10.1186/s12957-017-1128-1) [Medline](#)

133. D. Liu, J. Li, F. Qi, H. You, Semaphorins and their receptors in pancreatic cancer: Mechanisms and therapeutic opportunities. *Front. Oncol.* **12**, 1106762 (2023).
[doi:10.3389/fonc.2022.1106762](https://doi.org/10.3389/fonc.2022.1106762) [Medline](#)

134. H. G. Parker, D. L. Dreger, M. Rimbault, B. W. Davis, A. B. Mullen, G. Carpintero-Ramirez, E. A. Ostrander, Genomic Analyses Reveal the Influence of Geographic Origin, Migration, and Hybridization on Modern Dog Breed Development. *Cell Rep.* **19**, 697–708 (2017). [doi:10.1016/j.celrep.2017.03.079](https://doi.org/10.1016/j.celrep.2017.03.079) [Medline](#)

135. C. A. Pugh, B. M. C. Bronsvort, I. G. Handel, K. M. Summers, D. N. Clements, Dogslife: A cohort study of Labrador Retrievers in the UK. *Prev. Vet. Med.* **122**, 426–435 (2015). [doi:10.1016/j.prevetmed.2015.06.020](https://doi.org/10.1016/j.prevetmed.2015.06.020) [Medline](#)

5 **Acknowledgments:** We would like to thank Stephen O’Rahilly for valuable mentorship to ER and insightful comments on the manuscript; Jennifer Meadows for training in GWAS approaches early in the project; and Felix Day for helpful input developing our canine and comparative genomics methods. We are indebted to the owners who volunteered for the study and to the dogs who took part, the Northern England Flatcoated Retriever Association and Kennel Club for
10 helping with recruitment, Guide Dogs UK for sample contributions, the Golden Retriever Lifetime Study for use of their data; Elaine Withers and Carole Cornwell for sample recruitment administration; Nai-Cheih Liu for gathering phenotypes in pugs; the Institute of Metabolic Science Genomics and Histology Core Facilities, particularly James Warner; Nevin Lambert for the gift of the Venus-Kras plasmid; and Bryan Roth for the gift of the Tango-MC4R plasmid. We
15 thank the physicians who referred people to the Genetics of Obesity Study (GOOS) and the patients and families for their involvement. This work was conducted using the UK Biobank under application 9905. BCS chart in Fig. S1 is part of the Global Nutrition Committee Toolkit provided courtesy of the World Small Animal Veterinary Association.

20 **Funding:**

Wellcome Trust Clinical Research Career Development Fellowship 205187/Z/16/Z (ER)

UKRI Biotechnology and Biological Sciences Research Council (BBSRC) DTP
BB/M011194/1 (NJW, AM, GKCD)

Dogs Trust Canine Welfare Grants (ER, JW, MD, GL)

25 Morris Animal Foundation D22CA-406 (ER, AM)

UKRI Medical Research Council unit programmes MC_UU_00006/2 (KAK, KKO, JRBP), MC_UU_00014/1 (BYHL, KR, MM, GSHY)

France Génomique’ consortium ANR-10-INBS-009 (PF)

European Genomic Institute for Diabetes (EGID) (PF)

30 French National Center for Precision Diabetic Medicine – PreciDIAB (PF)

Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society.
224155/Z/21/Z (CMG)

Wellcome Principal Research Fellowship 207462/Z/17/Z (ISF)

National Institute for Health and Care Research (NIHR) Cambridge Biomedical Research
35 Centre (ISF)

Botnar Fondation (ISF)

Bernard Wolfe Health Neuroscience Endowment (ISF)

Leducq Fondation grant (ISF)

NIHR Senior Investigator Award (ISF)

Kennel Club Charitable Trust (JFL, DRS)
Royal Society grant RGS\R1\221094 (ER)

Author contributions:

5 Conceptualization: ER.
 Methodology: ER, NJW, AMo, ES, SS, CG.
 Investigation: ER, NJW, AMc, AMo, KK, AJ, GD, ES, AM, SS, BL, NS, JC, SSK, RMZ,
 JW, MD, GL, KR, DW, RB, MM, AJG, JL, DS, JT, CG.
 Visualization: ER, NJW, AMc, KK, GD, JT, CG.
10 Funding acquisition: ER.
 Project administration: ER, JW.
 Supervision: ER, PF, ISF, KO, GSHY, JP, CG.
 Writing – original draft: ER, NJW, AMc, AMo, KK, ES, JC, CG.
 Writing – review & editing: ER, NJW, AMc, KK, GD, SS, AJG, DS, ISF, KO, GSHY,
15 JT, JP, CG.

20 **Competing interests:** JRBP is an employee of Insmad Innovation UK and holds stock/stock
options in Insmad Inc. JRBP also receives research funding from GSK and engages in paid
consultancy for WW International Inc. GSHY receives grant funding from Novo Nordisk and
consults for both Novo Nordisk and Eli Lilly. ER has engaged in paid consultancy for Purina
Ltd. ISF has consulted for a number of companies developing weight loss drugs including Eli
Lilly, Novo Nordisk and Rhythm Pharmaceuticals and investors (Goldman Sachs, SV
Health). AJG's academic post at the University of Liverpool is financially supported by
25 Royal Canin. AJG has also received financial remuneration for providing educational
material, speaking at conferences and consultancy work for Mars Petcare. All other authors
declare that they have no competing interests.

30 **Data and materials availability:** All genetic and phenotypic data relating to this project are
available on Dryad (73) with the exception of the GRLS data which are available from the
GRLS data commons site (<https://datacommons.morrisanimalfoundation.org/>). Code used for
imputation is available on Zenodo (74). Research materials can be provided by E.R. pending
scientific review and a completed material transfer agreement. Requests for materials should
be submitted to E.R. (er311@cam.ac.uk).

Supplementary Materials:

35 Materials and Methods
 Supplementary Text
 Supplementary Figures S1 – S11
 Supplementary Tables S1 – S15
40 References (74 – 137)

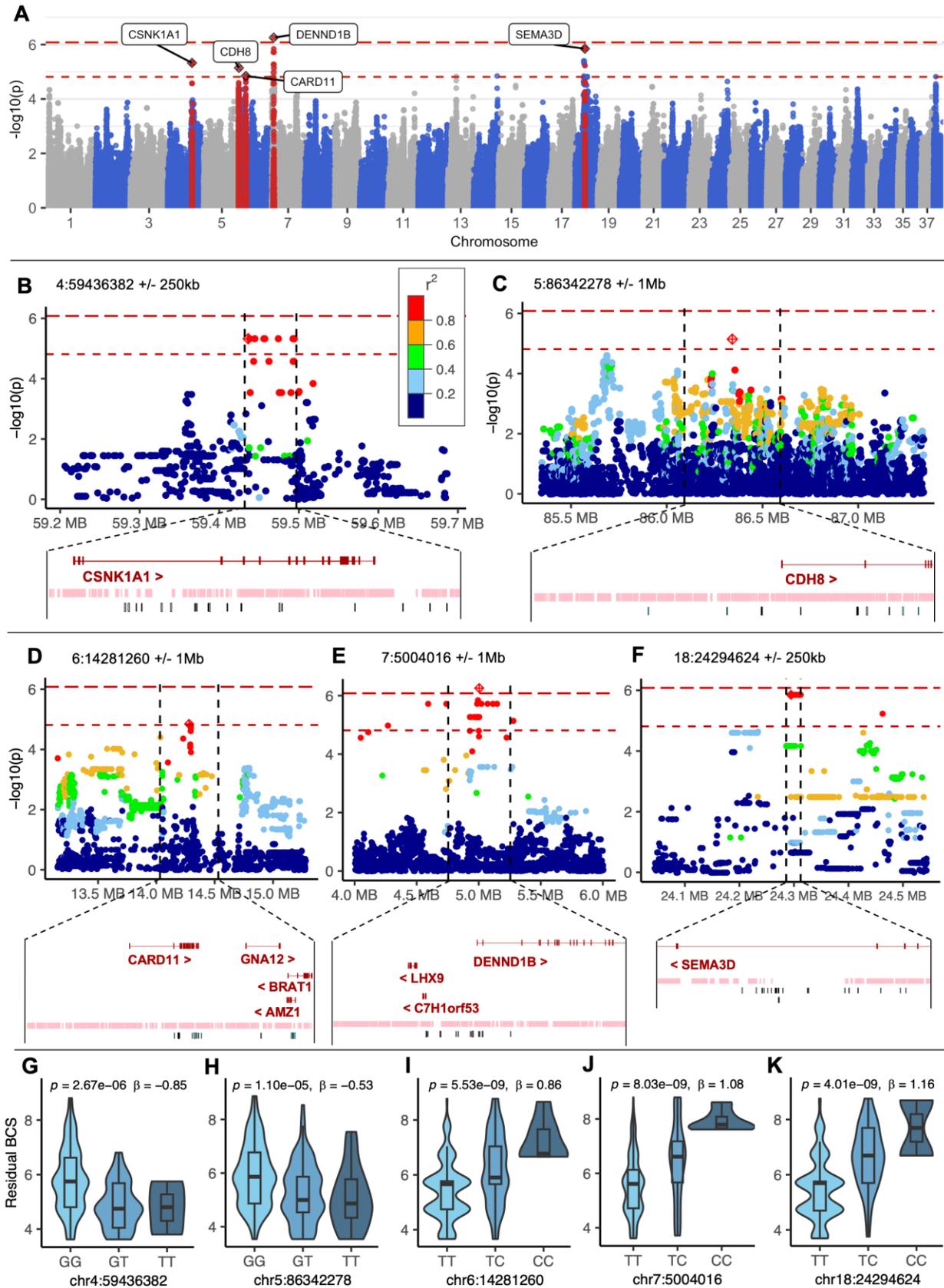


Fig. 1: GWAS in Labrador retrievers identifies multiple obesity genes. (A) Manhattan plot for GWAS for body condition score (BCS) in Labrador retrievers ($n = 241$). Suggestive significance shown with open dashes, Bonferroni-corrected significance $p < 8.31 \times 10^{-7}$ shown with closed dashes. Five independent loci which harbored protein coding genes are labelled with the most proximal protein coding genes (lead SNP at 4:59436382, 5:86342278, 6:14281260, 7:5004016, 18:24294624). Regional Manhattan plots shown for (B) Chr. 4, (C) Chr. 5, (D) Chr. 6, (E) Chr. 7, (F) Chr. 18 are colored by r^2 measure of linkage disequilibrium. For chromosomes 4 and 18, the plot extends ± 250 kb of the lead SNP and the funnel indicates the boundaries of the mapped haplotype. For other regions, plots extend ± 1 Mb and genes in the funnel are those lying ± 250 kb of the lead SNP. The lead SNP for each locus is indicated by a diamond, with genes within the region annotated below. Similarity with human genome is indicated by a LASTZ pairwise alignment with GRCh38 shown as the pink track. Variants identified from WGS as segregating with the lead SNP $\geq 70\%$ of the time are aligned in black below. Partial regression violin plots showing relationship between BCS and lead variant genotype detail large effect sizes at each locus on (G) Chr. 4, (H) Chr. 5, (I) Chr. 6, (J) Chr. 7, (K) Chr. 18. Chr., Chromosome.

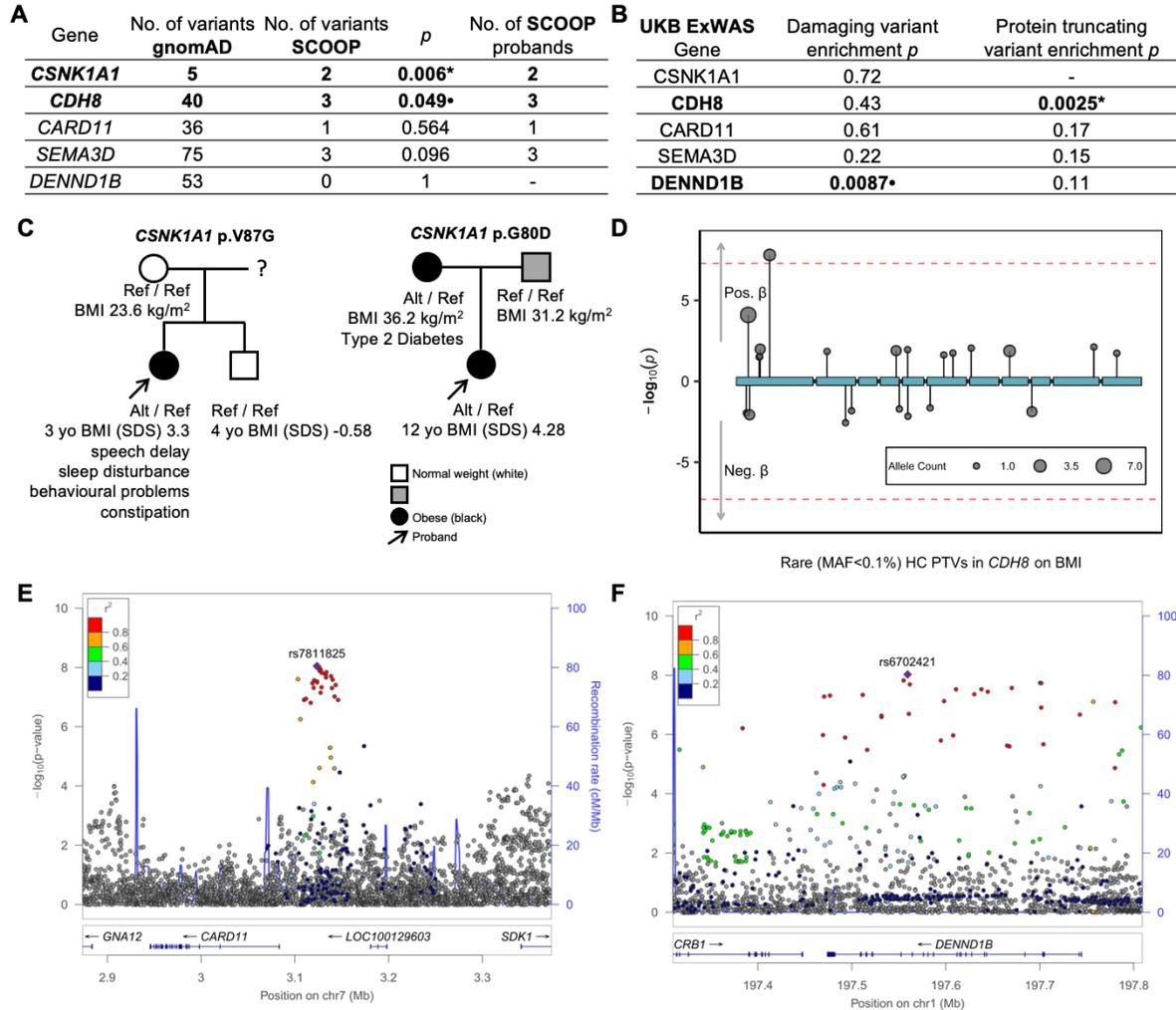


Fig. 2: Canine obesity-associated genes are also associated with human obesity. (A)

CSNK1A1 and *CDH8* are enriched for rare, deleterious (gnomAD minor allele count ≤ 3 , CADD ≥ 25) variants in the SCOOP cohort of 982 children with severe, early onset obesity compared to 56,885 controls from gnomAD (one sided Fisher's exact test). (B) ExWAS analysis on exome sequences of $\sim 500k$ individuals in UK Biobank showed rare protein truncating variants in *CDH8* were associated with BMI and damaging (protein truncating and high CADD) variants in *DENND1B* were nominally associated with BMI (Fig. S3). (C) Genotype segregated with obesity phenotype in pedigrees from families of probands with severe, early onset obesity for two severe, deleterious variants in *CSNK1A1* identified in SCOOP and *DENND1B* p.R501C identified in SOPP (Fig. S3). (D) Lollipop plot shows protein truncating variants in *CDH8* were associated with increased BMI in the UKB ExWAS. Regional Manhattan plots for (E) *CARD11* and (F) *DENND1B* show that in a GWAS of $\sim 800k$ individuals there were associations with BMI at two human loci orthologous to canine GWAS loci, at which these were called as the likely effector genes using the GWAS2Gene pipeline. Significance: $p < 0.00023$, '*'; nominal $p < 0.05$, '•'; no variants, '-'. PTV - protein truncating variants; HC – high confidence; MAF – minor allele frequency; SCOOP - Severe Childhood Onset Obesity Project.

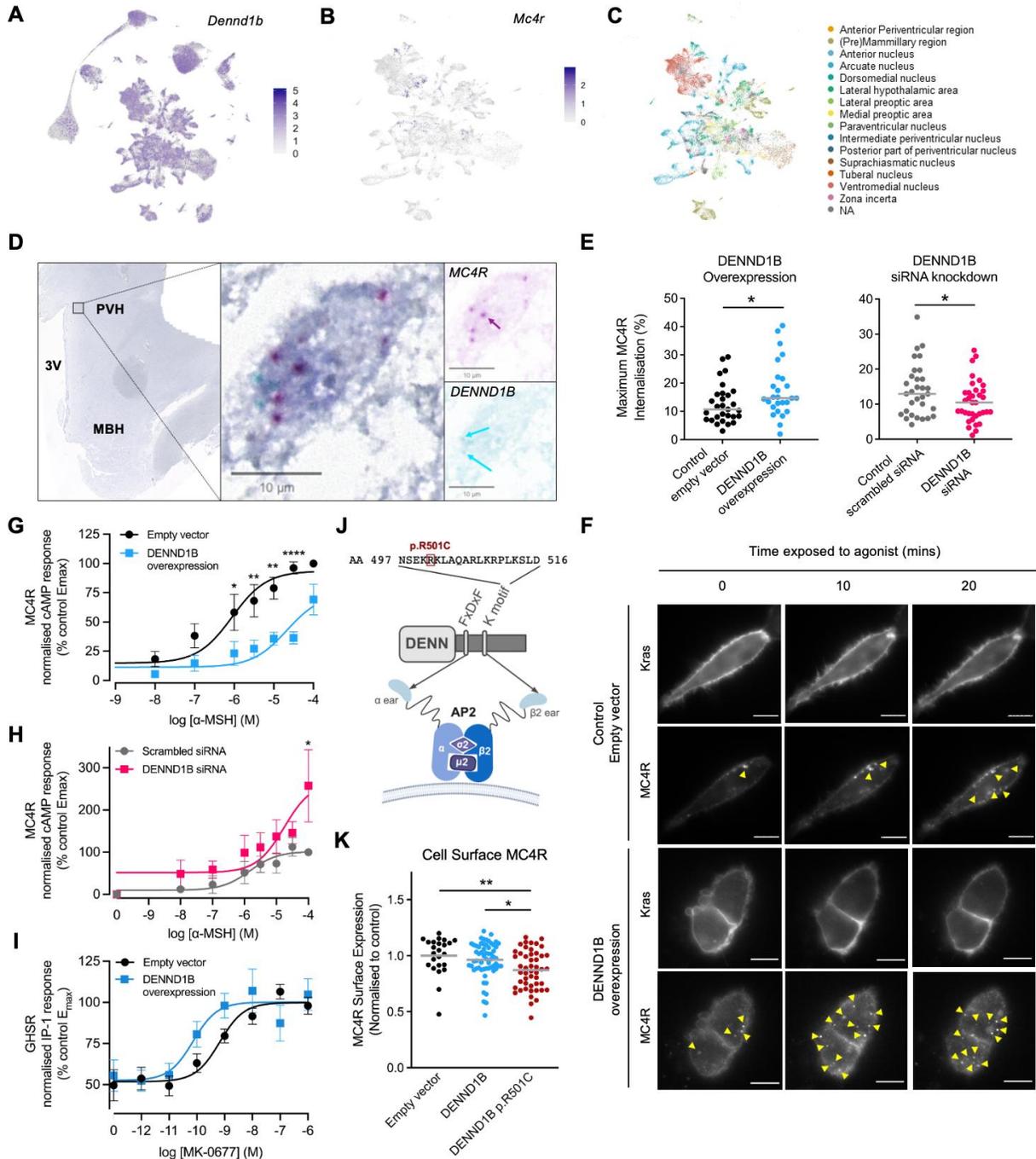


Fig. 3: *DENND1B* is co-expressed with *MC4R* and regulates its signaling. (A) UMAP plots showing log-normalized expression of (A) *Dennd1b* in murine hypothalamus and (B) *Mc4r* in the subset of *Dennd1b* positive neurons, highlighting co-expression. (C) *Dennd1b* positive neurons colored by hypothalamic region. (D) Coronal section of human hypothalamus stained for *MC4R* (red) and *DENND1B* (green) showing dual positive neurons within PVH. (E) *DENND1B* overexpression (blue) enhances ligand-stimulated *MC4R* internalization, while *DENND1B* siRNA knockdown (pink) reduces it. (F) HILO images of HEK293 expressing

MC4R and plasma membrane marker Kras showing colocalization at baseline and increased MC4R internalization on ligand stimulation with *DENND1B* overexpression. Scale 5 μ m. Ligand-induced G_s cAMP response downstream of MC4R in HEK293 cells is (G) reduced by *DENND1B* overexpression and (H) increased at maximal concentrations by *DENND1B* knockdown. (I) Overexpressing DENND1B causes increased signaling in the canonical IP-1 pathway downstream of GHSR. (J) DENND1B contains both the normal DENN protein AP-2 α ear-binding motif (FxDxF) and an AP-2 β ear-binding motif whose sequence is shown with *DENND1B* p.R501C highlighted. (K) Cell surface expression of MC4R during expression of DENND1B wild type, p.R501C, and empty vector. Significance: $p \leq 0.05$ ‘*’, $p \leq 0.01$ ‘**’, $p \leq 0.0001$ ‘****’. MBH, mediobasal hypothalamus; 3V, third ventricle; PVH, Paraventricular hypothalamic nucleus. Figure 3J created using BioRender.com.

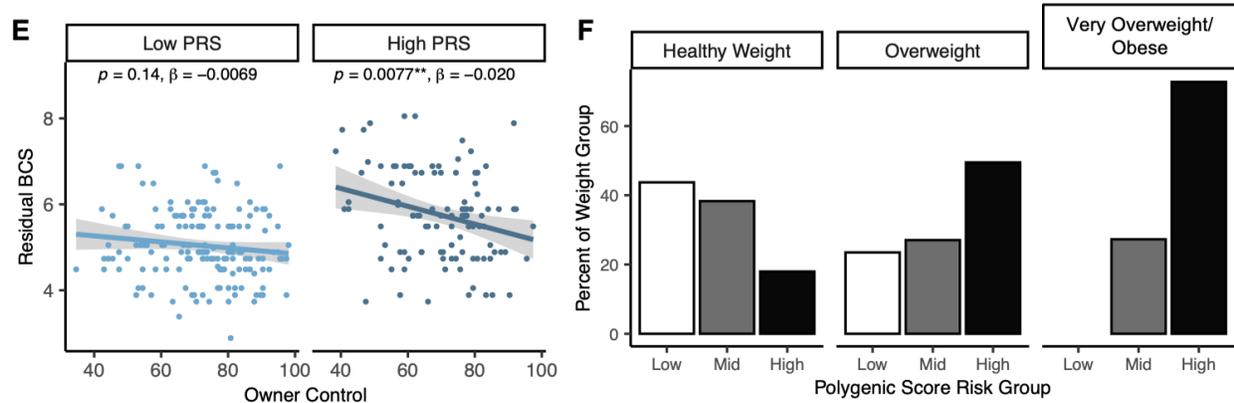
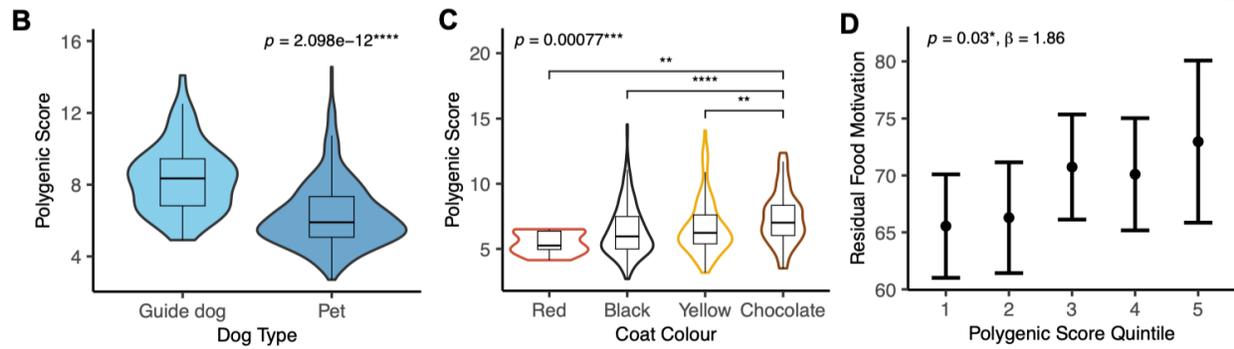
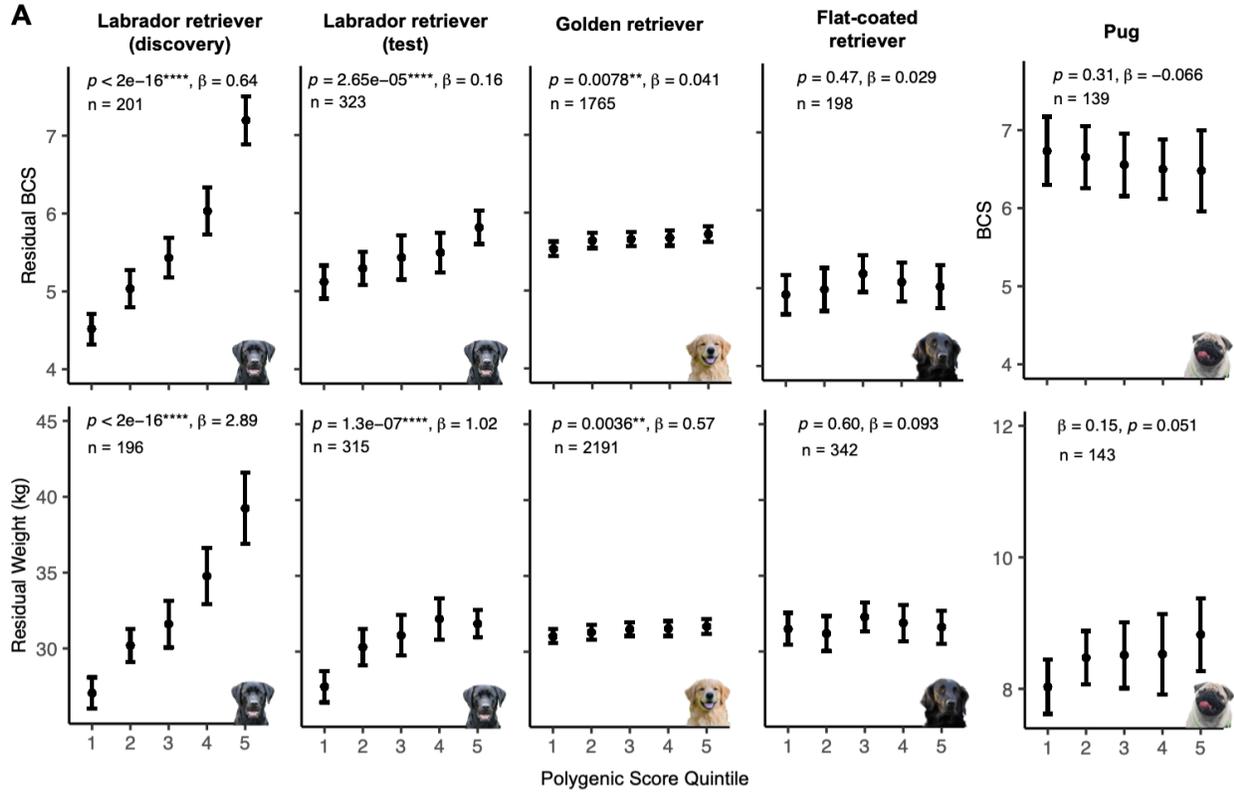


Fig. 4: Polygenic risk scores (PRS) predict obesity and provide insight into complex trait expression. A PRS was constructed comprising 16 SNPs weighted for GWAS effect size. (A) The PRS predicted BCS and weight in independent populations of Labrador retrievers. In golden retrievers, the PRS constructed in Labradors predicted BCS and weight, albeit less strongly. In more distantly related breeds, flat-coated retrievers and pugs, the PRS had no predictive value. Known risk factors for obesity in the Labrador population were explained by differences in PRS, with (B) higher PRS in obesity-prone assistance dogs and (C) chocolate-colored Labradors. (D) The PRS predicted food motivation score in Labrador retrievers (n = 298). (E) Gene-environment interaction: dogs with low PRS were resistant to obesity irrespective of owner control of diet and exercise, but management of dogs with high PRS does significantly affect obesity outcome (significant in regression model, $p = 0.0077$). (F) Prevalence of high PRS dogs increased with obesity category (healthy BCS < 6/9; overweight BCS 6-7/9; very overweight BCS >7/9) and there were no low-risk dogs in the extremely overweight group (PRS grouped by tertiles). Significance levels: $p \leq 0.05$ ‘*’, $p \leq 0.01$ ‘**’, $p \leq 0.001$ ‘***’, $p \leq 0.0001$ ‘****’.

5

10

15



Supplementary Materials for

Canine genome-wide association study identifies *DENND1B* as an obesity gene in dogs and humans

Natalie J. Wallis, Alyce McClellan, Alexander Mörseburg, Katherine A. Kentistou, Aqfan Jamaluddin, Georgina K.C. Dowsett, Ellen Schofield, Anna Morros-Nuevo, Sadia Saeed, Brian Y.H. Lam, Natasha T. Sumanasekera, Justine Chan, Sambhavi S.Kumar, Rey M Zhang, Jodie F. Wainwright, Marie Dittmann, Gabriella Lakatos, Kara Rainbow, David Withers, Rebecca Bounds, Marcella Ma, Alexander J. German, Jane Ladlow, David Sargan, Philippe Froguel, I. Sadaf Farooqi, Ken K. Ong, Giles S.H. Yeo, John A. Tadross, John R.B. Perry, Caroline M. Gorvin, Eleanor Raffan.

Corresponding author: er311@cam.ac.uk

The PDF file includes:

Materials and Methods

Supplementary Text

Figs. S1 to S11

(Tables S1 to S14 are presented in separate Excel file.)

References (74-137)

Materials and Methods

Ethics statement: Dogs

The research was approved by the Ethical Review Committee of the Department of Veterinary Medicine, University of Cambridge (CR73 and CR125), with sample collection at other centers also approved by local ethical review committees: Animal Health Trust Research Ethics Committee, and the University of Liverpool Research Ethics Committee (RETH000353). All dog owners gave full written consent to participate in the research.

Ethics statement: Humans

Data from the UK Biobank were accessed under application 9905. The Genetics of Obesity Study (GOOS) work was approved by the Multi-Regional Ethics Committee and the Cambridge Local Research Ethics Committee (MREC 97/21 and REC number 03/103). Each subject (or their parent for those under 16 years) provided written informed consent; minors provided oral consent. An anonymized human hypothalamic tissue sample was provided by the Cambridge Brain Bank for RNAscope analysis. Subjects were approached in life for written consent for brain banking, and all tissue donations were collected and stored following legal and ethical guidelines (NHS reference number 20/EE/0283). All studies were conducted in accordance with the Declaration of Helsinki.

Labrador retriever data collection

Labrador retriever owners were recruited by disseminating invitations to participate through veterinary practices, social media platforms, and emails to owners of relevant breeds via the Kennel Club, between 2013 - 2020. Additional dogs were pets which attended a specialist weight management clinic at the University of Liverpool School of Veterinary Science (n = 18). A population of Labrador retriever assistance dogs were also obtained through collaboration with Guide Dogs UK. Some recruitment targeted owners to volunteer if they had healthy weight or particularly overweight dogs making it unsuitable for estimating obesity prevalence. Demographic data were collected including breed, sex, neuter status, age, and coat color. Full clinical veterinary histories were also obtained and assessed by a trained veterinary professional.

Labrador retrievers were included if they met previously described criteria (28), outlined briefly here. Dogs were included if they had complete information on sex, neuter status, age and BCS. Clinically trained researchers reviewed their medical histories and excluded dogs affected by systemic disease, chronic orthopedic disease or on long-term medications known to cause weight gain or loss, such as corticosteroids and anti-seizure medication. Only Labrador retrievers over one year of age (at time of sampling) were included to avoid confounding by early life fluctuations in fat mass. Dogs with BCS <3 were also removed since this level of underweight may indicate poor underlying health. The final study population comprised 521 pet and 70 assistance Labrador retrievers.

Obesity was measured using the Body Condition Score (BCS), a validated measure of canine adiposity, measured using a 9-point BCS (25–27). Dogs are assigned to one of the body condition score categories according to how they match a set of descriptors and images and using a combination of visual and haptic assessment. The BCS chart used is shown in Fig. S1. Body weight was measured and BCS assigned by veterinary professionals who were provided with specific instructions about how to perform the assessment.

We used the Dog Obesity Risk and Appetite questionnaire to measure dogs' food motivation and owner control of food and exercise in the home environment. This validated, owner reported measure (24) contains 31 items which are used to obtain an 'owner control score' (combining factor scores on 'owner intervention', 'restriction of human food' and 'exercise') and food motivation score (combining factor scores on 'Interest in food', 'lack of fussiness' and 'Responsiveness and satiety'). Each statement is on a Likert scale scored numerically as 1-4 ('Not at all true' = 1, 'Somewhat true' = 2, 'Mainly true' = 3, 'Definitely true' = 4) or 1-5 ('Never' = 1, 'Rarely' = 2, 'Sometimes' = 3, 'Often' = 4, 'Always' = 5). Before factor calculation, some statements' scores are reversed. For example, for a statement like 'my dog would eat anything', stronger agreement indicates a high food motivation score but for the statement 'my dog takes his/her time to eat a meal', stronger agreement indicates lower food-motivation and so the scores are reversed. The mean response to items within each category is calculated and results reported as a percentage of the total maximum score. See Table S1 for details of DORA Questionnaire scoring items contributing to owner control and food motivation scores.

DNA sampling

The majority of the canine DNA samples were extracted from saliva collected using Performagene (PG-100) oral sponge swab kits (DNA Genotek). DNA was isolated according to the manufacturer's instructions using the Performagene PG-AC purification protocol (DNA Genotek). In dogs seen for clinical treatment, residual EDTA blood samples were stored at -20°C until shipping for analysis, and extracted using the DNeasy Blood & Tissue Kit (Qiagen, UK, Cat. No. 69504) according to the manufacturer's instructions.

Genotyping and imputation overview

Genome-level imputed data aligned to the CanFam3.1 reference assembly (GCA_000002285.2) were generated for the Labrador retrievers through one of two routes. The primary Labrador genetic dataset used for GWAS was generated via Route 1 (array genotyping followed by imputation) and the secondary or 'test' dataset used for polygenic risk score replication used a combination of Route 1 and Route 2 (low pass sequencing and imputation).

Array genotyping with imputation (route 1)

For 391 Labrador retriever participants DNA samples were genotyped using the 220k Canine HD BeadChip single nucleotide polymorphisms (SNPs) array (Illumina, San Diego, CA, USA). Post-QC genotypic data were imputed to genome level using an in-house pipeline which was made up of three wrapper scripts: generating an imputation reference panel, preparing the genotypic data (to be used for the GWAS) and the imputation process itself. The GitHub pipeline can be accessed via Zenodo (74). It was developed to run on the University of Cambridge High Performance Computing (HPC) system, with a SLURM workload manager so is not directly transferable but all the programs used within the pipeline are freely available to download and run.

Quality control for array genotypes was implemented using PLINK v.1.9. (61). SNPs residing on X and Y chromosomes were removed. Subsequent filters excluded markers with >3% genotype calls missing (--geno 0.03), individuals with >10% genotype calls missing (--

mind 0.1) and markers with minor allele frequency (MAF) <1% in the population (--maf 0.01). Ambiguous SNPs (A/T or C/G) were removed to limit strand and minor-allele confusion. Variants with extreme deviation from Hardy Weinberg equilibrium (HWE) were also removed (-hwe 0.00005). This data were also phased using SHAPEIT v2.r904 (75) using eight threads (-T 8) a window size of 2MB (--window 2) and effective population size of 200 (--effective-size 200).

To create a canine imputation reference panel, publicly available datasets were supplemented with additional genome sequence from Labrador retrievers generated locally all of which were aligned to CanFam3.1 genome build. Variant Call Format (VCF) files from a multibreed dataset were used to form most of the panel as described previously (76). The final imputation panel represented 676 individuals of 91 breeds, including 31 Labrador retrievers (BioProject accession PRJNA648123 and PRJNA726547). The panel was enriched with whole genome sequence from 7 Labradors extracted from the European Nucleotide Archive (ENA <https://www.ebi.ac.uk/ena/> accession codes SRR7120183, SRR13340562, SRR13340566, SRR13340565, SRR13340564, SRR13340563, SRR13340570) and 5 from in-house WGS data (BioSample IDs SAMEA115942716, SAMEA115942718, SAMEA115942720, SAMEA115942723, SAMEA115942717, SAMEA115942719, SAMEA115942721, SAMEA115942722).

The reference panel was built using a combination of BEDtools v.2.20.1 (77), BCFtools v.1.9 (78), and PLINK v.1.9 (61). VCF files were filtered to include only variants with quality > 20 (QUAL > 20, equating to 99% probability) and duplicate SNPs were removed. All datasets were merged and matched for REF/ALT alleles; ambiguous SNPs (A/T, G/C) were removed from the panel. Variants which had missing data in one or two of the datasets were called as reference using the BCFtools v.1.9 (78) commands --force-samples --missing-to-ref. Only SNP variants with a read depth of >10 were used in the panel; other variant types were removed to improve imputation efficiency. The merged panel underwent PLINK QC excluding markers with >3% genotype calls missing (--geno 0.03), individuals with >10% genotype calls missing (--mind 0.1) and markers with minor allele frequency (MAF) <1% in the population (--maf 0.01). Prior to imputation, the panel was phased using SHAPEIT v2.r904 (75) using 20 threads (-T 20), a window size of 2MB (--window 2) and effective population size of 200 (--effective-size 200).

We used package IMPUTE2 v2 (79) for imputation of the genotypic data described above. A genomic interval was assigned from length 1 to max length of chromosome using the -int command. Effective population size was set at 200, using command -Ne 200. Analysis of regions over 7Mb was permitted using the -allow_large_regions command.

Following imputation, data were converted from dosage genotypes (probabilistic) to discrete allele calls (0/1/2) for each SNP. Although this conversion may lead to loss of information for more rare variants, it improves the power for the GWAS and allows for easier interpretation of data downstream. As IMPUTE2 generates an imputation certainty per SNP, we were able to filter the SNPs for those with a high imputation certainty of ≥ 0.7 (where certainty typically ranges from 0-1). We converted the data to PLINK format and performed QC excluding markers with >5% genotype calls missing (--geno 0.05), individuals with >5%

genotype calls missing (--mind 0.05) and markers with minor allele frequency (MAF) <5% in the population (--maf 0.05).

SkimSEEK™ genotyping (route 2)

Low-pass sequencing with imputation was performed using skimSEEK™ technology (Canine skimSEEK v2.0, Neogen Europe Ltd). The imputation panel consisted of the same 676 dogs representing 91 breeds and 53 million variants that formed the majority of our in house imputation panel. Imputation was based on an underlying statistical model described previously on the basis of the ‘effective coverage’ model, described in the first report of this method (80) and validated elsewhere (81). Variants of interest for further study could then be extracted from this dataset as appropriate. This method was used for 150 Labrador retriever dogs included in the secondary or ‘test’ data set.

Genotyping of the POMC deletion

Dogs were genotyped for a POMC variant of interest - frameshift mutation (17:19431807-19431821:GCGCCGGCCCGGGA>-, p.P187fs) (19) using custom TaqMan™ assays (ThermoFisher Scientific, UK). Custom-designed primers and probes were generated (Forward primer AGGCCTTCCCCGTCGAGTTC; Reverse primer TACTCCAGGTCGGCCAGCG; Wild-type probe AGGGCCCCGGCCGCG with VIC fluorophore and MGB quencher; Deletion probe TCGGCCCGGGCGT with FAM fluorophore and MGB quencher). For the genotyping reaction, the TaqMan gene expression master mix (ThermoFisher CAT#4369016, UK) was used, with the addition of 3% DMSO (ThermoFisher CAT #D12345). Primers were used at a concentration of 0.4µM, probes at 0.1µM, and genomic DNA at 0.4ng/µl. The thermocycling process was performed on an Applied Biosystems (Cat#4329001) 7900HT Fast Real-Time PCR System, following this profile: an initial step of 2 minutes at 50°C, followed by 10 minutes at 95°C, and then 40 cycles consisting of 15 seconds at 95°C and 1 minute and 30 seconds at 65°C. For analysis purposes, a cycle threshold (Ct) of 0.04 was set for the wild-type probe, and a Ct of 0.09 was set for the deletion probe, to account for the different lengths of the DNA fragments being amplified. If the difference between the Ct values obtained from the two probes was greater than 0.5, the genotype was confirmed through agarose gel separation.

Data from other dog breeds

Golden retriever data were obtained from the Morris Animal Foundation (Denver, Colorado, USA) having been generated as part of the Golden Retriever Lifetime Study (GRLS) (82). Golden retrievers were genotyped on the 1.1 million SNP Axiom Canine Genotyping Array Set A (ThermoFisher Scientific, Massachusetts, US). A population of pugs and flat-coated retrievers were recruited and phenotyped as described for the Labrador retrievers. Pugs were genotyped for >170k markers using the Canine HD BeadChip SNP array (Illumina, San Diego, CA, USA) and imputed in-house using the methods described for Labradors above with the modification that the imputation panel was enriched with pug whole genomes instead of Labradors.

Flat-coated retrievers were genotyped for specific variants of interest using Matrix-associated laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) assay. Assay development and genotyping was performed at Neogen Genomics (Lincoln,

Nebraska, USA), employing the MassARRAY platform along with iPLEX GOLD chemistry in accordance with the manufacturer's protocol from Agena Bioscience (San Diego, California, USA). To design multiplex assays, the Agena Design Suite software provided by the manufacturer was used. A total of 24 markers of interest were split into two separate 'plex' pools of 12 and 11 markers each. The two multiplex assays were run on customer supplied genomic DNA, data generated, and quality check metrics applied, following which the final genotype calls were obtained.

Demographic data (sex, age, neuter status, breed, et cetera) and obesity-related phenotypic information were gathered for these additional populations. Obesity-related phenotypes for all datasets were measured by veterinary professionals or trained academic researchers. Phenotypic data underwent QC steps similar to that described above for Labradors (28). Dogs were removed if their age was <1 or >14 years old and if their BCS ≤ 3 . Only individuals with complete information on relevant demographic data and BCS were included.

Genome wide association study (GWAS)

For the primary GWAS for BCS, inclusion criteria were more rigorous to ensure only dogs with the most robust phenotype data were included. Labradors were included only if they were 1-10 years old (to eliminate old age as a confounding variable); pets (due to difference in assistance dog environment and genetic divergence); had comprehensive medical histories; and their owners had not answered positively the question 'My dog regularly sees the vet for health problems (not including check-ups/vaccinations)' in the DORA questionnaire. Additionally, only dogs who underwent direct genotyping followed by in-house imputation were included to avoid batch effects on genotypic data. The final GWAS cohort comprised 241 Labradors. The remaining pet and assistance Labradors were used as a replication cohort for polygenic risk score development.

Dogs were assessed for pairwise relatedness using Identity by Descent (IBD) pi-hat estimates generated by PLINK v.1.9 (61) --genome command. Where estimated pi-hat ≥ 0.5 indicated 1st degree relative (parent/sibling), one of such a pair of dogs was removed. Using Genome-wide Efficient Mixed Model Analysis (GEMMA) software v0.98.1 (83), we then generated a relatedness matrix which was transformed to a distance matrix using R v.4.2.2 (62). Multidimensional scaling (MDS) was used to visualize any clustering of individuals. If distinct clustering was observed, indicating population stratification, only one cluster was taken forward.

Regression modelling was used to determine population-specific covariates relevant to BCS in the GWAS dogs. Using Akaike's Information Criterion (AIC), a stepwise model selection method, we performed minimal model selection for BCS. In the initial model, we included relevant risk factors based on prior knowledge, along with interactions. These risk factors comprised: sex, neuter status, sex:neuter status interaction term, age:sex interaction term, and age:neuter status interaction term. After implementing AIC, the minimal linear model for BCS was: $BCS \sim \text{sex} * \text{neuter status}$. Three covariates were therefore used for subsequent analyses in this cohort: sex, neuter status and sex:neuter status interaction term.

Using Genome-wide Complex Trait Analysis (GCTA) v.1.93.2 (60), we performed genome-based restricted maximum likelihood ratio (GREML) analysis (GREML_LDMS),

stratified by linkage disequilibrium (LD) and minor allele frequency (MAF). This is a method to estimate heritability and generates a multi-genetic relatedness matrix (multi-GRM) for imputed data or whole genome sequenced (WGS) data. This method is more suitable for imputed data compared with a standard GRM, as it corrects for the LD bias in the estimated SNP-based heritability (63). First, segment-based LD score was calculated using an LD-score region (--ld-score-region) of 200 kilobases (kb) and overlap of 100kb between adjacent segments. Second, individual SNPs were stratified by segment-based LD scores in R v.4.2.2 (62). Third, multiple GRMs were generated on the stratified SNPs using GCTA v.1.93.2 (60) command --make-grm. These GRMs, along with the covariates, were used via --mgrm and --covar to perform restricted maximum likelihood (REML) analysis with --reml to get an estimate for heritability.

The GWAS analysis used a mixed linear model association (MLMA) analysis from GCTA v.1.93.2 (60), implementing the --mlma-loco option. This MLMA analysis excludes the chromosome on which the individual SNP is located when calculating the GRM. The multi-GRM generated during the GREML_LDMS analysis was used as a random effect in the GWAS model using the -mgrm command, as were the identified covariates using the -covar option.

For the threshold(s) of genome-wide significance we used a stringent Bonferroni corrected threshold of $p = 8.31 \times 10^{-7}$. Bonferroni correction was performed on LD-pruned data using 1000bp window size, 2 bp step size and variance inflation factor (VIF) of 3.33'. VIF threshold was based upon calculation: $1/(1-r^2)$ where $r^2 = 0.7$. Pruning was implemented using PLINK v.1.9 (61) using command --indep 1000 2 3.33'. This Bonferroni-corrected genome-wide significance threshold is notably conservative since 'independent' SNPs were defined using a relatively high LD cut-off ($r^2 < 0.7$). A suggestive threshold was defined at $p = 1.54 \times 10^{-4.81}$ where SNP associations deviated from the 95% confidence interval (CI) on the quantile-quantile (QQ) plot, an approach previously applied in canine GWAS studies (29, 30).

For the lead SNP at each locus, the allele with a positive β effect on BCS outcome was identified as the 'risk' allele and the allele associated with a negative β effect on BCS was identified as the 'non-risk' allele.

Stepwise conditional analysis was performed on the GWAS output on a per-chromosome basis, to identify independently associated loci. The genotype of the most highly associated SNP was extracted and the GWAS MLMA rerun using the genotype as a covariate. This process was repeated on a per-chromosome basis until no more variants on a chromosome were significant. SNPs which became significant during the conditional analysis process but did not pass the original threshold were not taken forward as candidate loci. Additionally, SNPs which lost their association when conducting conditional analysis on other chromosomes were regarded as confounded and removed.

SNPs which were corroborated by multiple other proximal SNPs with above average significance were prioritized in further analysis. This was observed by a regional clustering of associated SNPs on the resulting GWAS Manhattan plots, indicating multiple variants in high LD and similarly associated with the trait. SNPs which passed the significance threshold but were not accompanied by such clustering were not taken forward as solitary SNPs may be indicative of a genotyping artefact.

To ensure that the imputation process and conversion to discrete genotype calls was not generating artefactual loci, the GWAS was rerun, first using only directly genotyped SNPs and, second, using the imputed ‘probabilistic’ genotypes (prior to converting to 0/1/2 discrete calls). Only loci which demonstrated significant association across all three GWAS were taken forward.

Identifying genes and variants at associated loci

Regions of interests were first defined based on pairwise r^2 measure of linkage disequilibrium with the lead SNP ($r^2 \geq 0.8$). Using these inclusive regions, gene lists were extracted from Ensembl Biomart (Archive release 104, May 2021) (84, 85) to extract regional candidate genes from the CanFam3.1 genome. Due to incomplete annotation of the canine genome, syntenic genetic regions were identified using the UCSC liftOver tool (86) in the human genome (GRCh38 - hg38) and any additional annotated human protein coding genes in the syntenic region were added to the candidate gene list. Additionally, regions were checked in alternative genome builds CanFam4 (GCA_000002285.3) and CFam_1.0 (GCA_014441545.1) to ensure annotated genes were consistent.

Since our GWAS used only common SNP and no indels or structural variants, we interrogated WGS data from 25 Labrador retrievers using the Integrative Genomic Viewer (IGV) tool (87) to identify additional candidate variants within the initial regions of interest. Dogs were categorized as ‘risk’ or ‘control’ based on whether they carry the GWAS SNP of interest in each region. WGS data were then compared between the two groups to identify additional candidate variants (including indels) which co-segregated at a confidence of $\geq 70\%$ between the groups. For this, confidence was calculated as a percentage of the dog’s homozygous risk for the tagging SNP that were also homozygous for the variant in question, or from the heterozygous risk dogs if homozygous risk dogs were not available. All candidate variants were annotated using Ensembl’s Variant Effect Predictor (VEP) v.109.2 (88). Since there is limited annotation on the CanFam3.1 genome, we tested whether human genetic positions syntenic to canine variants were annotated as regulatory elements or epigenetic hotspots.

Fine mapping

The regions defined by being in high LD with the lead variant were further refined by performing haplotype mapping and variant phasing. Where possible, haplotype mapping using PLINK v.1.9 (60) command `-blocks` identified the haplotype in which the top SNP resided using default parameters (using the method of Gabriel et al. 2002 (89) of defining haplotype blocks, using 95% confidence bounds on the D prime (D’) measure of allelic association to estimate recombination between SNPs). Using the bounds of the inferred haplotype, regions of interest were refined. Variants 100% ‘in-phase’ with the top SNP were also identified and were encompassed into the refined region of interest in cases where they fell out of these haplotype regions. Protein coding genes within those refined regions were identified and further investigated. If refined regions did not harbor (or were not proximal to) a protein-coding gene, they were not explored further.

Investigation of candidate genes

Candidate genes were input into STRING v.11.5 (82)

and PANTHER v.17.0 databases (91). The STRING database was used to investigate protein-protein interactions in the candidate genes, encompassing both direct (physical) and indirect (functional) associations with a confidence threshold of 0.4 ('medium') as the cut off. The PANTHER database was used for the identification of pathways with functional links to the candidate gene list, testing in mouse, dog and human datasets using Fishers exact test with Bonferroni correction for multiple testing.

Publicly available cross-species databases were utilized to investigate candidate genes and to hypothesize mechanisms of action for the variants/genes of interest. Mouse knockout and targeted mutation models were investigated through the International Mouse Phenotyping Consortium (IMPC) (92) and the Mouse Genome informatics Database (MGD) from the Jackson laboratory (93). Summary statistics for previous human GWAS and phenome-wide association studies (PheWAS) were also explored to identify any existing association(s) with obesity and/or related traits. NHGRI-EBI GWAS catalog (94) was interrogated and if variants in candidate genes were identified as associated with traits of interest then they were explored further. GWAS Atlas (95) phenome wide association analysis (PheWAS) results uses GWAS catalog statistics to analyze several phenotypes against a single gene/variant. The *p*-values for PheWAS associations with phenotypes of interest were compared to a Bonferroni-corrected significance threshold. This threshold was calculated using the PheWAS significance threshold of 0.05 and correcting for the number of phenotypes associated with each gene. Genebase v. 0.13.0 (96) is a resource of exome-based association statistics from UK Biobank (UKB) in which we explored whether there were any associations with endocrine/metabolic traits, or physical traits of interest. The Online Mendelian Inheritance in Man (OMIM) catalogue was explored to check the clinical synopsis for candidate genes. More generally, we performed wide-scale literature searching for protein function and candidate gene empirical study.

To investigate if these genes have previously been implicated in other complex disorders, we obtained gene-disease association data from the Developmental Disorders Genotype-to-Phenotype (DDG2P) database (97), a curated resource linking genetic variants with developmental disorders based on evidence from clinical and genetic studies (downloaded from <https://www.deciphergenomics.org/ddd/ddgenes> on 28/10/2024).

Expression analysis for genes of interest

We evaluated tissue-specific expression of candidate genes by employing a multi-species approach to publicly available RNA sequencing data. We tested whether candidate genes were enriched in relevant tissues (brain, adipose tissue) or ubiquitously expressed across tissue types using the Human Protein Atlas (HPA) database v23.0 (98) and the Genotype-Tissue Expression (GTEx) portal v.8 (99) for human data and the MGI - Gene Expression Database (MGI-GXD) from the Jackson laboratory (100) for murine expression.

Canine bulk RNA sequencing data obtained from the BarkBase data repository (41) were used to examine expression across tissues, with data from 5 individuals analyzed across all available tissues, of which hypothalamus was not included (adipose, adrenal gland, brain cortex and cerebellum, kidney, liver, pancreas, pituitary gland, skin, stomach, thyroid gland). Candidate genes of interest were extracted from the BarkBase files and mean expression (in FPKM -

Fragments per kilobase of transcript per million fragments mapped) calculated on a per tissue basis, combining all donors.

To understand gene expression with a higher degree of resolution we interrogated single cell/single nucleus RNA sequencing (scRNAseq/snRNAseq) from the mouse and human brain, focusing on the hypothalamus because of its well characterized role in energy homeostasis. The scRNAseq murine HypoMap (42) represents eighteen datasets encompassing ~385,000 cells from 29 samples from the mouse hypothalamus. To look at hypothalamic data from the human, we utilised the snRNAseq dataset from human HYPOMAP (43). In both datasets, candidate genes were explored for brain region expression specificity (particularly nuclei implicated in energy homeostasis), co-expression with known obesity genes (specifically, those in the leptin-melanocortin pathway) in R v.4.2.2 (62) and visualized using Seurat v.4.0.2 (101). Co-expression of receptors with *Dennd1b/DENND1B* was visualized by taking a subset of neurons expressing >1 *Dennd1b/DENND1B* counts.

RNAscope

Simultaneous detection of Human *MC4R* and *DENND1B* mRNA was performed on formalin fixed paraffin embedded hypothalamic sections using Advanced Cell Diagnostics (ACD) RNAscope® 2.5 LS Duplex Reagent Kit, and RNAscope® probes 1239438-C1 and 538798-C2 (ACD, Hayward, CA, USA). The assay was performed as previously described (70) with one modification, that the time in Amp5 time was extended from 15 to 20 minutes. Slides were imaged on a Slide Scanner Axio Scan.Z1 microscope (Zeiss). Images were taken in regions where positive cells were detected using a 40x air objective and sharpened using the Unsharp Masking processing in ZEN Blue (Zeiss). CZI files were read into QuPath v0.5.1 for analysis. MC4R and DENND1B positive cells were detected by manual inspection and color deconvolution.

Cell culture and transfection

Adherent HEK293 (AdHEK) cells were obtained and cultured as previously described (71). AdHEK were maintained in DMEM-Glutamax media (Sigma-Aldrich) with 10% fetal bovine serum (FBS, Sigma) at 37°C, 5% CO₂. Full-length *DENND1B* cDNA was PCR amplified (Reagents from New England Biolabs) from human cDNA and inserted into a pcDNA3.1 expression construct by restriction enzyme cloning. Human SNAP-GHSR was purchased from Cisbio (UK). For the human SNAP-MC4R construct, MC4R was PCR amplified from a template of Tango-MC4R (Addgene #66430, from Bryan Roth (102)) using reagents from Promega and sequence verified by Source Bioscience. Single nucleotide polymorphisms were introduced by construct high-fidelity PCR amplification using primers containing the polymorphism at the center. Cycle conditions were: 95°C for 30 seconds, followed by 12 cycles of 95°C for 30 seconds, 55°C for 1 minute, and 72°C for 7.5 minutes, followed by hold at 4°C. The original DNA in the PCR product was degraded by DpnI digestion. Resulting expression constructs were sequence-verified (service provided by Department of Biochemistry, University of Cambridge). The pGloSensor-22F plasmid was purchased from Promega. Human DENND1B knockdown was achieved by transfection of 3 unique 27mer siRNA duplexes (SR316077, OriGene Technologies). All transfections were performed using Lipofectamine 2000 (LifeTechnologies) according to manufacturer's instructions.

Gene expression analysis

Overexpression and knockdown of *DENND1B* was assessed using quantitative real-time PCR (qPCR). RNA was extracted using Tri-Reagent (Sigma-Aldrich) and purified with the QIAwave RNA mini kit (Qiagen). RNA was treated with the TURBO DNA-free kit (Thermo Fisher Scientific) and 1 µg made into cDNA using the High-Capacity cDNA Reverse Transcription kit (Thermo). *DENND1B* and *GAPDH* specific primers were designed to achieve an amplicon with a melting temperature between 58°C-62°C, size of approximately 100bp and no secondary structure. Duplicate qPCR reactions of 2 µl cDNA were performed with Fast SYBR Green Master Mix using pre-set quantitative PCR cycling methods on the 7500 Fast Real-Time PCR System (Thermo). *GAPDH* RNA levels were used to normalize gene expression using the 2^{-ΔΔCt} method (103) and visualized as a percentage of the un-transfected cells. Two technical replicates were combined, and an unpaired student's t-test used to determine significant differences in gene expression.

cAMP GloSensor Assays

AdHEK were seeded in six-well plates and 24-hours later transfected with 100 ng pGloSensor-22F and 1000 ng SNAP receptor (human MC4R). When overexpressing *DENND1B*, 1000 ng pcDNA3.1 (empty vector, *DENND1B* WT or *DENND1B* R501C) were co-transfected. For knockdown, 24-hours after DNA transfection, a combination of 3 unique siRNA (10 nM) were transfected. Forty-eight hours after DNA transfection, cells were re-plated in 96-well white plates (Nunc) in FluoroBrite DMEM media (ThermoScientific), supplemented with 10% FBS. After 6 hours, cells were incubated in 40 µL of equilibration media consisting of FluoroBrite DMEM media containing 10% FBS, 2% (v/v) dilution of the GloSensor cAMP Reagent stock solution according to manufacturer's instructions and 0.5 mM IBMX. Cells were incubated for 1 hour at 37°C before basal luminescence was read on a FLUOstar® Omega (BMGLabtech) or a Glomax (Promega, Madison, USA) plate reader for 6 minutes before addition of receptor agonist. For MC4R assays, α-MSH (Bachem AG) was added (100 µM – 10 nM). As a positive control, we used 10 µM forskolin (Sigma). For GHSR assays, MK-0677 (Tocris, Abingdon, UK), a high affinity ghrelin receptor agonist, was added (10 µM - 10nM). Plates were read for a further 25 minutes. Statistics were performed as for the IP-1 experiments below.

IP-1 Gq assay

IP-1 assays were performed with the Cisbio IP-One Gq HTRF kit (Revvity, Codolet, France). Cells were co-transfected with 500 µg of SNAP-tagged GHSR and either 500 µg of pcDNA3.1, *DENND1B*-WT or *DENND1B*-R501C variant. Forty-eight hours after transfection, cells were washed once in 1X PBS and resuspended in FluoroBrite DMEM (Thermo Fisher Scientific) supplemented with 10% FBS and 100 mM lithium chloride at 500 µL/well. The cells were then replated into 384-well plates at 7 µL/ well. MK-0677 (Tocris) agonist dilutions were made in the stimulation buffer provided with the IP-One kit. MK-0677 was added to cells at 7 µL/well and incubated at 37°C for 1 hour. Cells were lysed according to the IP-One kit protocol. IP-1 signal was measured on a BMG Labtech PHERAstar microplate reader. The supplied IP-1 standard concentration range was run in tandem with each experimental repeat. Raw values were interpolated from the IP-1 standard curve. Non-linear log regression fit (dose-response – stimulation curve; three parameters logistic equation) was fitted to the datasets using GraphPad Prism 9. Independent experiments were converted into a combined concentration-response graph

by normalizing data to the fitted maximal response of GHSR co-transfected with pcDNA3.1. Mean pEC50 values from at least four independent experiments before normalization were presented as mean \pm SEM. pEC50 values from independent experiments were grouped and compared by one-way ANOVA.

HILO imaging

AdHEK cells were seeded on 24mm coverslips (VWR) and transfected with 500ng of each plasmid 48 hours prior to experiments. Venus-Kras (obtained from Nevin Lambert, Augusta University) was co-transfected at 500ng per well as a marker of the cell plasma membrane. For knockdown, a combination of 3 unique siRNA (10 nM) were also transfected 24 hours prior to experiments. SNAP-Surface Alexa Fluor 647 (NEB) was diluted 1:1000 in FluoroBrite complete media and applied to cells for 20-minutes, before washing and imaging. Coverslips were mounted onto plastic imaging chambers with a rubber seal and filled with imaging medium (HBSS with 10mM HEPES, Sigma). HILO images were acquired on a custom-built TIRF microscope (Cairn Research) comprising an Eclipse Ti2 (Nikon) equipped with an EMCCD camera (iXon Ultra, Andor), a 488 nm diode laser, a hardware Perfect Focus System, a TIRF iLas2 module, and a 100 \times oil-immersion objective (NA 1.49, Nikon). The objective and samples were maintained at 37°C in a heated enclosure. Images were acquired on MetaMorph software (Molecular Devices) using a frame exposure of 50–200 ms with an image acquired before ligand stimulation and a subsequent image taken every 30s thereafter, up to 20 min. Images were analyzed using ImageJ. Images were stacked, regions of interest corresponding to cell outlines assigned for each cell, and intracellular fluorescence intensity quantified. Data were normalized to the cytoplasmic intensity at time 0. Colocalization between SNAP-labelled receptors and Venus-Kras was measured using the ImageJ plugin JACoP (72). Regions of interest were selected and cropped so that images contained single cells, then JACoP thresholds for each channel were set using the Costes' automatic thresholding (72). Receptor surface expression was compared using a one-way ANOVA and receptor internalization using a Mann-Whitney test.

Human comparative genomics

Canine regions of interest (ROI) bounded by the limits of SNP with $r^2 \geq 0.8$ with the lead SNP at each locus were mapped to orthologous regions in the human genome using the UCSC liftOver tool (86) and integrated with human genome-wide association study (GWAS) and exome-wide association study (ExWAS) data for BMI and also GWAS data of other metabolic traits.

Human GWAS gene prioritization

We interrogated human GWAS data on body mass index (BMI) and waist-hip ratio (WHR) adjusted for BMI in up to 806,834 and 694,649 individuals respectively from the GIANT study (31, 104) and circulating high density lipoprotein (HDL) and triglyceride levels in up to 1,253,277 European individuals from the Global Lipids Genetics Consortium (GLGC) study (105) (up to n = 1,253,277).

To identify independent GWAS signals and prioritize causal gene candidates at the resulting loci, we used the “GWAS to Genes” (G2G) pipeline as described by Kentistou et al. 2023 (39) and discussed in brief below. GWAS summary statistics were filtered to retain variants with a MAF > 0.1%. Quasi-independent genome-wide significant (GWS, at a multiple test

corrected p -value threshold of 5×10^{-8}) signals were initially selected in 1Mb windows and secondary signals within these loci were further selected via conditional analysis in GCTA v.1.93.2 (60), using an LD reference derived from the UKB study. Primary signals were then supplemented with unlinked ($r^2 < 5\%$) secondary signals, whose association statistics did not overtly change in the conditional models and signals were mapped to proximal NCBI RefSeq genes, within 500kb windows.

Independent signals and closely linked SNPs ($r^2 > 0.8$) within the associated loci were annotated if they were coding variants within the identified genes or if they mapped within known enhancers of the identified genes, via activity-by-contact (ABC) enhancer maps (38). Signals were also annotated with their physically closest gene. Gene-level associations were then determined via Multi-marker Analysis of GenoMic Annotation (MAGMA) (106) by collapsing all coding variants within a gene. Genes with FDR-corrected MAGMA $p < 0.05$ were considered associated with the GWAS outcomes. Colocalization analyses between the GWAS and expression- or protein- quantitative trait loci (eQTL or pQTL) data were also performed via SMR-HEIDI (summary data-based Mendelian randomization-heterogeneity in dependent instruments) method v0.68, (107) and the ABF function within the R package “coloc” v5.1.0 (108). For the former, we considered gene expression to be influenced by the same GWAS outcome variants if the FDR-corrected SMR test $p < 0.05$ and HEIDI test $p > 0.001$. For the latter, loci exhibiting an H4 PP > 0.75 were considered to show evidence of colocalization. For eQTL analyses, these were applied for specifically enriched tissues (via LDSC-SEG) (109), as well as cross-tissue meta-analyzed GTEx eQTL data v.7 (110) and data from blood eQTL (111) and Brain-eMeta (112) studies. Lastly, genes residing within each locus underwent prioritization using the gene-level polygenic priority score (PoPS) method (113). Causal candidate genes were then prioritized by overlaying all of the above information and scoring the strength of evidence observed. For further details about the specific application of this method, see Kentistou et al. 2023 (39).

BMI GWAS signals identified as outlined above, were mapped (if present) to the human genomic regions syntenic to the Labrador retriever GWAS ROIs. Orthologous genes mapping within each ROI were also queried for the strength of G2G evidence in the two anthropometric (31, 104) and two lipid (105) GWAS described above.

Rare exome variant association in UK Biobank (UKBB)

We performed exome sequencing-based rare variant burden analyses, as described in Gardner et al. 2022 (114) and described in short below. We queried population-level VCF files data for 454,787 individuals from the UKBB study via the UKBB Research Access Platform (<https://ukbiobank.dnanexus.com>; application number 9905). BMI for all participants was obtained from the UKBB data showcase (field 21001).

We excluded individuals with excess heterozygosity or autosomal variant missingness $\geq 5\%$, based on the available genotyping array data, or those who were not included in the subset of phased samples as defined in Bycroft et al. 2018 (115). We also excluded participants who were not of broadly European genetic ancestry, leaving a total of 421,065 individuals (including related participants) for further analysis. Using BCFTools v.1.9 (78) multi-allelic variants were split and left-normalized, and all variants filtered using a missingness based approach. SNV

genotypes with depth < 7 and genotype quality < 20 or indel genotypes with a depth < 10 and genotype quality < 20 were set to missing (./). We also tested for an expected reference and alternate allele balance of 50% for heterozygous SNVs using a binomial test; SNV genotypes with a binomial test *p*. value $\leq 1 \times 10^{-3}$ were set to missing. Following genotype filtering, variants with > 50% missing genotypes were excluded from further analysis. Variants were then annotated with the Ensembl VEP v.109.2 (88) with the ‘everything’ flag and the LOFTEE plugin (36). For each variant we prioritized a single MANE v.0.97 or VEP canonical ENSEMBL transcript and most damaging consequence as defined by VEP defaults. To define Protein Truncating Variants (PTVs), we grouped high confidence (HC, as defined by LOFTEE) stop gained, splice donor/acceptor, and frameshift consequences. All variants were subsequently annotated using CADD v1.650 (116). After excluding individuals with missing data, 419,692 individuals with BMI measures remained for downstream analysis.

To assess the association between rare variant burden and BMI we implemented BOLT-LMM v2.3.551 (64), using a set of dummy genotypes representing the per-gene carrier status. For the latter, we collapsed variants with a minor allele frequency (MAF) < 0.1% across each gene and defined carriers of variants as those with a qualifying high confidence PTV (HC PTV) as defined by VEP and LOFTEE or “Damaging” variants (DMG), including missense variants with a CADD score ≥ 25 and the aforementioned HC PTVs. Genes with fewer than ten carriers were excluded. BOLT-LMM was run with default settings and the ‘lmmInfOnly’ flag. All analyses were controlled for sex, age, age-squared, WES batch, and the first ten genetic ancestry principal components as calculated in Bycroft et al., 2018 (115).

Gene-level BOLT association summary statistics were then extracted for genes falling within the identified ROIs. We identified 16 orthologous genes with at least 10 carriers of qualifying exome variants, setting the multiple-test corrected threshold at $p \leq 0.0031$ (0.05/16). Any downstream sensitivity analyses (such as leave-one-out tests) were performed in the same individuals as the BOLT associations, but using linear models in R.

Rare variant enrichment analysis in SCOOP

A case-versus-control, rare variant enrichment analysis was performed using independent case and control datasets. The Severe Childhood Onset Obesity Project (SCOOP) dataset is a subset of the Genetics of Obesity Study (GOOS) of UK British origin. This dataset was used as the case dataset and tested against the control dataset obtained from the Genome Aggregation Database (gnomAD) resource (36).

The UK10K SCOOP dataset (n = 982) was obtained from the European Genome-Phenome Archive (EGA; <https://ega-archive.org>; Study ID: EGAS00001000124; Dataset ID: EGAD00001000432, downloaded on 02/03/2021). Exome sequencing and variant calling for this dataset are described previously (117, 118). Only variants that were both polymorphic in SCOOP and marked with a “PASS” flag from GATK (v.1.3-21) were retained. Coordinates were lifted over from build 37 to build 38 using the LiftOverVcf module from GATK.

GnomAD exome r2.1.1 variants (GRCh38 liftOver) were downloaded from the gnomAD webportal (36) (downloaded on 25/10/2022). Gene annotations were based on Ensembl Release 108 (downloaded on 25/10/2022). Again, we used BCFTools v.0.1.18-r579 (78) to subset the

variant data to the genes of interest. Only very rare variants were retained in gnomAD v2.1.1 non-Finnish Europeans. We used a minor allele count ($MAC \leq 3$), equal to an allele frequency of $\sim 0.0026\%$.

The syntenic ROI that were highlighted through the canine analysis were extracted from both case and control datasets using BCFTools v.0.1.18-r579 (78) Ensembl VEP, including the combined annotation dependent depletion (CADD) plugin, was used to annotate all variants. Genotype counts were extracted with PLINK v2.0 and a missingness filter < 0.1 at site-level, applied to both datasets.

The final data analysis steps were run using R 4.1.1. For this analysis, we only retained deleterious variants, defined as moderate/high impact predicted by VEP, and/or $CADD > 25$. We generated cumulative minor allele counts for each of the genes in the regions of interest. For each gene, we applied a one-sided Fisher's test to determine whether there was an excess burden of very rare deleterious variants in the gene of interest in the SCOOP cohort vs gnomAD v2.1.1 non-Finnish Europeans. Alpha significance level corrected for multiple testing was set at $p = 0.01$ ($0.05/5$) and a suggestive significance level at $p = 0.05$.

If suggestive or significant enrichment in the case cohort was identified in any of the genes, we explored the specific variants and carriers further. Familial genotypes and phenotypes were obtained as were clinical phenotypes of the probands. Where possible, we investigated if the genotypes of the variants of interest segregated with obesity status within affected families. Obesity was assessed using the age-adjusted BMI chart relevant to each individual.

Severe Obesity in Pakistani Population (SOPP) cohort

We also investigated genes of interest in the Severe Obesity in Pakistani Population (SOPP) cohort. The cohort is comprised of well-characterized individuals originating from a distinct inbred population who reside in relatively uniform environmental conditions. The selection criteria encompass a BMI or BMI SDS (standard deviation score/Z score compared to WHO global reference data (65, 66) for age >35 or 3.5, respectively; the onset of obesity prior to 5 years of age, coupled with pronounced hyperphagia. Additionally, it is required that the participants exhibit either first- or second-degree parental consanguinity, with parents being either of normal weight or overweight, but explicitly excluding obesity. In all the cases, comprehensive anthropometric measurements are systematically conducted for all participants, alongside the detailed documentation of patient and/or family medical histories. We hypothesized that affected patients would be homozygous for deleterious variants in the genes of interest.

Ancestry diverse public datasets including gnomAD v.2.1.1, NCBI (67), TopMED (68), and NIH ClinVar (69) were used as comparator populations. To gain insight into mechanism of action for candidate variants (in addition to VEP annotation) they were examined using UniProt (119) and AlphaFold (120). UniProt allowed for identification of protein domains impacted by coding variants.

Testing for Purifying Selection

The two constraint metrics pLI (121) and LOUEF (36) were used to evaluate the intolerance of a gene to high confidence LoF mutations in the gnomAD database. We accessed the gnomAD v4.1 dataset (downloaded on 25/10/2024) and implemented thresholds suggested by the developers of each metric (<0.4 for LOUEF and >0.9 for pLI).

Canine polygenic scoring

Using the GWAS results, a polygenic score method was created using the ‘clumping and thresholding’ technique (46). First, an arbitrary p value significance threshold was allocated to the GWAS output ($p < 0.0004$) retaining only SNPs which passed this threshold ($n = 481$). Variants which had a pairwise r^2 measure of LD ≥ 0.7 were treated as the same locus, with the most significant ‘lead’ SNP per locus taken forward to the PRS calculation. Using this approach, 23 SNPs contributed to the primary PRS. Polygenic score was calculated as the sum of the product of each ‘risk’ allele genotype weighted by effect size from GWAS output.

$$\text{PRS} = (\text{Risk allele count}_{\text{SNP1}} * \beta \text{ effect size}_{\text{SNP1}}) + (\text{risk allele count}_{\text{SNP2}} * \beta \text{ effect size}_{\text{SNP2}}) + \text{etc...}$$

PRS was then refined in the secondary/test set of Labrador retrievers ($n = 350$). We retained only SNPs which retained a positive effect size for BCS outcome in the test set (significant association was not a requirement). Sixteen SNPs remained, contributing to the PRS method.

We subsequently tested the PRS method in multiple purebred datasets for prediction of obesity-related phenotypes. Genotypes for the sixteen contributing SNPs were extracted from the multiple datasets using PLINK v.1.9 (61) command `--extract`. If genetic datasets were missing some variants, alternative variants were identified by finding a proximal SNP in high LD ($r^2 > 70\%$) with the original, with a similar effect size and p value in the original GWAS output. If no suitable replacement SNP could be identified, the PRS was calculated with only the retained variants. Polygenic score per individual was calculated as described above.

Polygenic score predictive value

Importantly, prediction of obesity-related traits by the PRS was tested specifically for each purebred population, since specific breed groups act as closed populations. The Labrador retriever discovery ($n = 241$) and test ($n = 350$) populations were treated separately for this analysis. Minimal modelling was performed with known risk factors as described in the Genome Wide Association Study (GWAS) section above. After implementing AIC, the remaining covariates in the minimal model were specific to each population. The polygenic score was then added to the relevant model and tested for its predictive value using the linear model command `lm()` in R v.4.2.2 (62). To test whether individual variants were associated in the different breed groups, analysis was conducted in a similar manner.

For the GRLS, FCR and Pug populations, polygenic score was used to predict BCS and body weight (kg). For GRLS dogs the data were longitudinal so a single point in time was chosen for each participant. For the BCS phenotype, this entailed identifying the youngest age at which the dog had its highest BCS recorded. For body weight, the earliest weight was used. For the pug and FCR populations, phenotypic data were only recorded once.

Of the combined cohort of Labrador retrievers (n = 591), 298 had a food motivation phenotype. After minimal model selection for food motivation (as above), polygenic score was tested for prediction of the food motivation phenotype in these dogs. A linear model was used via the `lm()` command in R v.4.2.2 (62), accounting for relevant covariates. To calculate variance explained by PRS and individual PRS SNPs, the R^2 goodness-of-fit value from the linear regression model was used.

Polygenic score distribution in Labrador retrievers

To assess the distribution of polygenic score in Labrador retrievers and how it related to obesity status, polygenic score was binned into tertials to assign ‘low’, ‘middle’ and ‘high’ PRS groups. BCS was then categorized into three groups: healthy weight (BCS \leq 5), overweight (BCS 6-7), very overweight/obese (BCS $>$ 7).

To test whether genetic factors underlie known risk factors for obesity in Labradors (chocolate coat color and being from the assistance dog group), polygenic score distribution was compared between different demographic groups in the combined set of Labradors (n = 591), incorporating covariates from minimal model selection in that group.

To test whether there was a significant difference in polygenic scores between the categorical coat color groups, ANOVA was performed (accounting for relevant covariates in the minimal model). Assumptions for a parametric ANOVA test were not met, but an alternative non-parametric test allowing for covariates is not available. Therefore, a non-parametric Kruskal-Wallis test was also performed to confirm the ANOVA association. For post-hoc, pairwise comparisons between coat color groups, non-parametric Wilcoxon rank sum tests were used. To assess if there was significant difference in polygenic score between assistance dogs and pet dogs, a Welch’s t-test was performed. However, as above, not all assumptions for a parametric t-test were met so an additional analysis was performed using the non-parametric Wilcoxon rank sum test to confirm association.

High and low risk polygenic scores

Based on polygenic score distribution in the combined set of Labrador retrievers, polygenic score was re-grouped into quintiles. ‘Low risk’ and ‘high risk’ polygenic score group was then assigned based on lower two/upper two quintiles, respectively. Those individuals falling into the third/middle quintile were removed, with the aim of comparing only the ends of the polygenic risk distribution.

The interaction between owner management of the food and exercise environment on obesity outcome was also compared between high and low polygenic risk groups. Owner control was calculated using DORA questionnaire factor scores, as described above. Stepwise minimal model selection was implemented with inclusion of PRS and owner control score interaction. The owner control score was then tested for association in a linear model for BCS in the low-risk vs the high-risk polygenic score groups separately. Mean owner control was also compared between the two data subsets to ensure this was not confounding the findings.

Visualization and statistics

All data analysis and visualization were performed using R/R Studio v.4.2.2 (62). Packages tidyverse v.2.0.0 (122), data.table v.1.14.8, easystats v.0.7.0, ggplot2 v.3.4.4 were used for data analysis and visualization. Specifically, visualization of statistical assumptions was implemented using easystats. Decimal data are reported to two/three significant figures (sf) or two decimal places (dp) as appropriate.

For visualization of phenotype-genotype association and for polygenic score analysis across datasets, residual phenotypes were used. This is a method to better visualize cause and effect for a specific variable by adjusting the phenotype for covariate effects. The command `lm()` is used to perform a linear model and `summary()` used to determine the effect size (β) for all factors. The phenotype of interest (to be plotted) can then be adjusted for the relevant covariates per individual based on the equation: Residual phenotype = phenotype - (β effect covariate A * (covariate A - mean covariate A)) - (β covariate B * (covariate B - mean covariate B)).

Supplementary Text

Lead SNP in canine GWAS and definition of genes of interest

Conditional analysis and haplotype mapping identified seven independent loci, of which five contained protein coding genes (Fig. 1B-K, Table S3). Specifically rs850596879 lies within an intron of *CSNK1A1*, and had MAF = 0.09, $p = 4.73 \times 10^{-6}$, effect size $\beta = 0.80$ BCS; rs853159627 is upstream of *CDH8* and had MAF = 0.25, $p = 7.18 \times 10^{-6}$, $\beta = 0.52$ BCS; rs8831037 is in an intron of *CARD11* and had MAF = 0.16, $p = 1.43 \times 10^{-5}$, $\beta = 0.66$ BCS, with *CNA12*, *BRAT1* and *AMZ1* also present in the mapped region; rs24430444 is in an intron of *DENND1B* and had MAF = 0.06, $p = 5.54 \times 10^{-7}$, $\beta = 0.94$ BCS with *LHX9* also present in the mapped region; rs22632455 is in an intron of *SEMA3D* and had MAF = 0.14, $p = 1.41 \times 10^{-6}$, $\beta = 0.96$ BCS.

Human genetic associations with BMI for canine candidate genes

Human genomic regions orthologous to the fine mapped obesity associated canine loci were identified and interrogated for the presence of GWAS associations for BMI or related metabolic traits. Protein coding genes within those regions were considered as candidate genes and tested for their association with obesity in a GWAS for BMI on 806,834 participants from the GIANT study (31); an exome-wide association study (ExWAS) of rare (MAF < 0.1%) deleterious exome variants from 454,787 individuals from the UK Biobank study (UKB) (32–34); the Severe Childhood Onset Obesity Project (SCOOP-UK) (35) (n = 982) in which we tested for enrichment of very rare (MAF < 0.0026%), predicted deleterious (CADD ≥ 25) variants compared to reference exomes of similar ancestry (gnomAD v2.1.1, n = 56,885); and the Severe Obesity in Pakistani Population (SOPP) cohort of patients with severe, early onset obesity from a highly consanguineous population, all with unaffected parents and no prior genetic diagnosis, in which we focused on homozygous, predicted deleterious variants (44, 45).

In the GIANT GWAS there were associations with human BMI for variants in *DENND1B* (further details in main text) and *CARD11* (rs7811825, MAF = 0.14, $p = 9.02 \times 10^{-9}$, $\beta = -0.0146$ kg/m², Table S7). In each case, the human effect sizes were small and the statistical association significant but modest, meaning these associations were likely to be overlooked as having potential to meaningfully influence BMI at a population level. *CARD11* encodes a signaling scaffold protein which is best understood for its role in the adaptive immune response (123) and it is not clear how this might link mechanistically to increased BMI.

Rare variant enrichment was evident in the SCOOP cohort for *CSNK1A1* ($p = 0.006$, Fig. 2A, C) due to the presence of two variants which were absent from gnomAD (Table S11), map to the protein kinase domain and have high CADD scores (>27). A heterozygous variant *CSNK1A1* p.V87G was present in a proband with a complex phenotype that included severe, early onset obesity, speech delay, constipation, sleep disturbance and behavioral problems. At presentation at 3 years 4 months old, her weight was 23.3 kg, BMI 22.13 kg/m², BMI standard deviation score (SDS) 3.33. The two other family members from which DNA was available had the reference allele and were normal weight (mother BMI 23.6 kg/m² and brother age 4 years 4 months BMI 15 kg/m², BMI SDS -0.58). A second heterozygous variant *CSNK1A1* p.G80D was identified in a patient with severe, early onset obesity who presented aged 12 years with weight 83.6 kg, BMI 48.13 kg/m², BMI SDS 4.28. Her heterozygous, mother was also obese (BMI 36.2 kg/m²) and had type 2 diabetes. The father harbored only the reference allele and was borderline obese (BMI 31.2 kg/m²). Further information is available in Table S14. *CSNK1A1* is involved in

the development and function of hypothalamic POMC neurons (124) and has previously been implicated in regulation of the adipokine adiponectin (125), both of which provide plausible, if speculative, mechanistic links to energy homeostasis.

Rare variants were enriched within *CDH8* in both the UKB BMI analysis and SCOOP (Fig. 2B, D). In UKB the combined effect size of rare protein truncating variants (2.21 kg/m^2 , $p = 0.0025$, Fig. 2B) was predominantly driven by one frameshift variant of large effect (16:61653821:CA>C, L729X carried by three individuals, $\beta = 12.76 \pm 2.54 \text{ kg/m}^2$, $p = 4.90 \times 10^{-7}$, Fig. 2E) and the association was attenuated after the exclusion of this variant. In SCOOP there was a nominally significant enrichment for rare, predicted deleterious variants in *CDH8*, each in a single heterozygous proband ($p = 0.045$, Fig. 2A), specifically *CDH8* p.R479* (also detected in a single heterozygous individual in gnomAD, Table S11), *CDH8* p.R431W and p.R126W (both absent from gnomAD, Table S11). Pedigree analysis of affected families showed imperfect segregation of these variants with obesity (Table S14).

CDH8 encodes a brain-specific cadherin which has been shown to have a role in forming the anatomical and functional features critical for the formation of action-outcome associations (126). There are also >90 GWAS associations with *CDH8* in GWAS catalog (94), many of which are for traits related to behavior or educational attainment. Given the importance of the central nervous system in maintaining energy homeostasis (2), and the well described links between obesity and both eating behavior and educational attainment (59), the links we report between BCS/BMI and obesity seem credible.

At the fifth locus, the canine association lay within *SEMA3D* for which we found no new human obesity association. The gene is a member of the Semaphorin 3 gene family previously implicated in causing severe, early onset human obesity and which have been shown to affect the development of hypothalamic neurons (127). In the same study, knockout of *SEMA3D* caused an increased body fat percentage in zebrafish. A single patient from the SOPP cohort was homozygous for a deleterious *SEMA3D* variant which was rare in control populations (*SEMA3D* p.D380H, Table S11). She had a BMI of 36 at 15 years old and her normal weight parents were both heterozygous for the variant.

We considered whether the rarity of deleterious variants in the five candidate genes might indicate they were under purifying selection, and whether they might have been overlooked as obesity genes because they caused rare, complex syndromes. We used pLI (121) and LOUEF (36) constraint metrics to demonstrate that three of the genes of interest (*CARD11*, *CDH8*, *CSNK1A1*) show evidence of purifying selection but the others (*DENND1B* and *SEMA3D*) do not (Table S15). None of the five genes are currently listed as known developmental disorder genes in the Developmental Disorders Genotype-to-Phenotype (DDG2P) database (97), a curated resource linking genetic variants with developmental disorders based on evidence from clinical and genetic studies. We have yet no explanation for the rarity of the *DENND1B* or *SEMA3D* mutations reported and further phenotypic characterization in animal models and other patient cohorts is needed.

CARD11 is a known cancer causing gene, from tier 1 of the Cancer Gene Census (128). *CSNK1A1*, *DENND1B* and *SEMA3D* have all in the past been implicated in cancer pathogenesis, for example by being differentially regulated in certain cancer types, but are not listed in the Cancer Gene Census (129–133). We are not aware of any association between *CDH8* and cancer, although other members of the cadherin family are frequently implicated in metastatic

progression in cancer. In this context, it is of note that there was no history of cancer in the human variant carriers for any of the genes, nor any dogs included in the GWAS or follow-on studies (for which absence of systemic disease was an inclusion criteria).

Canine polygenic risk score application and utility

The Labrador PRS predicted BCS and body weight in a closely related breed, the golden retriever ($n = 1765$), although to a lesser extent. In more distantly related breeds (flat-coated retrievers, $n = 196$, pugs, $n = 139$), it was not predictive (Fig. 4A) (134). This reflected that in golden retrievers, 4 SNPs retained a significant effect on BCS, but none were significant in flat-coated retrievers or pugs (Table S13).

Polygenic risk score contributes to known within-breed variation

Labrador retrievers can be registered as one of three colors, black, yellow, and chocolate, with chocolate dogs being the most obesity-prone and food motivated (28). Chocolate Labradors in the discovery GWAS had higher BCS than other colors (mean BCS = 6.2 versus 5.5, $p < 0.05$, Table S2), as has been reported previously (28, 135). We considered whether this might be because the owners of chocolate colored dogs managed them less strictly (for instance, because they are popular family dogs). Using the DORA questionnaire (24) to quantify owner control of diet and exercise we showed there was no significant difference in owner management of diet and exercise between dogs of different coat colors (ANOVA $p = 0.95$). However, chocolate colored dogs had significantly higher PRS than dogs of other colors (mean PRS: red = 5.46, black = 6.41, yellow = 6.66, chocolate = 7.38, Fig. 4C). MDS plots reveal subtle clustering of chocolate dogs within the pet population which may explain this higher PRS (Fig. S11A).

Purpose-bred assistance Labradors had higher BCS than those kept as pets (mean BCS: assistance = 5.99, pet = 5.48, $p < 0.001$). This was associated with significantly higher PRS (mean PRS assistance = 8.36, mean PRS pet = 6.32, $p = 2 \times 10^{-12}$, Fig. 4B). When assistance dogs were included in MDS plots, stratification is evident as compared to the pet Labradors (Fig. S11B). No assistance dogs were included in the discovery GWAS.

PRS, food motivation and gene environment interaction

Healthy weight Labradors were predominantly from the middle or lower tertiles of PRS. The majority of dogs which were obese were from the highest PRS tertile, with the remainder from the middle tertile (Fig. 4F). In 248 Labradors, the PRS positively predicted food motivation score (quantified as per materials and methods, Fig. 4D, $p = 0.03$) suggesting increased food-seeking behavior drives higher food intake and BCS in genetically predisposed dogs. Using a validated, owner-reported questionnaire we captured information regarding dogs' activity levels (during 'walks', whilst working, performing agility etc) and about the stringency with which owners controlled their dogs' access to food (24). For dogs in the highest two quintiles of PRS, more assiduous management of diet and exercise mitigated the development of increased BCS ($\beta = -0.020$, $p = 0.0077$) but it had no significant effect on BCS ($\beta = -0.007$, $p = 0.14$) in dogs in the lowest two quintiles (Fig. 4E). Reflecting this, PRS and owner control of the food and exercise environment displayed a significant interaction within the linear model ($p = 0.0028$).

Body Condition Score



UNDER IDEAL

- 1 Ribs, lumbar vertebrae, pelvic bones and all bony prominences evident from a distance. No discernible body fat. Obvious loss of muscle mass.
- 2 Ribs, lumbar vertebrae and pelvic bones easily visible. No palpable fat. Some evidence of other bony prominences. Minimal loss of muscle mass.
- 3 Ribs easily palpated and may be visible with no palpable fat. Tops of lumbar vertebrae visible. Pelvic bones becoming prominent. Obvious waist and abdominal tuck.

IDEAL

- 4 Ribs easily palpable, with minimal fat covering. Waist easily noted, viewed from above. Abdominal tuck evident.
- 5 Ribs palpable without excess fat covering. Waist observed behind ribs when viewed from above. Abdomen tucked up when viewed from side.

OVER IDEAL

- 6 Ribs palpable with slight excess fat covering. Waist is discernible viewed from above but is not prominent. Abdominal tuck apparent.
- 7 Ribs palpable with difficulty, heavy fat cover. Noticeable fat deposits over lumbar area and base of tail. Waist absent or barely visible. Abdominal tuck may be present.
- 8 Ribs not palpable under very heavy fat cover, or palpable only with significant pressure. Heavy fat deposits over lumbar area and base of tail. Waist absent. No abdominal tuck. Obvious abdominal distention may be present.
- 9 Massive fat deposits over thorax, spine and base of tail. Waist and abdominal tuck absent. Fat deposits on neck and limbs. Obvious abdominal distention.

German A, et al. Comparison of a bioimpedance monitor with dual-energy x-ray absorptiometry for noninvasive estimation of percentage body fat in dogs. *AJVR* 2010;71:393-398.
 Jesuitte L, et al. Effect of breed on body composition and comparison between various methods to estimate body composition in dogs. *Res Vet Sci* 2010;88:227-232.
 Kealy RD, et al. Effects of diet restriction on life span and age-related changes in dogs. *JAVMA* 2002;282:1315-1320.
 Lallierme DP. Development and validation of a body condition score system for dogs. *Canine Pract* 1987;22:10-15.

©2013. All rights reserved.



Fig. S1.

Body condition scoring descriptors and images used for scoring adiposity in dogs. The chart in is part of the Global Nutrition Committee Toolkit provided courtesy of the World Small Animal Veterinary Association.

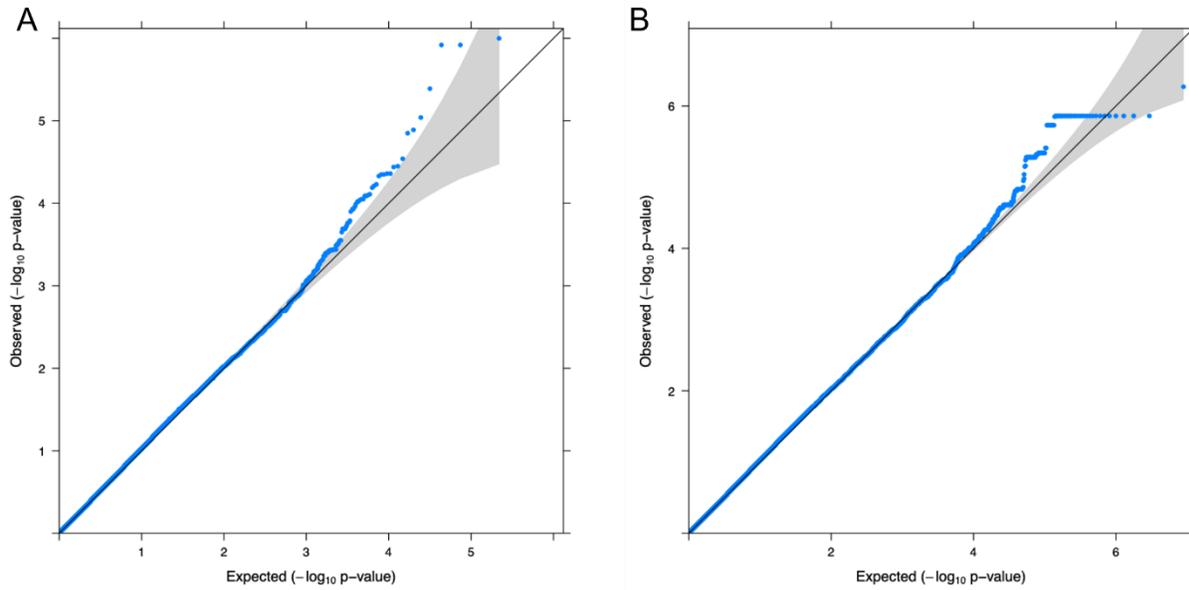


Fig. S2.

Quantile-Quantile (QQ) plot resulting from the GWAS in Labrador retrievers for body condition score (BCS) in 241 individuals. Generated LMM GWAS output applying GCTA MLMA-LOCO, using only genotypic data which were (A) directly genotyped or (B) including imputed genotypes.

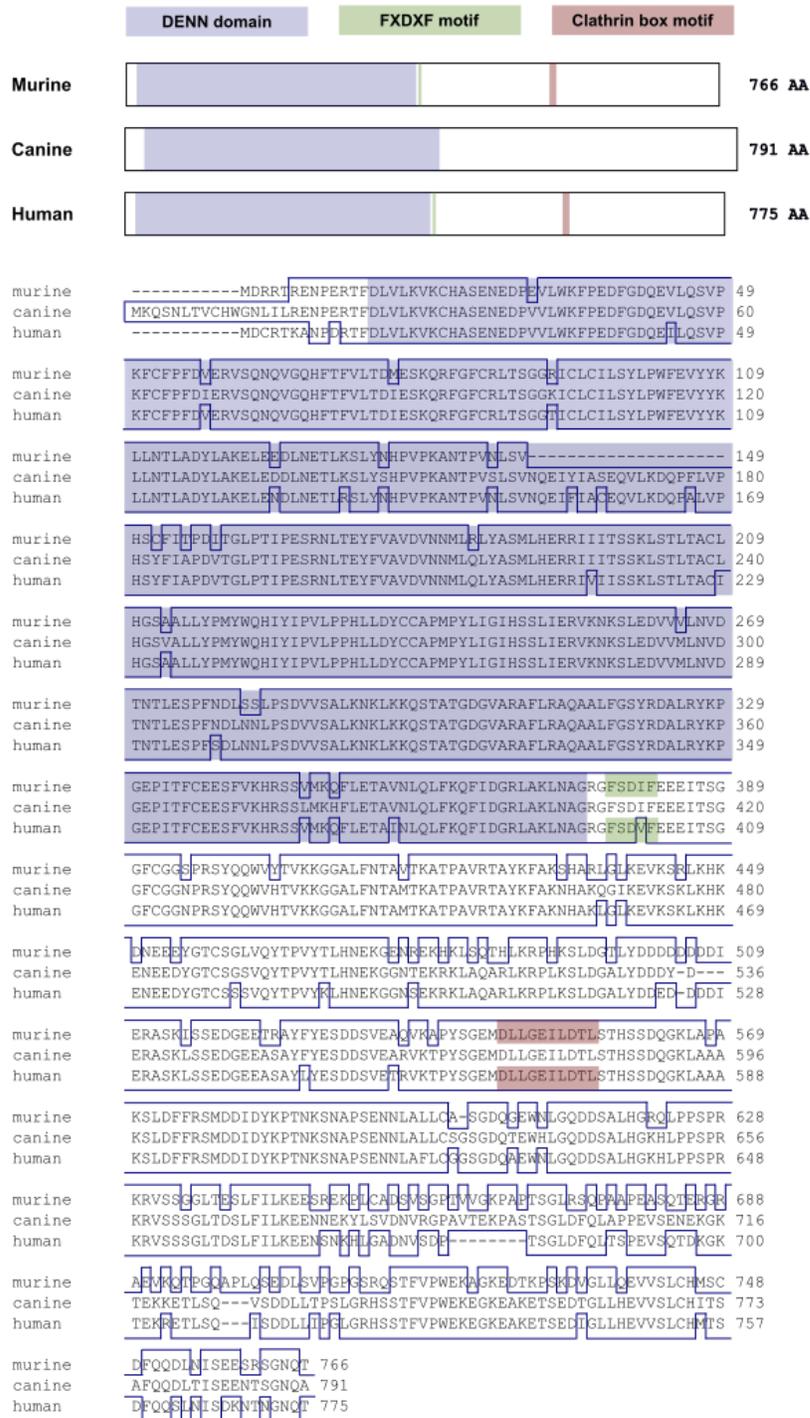


Fig. S3.

There is high homology between canine DENND1B and human and murine forms. Protein alignment of all three species is shown. Overall, canine has highest homology with human (89.4%) compared to mouse (82.7%), with near perfect homology in the functionally important DENN domain and within the FXDXF and clathrin box motifs.

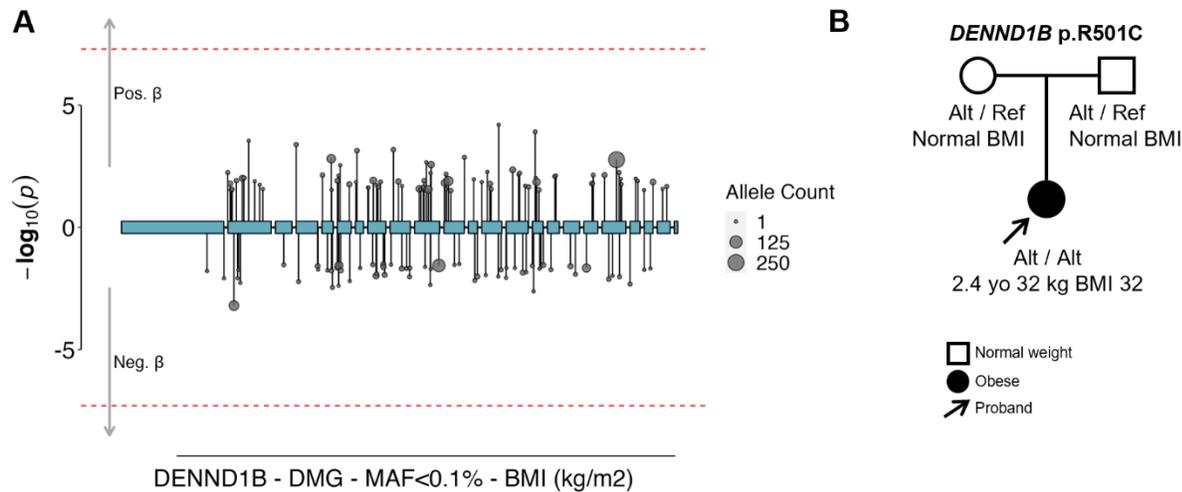


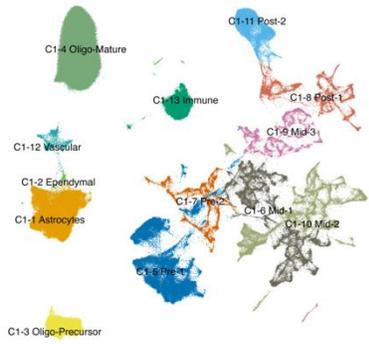
Fig. S4.

DENND1B associations with human obesity. (A) There was a nominally significant enrichment of damaging mutations in *DENND1B* with BMI in UK Biobank ($p = 0.0087$). Qualifying variants had a minor allele frequency (MAF) < 0.1% and were annotated as either high-confidence protein truncating variants or missense variants with a high CADD score (≥ 25). Each variant association is represented by a circle and vertical line, the line length indicates the p -value ($-\log_{10}$), in the direction of its effect on BMI in carriers of the rare allele, and the circle size indicates the number of carriers of each variant (allele count). The horizontal red lines indicate p -values smaller than 1.6×10^{-6} . Exons are indicated by the blue boxes. (B) *DENND1B* p.R501C was identified in a single proband in the Severe Obesity Pakistani Population (SOPP) cohort whose normal weight parents were heterozygous for the variant. No alternative genetic diagnosis was revealed by exome sequencing.

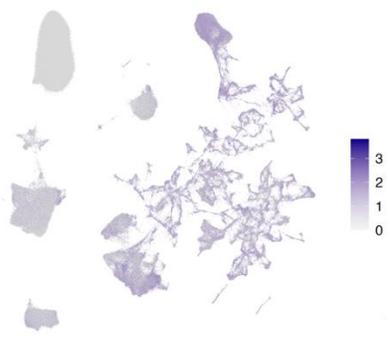


Fig. S5.

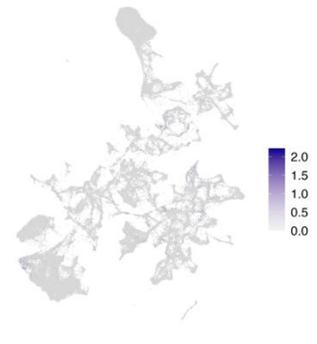
Dennd1b is widely expressed in the murine hypothalamus (HypoMap) and is co-expressed with multiple receptors involved in energy homeostasis. UMAP plots show co-expression of *Dennd1b* with *Ghsr*, *Mc3r*, *Npy1r*, *Npy5r*, *Lepr*, *Insr*, *Htr1b*, *Htr2c*, and *Glp1r*. UMAP plots show expression of each receptor in the subset of neurons from HypoMap that expressed *Dennd1b* transcripts.



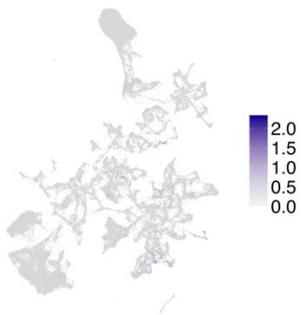
DENND1B



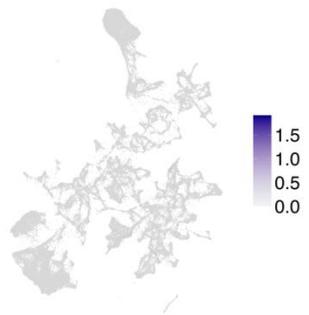
MC4R



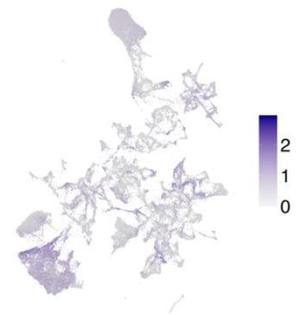
GHSR



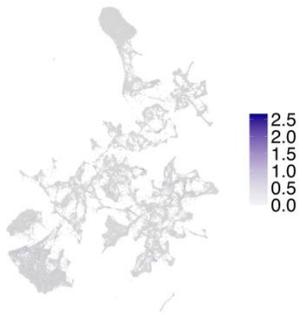
MC3R



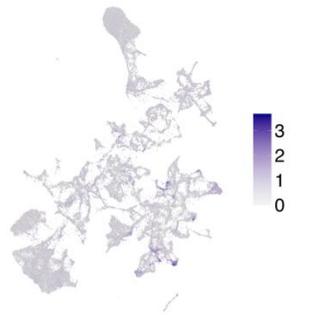
NPY1R



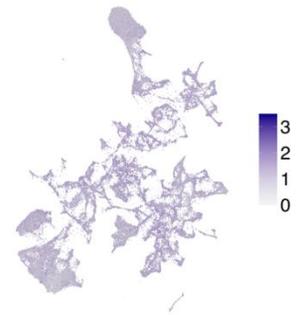
NPY5R



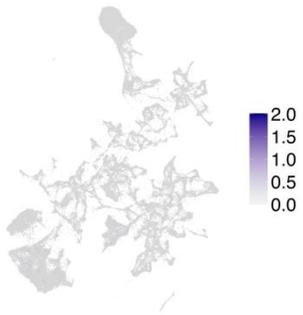
LEPR



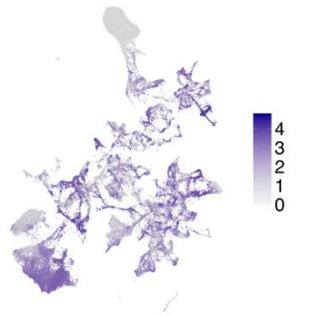
INSR



HTR1B



HTR2C



GLP1R

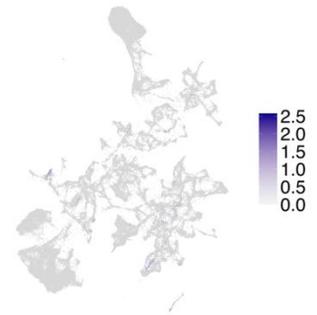


Fig. S6.

RNAseq shows *DENND1B* is widely expressed in the human hypothalamus and is co-expressed with *MC4R*. UMAP plots show human HYPOMAP single nucleus RNA sequencing data extracted from Tadross et al., 2023 (43) colored by general cell type (C1 clustering level), log-normalized expression of *DENND1B* in the human hypothalamus and log-normalized expression of *MC4R* and other hypothalamic receptors in the *DENND1B*⁺ neuronal subset.

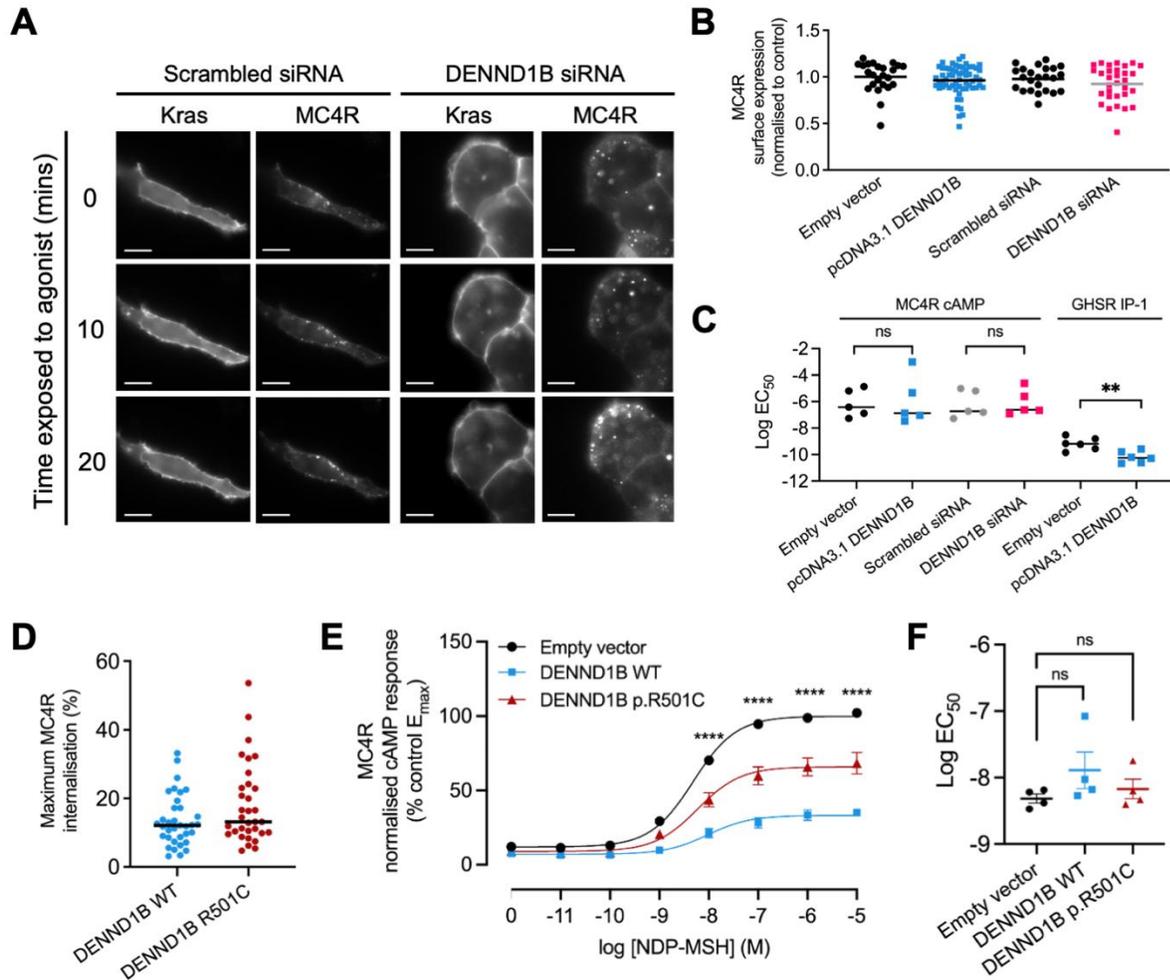


Fig. S7

Additional data on DENND1B effect on MC4R. (A) Representative HILO images of HEK293 with DENND1B knocked down (KD) using siRNA or control condition (scrambled siRNA). MC4R (SNAP-647) and plasma membrane marker Kras (Venus) were expressed and show colocalization at time point 0. Decreased ligand-induced MC4R internalization was observed in conditions of DENND1B knockdown ($p = 0.044$). Scale bar = 5 μm . (B) Basal MC4R surface expression (via colocalization with Kras) was quantified at time point 0 in control cells (empty vector – EV, scrambled siRNA), cells overexpressing DENND1B and cells in which DENND1B was knocked down using siRNA. Control and DENND1B conditions were not significantly different from one another (pairwise comparisons to internal control). (C) MC4R cAMP and GHSR IP-1 EC_{50} for dose response curves. No significant difference was observed between the control and either DENND1B overexpression or knockdown for MC4R. Overexpression of DENND1B significantly reduced the GHSR IP-1 log EC_{50} compared to the control ($p = 0.004^{**}$). Data from 5-6 independent experiments. Conditions compared using a student's t test. (D) There was no difference in maximum internalization of MC4R after overexpression of DENND1B WT or p.R501C (determined by localization to cell surface Kras in HILO images after 20 min α -MSH-stimulation). Results expressed as the percentage of the unstimulated receptor surface expression. (E) Ligand-induced cAMP response of MC4R in HEK293 cells after

overexpression of WT DENND1B or DENND1B p.R501C. Each condition was significantly different to the other at the highest four concentrations of ligand (two-way ANOVA with multiple comparisons, $p < 0.0001$ ****). (F) There was no difference between the log EC₅₀ of the dose response for the same experiments (4 independent experiments, one-way ANOVA multiple comparisons).

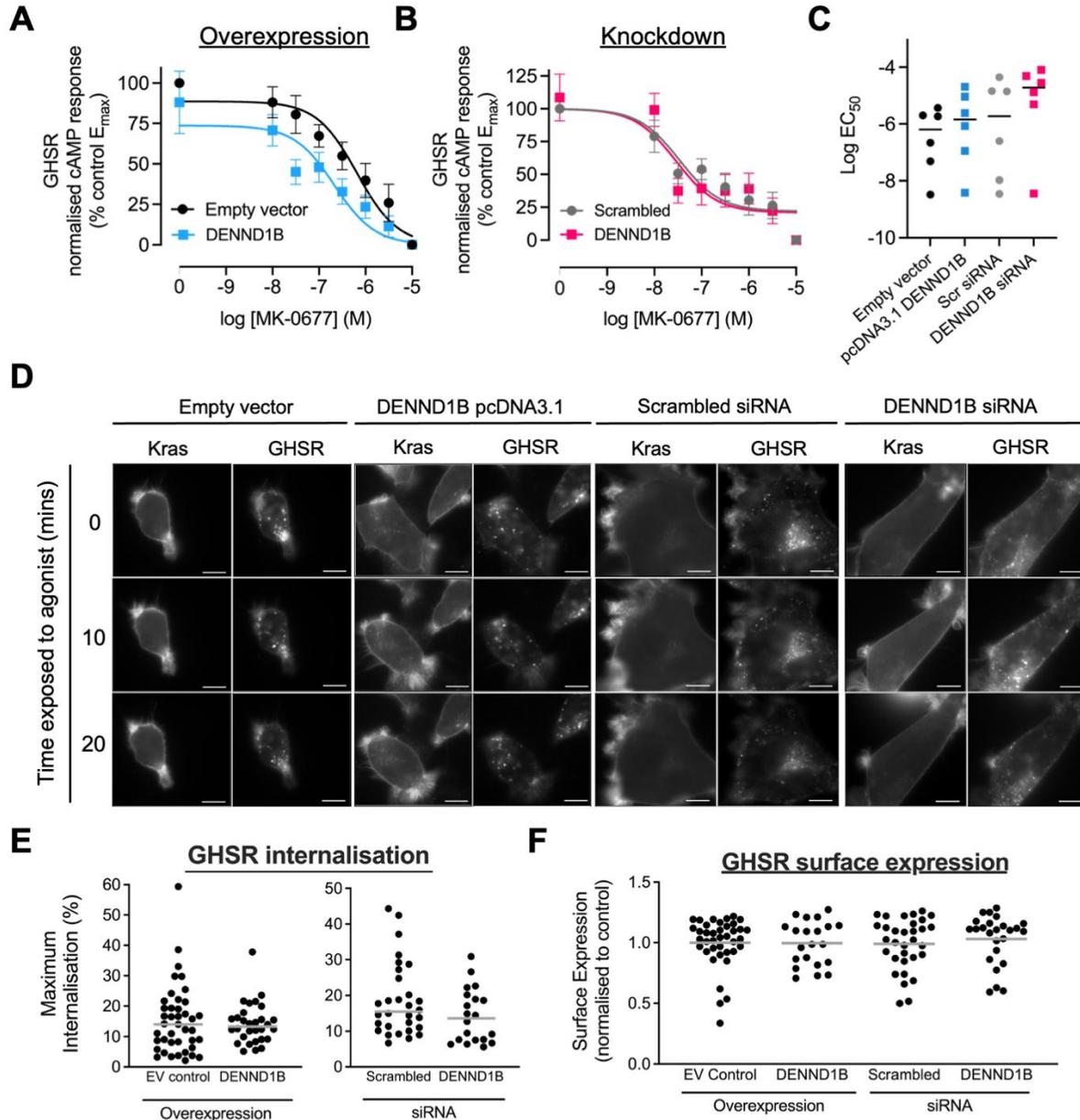


Fig. S8

Altering DENDD1B expression had no effect on GHSR cAMP signaling internalization or cell surface expression. GHSR-activated decrease in cAMP production in HEK293 cells expressing GHSR is shown in conditions of (A) DENND1B overexpression or (b) DENND1B siRNA knockdown (KD), after prior stimulation with forskolin. (C) Log EC_{50} from 6 independent experiments were compared (Student's t test, no significant differences). (D) Representative HILO images of HEK293 cells expressing GHSR (SNAP-647) and plasma membrane marker Kras (Venus) showing colocalization at time point 0, with GHSR internalisation following agonist stimulation. Scale bar = 5 μ m. There was no difference between (E) internalisation or (F) basal GHSR surface expression at time point 0 between control cells (empty vector – EV, scrambled siRNA) versus cells overexpressing DENND1B or with DENND1B knockdown using siRNA (pairwise comparisons).

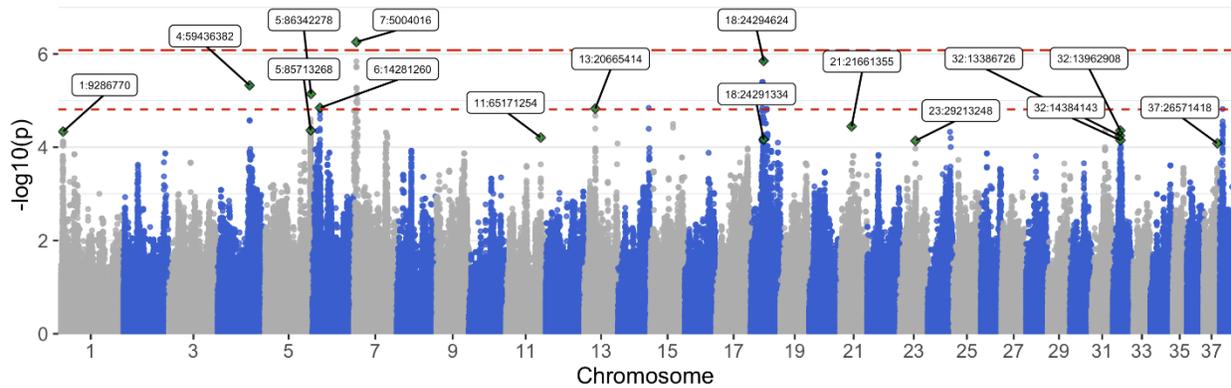


Fig. S9

Using the GWAS results, a polygenic score method was created using the ‘clumping and thresholding’ technique. First, an arbitrary p value significance threshold was allocated to the GWAS output ($p < 0.0004$) retaining only SNPs which passed this threshold ($n = 481$). Variants which had a pairwise r^2 measure of LD ≥ 0.7 were treated as the same locus, with the most significant ‘lead’ SNP per locus taken forward to the PRS calculation. Using this approach, 23 SNPs contributed to the primary PRS. Polygenic score was calculated as the sum of the product of each ‘risk’ allele genotype weighted by effect size from GWAS output. PRS was then refined in the secondary/test set of Labrador retrievers ($n = 350$). We retained only SNPs which retained a positive effect size for BCS outcome in the test set (significant association was not a requirement). The 16 remaining SNP retained in the final PRS are shown.

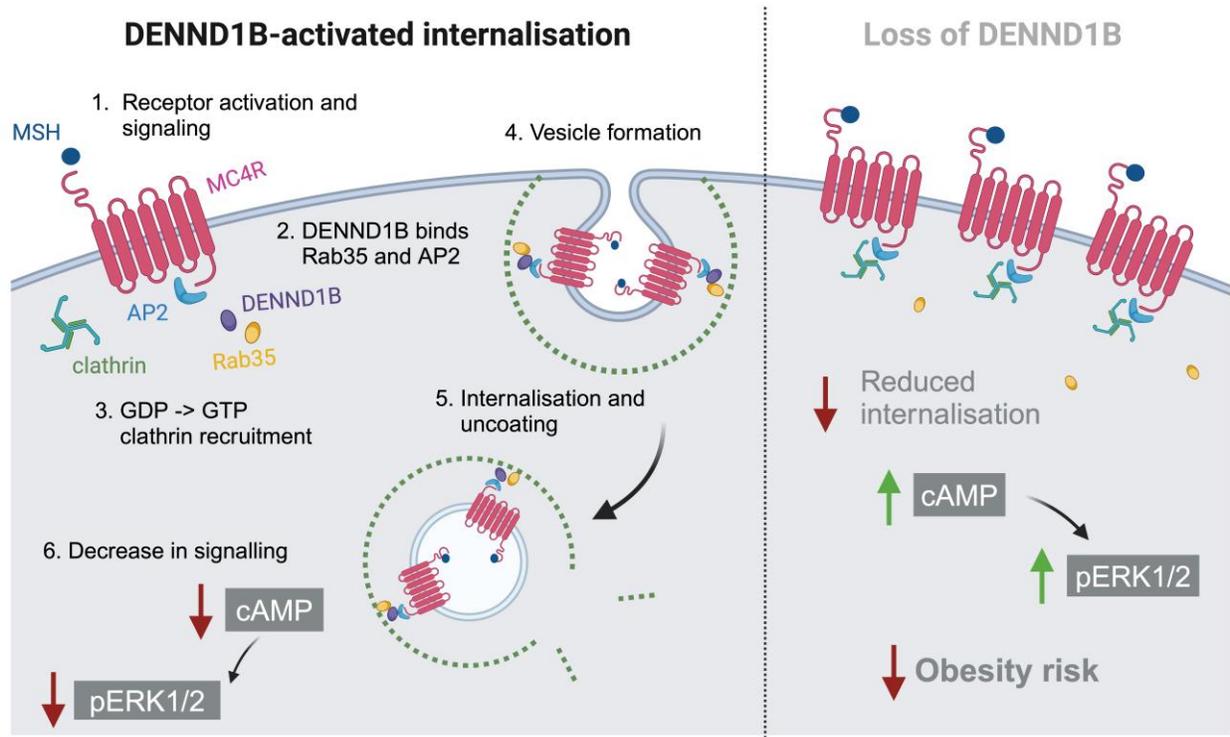


Fig. S10

Proposed mechanism linking regulation of obesity risk to DENND1B to energy homeostasis by regulation of MC4R signaling. Hypothalamic MC4R integrate signals from POMC and AgRP neurons to regulate energy homeostasis, exerting an anorexigenic effect. (1) α -MSH binds MC4R, activating signaling via Gs protein recruitment. Signaling is inactivated by β -arrestin recruitment triggering recruitment of clathrin machinery (AP-2 and clathrin). (2) DENND1B binds AP-2 and Rab35 and DENND1B GEF activity converts GDP to GTP and activates Rab35. (3) This initiates Rab35-dependent (4) vesicle formation and (5) MC4R endocytosis. (6) This downmodulation of MC4R surface expression reduces its signaling. The internalized MC4R vesicle is then clathrin uncoated. Right hand side: Reduced DENND1B impairs ligand-activated internalization and prolongs anorexigenic signaling to decrease risk of obesity. AP-2 – AP2 adaptor complex, GDP – guanosine diphosphate, GTP – Guanosine triphosphate, GEF – Guanine nucleotide exchange factors, cAMP – Cyclic adenosine monophosphate, pERK1/2 - active extracellular signal-regulated kinase 1/2. Figure created using BioRender.com.

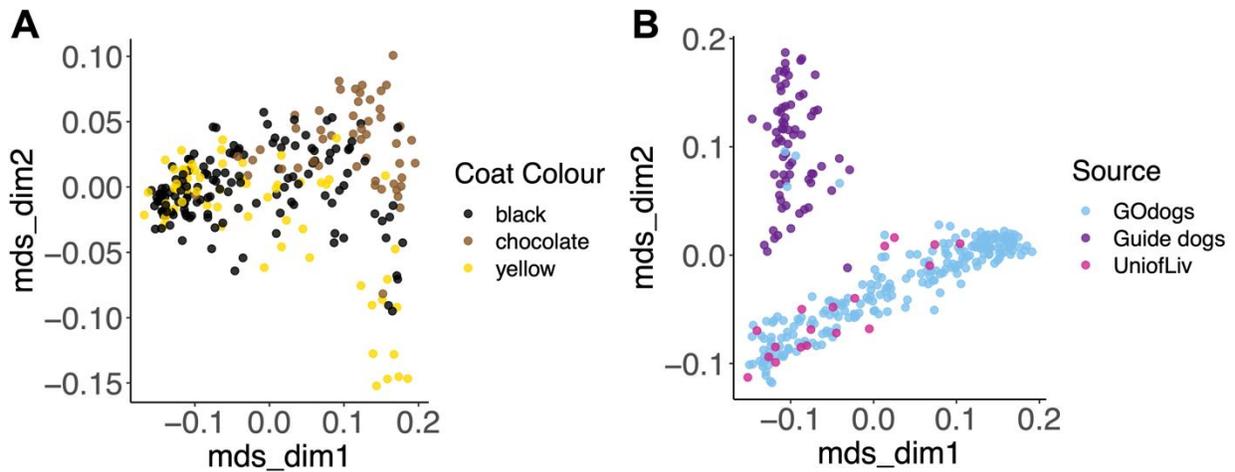


Fig. S11

MDS plots of Labrador population. Dimensional reduction plots presenting the genetic structure of the Labrador population in two dimensions. (A) The discovery GWAS set of dogs ($n = 241$) with points colored by coat color (black, chocolate, yellow). (B) Points are colored by source of data (GOdogs research group internal recruitment; assistance dogs collected through collaboration with Guide Dogs UK; UniofLiv, pet dogs through collaboration with the University of Liverpool). Only dogs which were identified as pets and which clustered with the larger pet population cluster were included in the discovery GWAS. The four pet dogs which clustered with assistance dogs were discovered (after owner interviews) to be retired or ‘failed’ guide dogs and were excluded from the discovery GWAS. MDS – multi-dimensional scaling; mds_dim – MDS dimension.

Table S1 – S15

Table S1: Items included in the Dog Obesity Risk Assessment (DORA) Questionnaire which contributed to owner control and food motivation scores (24). Owners respond on a 4 or 5 point Likert scale according to their level of agreement with each statement and responses are translated into a percentage. Overall food motivation score or owner control score are calculated as the mean response to the items within that category. For some items, the scores are reversed (as indicated) to ensure item scores are in the congruent direction for the overall factor score.

Table S2: Summary statistics for phenotypes in the Labrador populations reported in the study. In the discovery GWAS all dogs were kept as companions. In the wider population (used for examining the utility of PRS) there were also assistance dogs in the population. For each group, the summary statistics are provided for all dogs, and subgroups by role and color. 'Red' is not a recognized color for this breed in the Kennel Club registration system (considered officially as yellow) but was reported by owners and so is listed separately where it occurred. PRS, polygenic risk score; SD, standard deviation; BCS, body condition score.

Table S3: Regions of interest from BCS GWAS. Regions of interest tagged by lead SNP variant (in format chromosome:position) from GWAS for BCS via two methods: r^2 measure of LD regions ≥ 0.8 (80%) or haplotype mapping using PLINK. The 'refined' region was either the mapped haplotype plus in-phase SNPs or, if the SNP did not map to a haploblock, the $r^2 \geq 0.8$ region.

Table S4: Candidate causative variants identified through analysis of high coverage whole genome sequencing data from 25 Labrador retrievers. Labrador WGS ($>30x$) BAM files were visualized in IGV to detect structural variants using the 'soft clip' function which identified multiple regions suggestive of the presence of a moderate to large insertion or deletion mutation which are denoted below as 'poorly defined indels'. GWAS region of interest is categorized by top SNP from each locus. Distance from tagging SNP is determined by position of SNP or first bp of insertion/deletion. Ref/Alt indicates reference or alternative base on the Canfam3.1 genome. Control/risk label is determined based on whether the variant is inherited with the risk or non-risk allele for the lead GWAS SNP for BCS in the Labrador GWAS (for $>70\%$ of alleles).

Table S5: Results of investigation of positional candidate genes for evidence of effects related to obesity. PheWAS data were interrogated in GWAS Atlas (<https://atlas.ctglab.nl/PheWAS>). To investigate the effect of gene knockout or other genetic manipulation in mouse, we used web portals of the International Mouse Phenotyping Consortium (<https://www.mousephenotype.org/>), the Mouse Phenome Database (<https://phenome.jax.org/>) and performed searches using PubMed. The Online Mendelian Inheritance in Man (OMIM) database was searched to determine whether mutations in each gene were known to cause Mendelian disease in humans. The GTEx portal (<https://www.gtexportal.org/>) and Human Protein Atlas (<https://www.proteinatlas.org/>) were interrogated to determine tissues in which gene expression was enriched. - indicates no phenotype of interest. N/A indicates data not available.

Table S6: Summary of human BMI associations for genes identified on the canine GWAS.

Table S7: Predicted causative gene for obesity-related traits in human loci syntenic to canine obesity GWAS loci. We interrogated human GWAS data on body mass index (BMI) and waist-hip ratio (WHR) adjusted for BMI in up to 806,834 and 694,649 individuals respectively from the GIANT study and circulating high density lipoprotein (HDL) and triglyceride levels in up to 1,253,277 European individuals from the Global Lipids Genetics Consortium (GLGC) study (up to $n = 1,253,277$). G2G uses a variety of information to predict causative gene at each locus. G2G – GWAS to Genes; GWAS – genome wide association study; BMI – body mass index; HDL – high density lipoprotein; WHR – waist-hip ratio; PoPS – polygenic priority score; pQTL – protein quantitative trait loci; eQTL – expression quantitative trait loci; ABC – activity by contact; MAGMA – Multi-marker Analysis of GenoMic Annotation; LD – linkage disequilibrium.

Table S8: Canine GWAS obesity candidate genes are expressed in a range of canine tissues. Data obtained from the BarkBase database (41) and expressed in fragments per kilobase of transcript per million read pairs (FPKM). NA not in database.

Table S9: Candidate gene expression in single cell RNA sequencing from mouse and human hypothalamus. We interrogated HypoMap, a unified single cell gene expression atlas of the mouse hypothalamus (42) and snRNAseq of the human hypothalamus in data from the human HYPOMAP (43). Results shown for each gene of interest show expression in all cell types, all neurons, Pomc-expressing cells and Mc4r-expressing cells. GHSR is included since it was used, like MC4R, in molecular studies of DENND1B function. Presented as a percentage of all cells in subgroup to 2 decimal places.

Table S10: Co-expression of DENND1B and other candidate genes with POMC, MC4R and GHSR in the human hypothalamus. In snRNAseq data from the human hypothalamus, cells were clustered according to gene expression patterns (42). The table shows cell type and gene markers defining each cluster along with the number of nuclei contributing to that data. Data shown: % <gene>+ (% of cells / cluster expressing gene); AverageExpression_<gene> (Average log-normalized expression of gene / cluster); % <gene1>+<gene2>+ (Percentage of cells / cluster co-expressing both genes).

Table S11: Rare variants predicted to be deleterious and found to be associated with obesity in either SCOOP or SOPP data sets. To test if regions and genes identified on the canine GWAS were also relevant to human obesity, we examined regions orthologous to those mapped in dogs in human data sets. Specifically, we tested for enrichment of rare (MAF <math><0.0026\%</math>), predicted deleterious (CADD ≥ 25) variants in the Severe Childhood Onset Obesity Project (SCOOP-UK) (35) ($n = 982$), compared to reference exomes of similar ancestry (gnomAD v2.1.1, $n = 56,885$); and investigated the Severe Obesity in Pakistani Population (SOPP) cohort of patients with severe, early onset obesity from a highly consanguineous population, all with unaffected parents and no prior genetic diagnosis (44, 45), in which we focused on homozygous, predicted deleterious variants. The table shows variants identified using that approach and their frequency in different control populations including gNomad Non-Finnish European (NFE), NCBI Allele Frequency Aggregator (ALFA), Trans-Omics for Precision Medicine (TOPMed) participants and the 1000 Genomes database.

Table S12: Predictive value of PRS demonstrated by improved goodness of fit for obesity-related traits in Labrador retrievers when it is included in the model. The R^2 measure for goodness-of-fit within linear regression models is displayed for the combined Labrador retriever population and subgroups, for models predicting three obesity-related traits: BCS, body weight (kg) and food motivation. R^2 is displayed in its adjusted and unadjusted forms. All numbers rounded to 2sf. Food motivation on a scale of 1-100%. PRS – polygenic risk score; R^2 – linear regression goodness-of-fit measure; BCS – body condition score on a 9-point scale.

Table S13: PRS variant associations across purebred groups. Results of linear regression for each PRS SNP variant that contributed to the Labrador retriever PRS in other breed and population-specific minimal models. Data presented for each purebred cohort and phenotype combination as effect size β and significance p . Phenotypes are BCS, body weight (kg) and DORA food motivation.

Table S14: Clinical information and genotypes of probands and available family members from human patients with variants in CDH8 and CSNK1A1. Families delineated by background color. BMI SDS, body mass index standard deviation score (Z score).

Table S15: Results of tests for purifying selection. Two constraint metrics were used to evaluate the intolerance of a gene to high confidence predicted loss of function (pLoF) mutations in the gnomAD database. Results beyond the thresholds suggested by the developers of each metric (<0.4 for LOUEF and >0.9 for pLI) are in bold text. o/e, observed/expected.