doi: 10.1111/obr.12604

Genetics

Ethnic and population differences in the genetic predisposition to human obesity

C. Stryjecki[1], A. Alyass[1] and D. Meyre[1,2]

[1]Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada, and [2]Department of Pathology and Molecular Medicine, McMaster University, Hamilton, ON, Canada

Received 10 April 2017; revised 17 June 2017; accepted 2 August 2017

Address for correspondence: Dr D Meyre, Department of Health Research Methods, Evidence, and Impact, McMaster University, Room 3205, Michael DeGroote Centre for Learning and Discovery, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada.
E-mail: meyred@mcmaster.ca

Summary

Obesity rates have escalated to the point of a global pandemic with varying prevalence across ethnic groups. These differences are partially explained by lifestyle factors in addition to genetic predisposition to obesity. This review provides a comprehensive examination of the ethnic differences in the genetic architecture of obesity. Using examples from evolution, heritability, admixture, monogenic and polygenic studies of obesity, we provide explanations for ethnic differences in the prevalence of obesity. The debate over definitions of race and ethnicity, the advantages and limitations of multi-ethnic studies and future directions of research are also discussed. Multi-ethnic studies have great potential to provide a better understanding of ethnic differences in the prevalence of obesity that may result in more targeted and personalized obesity treatments.

Keywords: ethnic diversity, genetic susceptibility, obesity.

Abbreviations: BBS, Bardet–Biedl syndrome; BMI, body mass index; CNV, copy number variants; G × E, gene × environment interaction; G × G, gene × gene interaction; GWAS, genome wide association study; LD, linkage disequilibrium; *LEP*, leptin; *LEPR*, leptin receptor; *MC4R*, melanocortin 4 receptor; PCA, principal component analysis; *PCSK1*, proprotein/prohormone convertase 1; *POMC*, proopiomelanocortin; PWS, Prader–Willi syndrome; RAF, risk allele frequencies; snoRNAs, five small nucleolar RNAs; SNPs, single nucleotide polymorphisms; T2D, type 2 diabetes.

Introduction

Obesity rates have escalated to the point of a global epidemic over the last three decades. According to the World Health Organization, approximately 600 million adults worldwide were classified as obese in 2014 (body mass index, BMI ≥ 30 kg m^{-2}), while in parallel, the worldwide prevalence of childhood overweight and obesity has increased from 4.2% in 1990 to 6.7% in 2010 and is expected to reach 9.1% by 2020 (1). Obesity is associated with several comorbidities including type 2 diabetes (T2D), cardiovascular disease and some forms of cancer (2). Furthermore, childhood obesity is associated with more serious health outcomes later in life (3). Ultimately, severe forms of obesity reduce life expectancy by 13 and 8 years for men and women respectively (4).

Notable differences in the prevalence of obesity have been observed across diverse ethnic groups. In the USA alone, 21.8% of Caucasians, 34.8% of African Americans, 28.3% of Hispanics, 34.3% of Native Americans and 33.0% of Pacific Islanders over the age of 30 were considered to be obese between 2001 and 2002 (5); in contrast, only 4.8% of Asian Americans (individuals of Chinese, Filipina, Asian Indian, Vietnamese, Korean, Japanese and other Asian ancestry) were found to be obese (5). More recently, the National Health and Nutrition Examination Survey found 30% of Caucasians, 45% of African Americans and 36.8% of Mexican American adults over the age of 20

to be obese in the USA between 2009 and 2010 (6). These data clearly demonstrate ethnic disparities in the prevalence of obesity despite living in the same country. These disparities may be due to differences in lifestyle, socio-economic status, access to health care, social marginalization or discrimination; however, these differences may also reflect ethnic differences in biological susceptibility for obesity (7). The Oslo Immigrant Health Study, e.g., found the highest prevalence of obesity among Turks (51%) and the lowest prevalence among the Vietnamese (2.7%) with differences in BMI remaining despite adjusting for socio-demographic and lifestyle factors (8).

Currently, a growing body of evidence demonstrates ethnic differences in the genetic predisposition to obesity; however, many of the genetic variants responsible for these differences remain unidentified. This review provides a comprehensive examination of the ethnic differences in the genetics of obesity, characterized by BMI. We summarize the debate over the definitions of race and ethnicity, offer possible explanations for ethnic differences in the prevalence of obesity and describe heritability and admixture studies of obesity-related traits. We outline ethnic differences in monogenic syndromic, non-syndromic and polygenic forms of obesity followed by a discussion of the advantages and limitations of using multi-ethnic study designs to better understand ethnic differences in the prevalence of obesity and the genetic aetiology of this disease. We also propose several innovative research strategies.

How do we define ethnicity?

'Race' and 'ethnicity' are controversial and misunderstood terms within the scientific community (9). Historically, race has been used to classify populations based on shared biological characteristics such as skin colour, while ethnicity generally takes into account cultural characteristics (10,11). However, both terms are complex, multifactorial concepts reflecting religion, history and ancestral geographical origins (11,12).

Indeed, 99.9% of the human genome is identical in every individual; however, evolutionary forces including genetic drift, natural selection and *de novo* mutations have led to slight genetic differences among populations (13). With the completion of the Human Genome Project, genetic variants associated with disease susceptibility have been identified with varying frequencies across populations (14). Genetic variants can now be used with a high degree of precision to differentiate individuals from various ethnic groups, providing further evidence for the link between race, ethnicity and biology (15). Yet, this link is often blurry due to (i) numerous non-genetic connotations of race and ethnicity; (ii) the high degree of genetic diversity and the presence of population isolates within a given ethnic group; (iii) the lack of defined boundaries between populations; (iv) the admixed composition of certain populations (e.g. the Mexican population) and (v) the fact that many people have ancestors from diverse regions of the world (12,16).

The lack of clear definitions poses serious problems for geneticists as definitions can vary between studies. To bypass the definition debate, some suggest using geographical location and ancestry rather than race as genetic variation does not support the existence of race *per se*; however, this practice is highly debated (12,13). Researchers are in need of guidelines to properly describe diverse populations that accounts for both ethnicity and geographical location to improve generalizability and protect against spurious associations (17). Self-reported ethnicity is a practical way to adjust for ethnicity in genetic association studies; however, as Serre *et al*. demonstrated, it is not sufficient to protect against population stratification (15). Instead, principal component analysis (PCA) methods including EIGENSTRAT have been developed to correct for population stratification and geographical differences between and within ethnic groups (Fig. 1) (18). PCA is so precise that Karakachoff *et al*. accurately determined one's geographical origin within a few hundred kilometres by in a sample of 1,684 individuals from Western France (19).

Origins of the ethnic differences in the prevalence of obesity

Mankind has historically been exposed to prolonged periods of starvation where abilities to effectively store energy in times of abundance would grant one a survival and reproductive advantages. This is the essence of the thrifty genotype hypothesis proposed by James Neel in 1962 (20). Neel suggests the human genome is enriched with

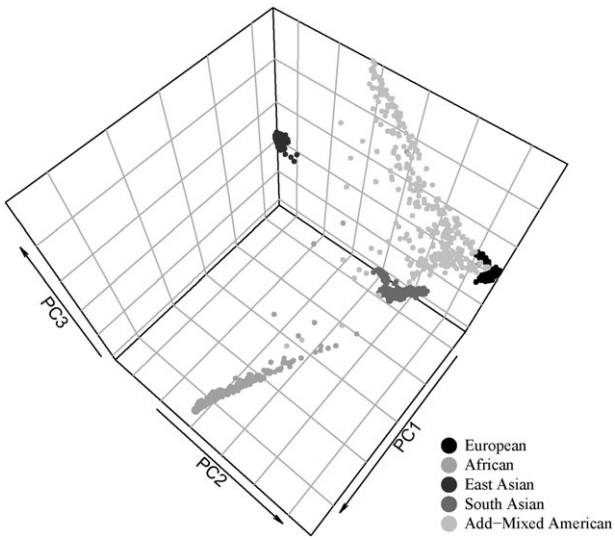

Figure 1 3D principle component analysis of different ethnic groups of the 1000 Genomes Project.

metabolically thrifty genes that provide a survival advantage during times of food shortage (20). However, these thrifty genes have been rendered detrimental by progress. Though highly controversial, this hypothesis can be applied today to explain the high and ethnic-dependent prevalence of obesity (21). Human ingenuity has mechanized many formerly labour-intensive processes resulting in a sedentary population reliant on automation. Food is no longer scarce, resulting in increased energy consumption. Today's industrialized countries see an improved quality of life at the expense of an evolutionary disadvantageous obesogenic environment (22).

The Pima Indians of Arizona, who have the highest reported prevalence of obesity (64% and 75% in men and women respectively), are a living example of the transition from a traditional to a modern, sedentary lifestyle (23). The Pima Indians were traditionally farmers but today, live a rural American lifestyle (23). It is believed that the migration of Pima Indian ancestors across the Bering land bridge and settling in the desert for 1000s of years may have selected for thrifty genes (23). These genes, however, no longer provide a survival advantage against starvation and may make this population more susceptible to obesity. In contrast, Europeans have benefited from a stable food supply and the availability of labour-saving devices for at least 300 years (24). Because they have reached food stability for centuries, Gerstein suggests that Europeans have 'purged' their thrifty genes because of their negative impact on cardio-metabolic health (24). It is possible that Europeans have been selected for an ability to thrive in a food-secure environment, while other more recently exposed populations have little resistance to obesity because they have not had enough time to genetically adapt (24).

Studies of rare and common variants predisposing to obesity have been carried out to test the validity of the thrifty genotype hypothesis. Analysis of several validated obesity variants provide some evidence of positive natural selection at the *FTO*, *NEGR1*, *SH2B1* and *FAIM2* loci in accordance with the thrifty genotype hypothesis (25,26). Recently, Wang *et al*. found that 9 out of 115 BMI single nucleotide polymorphisms (SNPs) were positively selected; however, five of these involved positive selection for the obesity protective allele (27). The lack of consistent signals for positive selection does not support the notion that genetically driven adiposity provided a survival or selective advantage (27).

A recent genome wide association study (GWAS) in 3,072 Samoans discovered a private mutation in *CREBRF* (rs12513649) strongly associated with BMI (28). The Samoans are a founder population with an extremely high prevalence of obesity. The *CREBRF* variant is common in Samoans (frequency of 0.30) but almost absent from other populations, demonstrating that rare variants can be highly prevalent in isolate populations (28). While *CREBRF* is presented as a 'thrifty' variant, the high frequency of this variant may be explained by a founder effect and a lack of natural selection pressures. Rare mutations in the melanocortin 4 receptor (*MC4R*) are the most common cause of monogenic obesity. Evolutionary analysis of non-synonymous deletions in *MC4R* in both humans and primates suggests a strong negative or purifying selection on *MC4R* to remove deleterious mutations from the population, which is in contrast to the thrifty genotype hypothesis (29). An analysis of common variants associated with obesity indicates either an absence of positive selection, positive selection for leanness promoting variants or positive selection for tall and slender stature among Europeans, providing further evidence that the thrifty genotype hypothesis, if true, may be context-dependent (25–27,30).

Beyond the thrifty genotype hypothesis, the 'predation release', 'drifty gene' and 'thrifty epigenotype' hypotheses may explain ethnic differences in the prevalence of obesity (31–33). The 'drifty gene' or 'predation release' hypotheses were put forward by John Speakman as an alternative to the long-standing thrifty genotype hypothesis (31,32). Speakman argues that when ancestral humans acquired the ability to use fire and tools and form organized societies, they subsequently removed the threat of predatory danger (31,32). In the absence of the predation selection pressure, genes promoting energy storage were allowed to drift (31,32). The 'thrifty epigenotype' hypothesis builds upon the thrifty genotype and phenotype hypotheses, arguing that all human posses a thrifty genome, but phenotypic expression can vary due to inherited epigenetic changes (33). Stoger argues that individuals born during times of famine carry epigenetic changes allowing for more efficient energy storage (33). Conversely, individuals born during times of food abundance will be less prone to obesity (33).

Societal conventions such as the practice of consanguineous marriages has resulted in a high prevalence of monogenic obesity in Pakistani children with 30% of the severe cases of obesity being due to genetic mutations in the genes encoding leptin (*LEP*) and *MC4R* (34). The practice of intra-caste marriages in India may also increase the average degree of homozygosity in the genome resulting in an increased incidence of autosomal recessive disorders including recessive forms of Mendelian obesity (35). Assortative marriages for BMI confer a higher genetic predisposition to obesity in the offspring generation; 50% of parents of extremely obese offspring had a BMI in the top 10% themselves (36). Consanguineous marriages and assortative marriages may therefore lead to genetic differences between countries with a divergent prevalence in obesity within a few generations.

Heritability and ethnic background

The familial aggregation of one's body size is not a recent concept. The strongest risk factor for childhood obesity is

parental obesity where a child's risk of obesity is 2.5-fold to 4.0-fold higher if one parent is obese and 10-fold higher if both parents are obese, compared with having both parents of normal weight (37). Knowing that familial resemblance can be explained by both shared environments and genetic factors, milestone twin and family studies have emerged over the past 35 years. Because monozygotic twins share all genetic makeup while dizygotic twins share only half, one would expect monozygotic twins to be more similar in terms of weight than dizygotic twins if body weight is influenced by genetic factors (38). In fact, estimates of heritability (defined as the proportion of phenotypic variation of a trait attributed to genetic variation) from twin and family studies range between 40% and 70% (39). Studies of twins reared apart and twins raised together found similar estimates of heritability for BMI, providing evidence that genetics have a stronger impact on weight than the environment (40).

We performed a random-effect meta-analysis of heritability estimates of BMI from 19 twin and 20 family studies from various ethnic groups. Heritability estimates were pooled on the logit scale, and standard errors were derived using the delta-method (Figs. 2 and 3). Our meta-analysis includes only studies involving adults as the genetic influences on BMI have been shown to increase in strength during childhood (40). Overall, heritability estimates in twin studies (h^2 = 0.72 [0.69–0.75]) were higher than those from family studies (h^2 = 0.46 [0.40–0.52]). Due to the limited number of twin studies from non-European populations, we were unable to assess ethnic differences in the heritability of BMI. Heritability estimates for BMI obtained from family studies were not significantly different in African (h^2 = 0.53

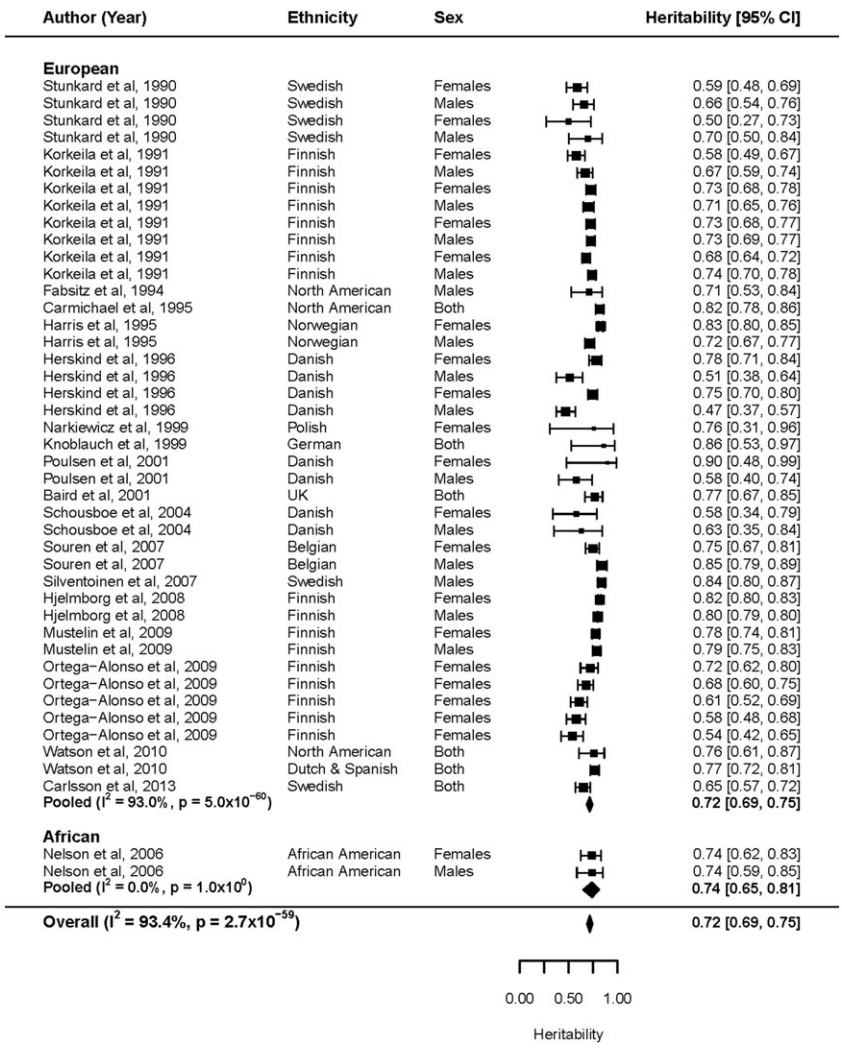

Figure 2 Meta-analysis of body mass index heritability estimates in twin studies.

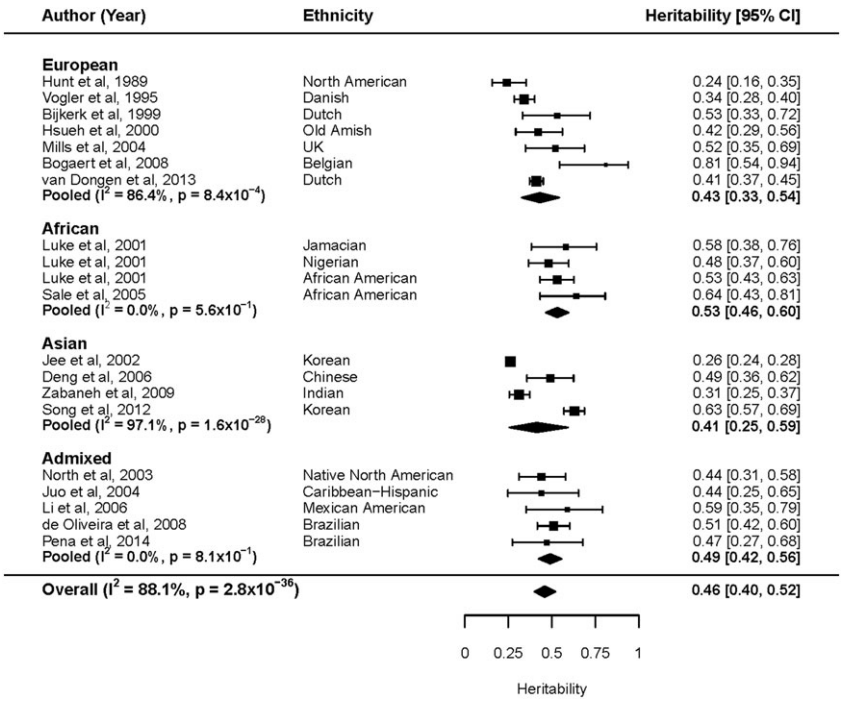

Figure 3 Meta- analysis of body mass index heritability estimates in family studies.

[0.46–0.60]), admixed (h^2 = 0.49 [0.42–0.56]) and Asian (h^2 = 0.41 [0.25–0.59]) populations, relative to Europeans (h^2 = 0.43 [0.33–0.54]).

Admixture studies and obesity-related traits

Most American ethnic groups present today are the result of the intermixing of European, African and Native American populations during the colonization of the New World (41); genetic variants from previously isolated populations were brought together in new combinations to establish the contemporary European, African, Hispanic and Native American gene pools. Consequently, populations may have inherited ethnic specific disease susceptibility genetic variants, affecting the likelihood of acquiring diseases (42,43).

Genetic admixture studies have been valuable in identifying differences in ethnicities that cannot be explained by environmental factors alone. Data from the 2003–2004 National Health and Nutrition Examination Survey found African Americans to be 1.5 times more likely to be obese than European Americans despite homogenous socioeconomic status, suggesting that differences in genetic background may account for ethnic differences in obesity risk (5). Using genome-wide admixture mapping in 15,280 African Americans, Cheng *et al.* identified inverse negative correlation with BMI and percentage of European ancestry (44). Similar associations with BMI and Native American admixture have been reported, suggesting that the European genome may contain fewer obesity risk alleles and/or may be enriched in obesity protective genetic factors (45).

Monogenic syndromic forms of obesity and ethnic diversity

Currently, obesity is a defining characteristic of 79 distinct Mendelian syndromes, providing further evidence for the role of genetics in the aetiology of obesity (46). Ethnic differences in the prevalence of these diseases and syndromic obesity are well documented for Alström syndrome, Bardet–Biedl syndrome (BBS), Cohen syndrome and Prader–Willi syndrome (PWS) are outlined in the succeeding texts.

Alström syndrome

Alström syndrome is a rare autosomal recessive disease affecting less than one in one million people in the general population (47). Clinical symptoms of Alström syndrome include childhood obesity, severe insulin resistance, hyperinsulinemia, impaired glucose tolerance and T2D, independent of the degree of obesity (47). Alström syndrome is the result of mutations in exons 8, 10 and 16 in the *ALMS1* gene on chromosome 2p13 (47). To date, 109 different mutations in *ALMS1* have been identified, mostly frameshift and non-sense mutations resulting in the premature truncation of ALMS1 (47).

Founder mutations for Alström syndrome have been observed in families of French Acadian and English descent. Genealogical analysis of large Acadian kindred including eight individuals with Alström syndrome confirmed that the affected individuals are from a common founder. In the early 17th century, the first Acadians migrated from France to Acadia, now known as Nova Scotia where they lived in relative isolation (48). The ancestry of the affected individuals was traced back to a small group of 17th century Acadians who emigrated from Northern France to Acadia (48). One ancestral pair common to both maternal and paternal lineages of all affected individuals was found, suggesting a founder effect for Alström syndrome in this population (48). Historical records also confirm the presence of Alström syndrome in this lineage; two half-sisters were reported to have been blind, hearing impaired, obese and chronically hyperglycemic (48). Further investigation identified a 19-base pair insertion in exon 16 of affected individuals, causing a frameshift and early termination at codon 3530 (49). Among 12 unrelated patients with Alström syndrome in the UK, a deletion in exon 16 was identified in five affected individuals. These individuals either resided or originated from Yorkshire, UK, suggesting the possibility of a founder effect (50). Founder effects for Alström syndrome have also been identified in Pakistani and Turkish families (51,52).

Lastly, four novel mutations in *ALMS1* were identified among six Saudi Arabian patients with Alström syndrome. These mutations are believed to have arisen independently at a rate similar to that of other populations due to the high prevalence of consanguinity in the Saudi population. Thus, the high degree of homozygosity in this population has led to the expression of this disease and provides evidence for the powerful effect of consanguinity in shaping the genetic landscape (53).

Bardet–Biedl syndrome

Bardet–Biedl syndrome is a rare autosomal recessive disease characterized by six cardinal manifestations: obesity, retinitis pigmentosa, renal anomalies, polydactyly, learning disabilities and urogenital tract defects (54). The prevalence of obesity among individuals with BBS is between 72 and 86% (54).

To date, 21 genes involved in BBS have been identified through various gene identification strategies (55–57). The majority of pathogenic mutations are found in *BBS1* and *BBS10* (54). The BBS proteins form a complex (BBSome) essential for ciliary function; this complex associates with the ciliary membrane and sorts and directs protein and vesicle trafficking (58). Interestingly, heterozygous carriers of BBS mutations have an increased risk of developing obesity than non-carriers despite not exhibiting other BBS phenotypes (59). Furthermore, associations between four common genetic variants in three BBS genes (*BBS2*, *BBS4* and *BBS6*) and common obesity have been identified in French–Caucasian populations suggesting that BBS genes may be associated with polygenic obesity risk (59). The association of SNPs at the *BBS4* locus and polygenic adult obesity has been recently confirmed by a large-scale GWAS (60).

Some BBS genes appear to have a greater ethnic specific frequency than others although no genes are found exclusively in one ethnic group. In northern Europeans, mutations in *BBS1* and *BBS10* are common, while mutations in *BBS4*, *BBS5* and *BBS8* are commonly seen in individuals of Middle Eastern and North African descent (54). The prevalence of BBS has also been found to vary between populations from 1 in 160,000 in Northern Europe to 1 in 13,500 and 1 in 180,000 in isolated communities in Kuwait and Newfoundland respectively (61,62). In Newfoundland, at least six BBS loci and eight different BBS mutations have been found in affected individuals, suggesting that the high prevalence of BBS cannot be due to a single founder (54). It is possible that consanguinity, large sibship sizes and a survival advantage for heterozygotes who have an enhanced ability to store fat, may contribute to the high prevalence of BBS in Newfoundland (62). In Tunisia, the prevalence of BBS was estimated to be 1 in 156,000, while the frequency in the North of the country was estimated to be 1 in 87,000 (63). The high prevalence of BBS in the Tunisian population may be due to the high rate of consanguinity (31%) (63).

Cohen syndrome

Cohen syndrome is a rare autosomal recessive disorder characterized by mental retardation, motor clumsiness, microcephaly, severe myopia, distinct facial features, childhood hypotonia and joint laxity (64). Truncal obesity is present but is not always a ubiquitous feature of this disease. Cohen syndrome is caused by a mutation in the *COH1* gene on chromosome 8q22. This gene encodes a protein of unknown function; however, domain structure and homologies suggest a role in vesicle-mediated sorting and intracellular protein transport (64). A recent study found *COH1* to code for a Golgi-associated matrix protein that is required for Golgi integrity (65).

To date, about 100 mutations in *COH1* have been identified with the majority resulting in a null allele; missense and frameshift mutations have also been identified but are less common (66). The best characterized mutation is the c.3348_3349delCT, which causes a frameshift at codon 1117, resulting in protein truncation at codon 1124 (67). This mutation is found in high frequencies in the Finnish population where Cohen syndrome is overrepresented and may explain the high levels of clinical homogeneity within this population (68). Overrepresentation of this allele in

the Finnish population provides evidence for a founder effect with a common ancestral mutation being responsible for most cases (66).

Other mutations in *COH1* in populations with a known founder effect have been identified. Cohen syndrome is frequently observed among Irish travellers (estimated 0.5 per 1000 Irish traveller children) where the c.4471G->T results in a null mutation (69). The c.11564delA deletion results in the deletion of exons 6–16 and was identified in 14 Greek patients originating from two small neighbouring islands where the incidence of Cohen syndrome is 1 in 110 (70). The c.11564delA has also been identified in two families from Central Italy and one in Southern Italy. The c.8459T->C variant among a population of Ohio Amish results in a null mutation where the prevalence of Cohen syndrome is as high as 1 in 500 (67). These findings suggest that extensive allelic heterogeneity is responsible for this disease (68). The prevalence of obesity among affected individuals shows ethnic variation with a prevalence of 80% among Irish travellers, 53% among Greek/Mediterranean individuals, 37.5% among the Amish and 17% among Finnish children (67).

Prader–Willi syndrome

Prader–Willi syndrome is characterized by short stature, small hands and feet, hypogonadism, mental retardation, obsessive–compulsive behaviours, early childhood-onset hyperphagia and obesity (71). Most morbidities and mortalities in PWS are the result of being severely obese (72). The majority of PWS cases (70–75%) are caused by a deletion of imprinted genes within the paternally inherited locus 15q11-q13 (73). Ten known paternally expressed loci are involved in PWS features and include *MKRN3*, *MAGEL2*, *NDN*, *NPAP1*, *SNURF-SNRPN* and five small nucleolar RNAs (snoRNAs) (58). A microdeletion in the HBII-85 snoRNA cluster in a child with PWS provides conclusive evidence for the role of snoRNAs in the aetiology of PWS (74). Other case of PWS are the result of maternal uniparental disomy (~30%), imprinting defects (>5%) or balanced translocations on 15q11-q13 (>1%) (72).

Prader–Willi syndrome cases have been reported worldwide and generally occur in about 1 in 15,000 births. In the USA, the prevalence of PWS has been reported between 1 in 16,000 to 1 in 25,000 (75,76). Elsewhere, the prevalence of PWS ranges from 1 in 8,000 in rural Sweden; to 1 in 16,000 in the San-in district of western Japan; 1 in 15,830 in Australia and 1 in 26,676 in Flanders Belgium (77–80). In the UK, the proposed true prevalence of PWS is 1 in 45,000 (81). Given that most cases of PWS (~70%) are the result of *de novo* deletions and epigenetic effects, the prevalence of PWS is not influenced by consanguinity or founder effects.

Monogenic/oligogenic non-syndromic forms of obesity and ethnic diversity

Obesity can show Mendelian patterns of inheritance due to homozygous/heterozygous compound loss-of-function mutations in five genes that are part of the LEP melanocortin pathway: *LEP*, leptin receptor (*LEPR*), proopiomelanocortin (*POMC*), proprotein/prohormone convertase 1 (*PCSK1*) and *MC4R* (82). This pathway is critical for regulating food intake and body weight; thus, complete inactivation of these genes results in severe hyperphagia and fully penetrant early-onset obesity (82). Loss-of-function mutations display recessive inheritance, and as will be shown in the succeeding texts, monogenic forms of obesity have been identified mainly in ethnic groups practising consanguinity and founder populations (34). Partial inactivation of these genes results in oligogenic forms of obesity (58).

Leptin and Leptin Receptor

Leptin is produced by adipose tissue and plays an essential role in regulating food intake and body weight (83). Cases of complete LEP deficiency are very rare with only 34 cases being reported worldwide and results in severe hyperphagia and early-onset obesity (83,84). The first cases of complete LEP deficiency were identified in Pakistani cousins from a highly consanguineous pedigree with severe obesity, and since then, other cases of LEP deficiency have been identified in this population. In a cohort of Pakistani children with early-onset severe obesity, 16.1% were found to have homozygous mutations in *LEP* (85). Of these, nine children were homozygous for the ΔG133 frameshift mutation and one child was homozygous for a 3-base pair deletion (85). Pakistani individuals with obesity heterozygous for the ΔG133 frameshift mutation have also been identified (83). The ΔG133 mutation is frequently identified in the Pakistani population, suggesting a possible founder mutation (85). Other mutations in *LEP* have been identified in individuals from Pakistan (n = 27), Turkey (n = 5), Turkmenistan (n = 2), Egypt (n = 1), Austria (n = 1) and China (n = 1), suggesting that they are ethnic-specific (86).

Mutations in *LEPR* result in a *LEPR* lacking transmembrane and intracellular domains due to abnormal transcript splicing (87). Like *LEP* mutations, mutations in *LEPR* cause extreme hyperphagia and early-onset obesity. The first *LEPR* mutations were identified in three Algerian siblings from a consanguineous family, and to date, only 13 cases of complete LEPR deficiency have been identified (82,87). Among those with hyperphagia and severe early-onset obesity, the prevalence of *LEPR* mutations was 3% (87). In a multi-ethnic cohort, homozygous frameshifts were found in Bangladeshi, Turkish and Iranian subjects, and homozygous non-sense mutations in Southern European

subjects and homozygous missense mutations were found in subjects of Turkish, Norwegian or British descent (87). Mutations in *LEPR* have also been identified in Pakistani, Turkmenian and Egyptian children and may be due to the high rate of consanguineous marriages (34,88,89). Recently, a *LEPR* frameshift mutation (p.P166CfsX7) was identified in six individuals with morbid obesity from Reunion Island and is suggestive of a founder effect (90).

Pro-opiomelanocortin

POMC is expressed in the pituitary gland, and sequential cleavage of POMC produces the melanocortin peptides adrenocorticotropin, alpha-melanocyte-stimulating hormone, beta-melanocyte-stimulating hormone and β-endorphin. Obesity is thought to be due to deficiency of alpha-melanocyte-stimulating hormone signalling at MC4R, resulting in a lack of appetite suppression (91). Humans homozygous for loss-of-function mutations in *POMC* develop severe obesity, adrenal dysfunction and red hair pigmentation (92). Heterozygous loss-of-function mutations in *POMC* result in a non-fully penetrant/oligogenic form of obesity (93). The most frequent mutation in *POMC* is R236G, which alters POMC processing and reduces its ability to activate MC4R (91). The frequency of the R236G mutation was found to be 0.6% among Danish individuals, 0.76% in British individuals and 1.65% among French individuals; R236G was also associated with early-onset obesity in British and French individuals (91). *POMC* mutations were found in 1.5% of Italian adults with obesity but were not associated with early-onset obesity (91). Other *POMC* mutations have been identified in children with severe obesity of German, UK Caucasian, French, Egyptian and Indian origin (94–97). Two children with obesity of Turkish and North African ancestry with *POMC* deficiency have been described despite lacking the characteristic red hair phenotype, suggesting an ethnic dependent clinical presentation (98,99). It can be assumed that other genetic variants act epistatically in populations practising consanguinity to maintain pigmentation, while pigmentation is more dependent on the presence of POMC-derived ligands in European populations (93).

Proprotein/prohormone convertase 1

Proprotein/prohormone convertase 1 is expressed in the brain, enteroendocrine cells and the neuroendocrine system and is responsible for processing precursor proteins (100). Mutations in *PCSK1* cause early-onset obesity, hyperphagia, postprandial hypoglycemia, diabetes insipidus, intestinal and endocrine dysfunctions (100). Nineteen patients with homozygous or compound heterozygous mutations in *PCSK1* have been identified. Two carriers of homozygous *PCSK1* mutations with obesity were identified in a North-African consanguineous family and in a Turkish family with possible consanguinity (101–103). An additional 13 carriers of homozygous *PCSK1* mutations have been reported from a multi-ethnic cohort with consanguineous families (101,104). Three carriers of *PCSK1* compound heterozygous mutations were identified in non-consanguineous Caucasian families (105–108). Haploinsufficient heterozygous *PCSK1* mutations result in non-fully penetrant obesity, with an estimated prevalence of 0.83% among European children and adults with severe obesity (100).

Melanocortin 4 receptor

Melanocortin 4 receptor is expressed mainly in the central nervous system where it regulates energy metabolism (85,109); roles of MC4R in controlling food intake and food choices have also been suggested (109). In humans, MC4R haploinsufficiency due to loss-of-function mutations is the most common cause of monogenic obesity (109). At the population level, only 20 individuals with a complete MC4R deficiency have been identified (82).

The prevalence of *MC4R* heterozygous, heterozygous compound and homozygous loss-of-function mutations reported in children and adults with obesity from various ethnic groups ranges from 0.5 to 5.8% (110,111). A study from the UK reported a prevalence of 5.8% in a sample of 500 individuals with severe early onset obesity, and a frequency of 4% was reported in a French study of adults with severe obesity (112,113). In contrast, low frequencies of *MC4R* mutations have been reported in German (1.9%), Greek (0.2%) and Italian (< 0.5%) populations (114–116). To date, only two individuals of Japanese descent with mutations in *MC4R* have been identified (117,118). In two small studies of Chinese adults and children with obesity, mutations in *MC4R* were identified in less than 1.5% of the cohort (119,120). A subsequent study of Chinese, Malay and Indian children and adolescents with severe obesity identified three individuals (1.3%) with *MC4R* mutations, suggesting that mutations in *MC4R* are not a major cause of obesity in Asian populations (121). A high frequency of homozygous loss-of-function mutations (3.2%) has been identified among Pakistanis, likely due to the high rate of consanguineous marriages (60–70%, one the highest rates in the world) in this population (85). Among Pima Indians, three private mutations in *MC4R* have been identified, which are not found in other populations (122). Of particular interest is the R165Q variant, which was found in 3% of Pima individuals with severe obesity and is one of the highest reported frequency for any *MC4R* variant (123).

While *MC4R* loss-of-function homozygous/heterozygous compound mutations lead to a fully penetrant form of early-onset morbid obesity, heterozygous individuals have a milder, non-fully penetrant form (82); among European heterozygotes, the penetrance of obesity ranges from 40 to

100% (124). In a Greek population, the penetrance of obesity was found to be 6.3%, which is relatively low in relation to the high penetrance seen in other European populations. This is of interest as the prevalence of obesity in the Greek population is 18%, one of the highest in Europe (115). It is possible that variants in other genes directly or indirectly involved in the melanocortin pathway may antagonistically interact with the loss-of-function mutations or the Mediterranean diet can attenuate the effects of *MC4R* loss of function, minimizing the penetrance of obesity (115). Similarly, Pakistani heterozygous carriers were found to have a normal weight, suggesting that a rural environment may mitigate the penetrance of *MC4R* mutations (85). Together, these results suggest that the penetrance of obesity due to *MC4R* heterozygous mutations may to a certain extent be dependent on the environment and lifestyle choices.

Polygenic forms of obesity and ethnic diversity

GWAS for obesity in European and non-European populations

Genetic predisposition to obesity is polygenic in nature in most cases and is attributed to the simultaneous presence of risk polymorphisms in multiple genes. Independently, polygenic variants have small to modest effects on the obese phenotype but together, give rise to a sizeable effect (82). Until recently, the genetic determinants of obesity were largely unknown until the emergence of statistically powerful GWAS that have revolutionized the search for genetic determinants of complex traits. GWAS searches the genome for several hundred-thousand SNPs and identifies SNPs that occur more frequently in individuals with a particular disease than in those without the disease (82). In 2007, common variation in intron 1 of *FTO* was associated with obesity in Europeans by four independent groups (125–128). To date, *FTO* is viewed as the main contributor to polygenic obesity in Europeans. Since 2007, the association of *FTO* on obesity has been extended to diverse ethnic groups including African-American, Hispanic, Pacific Islander and East Asian populations (129). Subsequent large meta-analyses of GWAS in predominately European populations have identified 142 polygenic loci associated with BMI and/or obesity (58). Although these loci have considerably smaller effects than *FTO*, they provide valuable insight into the genetic architecture of obesity. Pathway analysis of genes associated with BMI provides strong support for a role of the central nervous system, adipose tissue, the musculoskeletal system and digestive tract, highlighting the complex aetiology of obesity that encompasses biological pathways in multiple organ systems (58). Meta-analysis of GWAS of BMI in predominantly European children identified 12 loci previously associated with BMI in adults, demonstrating the shared genetic background between childhood and adult BMI (130).

In recent years, several GWAS for obesity traits have been conducted in non-European populations such as East Asians and Africans. GWAS in non-Europeans have been critical in confirming European obesity loci and identifying novel, ethnic-specific loci. In a recent meta-analysis of 86,757 individuals of Asian ancestry, Wen *et al.* confirmed seven previously reported BMI-associated loci in European populations (*FTO*, *SEC16B*, *MC4R*, *GIPR-QPCTL*, *ADCY3-DNAJC27*, *BDNF* and *MAP2K5*) and identified three novel loci associated with BMI (*CDKAL1*, *PCSK1* and *GP2*) (131). Another study of East Asians also confirmed *CDKAL1* as a novel BMI-susceptibility locus, with *KLF9* as an additional locus (132). In 2011, the first GWAS in a Filipino population (n ~ 1,7000 women) replicated GWAS signals for *MC4R*, *FTO* and *BDNF* (133). Among Asian Indians, Been *et al.* confirmed the association of *MC4R* (rs12970134) with BMI (134). Genome-wide heterogeneity of variance analysis in 14,131 Pakistani individuals identified an interaction with smoking status and a novel obesity variant in *FLJ33534* (rs140133294) on BMI (135). Meta-analysis of 9,881 African-Americans has demonstrated an association between *FTO* (rs3751812 and rs9941349) and obesity, while two smaller GWAS with individuals of African ancestry provide some evidence of replication of the association between *MC4R* (rs6567160 and rs17782313) and BMI (136,137). More recently, a large meta-analysis in over 30,000 individuals of African ancestry identified one new locus associated with BMI (*GALNT10*) and five previously identified European BMI loci (*MC4R*, *FTO*, *GNPDA2*, *ADCY3* and *SEC16B*) reached genome-wide significance (138). GWAS in Samoans identified a private variant in *CREBRF* (rs12513649), common in this population (frequency: 25.9%) and strongly associated with BMI (28). Taken together, GWAS in non-European populations are suggestive of a partial genetic overlap between obesity loci across various ethnic groups (Fig. 4).

Filling in the gaps of missing heritability

Despite the surge in GWAS and meta-analyses, most of the genetic variability in obesity remains unexplained. In a large meta-analysis by the GIANT consortium, 97 BMI-associated loci only explain 2.7% of the variance in BMI, suggesting that numerous additional variants associated with obesity remain unidentified (60). Possible explanations for this 'missing' heritability include lack of power to detect common variants with subtle effects and causal variants, poor coverage of rare variants, genetic heterogeneity, structural variants, epigenetics and gene–gene and gene–environment interactions. The majority (> 80%) of identified variants are located in non-coding regions and are

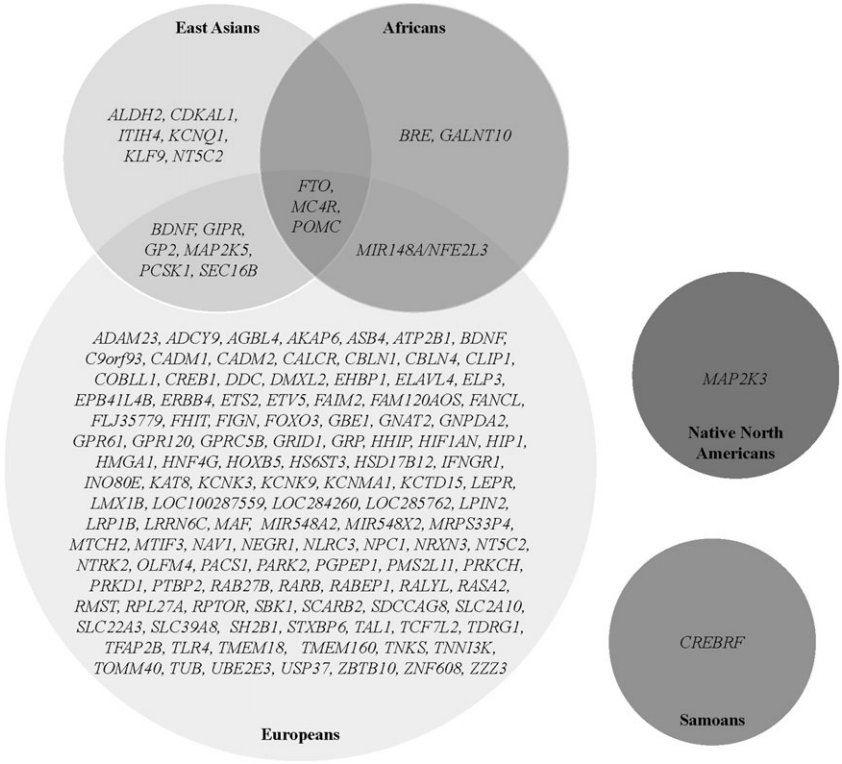

Figure 4 Obesity loci discovered through genome-wide association studies in European and non-European populations.

thought to be non-causal (139). GWAS index SNPs are assumed to be in linkage disequilibrium (LD) with the causal variant, making it difficult to distinguish the causal variant and elucidate its role in the development of obesity. Furthermore, if index SNPs are tested in a population with different ancestry, the LD between the index SNP and causal SNP may be weaker. This results in a weak, none or inverse association with the trait in the replication cohort and a wrongful conclusion that the SNP does not transfer across ancestries (140). Such is the case with *FTO* rs3751812 (a perfect surrogate for the previously reported *FTO* rs9939609), which replicates in both European and African–American populations, unlike *FTO* rs9939609, which only replicates in Europeans (141).

In evolutionary terms, populations of African ancestry are the most ancestral and have experienced more generations of LD decay, relative to European-ancestry and Asian-ancestry populations. Due to the accumulation of more recombination events, the African population has smaller regions of LD. Out-of-Africa migrations and genetic bottlenecks have reduced haplotype diversity in European-ancestry and Asian-ancestry populations, resulting in larger regions of LD. The weak LD in African populations can be leveraged for fine-mapping studies to pinpoint the causal variant (136). To date, most fine-mapping efforts have focused on the *FTO* locus due to its strong association with obesity-related traits (140). Recent mechanistic works suggest that the *FTO* rs1421085 variant may be a functional variant involved in white adipocyte browning and thermogenesis by disrupting ARID5B repression of IRX3/5, resulting in increased lipid storage and weight gain (142,143). However, initial attempts to fine-map the causal variant(s) in *FTO* in populations of African ancestry have yielded inconsistent results. As described earlier, Grant *et al*. determined that the rs3751812 variant was a better surrogate for the causal variant then rs9939609 in African–American children (141); these results were further confirmed by Hassanein *et al*. who fine-mapped the association between variation at the *FTO* locus and BMI in 9,881 African adults (136). In a large study of over 20,000 African Americans, Peters *et al*. densely fine-mapped the entire *FTO* gene and the adjacent *RPGRIP1L* gene to narrow down the functional variant (144). While they significantly reduced the number of functional candidates, they were unable to narrow in on the functional variant. More recently, Gong *et al*. fine-mapped 21 BMI-related loci in African Americans and found eight of the 21 associated with BMI in this population (145). These examples demonstrate the utility of fine-mapping in non-European populations to narrow down on the causal variant.

Admixture mapping in recently admixed populations is a powerful way to identify disease-causing variants and is well suited for the genetic investigation of complex diseases such as obesity (146). Knowing that the prevalence of

certain diseases and complex traits varies with ethnicity, admixture studies scan the genome for regions where the proportion of one ethnicity is significantly different than average. Admixture mapping has greater statistical power to identify variants with modest effects and has successfully reported associations between risk of obesity or increased BMI in West African and Native American populations (45). Furthermore, individuals of mixed ethnicities (Asian/Caucasian, Hawaiian/Caucasian, Hawaiian/Asian, Latina/Caucasian and Hawaiian/Asian/Caucasian) have been found to have an above average BMI than their parental ethnic groups, suggesting that differences in ancestral background may partially explain ethnic differences in the prevalence of obesity (45). Admixture mapping in African Americans has identified several chromosomal regions associated with BMI, including regions on chromosomes X (Xq25 and Xq13.1) (44), 1 (1q23.2 and 1q25.1) (41), 2 (2p23.3) (42), 3 (3q29) (147), 5 (5q14 and 5q13.3) (44,147), 11 (11q23.2) (41), 12 (12p13.31) (41) and 15 (15q26) (147). A fine-mapping study of four genomic regions reported in previous admixture analyses identified an association between SNP rs631465 in *F2RL1* and BMI in an African–American population (148).

The risk allele frequencies (RAF) of obesity loci identified through GWAS are generally high (allele frequency > 10%), but RAF can vary across populations (149). For example, the RAF of *FTO* rs3751812 in the 1000 Genomes Project demonstrate considerable ethnic differences, ranging from 0.05 in African, 0.17 in East Asian, 0.29 in South Asian and 0.41 in European populations.

Low frequency (~1–5%) and rare (<1%) variants contributing to polygenic obesity are not frequent enough to be captured by current genome-wide association approaches, nor are they penetrant enough to be identified through traditional linkage studies, yet they could explain part of the missing heritability of disease risk (150–152). Candidate gene approaches in populations of European descent have identified low-frequency loss-of-function coding non-synonymous variants in *GPR120* (R270H/rs116454156) and *PCSK1* (N221D/rs6232) associated with increased risk of obesity and low-frequency coding gain-of-function non-synonymous variants in *MC4R* (V103I/rs2229616 and I251L/rs52820871) associated with protection from obesity (153–155). No well-established association between low frequency/rare variants and obesity traits has been reported in non-European populations to date.

Genomic rearrangement due to deletions or duplications of chromosomal regions can give rise to copy number variants (CNV). Studies of CNV and obesity have been performed in predominately European populations in attempt to further explain the missing heritability of obesity. The rare heterozygous 16p11.2 deletion of at least 593 kb is well studied and is associated with severe early-onset obesity in Europeans (156). Genome-wide association meta-analyses of individuals of European ancestry identified deletions in regions near *NEGR1* and *GPRC5B* associated with BMI (157,158). A GWAS for early-onset extreme obesity in individuals of German ancestry identified a CNV near 11q11 associated with early-onset obesity (159). A GWAS for BMI in a small Chinese population also identified a region near 10q11.22 associated with BMI (160). Subsequent studies in European populations confirmed this association, demonstrating the utility of screening for CNV in non-European populations to identify novel variants implicated in obesity (161). Currently, the role of rare and common CNV in obesity remains relatively unexplored; however, it is unlikely that the CNV explain a significant portion of the missing heritability of obesity (162).

Convincing evidence for gene × gene (G × G) interactions has emerged for several obesity loci. A study in East Asians identified a significant G × G interaction between two new BMI associated SNPs in the *CDKAL1* and *GDF8* loci (132). Recent data indicate that obesity-predisposing variants interact with a variety of environmental, lifestyle and therapeutic treatments (163). Consistent gene × environment (G × E) interactions between *FTO*, level of physical activity and BMI or obesity have been described in 16 cross-sectional and intervention studies in European, East Asian and African populations (164); this was confirmed by a large meta-analysis of 218,166 adults predominately of European descent where physical activity reduced the risk of obesity by 27% (165). Using a quantitative measure of energy expenditure (metabolic equivalent score) to provide a more comprehensive assessment of physical activity, Reddon *et al*. demonstrated that physical activity can blunt the effects of *FTO* on adiposity (measured by BMI and body adiposity index) by 36–75% in a longitudinal multi-ethnic cohort (164). Populations in low-income and middle-income countries are undergoing rapid transitions from traditional to Western lifestyles. Taylor *et al*. investigated the influence of living in rural versus urban India on the role of *FTO* on obesity related traits (166). When genetic variants in these genes were analysed with regards to environment, a stronger association between *FTO*, weight and living in an urban environment was found in comparison with those living in a rural environment (166,167). A novel interaction between smoking status and the *FLJ33534* locus on BMI has recently been reported in a Pakistani population (135). G × G and G × E interaction remain largely unexplored due to statistical challenges associated with inadequate sample sizes but may explain some of the missing heritability (164).

Lastly, epigenetic changes may explain the missing heritability in obesity. Epigenetics is defined as changes in gene transcription and expression that do not involve changes to the underlying DNA sequence (168). Epigenetic modifications include DNA methylation, histone post-

translational modifications and chromatin remodelling or the inheritance of mRNAs that regulate gene expression (58). DNA methylation consists of the addition of methyl groups to cytosine residues and is typically associated with gene silencing (168). Candidate gene approaches and more recently, epigenome-wide association studies have identified changes in DNA methylation patterns in genes associated with BMI and obesity (169,170). After analysing 450 million CpG sites and subsequent validation in two replication cohorts, Dick *et al*. found an association with increased BMI and DNA methylation at the *HIF-3α* locus (171). More recently, methylation within a variably methylated region in *POMC* has been strongly associated with BMI in a multi-ethnic cohort (172). Higher methylation of sites within intron 1 of *FTO* and differential methylation of other genes has been observed, suggesting that *FTO* can influence methylation patterns of other genes (173). Of the 52-known obesity-associated SNPs, 28 have been associated with DNA methylation levels at 107 proximal CpG sites, suggesting that they affect multiple genes (174). Using an epigenome-wide association study, Wahl *et al*. recently found an association with BMI and changes in DNA methylation in 187 loci involved in lipid and lipoprotein metabolism, adipose tissue biology and insulin resistance (170). Replication has been problematic for epigenetic studies; except for *HIF-3α* where the association between *HIF-3α* methylation and BMI has been replicated in subsequent studies, many associations have not been successfully replicated in independent cohorts (175,176). Furthermore, it is unclear if changes in DNA methylation are a consequence of obesity rather than the cause (177). Overall, the field of epigenetics has provided novel insight into the complex genetic architecture of obesity but is unlikely to fully explain the missing heritability of obesity. Further studies, especially in non-European populations, are needed.

To conclude, GWAS and meta-analyses have made significant advances in identifying genetic variants associated with BMI and obesity; however, they have explained very little in the variance of BMI (60). Part of the missing heritability is hypothesized to be due to rare genetic variants with large effect sizes, which are not captured by GWAS; multiple common genetic variants with small effect sizes cannot be detected in very large GWAS, or heritability is overestimated due to environmental effects or genetic interactions (178,179). New methods of estimating heritability suggest that the heritability of BMI is likely 30–40%; therefore, there is little missing heritability (180). In order to fully understand the missing heritability, large sample sizes and new technologies are needed to discover more obesity-associated loci.

Advantages, limitations and future directions for multi-ethnic designs in obesity genetics

Advantages

While over 90% of obesity-susceptibility loci have been identified in European populations (Fig. 4), a growing number of GWAS are now being performed in populations of non-European ancestry in addition to replication and transferability studies (140). The inherent advantage of using multi-ethnic studies is identifying which genetic signals are shared across populations with distinct genetic ancestries or are ethnic specific (Table 1) (140). Multi-ethnic studies are also advantageous for identifying ethnic specific disease predisposing variants and private mutations. Moving beyond GWAS, other methods including whole-exome and whole-genome sequencing are better suited to assess rare variants and CNV and can be applied to multi-ethnic cohorts to reveal novel loci implicated in obesity.

As ethnic differences are observed in obesity predisposing genes, it is essential to assemble multi-ethnic designs to assess the ethnic-specific contribution of these genes (181). This study design is also a pre-requisite for identifying

Table 1 Summary of advantages and limitations of using multi-ethnic study designs in obesity genetics

Advantages	Limitations
Identify genetic variants shared across multiple ancestries, ethnic specific variants and private mutations	Definitions of ethnicity can vary between studies, limiting the generalizability of results
Differences in LD structure across diverse ethnic groups can be leveraged to pin-point causal variants	Weak LD structure in ancestral populations may result in weak, null or inverse associations, limiting transferability and replication
Reconstruct evolutive and non-evolutive forces (i.e. founder effect) that have shaped the genetic architecture of obesity susceptibility	Spurious results arise when population stratification and admixture are not considered and accounted for
Increased statistical power due to high frequencies of rare and deleterious variants in isolated populations or populations practising consanguinity	Small sample size of non-European replication cohorts and under-representation of ethnic minorities in multi-ethnic studies limits statistical power and the ability to detect associations
Identify novel G × G and G × E interactions due to the co-occurrence of interacting variants and unique environments	Achieving adequate statistical power to detect G × G and G × E interactions is difficult in multi-ethnic studies
Understanding of how societal practices (i.e. consanguineous marriages) influence genetic susceptibility for obesity	Unique societal practices can be population specific, limiting generalizability of results

LD, linkage disequilibrium; G × E, gene × environment interaction; G × G, gene × gene interaction.

causal variations using trans-ethnic fine-mapping approaches through candidate gene/locus resequencing or genotyping of custom arrays (i.e. metabochips) (144,182). Future large-scale trans-ethnic designs combining data from diverse ethnic groups with differences in LD structure can provide better resolution to identify causal variants for functional follow-up studies (140). The use of dense genome-wide SNP arrays (five million SNPs) in combination with whole-genome sequencing and imputation in multi-ethnic populations may lead to the identification of additional independent signals within GWAS loci and likely causal obesity predisposing variants for functional follow-up studies in one step in a near future (145,183,184).

Studying genetic differences in diverse ethnic groups is critical for reconstructing the evolutive and non-evolutive forces (i.e. genetic drift, migration and founder effect) that have shaped the genetic predisposition or protection from obesity in modern human populations (185). Interestingly, it may be advantageous to use small and historically isolated founder populations in genetic association studies due to their increased statistical power and high allele frequency of deleterious variants (28).

Multi-ethnic designs aid in the understanding of the impact of societal practices (i.e. intra-caste marriages, preferred consanguineous marriages, polygamy and assortative marriages) on the present and future genetic susceptibility to obesity (35). To overcome problems of limited statistical power, populations practising consanguinity can be used as this practice leads to an exceptionally high prevalence of rare homozygous mutations (85).

Multi-ethnic populations are also highly relevant for identifying G × G and G × E interactions. G × G interactions may be identified in certain ethnic groups due to the frequent co-occurrence of the interacting genetic variants (186). Similarly, novel G × E interactions may emerge when ethnic groups are exposed to unique combinations of lifestyle and environment (115).

Limitations

Early replication efforts aimed to directly replicate European candidate SNPs in independent cohorts reveal significant challenges as some disease-associated SNPs reaching genome-wide significance do not directly replicate in populations of different ancestries. Lack of replication across multiple ethnic groups can be attributed to several factors (187).

Limited statistical power due to relatively small sample sizes of non-European cohorts is a major challenge for replication studies (Table 1). Statistical power to detect an association is a function of sample size, the effect size and minor allele frequency. Often, replication cohorts are substantially smaller than the discovery population and ethnic minorities are under-represented in multi-ethnic studies. Variants identified in European GWAS generally have larger effect sizes and/or minor allele frequency, making them easier to discover (140). It is often unclear if the lack of significance in the replication cohort is the result of limited power/sample size or truly an absence of genetic association; thus, authors need to report power calculations when discussing the presence/absence of an association (188). The formation of large, international genomic consortia for replication in various ethnic groups may alleviate the problem of inadequate sample size (189).

Transferability may also be limited due to differences in genetic architecture. If the GWAS index SNP is tested in a population with different ancestry, the LD between the index SNP and causal SNP may be weaker than in the discovery population. The result is a weak, no or inverse (flip-flop) association with the trait in the replication cohort and a wrongful conclusion that the SNP does not transfer across ancestries (140). New methods are needed to assess differences in population allele frequencies and LD in order to determine which SNPs are expected (or not) to be replicated in other populations to avoid conducting replication studies in populations where the variant is too rare (188). For example, transferability studies in other ethnic groups may use a dense set of variants rather than testing the GWAS hit alone (141).

G × G and G × E interactions also challenge replication and transferability. Adequate statistical power is critical for interaction studies, and meta-analyses are recommended to reach sufficient power. Care must also be taken when selecting study design to adequately capture interactions with population based or nested case–control studies having a greater ability to detect interactions (190).

Population stratification and admixture must also be considered as polymorphism frequency may vary by ancestral origin. Using self-reported ethnicity is the simplest and most economical approach but does not adequately control for population stratification (15). Genetic classification of ancestry through ancestry informative markers provides a more objective and accurate method of defining ethnicity. Other methods including PCA using EIGENSOFT can precisely identify national and local ancestry (19). These corrections are required in genetic association studies to account for population stratification to minimize spurious associations without compromising power to detect true associations (191).

Lastly, BMI thresholds for obesity (BMI $\geq$ 30 kg m^{-2}) have been derived from European populations and correspond with an elevated risk for morbidity and mortality (192). The use of a single, universal threshold for obesity for non-European populations has been questioned with evidence, suggesting that Asian populations suffer from a greater risk of T2D, hypertension and dyslipidemia despite a low BMI (<25 kg m^{-2}) (193). Despite no formal

recommendation for ethnic-specific thresholds by the World Health Organization, the use of ethnic-specific thresholds has been proposed (193,194). Care must be taken when using multi-ethnic cohorts to avoid misclassifying non-European individuals for accurate estimates of the prevalence of obesity.

Conclusions

This review demonstrates the importance of a multi-ethnic perspective in the genetic elucidation of obesity. Identifying obesity predisposing genes in European populations has been undeniably successful, but non-European and multi-ethnic populations have been under-investigated so far. Multi-ethnic study designs have great potential to reconstruct the evolutionary history of genetic predisposition to obesity, isolate disease-causing variants and distinguish global from local G × G and G × E interactions. Novel epidemiological approaches including Mendelian randomization have been conducted predominately in European populations to determine the causal role of obesity loci in the pathology of obesity (195). While the results from Mendelian randomization studies in Europeans are considered universally valid, it is uncertain if these results hold true in different ethnic groups. Funding initiatives expanding gene identification efforts in non-European or isolated populations should be encouraged, especially in populations at high or low risk for obesity. Undoubtedly, such studies will enhance our understanding of the biological bases of obesity susceptibility and protection and encourage innovative prevention and treatment strategies. The observed unique ethnic patterns of genetic predisposition to obesity stress the limitations of a 'one size fits all' approach and emphasize the importance of ethnicity as we transition from big genetic data to precision medicine for all (196).

Conflict of interest statement

The authors have no conflict of interest to declare.

Acknowledgements

D. M. holds a Canada Research Chair in Genetics of Obesity.

References

1. de Onis M, Blossner M, Borghi E. Global prevalence and trends of overweight and obesity among preschool children. *Am J Clin Nutr* 2010; **92**: 1257–1264.
2. Must A, Spadano J, Coakley EH, Field AE, Colditz G, Dietz WH. The disease burden associated with overweight and obesity. *JAMA* 1999; **282**: 1523–1529.
3. Yanovski SZ, Yanovski JA. Obesity prevalence in the United States – up, down, or sideways? *N Engl J Med* 2011; **364**: 987–989.
4. Fontaine KR, Redden DT, Wang C, Westfall AO, Allison DB. Years of life lost due to obesity. *JAMA* 2003; **289**: 187–193.
5. Wang Y, Beydoun MA. The obesity epidemic in the United States – gender, age, socioeconomic, racial/ethnic, and geographic characteristics: a systematic review and meta-regression analysis. *Epidemiol Rev* 2007; **29**: 6–28.
6. Ogden CL, Carroll MD, Curtin LR, McDowell MA, Tabak CJ, Flegal KM. Prevalence of overweight and obesity in the United States, 1999–2004. *JAMA* 2006; **295**: 1549–1555.
7. Sankar P, Cho MK, Condit CM *et al*. Genetic research and health disparities. *JAMA* 2004; **291**: 2985–2989.
8. Kumar BN, Meyer HE, Wandel M, Dalen I, Holmboe-Ottesen G. Ethnic differences in obesity among immigrants from developing countries, in Oslo, Norway. *Int J Obes (Lond)* 2006; **30**: 684–690.
9. Yudell M, Roberts D, DeSalle R, Tishkoff S, Science and Society. Taking race out of human genetics. *Science* 2016; **351**: 564–565.
10. Williams DR. Race and health: basic questions, emerging directions. *Ann Epidemiol* 1997; **7**: 322–333.
11. Senior PA, Bhopal R. Ethnicity as a variable in epidemiological research. *BMJ* 1994; **309**: 327–330.
12. Collins FS. What we do and don't know about 'race', 'ethnicity', genetics and health at the dawn of the genome era. *Nat Genet* 2004; **36**: S13–S15.
13. Bonham VL, Warshauer-Baker E, Collins FS. Race and ethnicity in the genome era: the complexity of the constructs. *Am Psychol* 2005; **60**: 9–15.
14. Harford KA, Reynolds CM, McGillicuddy FC, Roche HM. Fats, inflammation and insulin resistance: insights to the role of macrophage and T-cell accumulation in adipose tissue. *Proc Nutr Soc* 2011; **70**: 408–417.
15. Serre D, Montpetit A, Pare G *et al*. Correction of population stratification in large multi-ethnic association studies. *PLoS One* 2008; **3**: e1382.
16. Arcos-Burgos M, Muenke M. Genetics of population isolates. *Clin Genet* 2002; **61**: 233–247.
17. Fujimura JH, Rajagopalan R. Different differences: the use of 'genetic ancestry' versus race in biomedical human genetic research. *Soc Stud Sci* 2011; **41**: 5–30.
18. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006; **38**: 904–909.
19. Karakachoff M, Duforet-Frebourg N, Simonet F *et al*. Fine-scale human genetic structure in Western France. *European journal of human genetics : EJHG* 2015; **23**: 831–836.
20. Neel JV. Diabetes mellitus: a "thrifty" genotype rendered detrimental by "progress"? *Am J Hum Genet* 1962; **14**: 353–362.
21. Speakman JR. Evolutionary perspectives on the obesity epidemic: adaptive, maladaptive, and neutral viewpoints. *Annu Rev Nutr* 2013; **33**: 289–317.
22. Lee YS. The role of genes in the current obesity epidemic. *Ann Acad Med Singapore* 2009; **38**: 45–43.
23. Schulz LO, Bennett PH, Ravussin E *et al*. Effects of traditional and western environments on prevalence of type 2 diabetes in Pima Indians in Mexico and the U.S. *Diabetes Care* 2006; **29**: 1866–1871.
24. Gerstein HC, Waltman L. Why don't pigs get diabetes? Explanations for variations in diabetes susceptibility in human populations living in a diabetogenic environment. *CMAJ* 2006; **174**: 25–26.
25. Southam L, Soranzo N, Montgomery SB *et al*. Is the thrifty genotype hypothesis supported by evidence based on confirmed type 2 diabetes- and obesity-susceptibility variants? *Diabetologia* 2009; **52**: 1846–1851.

26. Klimentidis YC, Abrams M, Wang J, Fernandez JR, Allison DB. Natural selection at genomic regions associated with obesity and type-2 diabetes: East Asians and sub-Saharan Africans exhibit high levels of differentiation at type-2 diabetes regions. *Hum Genet* 2011; **129**: 407–418.
27. Wang G, Speakman JR. Analysis of positive selection at single nucleotide polymorphisms associated with body mass index does not support the "thrifty gene" hypothesis. *Cell Metab* 2016; **24**: 531–541.
28. Minster RL, Hawley NL, Su CT *et al*. A thrifty variant in CREBRF strongly influences body mass index in Samoans. *Nat Genet* 2016; **48**: 1049–1054.
29. Hughes DA, Hinney A, Brumm H *et al*. Increased constraints on MC4R during primate and human evolution. *Hum Genet* 2009; **124**: 633–647.
30. Robinson MR, Hemani G, Medina-Gomez C *et al*. Population genetic differentiation of height and body mass index across Europe. *Nat Genet* 2015; **47**: 1357–1362.
31. Speakman JR. A nonadaptive scenario explaining the genetic predisposition to obesity: the "predation release" hypothesis. *Cell Metab* 2007; **6**: 5–12.
32. Speakman JR. Thrifty genes for obesity, an attractive but flawed idea, and an alternative perspective: the 'drifty gene' hypothesis. *Int J Obes (Lond)* 2008; **32**: 1611–1617.
33. Stoger R. The thrifty epigenotype: an acquired and heritable predisposition for obesity and diabetes? *BioEssays : news and reviews in molecular, cellular and developmental biology* 2008; **30**: 156–166.
34. Saeed S, Bonnefond A, Manzoor J *et al*. Genetic variants in LEP, LEPR, and MC4R explain 30% of severe obesity in children from a consanguineous population. *Obesity* 2015; **23**: 1687–1695.
35. Reich D, Thangaraj K, Patterson N, Price AL, Singh L. Reconstructing Indian population history. *Nature* 2009; **461**: 489–494.
36. Hebebrand J, Wulftange H, Goerg T *et al*. Epidemic obesity: are genetic factors involved via increased rates of assortative mating? *Int J Obes Relat Metab Disord* 2000; **24**: 345–353.
37. Reilly JJ, Armstrong J, Dorosty AR *et al*. Early life risk factors for obesity in childhood: cohort study. *BMJ* 2005; **330**: 1357.
38. Wardle J, Carnell S, Haworth CM, Plomin R. Evidence for a strong genetic influence on childhood adiposity despite the force of the obesogenic environment. *Am J Clin Nutr* 2008; **87**: 398–404.
39. O'Rahilly S, Farooqi IS. In: De Groot LJ, Chrousos G, Dungan K, Feingold KR, Grossman A, Hershman JM et al. (eds). The Genetics of Obesity in Humans. Endotext: South Dartmouth (MA), 2000.
40. Hinney A, Vogel CI, Hebebrand J. From monogenic to polygenic obesity: recent advances. *Eur Child Adolesc Psychiatry* 2010; **19**: 297–310.
41. Fernandez JR, Shriver MD, Beasley TM *et al*. Association of African genetic admixture with resting metabolic rate and obesity among women. *Obes Res* 2003; **11**: 904–911.
42. Cheng CY, Reich D, Coresh J *et al*. Admixture mapping of obesity-related traits in African Americans: the Atherosclerosis Risk in Communities (ARIC) Study. *Obesity* 2010; **18**: 563–572.
43. Nassir R, Qi L, Kosoy R *et al*. Relationship between adiposity and admixture in African-American and Hispanic-American women. *Int J Obes (Lond)* 2012; **36**: 304–313.
44. Cheng CY, Kao WH, Patterson N *et al*. Admixture mapping of 15,280 African Americans identifies obesity susceptibility loci on chromosomes 5 and X. *PLoS Genet* 2009; **5** e1000490.
45. Fernandez JR, Pearson KE, Kell KP, Bohan Brown MM. Genetic admixture and obesity: recent perspectives and future applications. *Hum Hered* 2013; **75**: 98–105.
46. Kaur Y, de Souza RJ, Gibson WT, Meyre D. A systematic review of genetic syndromes with obesity. *Obes Rev* 2017; **18**: 603–634.
47. Marshall JD, Maffei P, Collin GB, Naggert JK. Alstrom syndrome: genetics and clinical overview. *Curr Genomics* 2011; **12**: 225–235.
48. Marshall JD, Ludman MD, Shea SE *et al*. Genealogy, natural history, and phenotype of Alstrom syndrome in a large Acadian kindred and three additional families. *Am J Med Genet* 1997; **73**: 150–161.
49. Collin GB, Marshall JD, Ikeda A *et al*. Mutations in ALMS1 cause obesity, type 2 diabetes and neurosensory degeneration in Alstrom syndrome. *Nat Genet* 2002; **31**: 74–78.
50. Minton JA, Owen KR, Ricketts CJ *et al*. Syndromic obesity and diabetes: changes in body composition with age and mutation analysis of ALMS1 in 12 United Kingdom kindreds with Alstrom syndrome. *J Clin Endocrinol Metab* 2006; **91**: 3110–3116.
51. Bond J, Flintoff K, Higgins J *et al*. The importance of seeking ALMS1 mutations in infants with dilated cardiomyopathy. *J Med Genet* 2005; **42**: e10.
52. Marshall JD, Hinman EG, Collin GB *et al*. Spectrum of ALMS1 variants and evaluation of genotype-phenotype correlations in Alstrom syndrome. *Hum Mutat* 2007; **28**: 1114–1123.
53. Aldahmesh MA, Abu-Safieh L, Khan AO *et al*. Allelic heterogeneity in inbred populations: the Saudi experience with Alstrom syndrome as an illustrative example. *Am J Med Genet A* 2009; **149A**: 662–665.
54. Forsythe E, Beales PL. Bardet-Biedl syndrome. *European journal of human genetics : EJHG* 2013; **21**: 8–13.
55. Heon E, Kim G, Qin S *et al*. Mutations in C8ORF37 cause Bardet Biedl syndrome (BBS21). *Hum Mol Genet* 2016; **25**: 2283–2294.
56. Novas R, Cardenas-Rodriguez M, Irigoin F, Badano JL. Bardet-Biedl syndrome: Is it only cilia dysfunction? *FEBS Lett* 2015; **589**: 3479–3491.
57. Schaefer E, Stoetzel C, Scheidecker S *et al*. Identification of a novel mutation confirms the implication of IFT172 (BBS20) in Bardet-Biedl syndrome. *J Hum Genet* 2016; **61**: 447–450.
58. Pigeyre M, Yazdi FT, Kaur Y, Meyre D. Recent progress in genetics, epigenetics and metagenomics unveils the pathophysiology of human obesity. *Clin Sci (Lond)* 2016; **130**: 943–986.
59. Guo DF, Rahmouni K. Molecular basis of the obesity associated with Bardet-Biedl syndrome. *Trends in endocrinology and metabolism: TEM* 2011; **22**: 286–293.
60. Locke AE, Kahali B, Berndt SI *et al*. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015; **518**: 197–206.
61. Farag TI, Teebi AS. High incidence of Bardet Biedl syndrome among the Bedouin. *Clin Genet* 1989; **36**: 463–464.
62. Moore SJ, Green JS, Fan Y *et al*. Clinical and genetic epidemiology of Bardet-Biedl syndrome in Newfoundland: a 22-year prospective, population-based, cohort study. *Am J Med Genet A* 2005; **132**: 352–360.
63. M'Hamdi O, Ouertani I, Maazoul F, Chaabouni-Bouhamed H. Prevalence of Bardet-Biedl syndrome in Tunisia. *J Community Genet* 2011; **2**: 97–99.
64. Kolehmainen J, Black GC, Saarinen A *et al*. Cohen syndrome is caused by mutations in a novel gene, COH1, encoding a transmembrane protein with a presumed role in vesicle-mediated sorting and intracellular protein transport. *Am J Hum Genet* 2003; **72**: 1359–1369.
65. Seifert W, Kuhnisch J, Maritzen T, Horn D, Haucke V, Hennies HC. Cohen syndrome-associated protein, COH1, is a novel, giant

Golgi matrix protein required for Golgi integrity. *J Biol Chem* 2011; **286**: 37665–37675.
66. Parri V, Katzaki E, Uliana V *et al*. High frequency of COH1 intragenic deletions and duplications detected by MLPA in patients with Cohen syndrome. *European journal of human genetics : EJHG* 2010; **18**: 1133–1140.
67. Douzgou S, Petersen MB. Clinical variability of genetic isolates of Cohen syndrome. *Clin Genet* 2011; **79**: 501–506.
68. Hennies HC, Rauch A, Seifert W *et al*. Allelic heterogeneity in the COH1 gene explains clinical variability in Cohen syndrome. *Am J Hum Genet* 2004; **75**: 138–145.
69. Murphy AM, Flanagan O, Dunne K, Lynch SA. High prevalence of Cohen syndrome among Irish travellers. *Clin Dysmorphol* 2007; **16**: 257–259.
70. Bugiani M, Gyftodimou Y, Tsimpouka P *et al*. Cohen syndrome resulting from a novel large intragenic COH1 deletion segregating in an isolated Greek island population. *Am J Med Genet A* 2008; **146A**: 2221–2226.
71. Cassidy SB, Driscoll DJ. Prader-Willi syndrome. *European journal of human genetics : EJHG* 2009; **17**: 3–13.
72. Stefan M, Nicholls RD. What have rare genetic syndromes taught us about the pathophysiology of the common forms of obesity? *Curr Diab Rep* 2004; **4**: 143–150.
73. Robinson WP, Bottani A, Xie YG *et al*. Molecular, cytogenetic, and clinical investigations of Prader-Willi syndrome patients. *Am J Hum Genet* 1991; **49**: 1219–1234.
74. Sahoo T, del Gaudio D, German JR *et al*. Prader-Willi phenotype caused by paternal deficiency for the HBII-85 C/D box small nucleolar RNA cluster. *Nat Genet* 2008; **40**: 719–721.
75. Burd L, Vesely B, Martsolf J, Kerbeshian J. Prevalence study of Prader-Willi syndrome in North Dakota. *Am J Med Genet* 1990; **37**: 97–99.
76. Butler MG. Prader-Willi syndrome: current understanding of cause and diagnosis. *Am J Med Genet* 1990; **35**: 319–332.
77. Akefeldt A, Gillberg C, Larsson C. Prader-Willi syndrome in a Swedish rural county: epidemiological aspects. *Dev Med Child Neurol* 1991; **33**: 715–721.
78. Ehara H, Ohno K, Takeshita K. Frequency of the Prader-Willi syndrome in the San-in district. *Japan Brain & development* 1995; **17**: 324–326.
79. Vogels A, Van Den Ende J, Keymolen K *et al*. Minimum prevalence, birth incidence and cause of death for Prader-Willi syndrome in Flanders. *European journal of human genetics : EJHG* 2004; **12**: 238–240.
80. Lionti T, Reid SM, White SM, Rowell MM. A population-based profile of 160 Australians with Prader-Willi syndrome: trends in diagnosis, birth prevalence and birth characteristics. *Am J Med Genet A* 2015; **167A**: 371–378.
81. Whittington JE, Holland AJ, Webb T, Butler J, Clarke D, Boer H. Population prevalence and estimated birth incidence and mortality rate for people with Prader-Willi syndrome in one UK Health Region. *J Med Genet* 2001; **38**: 792–798.
82. Choquet H, Meyre D. Molecular basis of obesity: current status and future prospects. *Curr Genomics* 2011; **12**: 154–168.
83. Farooqi IS, Keogh JM, Kamath S *et al*. Partial leptin deficiency and human adiposity. *Nature* 2001; **414**: 34–35.
84. Funcke JB, von Schnurbein J, Lennerz B *et al*. Monogenic forms of childhood obesity due to mutations in the leptin gene. *Mol Cell Pediatr* 2014; **1**: 3.
85. Saeed S, Butt TA, Anwer M, Arslan M, Froguel P. High prevalence of leptin and melanocortin-4 receptor gene mutations in children with severe obesity from Pakistani consanguineous families. *Mol Genet Metab* 2012; **106**: 121–126.
86. Paz-Filho G, Mastronardi CA, Licinio J. Leptin treatment: facts and expectations. *Metabolism* 2015; **64**: 146–156.
87. Farooqi IS, Wangensteen T, Collins S *et al*. Clinical and molecular genetic spectrum of congenital deficiency of the leptin receptor. *N Engl J Med* 2007; **356**: 237–247.
88. Andiran N, Celik N, Andiran F. Homozygosity for two missense mutations in the leptin receptor gene (P316:W646C) in a Turkmenian girl with severe early-onset obesity. *Journal of pediatric endocrinology & metabolism : JPEM* 2011; **24**: 1043–1045.
89. Mazen I, El-Gammal M, Abdel-Hamid M, Farooqi IS, Amr K. Homozygosity for a novel missense mutation in the leptin receptor gene (P316T) in two Egyptian cousins with severe early onset obesity. *Mol Genet Metab* 2011; **102**: 461–464.
90. Huvenne H, Le Beyec J, Pepin D *et al*. Seven novel deleterious LEPR mutations found in early-onset obesity: a DeltaExon6-8 shared by subjects from Reunion Island, France, suggests a founder effect. *J Clin Endocrinol Metab* 2015; **100**: E757–E766.
91. Buono P, Pasanisi F, Nardelli C *et al*. Six novel mutations in the proopiomelanocortin and melanocortin receptor 4 genes in severely obese adults living in southern Italy. *Clin Chem* 2005; **51**: 1358–1364.
92. Challis BG, Pritchard LE, Creemers JW *et al*. A missense mutation disrupting a dibasic prohormone processing site in pro-opiomelanocortin (POMC) increases susceptibility to early-onset obesity through a novel molecular mechanism. *Hum Mol Genet* 2002; **11**: 1997–2004.
93. Farooqi IS, Drop S, Clements A *et al*. Heterozygosity for a POMC-null mutation and increased obesity risk in humans. *Diabetes* 2006; **55**: 2549–2553.
94. Krude H, Biebermann H, Luck W, Horn R, Brabant G, Gruters A. Severe early-onset obesity, adrenal insufficiency and red hair pigmentation caused by POMC mutations in humans. *Nat Genet* 1998; **19**: 155–157.
95. Lee YS, Challis BG, Thompson DA *et al*. A POMC variant implicates beta-melanocyte-stimulating hormone in the control of human energy balance. *Cell Metab* 2006; **3**: 135–140.
96. Dubern B, Lubrano-Berthelier C, Mencarelli M *et al*. Mutational analysis of the pro-opiomelanocortin gene in French obese children led to the identification of a novel deleterious heterozygous mutation located in the alpha-melanocyte stimulating hormone domain. *Pediatr Res* 2008; **63**: 211–216.
97. Hung CN, Poon WT, Lee CY, Law CY, Chan AY. A case of early-onset obesity, hypocortisolism, and skin pigmentation problem due to a novel homozygous mutation in the pro-opiomelanocortin (POMC) gene in an Indian boy. *Journal of pediatric endocrinology & metabolism : JPEM* 2012; **25**: 175–179.
98. Clement K, Dubern B, Mencarelli M *et al*. Unexpected endocrine features and normal pigmentation in a young adult patient carrying a novel homozygous mutation in the POMC gene. *J Clin Endocrinol Metab* 2008; **93**: 4955–4962.
99. Cirillo G, Marini R, Ito S *et al*. Lack of red hair phenotype in a North-African obese child homozygous for a novel POMC null mutation: nonsense-mediated decay RNA evaluation and hair pigment chemical analysis. *Br J Dermatol* 2012; **167**: 1393–1395.
100. Creemers JW, Choquet H, Stijnen P *et al*. Heterozygous mutations causing partial prohormone convertase 1 deficiency contribute to human obesity. *Diabetes* 2012; **61**: 383–390.
101. Martin MG, Lindberg I, Solorzano-Vargas RS *et al*. Congenital proprotein convertase 1/3 deficiency causes malabsorptive diarrhea and other endocrinopathies in a pediatric cohort. *Gastroenterology* 2013; **145**: 138–148.
102. Farooqi IS, Volders K, Stanhope R *et al*. Hyperphagia and early-onset obesity due to a novel homozygous missense mutation

in prohormone convertase 1/3. *J Clin Endocrinol Metab* 2007; **92**: 3369–3373.
103. Harter B, Fuchs I, Muller T, Akbulut UE, Cakir M, Janecke AR. Early clinical diagnosis of PC1/3 deficiency in a patient with a novel homozygous PCSK1 splice-site mutation. *J Pediatr Gastroenterol Nutr* 2016; **62**: 577–580.
104. Yourshaw M, Solorzano-Vargas RS, Pickett LA *et al*. Exome sequencing finds a novel PCSK1 mutation in a child with generalized malabsorptive diarrhea and diabetes insipidus. *J Pediatr Gastroenterol Nutr* 2013; **57**: 759–767.
105. O'Rahilly S, Gray H, Humphreys PJ *et al*. Brief report: impaired processing of prohormones associated with abnormalities of glucose homeostasis and adrenal function. *N Engl J Med* 1995; **333**: 1386–1390.
106. Jackson RS, Creemers JW, Ohagi S *et al*. Obesity and impaired prohormone processing associated with mutations in the human prohormone convertase 1 gene. *Nat Genet* 1997; **16**: 303–306.
107. Jackson RS, Creemers JW, Farooqi IS *et al*. Small-intestinal dysfunction accompanies the complex endocrinopathy of human proprotein convertase 1 deficiency. *J Clin Invest* 2003; **112**: 1550–1560.
108. Frank GR, Fox J, Candela N *et al*. Severe obesity and diabetes insipidus in a patient with PCSK1 deficiency. *Mol Genet Metab* 2013; **110**: 191–194.
109. Xi B, Chandak GR, Shen Y, Wang Q, Zhou D. Association between common polymorphism near the MC4R gene and obesity risk: a systematic review and meta-analysis. *PLoS One* 2012; **7**: e45731.
110. Dubern B, Bisbis S, Talbaoui H *et al*. Homozygous null mutation of the melanocortin-4 receptor and severe early-onset obesity. *J Pediatr* 2007; **150**: 613–617 17 e1.
111. Vollbach H, Brandt S, Lahr G *et al*. Prevalence and phenotypic characterization of MC4R variants in a large pediatric cohort. *Int J Obes (Lond)* 2017; **41**: 13–22.
112. Vaisse C, Clement K, Durand E, Hercberg S, Guy-Grand B, Froguel P. Melanocortin-4 receptor mutations are a frequent and heterogeneous cause of morbid obesity. *J Clin Invest* 2000; **106**: 253–262.
113. Farooqi IS, Yeo GS, Keogh JM *et al*. Dominant and recessive inheritance of morbid obesity associated with melanocortin 4 receptor deficiency. *J Clin Invest* 2000; **106**: 271–279.
114. Hinney A, Hohmann S, Geller F *et al*. Melanocortin-4 receptor gene: case-control study and transmission disequilibrium test confirm that functionally relevant mutations are compatible with a major gene effect for extreme obesity. *J Clin Endocrinol Metab* 2003; **88**: 4258–4267.
115. Rouskas K, Meyre D, Stutzmann F *et al*. Loss-of-function mutations in MC4R are very rare in the Greek severely obese adult population. *Obesity* 2012; **20**: 2278–2282.
116. Miraglia Del Giudice E, Cirillo G, Nigro V *et al*. Low frequency of melanocortin-4 receptor (MC4R) mutations in a Mediterranean population with early-onset obesity. *Int J Obes Relat Metab Disord* 2002; **26**: 647–651.
117. Ohshiro Y, Sanke T, Ueda K *et al*. Molecular scanning for mutations in the melanocortin-4 receptor gene in obese/diabetic Japanese. *Ann Hum Genet* 1999; **63**: 483–487.
118. Kobayashi H, Ogawa Y, Shintani M *et al*. A Novel homozygous missense mutation of melanocortin-4 receptor (MC4R) in a Japanese woman with severe obesity. *Diabetes* 2002; **51**: 243–246.
119. Rong R, Tao YX, Cheung BM, Xu A, Cheung GC, Lam KS. Identification and functional characterization of three novel human melanocortin-4 receptor gene variants in an obese Chinese population. *Clin Endocrinol (Oxf)* 2006; **65**: 198–205.
120. Wang CL, Liang L, Wang HJ, Fu JF, Hebebrand J, Hinney A. Several mutations in the melanocortin 4 receptor gene are associated with obesity in Chinese children and adolescents. *J Endocrinol Invest* 2006; **29**: 894–898.
121. Lee YS, Poh LK, Kek BL, Loke KY. Novel melanocortin 4 receptor gene mutations in severely obese children. *Clin Endocrinol (Oxf)* 2008; **68**: 529–535.
122. Thearle MS, Muller YL, Hanson RL *et al*. Greater impact of melanocortin-4 receptor deficiency on rates of growth and risk of type 2 diabetes during childhood compared with adulthood in Pima Indians. *Diabetes* 2012; **61**: 250–257.
123. Ma L, Tataranni PA, Bogardus C, Baier LJ. Melanocortin 4 receptor gene variation is associated with severe obesity in Pima Indians. *Diabetes* 2004; **53**: 2696–2699.
124. Stutzmann F, Tan K, Vatin V *et al*. Prevalence of melanocortin-4 receptor deficiency in Europeans and their age-dependent penetrance in multigenerational pedigrees. *Diabetes* 2008; **57**: 2511–2518.
125. Frayling TM. Genome-wide association studies provide new insights into type 2 diabetes aetiology. *Nat Rev Genet* 2007; **8**: 657–662.
126. Dina C, Meyre D, Gallina S *et al*. Variation in FTO contributes to childhood obesity and severe adult obesity. *Nat Genet* 2007; **39**: 724–726.
127. Scuteri A, Sanna S, Chen WM *et al*. Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet* 2007; **3**: e115.
128. Hinney A, Nguyen TT, Scherag A *et al*. Genome wide association (GWA) study for early onset extreme obesity supports the role of fat mass and obesity associated gene (FTO) variants. *PLoS One* 2007; **2**: e1361.
129. Fesinmeyer MD, North KE, Ritchie MD *et al*. Genetic risk factors for BMI and obesity in an ethnically diverse population: results from the population architecture using genomics and epidemiology (PAGE) study. *Obesity* 2013; **21**: 835–846.
130. Felix JF, Bradfield JP, Monnereau C *et al*. Genome-wide association analysis identifies three new susceptibility loci for childhood body mass index. *Hum Mol Genet* 2016; **25**: 389–403.
131. Wen W, Zheng W, Okada Y *et al*. Meta-analysis of genome-wide association studies in East Asian-ancestry populations identifies four new loci for body mass index. *Hum Mol Genet* 2014; **23**: 5492–5504.
132. Okada Y, Kubo M, Ohmiya H *et al*. Common variants at CDKAL1 and KLF9 are associated with body mass index in east Asian populations. *Nat Genet* 2012; **44**: 302–306.
133. Croteau-Chonka DC, Marvelle AF, Lange EM *et al*. Genome-wide association study of anthropometric traits and evidence of interactions with age and study year in Filipino women. *Obesity* 2011; **19**: 1019–1027.
134. Been LF, Nath SK, Ralhan SK *et al*. Replication of association between a common variant near melanocortin-4 receptor gene and obesity-related traits in Asian Sikhs. *Obesity* 2010; **18**: 425–429.
135. Ahmad S, Zhao W, Renstrom F *et al*. A novel interaction between the FLJ33534 locus and smoking in obesity: a genome-wide study of 14 131 Pakistani adults. *Int J Obes (Lond)* 2016; **40**: 186–190.
136. Hassanein MT, Lyon HN, Nguyen TT *et al*. Fine mapping of the association with obesity at the FTO locus in African-derived populations. *Hum Mol Genet* 2010; **19**: 2907–2916.
137. Kang SJ, Chiang CW, Palmer CD *et al*. Genome-wide association of anthropometric traits in African- and African-derived populations. *Hum Mol Genet* 2010; **19**: 2725–2738.

138. Monda KL, Chen GK, Taylor KC *et al*. A meta-analysis identifies new loci associated with body mass index in individuals of African ancestry. *Nat Genet* 2013; **45**: 690–696.
139. Manolio TA, Collins FS, Cox NJ *et al*. Finding the missing heritability of complex diseases. *Nature* 2009; **461**: 747–753.
140. Lu Y, Loos RJ. Obesity genomics: assessing the transferability of susceptibility loci across diverse populations. *Genome Med* 2013; **5**: 55.
141. Grant SF, Li M, Bradfield JP *et al*. Association analysis of the FTO gene with obesity in children of Caucasian and African ancestry reveals a common tagging SNP. *PLoS One* 2008; **3**: e1746.
142. Smemo S, Tena JJ, Kim KH *et al*. Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature* 2014; **507**: 371–375.
143. Claussnitzer M, Dankel SN, Kim KH *et al*. FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. *N Engl J Med* 2015; **373**: 895–907.
144. Peters U, North KE, Sethupathy P *et al*. A systematic mapping approach of 16q12.2/FTO and BMI in more than 20,000 African Americans narrows in on the underlying functional variation: results from the Population Architecture using Genomics and Epidemiology (PAGE) study. *PLoS Genet* 2013; **9** e1003171.
145. Gong J, Schumacher F, Lim U *et al*. Fine Mapping and Identification of BMI Loci in African Americans. *Am J Hum Genet* 2013; **93**: 661–671.
146. McKeigue PM. Prospects for admixture mapping of complex traits. *Am J Hum Genet* 2005; **76**: 1–7.
147. Basu A, Tang H, Arnett D *et al*. Admixture mapping of quantitative trait loci for BMI in African Americans: evidence for loci on chromosomes 3q, 5q, and 15q. *Obesity (Silver Spring)* 2009; **17**: 1226–1231.
148. Shetty PB, Tang H, Tayo BO *et al*. Variants in CXADR and F2RL1 are associated with blood pressure and obesity in African-Americans in regions identified through admixture mapping. *J Hypertens* 2012; **30**: 1970–1976.
149. Sandholt CH, Hansen T, Pedersen O. Beyond the fourth wave of genome-wide obesity association studies. *Nutr Diabetes* 2012; **2**: e37.
150. Lessard S, Manning AK, Low-Kam C *et al*. Testing the role of predicted gene knockouts in human anthropometric trait variation. *Hum Mol Genet* 2016; **25**: 2082–2092.
151. Genomes Project C, Auton A, Brooks LD *et al*. A global reference for human genetic variation. *Nature* 2015; **526**: 68–74.
152. Saunders CL, Chiodini BD, Sham P *et al*. Meta-analysis of genome-wide linkage studies in BMI and obesity. *Obesity* 2007; **15**: 2263–2275.
153. Ichimura A, Hirasawa A, Poulain-Godefroy O *et al*. Dysfunction of lipid sensor GPR120 leads to obesity in both mouse and human. *Nature* 2012; **483**: 350–354.
154. Benzinou M, Creemers JW, Choquet H *et al*. Common nonsynonymous variants in PCSK1 confer risk of obesity. *Nat Genet* 2008; **40**: 943–945.
155. Stutzmann F, Vatin V, Cauchi S *et al*. Non-synonymous polymorphisms in melanocortin-4 receptor protect against obesity: the two facets of a Janus obesity gene. *Hum Mol Genet* 2007; **16**: 1837–1844.
156. Walters RG, Jacquemont S, Valsesia A *et al*. A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. *Nature* 2010; **463**: 671–675.
157. Willer CJ, Speliotes EK, Loos RJ *et al*. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat Genet* 2009; **41**: 25–34.
158. Speliotes EK, Willer CJ, Berndt SI *et al*. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* 2010; **42**: 937–948.
159. Jarick I, Vogel CI, Scherag S *et al*. Novel common copy number variation for early onset extreme obesity on chromosome 11q11 identified by a genome-wide analysis. *Hum Mol Genet* 2011; **20**: 840–852.
160. Sha BY, Yang TL, Zhao LJ *et al*. Genome-wide association study suggested copy number variation may be associated with body mass index in the Chinese population. *J Hum Genet* 2009; **54**: 199–202.
161. Aerts E, Beckers S, Zegers D *et al*. CNV analysis and mutation screening indicate an important role for the NPY4R gene in human obesity. *Obesity* 2016; **24**: 970–976.
162. Wellcome Trust Case Control C, Craddock N, Hurles ME *et al*. Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. *Nature* 2010; **464**: 713–720.
163. Reddon H, Gueant JL, Meyre D. The importance of gene-environment interactions in human obesity. *Clin Sci (Lond)* 2016; **130**: 1571–1597.
164. Reddon H, Gerstein HC, Engert JC *et al*. Physical activity and genetic predisposition to obesity in a multiethnic longitudinal study. *Sci Rep* 2016; **6**: 18672.
165. Kilpelainen TO, Qi L, Brage S *et al*. Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. *PLoS Med* 2011; **8** e1001116.
166. Taylor AE, Sandeep MN, Janipalli CS *et al*. Associations of FTO and MC4R variants with obesity traits in Indians and the role of rural/urban environment as a possible effect modifier. *Journal of obesity* 2011; **2011**: 307542.
167. Vasan SK, Fall T, Neville MJ *et al*. Associations of variants in FTO and near MC4R with obesity traits in South Asian Indians. *Obesity* 2012; **20**: 2268–2277.
168. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet* 2008; **9**: 465–476.
169. Albuquerque D, Stice E, Rodriguez-Lopez R, Manco L, Nobrega C. Current review of genetics of human obesity: from molecular mechanisms to an evolutionary perspective. *Mol Genet Genomics* 2015; **290**: 1191–1221.
170. Wahl S, Drong A, Lehne B *et al*. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* 2017; **541**: 81–86.
171. Dick KJ, Nelson CP, Tsaprouni L *et al*. DNA methylation and body-mass index: a genome-wide analysis. *Lancet* 2014; **383**: 1990–1998.
172. Kuhnen P, Handke D, Waterland RA *et al*. Interindividual variation in DNA methylation at a putative POMC metastable epiallele is associated with obesity. *Cell Metab* 2016; **24**: 502–509.
173. Bell CG, Finer S, Lindgren CM *et al*. Integrated genetic and epigenetic analysis identifies haplotype-specific methylation in the FTO type 2 diabetes and obesity susceptibility locus. *PLoS One* 2010; **5**: e14040.
174. Voisin S, Almen MS, Zheleznyakova GY *et al*. Many obesity-associated SNPs strongly associate with DNA methylation changes at proximal promoters and enhancers. *Genome Med* 2015; **7**: 103.
175. Demerath EW, Guan W, Grove ML *et al*. Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. *Hum Mol Genet* 2015; **24**: 4464–4479.
176. Pan H, Lin X, Wu Y *et al*. HIF3A association with adiposity: the story begins before birth. *Epigenomics* 2015; **7**: 937–950.

177. Mendelson MM, Marioni RE, Joehanes R *et al*. Association of body mass index with DNA methylation and gene expression in blood cells and relations to cardiometabolic disease: a Mendelian randomization approach. *PLoS Med* 2017; **14** e1002215.
178. Llewellyn CH, Trzaskowski M, Plomin R, Wardle J. Finding the missing heritability in pediatric obesity: the contribution of genome-wide complex trait analysis. *Int J Obes (Lond)* 2013; **37**: 1506–1509.
179. Yang J, Bakshi A, Zhu Z *et al*. Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat Genet* 2015; **47**: 1114–1120.
180. Zhang D, Li Z, Wang H *et al*. Interactions between obesity-related copy number variants and dietary behaviors in childhood obesity. *Nutrients* 2015; **7**: 3054–3066.
181. Fesinmeyer MD, North KE, Ritchie MD *et al*. Genetic risk factors for BMI and obesity in an ethnically diverse population: results from the population architecture using genomics and epidemiology (PAGE) study. *Obesity* 2012; **21**: 835–846.
182. Voight BF, Kang HM, Ding J *et al*. The metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. *PLoS Genet* 2012; **8** e1002793.
183. Holm H, Gudbjartsson DF, Sulem P *et al*. A rare variant in MYH6 is associated with high risk of sick sinus syndrome. *Nat Genet* 2011; **43**: 316–320.
184. Ong RT, Wang X, Liu X, Teo YY. Efficiency of trans-ethnic genome-wide meta-analysis and fine-mapping. *European journal of human genetics : EJHG* 2012; **20**: 1300–1307.
185. Helgason A, Palsson S, Thorleifsson G *et al*. Refining the impact of TCF7L2 gene variants on type 2 diabetes and adaptive evolution. *Nat Genet* 2007; **39**: 218–225.
186. Feng N, Young SF, Aguilera G *et al*. Co-occurrence of two partially inactivating polymorphisms of MC3R is associated with pediatric-onset obesity. *Diabetes* 2005; **54**: 2663–2667.
187. Hennig BJ, Fulford AJ, Sirugo G *et al*. FTO gene variation and measures of body mass in an African population. *BMC Med Genet* 2009; **10**: 21.
188. Li YR, Keating BJ. Trans-ethnic genome-wide association studies: advantages and challenges of mapping in diverse populations. *Genome Med* 2014; **6**: 91.
189. Rich SS, Concannon P, Erlich H *et al*. The Type 1 Diabetes Genetics Consortium. *Ann N Y Acad Sci* 2006; **1079**: 1–8.
190. Li A, Meyre D. Challenges in reproducibility of genetic association studies: lessons learned from the obesity field. *Int J Obes (Lond)* 2013; **37**: 559–567.
191. Barnholtz-Sloan JS, McEvoy B, Shriver MD, Rebbeck TR. Ancestry estimation and correction for population stratification in molecular epidemiologic association studies. *Cancer Epidemiol Biomarkers Prev* 2008; **17**: 471–477.
192. Berrington de Gonzalez A, Hartge P, Cerhan JR *et al*. Body-mass index and mortality among 1.46 million white adults. *N Engl J Med* 2010; **363**: 2211–2219.
193. Razak F, Anand SS, Shannon H *et al*. Defining obesity cut points in a multiethnic population. *Circulation* 2007; **115**: 2111–2118.
194. Consultation WHOE. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet* 2004; **363**: 157–163.
195. Franks PW, Atabaki-Pasdar N. Causal inference in obesity research. *J Intern Med* 2017; **281**: 222–232.
196. Alyass A, Turcotte M, Meyre D. From big data analysis to personalized medicine for all: challenges and opportunities. *BMC Med Genomics* 2015; **8**: 33.

