Chipping Away the ‘Missing Heritability’: GIANT Steps Forward in the Molecular Elucidation of Obesity – but Still Lots to Go

Johannes Hebebrand  Anna-Lena Volckmar  Nadja Knoll  Anke Hinney

Department of Child and Adolescent Psychiatry and Psychotherapy, University of Duisburg-Essen, Essen, Germany

Summary

Although heritability of human body weight is assumed to be high, only a small fraction of the variance can as yet be attributed to molecular genetic factors. Single monogenic forms of obesity have been identified. Functionally relevant coding mutations in the melanocortin-4 receptor gene occur in 1–6% of extremely obese children and adolescents and thus represent the most common major gene effect. Genome-wide association studies (GWAS) had previously identified 14 obesity loci in independent individuals (Speliotes et al: Association analyses of 249,796 individuals reveal eighteen new loci associated with body mass index. Nature Genetics; epub October 2010 [1]). Apart from confirming the 14 known loci, 18 novel BMI-associated loci (p < 5 x 10^-8) were identified. Several of the new loci point to genes involved in key hypothalamic pathways of energy balance. The identified variants mostly have small to very small effect sizes; only 1–2% of the BMI variance is explained. Currently, a consensus explanation for this ‘missing heritability’ in complex diseases has not yet emerged.

Keywords

Synthetic association · Major gene · Polygene

Monogenic Obesity and Major Gene Effects

The cloning and initial detection of mutations in the mouse and human leptin gene [2, 3] marked the beginning of the large-scale elucidation of DNA variation underlying inter-individual differences in body weight. Currently, several types of monogenic forms of obesity have been identified [4]. All of the respective mutations are rare, some have only been reported in single cases worldwide; the respective genes are all expressed in the hypothalamus. The elucidation of such mutations and the successful treatment of leptin-deficient extremely obese individuals [e.g. 5] has firmly established that the behavioural phenotype hyperphagia/overeating and, as a consequence, obesity can be caused by mutations in specific genes. These findings have provided further support to the notion that obese individuals are responsible for their excess body weight to a limited extent only [6].

Whereas the elucidation of each one of these monogenic forms of obesity has proven invaluable with respect to the insight into hypothalamic pathways involved in body weight regulation, the clinical implications of each of these findings is limited, because of the rare occurrence of the respective mutations. For instance, only single research groups worldwide have reported on patients with functionally relevant mutations in the leptin or leptin receptor gene. For most of the monogenic forms of obesity [e.g. 4], only single pedigrees with an affected individual have been reported. Furthermore, several of the detected novel monogenic forms of obesity actually do not entail idiopathic obesity; instead the mutations can be viewed as giving rise to specific syndromes, which apart from obesity also encompass other clinically recognizable features. For instance, mutations in both the leptin [3] and leptin receptor [7] genes entail hypothalamic hypogonadism; other endo-
crinological abnormalities and immune system dysfunction have also been associated with leptin deficiency [8]. In the original report of the clinical features of patients with mutations in the pro-opiomelanocortin gene (POMC), adrenal insufficiency and red hair were characteristic features in addition to extreme obesity [9]. Developmental delay was shown to co-occur with severe obesity in a patient with a mutation in the neurotrophin receptor TrkB [10].

The detection of mutations in the melanocortin-4 receptor gene (MC4R) [11–13] marked a turning point in the sense that for the first time the molecular basis of a subgroup of patients with idiopathic obesity was identified. Whereas hyperinsulinemia and an elevated height growth have been reported as additional features of carriers of MC4R mutations [14] the respective phenotype does not qualify as a readily distinct and clinically recognizable syndrome. The detection of the association between MC4R mutations and obesity is also noteworthy because the combined frequency of all functionally relevant mutations is in the range of 1–6% among extremely obese children and adolescents [see 15]. MC4R mutations thus represent the first genetic form of obesity that affects a sizeable proportion of the obese population. In some European case-control studies a single haplotype encompassing both a stop (Tyr35stop) and a non-translated exchange that would on its own lead to a non-synonymous variant (Asp37Val) occurs in approximately 0.5–1% of extremely obese children and adolescents [13, 15, 16].

It is debatable if the term ‘monogenic obesity’ should be used in the context of MC4R mutations. Whereas there is little doubt that the rare occurrence of homozygosity or compound heterozygosity for functionally relevant mutations entails extreme obesity [14, 15, 17], the effect sizes of mutations in heterozygous carriers warrant discussion. For one, the body weights of heterozygous knockout mice are elevated, but overlap with those of wild-type mice [18]. Particularly in male mice, this overlap has been shown to be substantial. Furthermore, work in humans early on indicated that functionally relevant mutations also occur in normal-weight or lean controls [19–24]. Within pedigrees, obesity and MC4R mutations do not completely co-segregate [21, 22]; both obesity in wild-type carriers and normal weight/leanness in mutation carriers were detected.

Two family-based studies have attempted to quantify the effect of functionally relevant MC4R mutations [20, 25] by comparing BMI of mutation and wild-type carriers of relatives ascertainment via the index cases. Both studies came to remarkably similar results: According to Dempfle et al. [20] and Stutzmann et al. [25] adult male carriers were 4 and 4.3 kg/m² heavier than their male relatives with the wild-type genotype. In females, the effect size of MC4R mutations is larger; the corresponding values were 9.5 and 8.7 kg/m². In both of these studies all functionally relevant mutations were analysed as one; however, it appears highly likely that the effect sizes of such mutations actually differ. In this context, mutations with a dominant negative effect [26, 27] may entail the largest effect sizes.

The discussion as to whether MC4R mutations entail co-dominantly inherited obesity with reduced penetrance or represent a major gene effect is important in both clinical and research terms. If such mutations are detected in a child, genetic counselling of the parents and potentially other family members, and in future of the child himself/herself requires precise knowledge of the risks for obesity in relationship to the carrier status. Currently, we believe that provision of the information that adult male and female carriers of average height are on average approximately 13 and 28 kg heavier than family members without MC4R mutations best serves to illustrate the implications of the carrier status on body weight. Obviously, this information neither addresses the risk of childhood obesity nor the effect size of the specific mutation segregating in the respective family.

In research terms precise clarification of the effect size issue is important because other forms of ‘monogenic obesity’ have also been described, for which effect sizes were not assessed. In our opinion, a conservative approach should be pursued which actually requires proof of a strong effect size before monogenic inheritance should be invoked. The mere detection of a mutation in a gene – even if it represents a solid candidate gene, for which evidence implicating its role in body weight regulation has accumulated – is not by itself sufficient to warrant labelling of the respective obesity phenotype as monogenic. It is important to realize that identification of mutations typically occurs via the ascertainment of extremely obese cases, who have first-degree relatives with a high a priori risk of obesity due to a combination and potentially an interaction of both genetic and environmental risk factors. Based on 201 extremely obese children and adolescents, who were sampled for molecular genetic studies, we found that >50% of mothers and fathers had a BMI ≥90th age centile; in one third of the families both parents had a BMI in the uppermost decile [28]. Thus, the detection of a mutation in an obese index patient and a first-degree relative with obesity is insufficient to conclude that the respective mutation underlies the familial obesity. If functional studies prove implications of a specific mutation, such a finding obviously supports the hypothesis of a monogenic or major gene effect. Nevertheless, the detection of functional consequences of a mutation/variant in itself cannot prove that the mutation indeed underlies or contributes to the obesity of the carrier; furthermore no statement as to the effect size is possible.

In the field of human genetics, an example of the effect of ascertainment on parameter estimates pertains to breast cancer gene 1 (BRCA1). The highest penetrance estimates were originally obtained in high-risk families with several females affected with breast cancer [29]. Lower estimates stemmed from studies based on females with breast cancer (cases) ascertained independently of family history [30]. The lowest estimates were obtained in a population-based study [31].
As for MC4R mutations we are aware of two epidemiological studies: We performed a mutation screen in a total of 4,068 probands representative of the general population of the region of Augsburg, Germany [19]. None of the 6 carriers of non-synonymous MC4R mutations, which in vitro lead to reduced receptor function, were obese. In a smaller United Kingdom population sample (1,100 subjects) a mutation scan identified no functionally relevant MC4R mutations among the normal-weight individuals; a non-synonymous mutation (Ala87Asp) leading to a reduced receptor function was observed in one obese individual (BMI of 31.5 kg/m²) [32].

**Polygenic Obesity**

Approximately 130 mutations have been detected in the human MC4R gene [33]; those leading to a reduced function entail a substantially elevated risk for the development of obesity. The coding variants in the MC4R that lead to a reduced receptor function seemingly support the ‘common disease – rare variant’ hypothesis, according to which a substantial proportion of the predisposition to complex disorders is due to rare variants [34]. Indeed, it seems likely to assume that other such loci exist in the genome that cannot readily be picked up via linkage or association analyses due to locus heterogeneity and the low overall combined rate of such mutations at a specific gene locus.

Over the past 15 years, the major focus has shifted from the detection of monogenic forms and major gene effects to the identification of common single nucleotide polymorphisms (SNPs) with substantially smaller effect sizes that explain inter-individual BMI variation at the population level. The results of recent genome-wide association studies (GWAS) underscore that the ‘common disease – common variant’ hypothesis is potentially also important. This hypothesis states that specific common alleles or variants predispose to common disorders and that such alleles/variants will be found in all human populations, in which the respective disease occurs. Indeed, for several complex traits and disorders common variants, mostly SNPs, have been detected in and outside of the coding and regulatory sequences of genes [e.g. 35–37]. It is mostly unclear how the detected variation alters the function of a gene. It appears likely that frequently, the respective SNP in itself merely tags a functionally relevant SNP, haplotype or variant (linkage disequilibrium). In addition, because such SNPs can be located far from a gene it is potentially not possible to definitely determine which gene is altered in its function [e.g. 1, 38, 39]. It appears very likely that many of such functional variants influence gene expression levels.

Three factors clearly differentiate obesity association studies in the pre and post GWAS era: (i) The switch from the candidate gene approach to whole-genome studies was the result of the advent of DNA chip technology, which based on findings for several complex disorders can be viewed as a quantum leap forward in the elucidation of the genetic mechanisms underlying complex disorders. (ii) In parallel, but not directly related to the introduction of DNA chip technology, the number of cases and controls used in association studies increased substantially; both the overall predominance of negative findings and the frequent failure to replicate associations based on candidate gene studies had been discussed as potentially reflecting a too low power. The transition to large-scale studies encompassing thousands of cases and controls (and meta-analyses) was boosted by the early results of GWAS which convincingly demonstrated that effect sizes are indeed small [e.g. 40]. (iii) Finally, population-based studies have entered the scene; the most recent GWAS meta-analysis [e.g. 1] emerged from both population-based and case-control samples.

In the pre-GWAS era, the MC4R was the first gene shown to unambiguously harbour a common variant with a small effect size; the original report already included a meta-analysis of all 12 case-control studies that had been published at that time [41]. The association of the minor allele, coding for Ile103, with leanness was confirmed in subsequent larger meta-analyses with up to almost 124,000 probands [1, 42–44]. Roughly 3–6% of different populations are heterozygous for this variant, which according to German population-based data entails a 0.5-kg/m² reduction in mean BMI [41]. This would qualify this SNP as the one with the strongest effect size of all currently known common variants including the SNPs in intron 1 of the fat mass and obesity-associated gene (FTO) [40] (for effect sizes see table 1). Because other SNPs in the 5’ region of the MC4R are in total linkage disequilibrium with the SNP coding for Val103Ile [1, 41], the functional role of the cSNP requires further elucidation. Functional characterizations of the variant and wild-type receptor have repeatedly revealed no differences [15, 22, 45]; a single study reported a two-fold decreased potency of the MC4R antagonist hAGRP(87–132) [46] for the Ile130 variant which is compatible with the assumption that the variant receptor entails an elevated function of the MC4R.

A second common coding MC4R variant Ile251Leu also protects from obesity [47]; the per-allele effect size was estimated to be similar to that of Val103Ile. In functional terms, Ile251Leu was also initially considered as ‘neutral’ [5, 22] but again in one in vitro study the Ile251Leu substitution was shown to increase MC4R basal activity [46] by alteration of downstream intracellular events associated with GPCR/cAMP signal transduction [47].

More recently, GWAS meta-analyses localized SNPs more than 150 kb downstream of the MC4R, which also account for inter-individual BMI variation [38, 48]. The MC4R is thus currently unique in that it harbours variants of considerably different effect sizes both within and outside the coding region and whose minor alleles both increase and decrease BMI.
Prior to the most recent GWAS meta-analysis, 10 loci with genome-wide significance defined via a p-value < 5 × 10⁻⁸ had been identified (FTO, MC4R, TMEM18, GNPDA2, BDNF, NEGR1, SH2B1, ETV5, MTC2, and KCTD15) [38, 39, 49, 50]. The most recent and up-to-date worldwide largest meta-analysis of the GIANT (Genetic Investigation of ANthropometric Traits) Consortium included 46 studies in total encompassing up to 123,865 individuals of European ancestry [1]. The 42 best independent loci (pair-wise linkage disequilibrium measured in r² < 0.1 and separated by at least 1 Mb) were followed-up in up to 125,931 additional individuals. All 19 loci which had rendered genome-wide significance in the initial meta-analysis were confirmed. In a joint analysis of the original data and the confirmatory study, a total of 32 SNPs revealed p-values < 5 × 10⁻⁸. It is noteworthy to point out that in the confirmation of these SNPs they had a nominal p-value of 2.4 × 10⁻² and was thus above a significant p-value (of 0.05) after a conservative Bonferroni adjustment for the 42 tested SNPs. Overall, the obesity-predisposing allele showed directionally consistent effects on body fat percentage at 31 of the 32 loci.

The frequencies of alleles predisposing to obesity ranged from 4 to 87%; 20 alleles had a frequency <50%. Effect sizes.
of the novel loci tended to be lower than those of the 10 loci with genome-wide significance of $5 \times 10^{-8}$ detected in previous meta-analyses. The per-allele change in BMI ranged from 0.06 (MTCH2, KCTD15, PTBP2, RPL27A, NUDT3) to 0.39 (FTO) kg/m²; a total of 10 SNPs showed per-allele changes $<0.1$ equivalent to less than 324 and 273 g in males and females of average heights (1.8 and 1.65 m, respectively). 5 SNPs with the lowest per-allele change of 0.06 each account for approximately one half of the aforementioned weights of individuals of average height.

As based on the results obtained for the Atherosclerosis Risk in Communities (ARIC), which was one of the largest population-based studies included in the GIANT analysis, the average BMI increment per risk allele was estimated at 0.17 kg/m². The difference in average BMI between individuals with ≥38 BMI-increasing alleles (1.5% of the ARIC sample) and those with ≤21 BMI-increasing alleles (2.2%) was 2.73 kg/m². The predictive value for obesity risk of the 32 variants combined was only modest. Non-additive effects of the SNPs were not detected [1].

The 32 identified variants explained a mere 1.45% of the BMI variance; this roughly corresponds to 2–4% of the genetic variance based on an estimated BMI heritability of 0.4–0.7. Speliotes et al. [1] estimated that there are approximately an additional 284 loci (95% CI: 132–510) with similar effect sizes as the detected 32, which together would account for roughly 4.5% of the variation in BMI or 6–11% of the genetic variation (again based on an estimated heritability of 0.4–0.7). Sub-analyses based on children and adolescents revealed a substantial overlap with the genetic predisposition to adult obesity; directionally consistent effects were found for 23 of the 32 SNPs.

Common missense SNPs were shown to be in linkage disequilibrium ($r^2 \geq 0.75$) with a quarter ($n=8$) of the 32 detected SNPs. One signal was in linkage disequilibrium with coding variants in 4 neighbouring genes. Two SNPs tagged copy-number variants (CNVs). Further studies are required to determine if these in functional terms plausible SNPs and CNVs indeed underlie the association signal.

Because many of the 32 detected genes are expressed in the central nervous system, the role of the brain in the regulation of body weight is seemingly substantiated. However, the high proportion of all genes ($\approx80\%$ in mouse brain [51]) expressed centrally needs to be considered prior to drawing firm conclusions. Pathway-based analyses, depending on the genes identified via the association signals, revealed evidence of enrichment for pathways involved in the platelet-derived growth factor signalling, translation elongation, hormone or nuclear hormone receptor binding, homeobox transcription, regulation of cellular metabolism, neurogenesis and neuron differentiation, protein phosphorylation and numerous other pathways related to growth, metabolism, immune and neuronal processes [1].

The Role of Copy-Number Variants in Obesity

CNVs are DNA segments at a size of by definition at least 1 kb to several megabases which are present in variable copy numbers. Several hundred genes map to regions of copy-number polymorphic (CNP) duplications. These genes might be highly variable among individuals. Because of their repetitive and multicyclone nature, these genes are considered inaccessible by most genotyping and sequencing technologies. Such regions are more likely to be subject to recurrent mutations and to be inadequately assayed by a correlated neighbouring SNP. Both common and infrequent CNVs have been related to several phenotypes including psoriasis [52], schizophrenia [53], autism [54] and developmental disorders [55]. ‘Pathological’ CNVs, which have been associated with higher effect sizes, were frequently found to have arisen de novo.

Willer et al. [39] reported genome-wide significant association with BMI of an SNP capturing a common 45-kb deletion near NEGR1 (neuronal growth regulator 1 precursor gene). This CNV was detected in a meta-analysis of 15 GWAS comprising more than 32,000 individuals. Recently, another common CNV (chromosome 10p11.22) was shown to be associated with BMI [56] in a Chinese sample (unadjusted p-value 0.011). Furthermore, Speliotes et al. [1] identified two SNPs that tag common CNVs, one of which was already discovered [39]. The other CNV is a deletion of 21 kb that lies 50 kb upstream of GPRCS5 (G-protein coupled receptor family C group 5 member B gene). The T-allele of rs12444979 tags the deletion allele ($r^2 = 1$). However an analysis of CNVs without the aid of tagging SNPs is lacking in the GIANT approach.

Large (>500 kb) variants (deletions, duplications and inversions) are individually rare but collectively common in the human population. These genomic imbalances represent a special class of rare variants that can potentially affect many genes and pathways in a single individual [e.g. 52, 53]. In contrast to the common CNV with relevance for BMI reported above, three recent reports [57–59] showed the importance of rare large CNVs for body weight regulation. Two studies [57, 58] depicted a genomic region on chromosome 16p11.2 which harbours highly penetrant microdeletions (about 500 kb) associated with (extreme) obesity. This region had previously been associated with autism and mental retardation [57, 58, 60]. In fact, some of the obese patients analysed by Bochukova et al. [57] additionally had developmental disorders. However, Walters and co-workers [58] found the association of these microdeletions and obesity also in individuals ascertained only for the obese phenotype, suggesting a possible direct association of the deletions at 16p11.2 with obesity, distinct from their implications for cognition.

Wang et al. [59] found that CNVs >1 Mb were over-represented in obese cases versus normal-weight controls (non-significant odds ratio of 1.5) with a stronger effect for larger CNVs (>2 Mb). Hence, large CNVs (especially deletions) seemingly confer a risk for obesity. As the effect was more
pronounced for CNVs disrupting genes, these genes should be further analysed [59]. However, the respective genes did not show up in the large-scale GWAS meta-analyses so far [1, 38, 39, 49, 50, 61]. Hence, they might be relevant for a small percentage of obese individuals only.

**Missing Heritability**

GWAS have provided valuable insights into the genetic basis of complex traits and disorders. However, a common underlying theme has been that the identified genetic variation merely explains a comparatively small fraction of the heritability of such complex phenotypes; the identified variants mostly have small to very small effect sizes. Currently, a consensus explanation for this ‘missing heritability’ in complex diseases has not yet emerged (for review see [62]). We conclude our review by focussing on BMI and obesity-related aspects pertaining to this discussion.

*Non-specific quantitative phenotype BMI:* Both fat mass and lean body mass contribute to body weight; in theory, both compartments could be regulated differentially in genetic terms. Indeed, the proportion of shared variance of BMI with adiposity as determined with hydrostatic techniques has been shown to average only 0.5 [63]. The relative contribution of the weights of bone mass and other distinct tissue compartments differs inter-individually. Accordingly, potentially many different genetic and environmental factors impinge on this complex quantitative phenotype.

*Measurement error:* For determination of BMI both weight and height are required. Heritability estimates for height have been shown to be higher than those for BMI [64]. Nevertheless, the rather slight difference does not readily explain why the GIANT Consortium picked up approximately 10 times more SNPs for body height than for BMI [65] as based on a GWAS sample to a major part overlapping with the one that formed the basis of the BMI meta-analysis conducted by Speliotes et al. [1]. The SNPs detected for body height explained approximately 14% of the variance.

BMI obviously depends on two measurements, thus increasing the risk for measurement errors. Depending on age, the correlations of body weight and height with BMI differ [66], thus potentially adding to the difficulty of detecting relevant gene variants. Finally, body height is much more stable over time than BMI. In adults weight can vary by a few percent on a short-term basis; seasonal effects on weight are well known [67]. These factors enhance the noise in attempts to detect genetic variation explaining BMI variance.

*Small effect sizes:* Speliotes et al. [1] calculated that a sample size of 730,000 probands would be required to pick up an additional 250 loci with effect sizes similar to the detected 32, which in total would then explain 4.5% of the BMI variation. The currently identified loci might however merely represent the tip of the iceberg: effect sizes of most common variants with an influence on BMI could very well be (substantially) lower. The bottom range of the per-allele effect size of the currently identified loci is in the magnitude of 0.06 kg/m²; conceivably many more variants exist with effect sizes equivalent to less than 100 g. In this context, it is worthwhile to point out that in transgenic rodent models an effect on body weight is usually only reported if it is pronounced; due to the lack of systematic statistical assessments of large numbers of rodents it is currently unclear if body weights of specific transgenics differ slightly (but significantly) from that of wild-type animals and if so, in what percentage of transgenic mouse models this indeed is the cases.

Very recently it was estimated that as much as 45% of the variance in body height can be explained at the molecular level. Yang et al. [68] estimated the proportion of variance for human height explained by 294,831 SNPs genotyped on 3,925 unrelated individuals using a linear model analysis. They analysed the SNPs simultaneously and validated their estimation method with simulations based on the observed genotype data. The remaining heritability was shown to be attributable to incomplete linkage disequilibrium between causal variants and genotyped SNPs [68]. Hence, most of the heritability is not missing but, because the individual effects are very small, has not previously been detected. While this percentage is also lower than what heritability estimates of approximately 0.9 would predict, BMI has proven to be much more elusive to a dissection at the molecular level.

*Infrequent to rare variants:* As pointed out above, it appears very likely that additional infrequent to rare variants similar to the mutations in the *MC4R* exist, which potentially have large effect sizes. Because the respective variants are not represented on the DNA chips used in the GWAS, they could easily have escaped detection, particularly if they are rare. One might assume that the infrequent SNP coding for the Val103Ile variant would not have been detected in the Speliotes et al. study [1], as rs7227255, which is in complete linkage disequilibrium (\(r^2 = 1\)) to Val103Ile, reached only a \(p = 6.56 \times 10^{-7}\) which is not significant at the genome-wide level. Approximately 3-6% of probands of European populations are heterozygous for the Val103Ile SNP; this heterozygosity rate, the respective effect size in combination with the p-value, provides a limit of the resolution of the current GWAS.

The BMI distribution in the general population is shifted to the right; however, a clear bimodal distribution has not been detected; such a bi-modality could potentially ensue as the net result of many rare variants with strong effect sizes. According to the recently proposed synthetic association [69, 70], GWAS signals of common non-functional SNPs outside of coding regions may be the result of a combination of rare coding/functional variants with stronger effects given that these rare variants arose on a haplotype which is tagged by the common SNP. An accumulation of rare variants has recently been described in GWAS-identified genes for hypertriglyceridae-
Resequencing revealed a significantly higher frequency of rare missense or nonsense variants in individuals with HTG compared to controls. The rare variants in these genes increased the proportion of genetic variation contributing to HTG. Thus, there might be an excess of rare variants in GWAS-derived genes [71], which is conform to the synthetic association hypothesis. Recently, it was proposed to use statistical power calculations for future GWAS by use of estimated distributions of effect sizes. For some traits, it was shown that they likely harbor additional loci within the spectrum of common variants. These loci could explain at least 15–20% of the known heritability of these traits [72].

**Heritability estimates:** Numerous twin, adoption and family studies have allowed for the calculation of heritability estimates for BMI, which mostly range from 0.4 to 0.7. Twin studies have consistently led to uniformly higher estimates; family studies typically derived at the lower estimates. Based on complex statistical modelling, single family studies have nevertheless also led to estimates in the range of 0.7 [73, 74]. To our knowledge the lowest heritability estimate of 0.05 was reported in a large family study of Bouchard and co-workers in 1988 [75]; in contrast, estimates in excess of 0.8 were derived from single twin studies [e.g. 76, 77]. In adults, mean BMI values are greater among men than among women, although the variance is greater among women. Whereas there is no clear-cut evidence for an effect of gender on heritability estimates, the sets of genes contributing to variation in BMI are presumably not identical among men and women [78].

In light of the current failure to detect a substantial proportion of BMI variance, particularly the high heritability estimates obtained in twin studies require scrutiny. These have previously mainly been explained via non-additive or epistatic effects [e.g. 79]. It should be recalled that the influence of the common environment differs between monozygotic and dizygotic twins; the ‘twinning reaction’, or mutual dependence between twins, is greater between identical twins than between fraternal twins [e.g. 80, 81]. Approximately, two thirds of monozygotic (MZ) twins share the same amnion, thus increasing the influence of the common intrauterine environment. Postnatally, MZ twins are usually treated much more similarly by parents, teachers and other individuals, they spend more time together, have more mutual friends; accordingly, common environment is more similar among MZ than dizygotic (DZ) twins. It is conceivable that the more similar common environment for MZ twins also influences eating behaviour and physical activity.

Speliotes et al. [1] base their calculations pertaining to explained variance on a heritability estimate of 0.4–0.7, which is below the estimate obtained in twin studies. However, some adoption and family studies have derived at considerably lower heritability estimates of approximately 0.2 [73]; if the true heritability of BMI is within this range, the explained genetic variance would increase by a factor of approximately 2.5.

**Developmental aspects:** Body weight increases by on average 100–400 g a year during most of adulthood [82]. Recent GWAS take this into account by adjusting BMI for age [e.g. 1]. Childhood BMI is only weakly correlated with adult BMI; longitudinal studies have consistently shown that the correlations between childhood and adult BMI increase with increasing age of the child. According to one of such studies, BMI at ages 1, 10 and 18 were correlated with BMI at age 35 by 0.15, 0.5 and 0.7, respectively [83]. Heritability estimates based on twins of specific ages indicate low heritability in infancy up to age 1 [84, 85]. With the exception of this early developmental stage, estimates are similarly high from childhood throughout early and middle adulthood [73, 77].

A recent review on 13 longitudinal studies found a strong genetic continuity in BMI from early childhood to onset of adulthood [84]. There was also evidence for an effect of a common environment on the tracking of BMI during childhood; in adults the influence of the non-shared environment substantially exceeded that of the shared environment. Heritability estimates based on the 9 longitudinal twin studies ranged from 0.57 to 0.86 for the trend of BMI from early adulthood to late middle age and were higher than estimates obtained from the single longitudinal adoption study and the three family studies [84].

Based on these developmental aspects, it is not surprising that parental-offspring BMI correlations are typically lower (usual range: 0.1–0.3) than the correlations between the offspring themselves (r ≈ 0.3–0.4) [73]. Maternal-offspring in comparison to paternal-offspring correlations have been found to be higher in several studies thus suggesting a specific maternal influence of genetic factors (e.g. imprinting, mitochondrial genome), intrauterine and/or postnatal environment [86]. Family-based heritability estimates are frequently based on more or less complex adjustments for age of the family members. In contrast, the same age of twins circumvents this problem in twin studies. Despite the secular trend towards higher BMI in recent history, there is no indication that heritability estimates have changed over time [e.g. 77, 84].

Current evidence suggests that common alleles detected in adults are important throughout different developmental stages [1]. A twin study has, however, suggested that a proportion of the genes involved in the regulation of BMI change with age [87]. Scherag and colleagues [61] recently detected an SNP between TNN5 (tankyrase, TRF1-interacting ankyrin-related ADP-ribose polymerase gene) and MSRA (methionine sulfoxide reductase A gene) which was associated with childhood but not adult obesity. Developmentally oriented molecular genetic studies have been initiated. For example, in accordance with the low heritability estimates for BMI in early life [84], a significant effect of the FTO genotypes was not detectable in children prior to age 4 in two birth cohort studies [88].

**Non-additive or epistatic mechanisms:** To date, no evidence has been found for non-additive or epistatic mechanisms at the molecular level; gene-environment interactions have not yet
been tested for in GWAS-based meta-analyses. The detected gene variants seemingly influence BMI in an additive manner [see 1]; evidently, lack of power could contribute to the lack of significance for gene-gene interactions. Non-additive genetic factors have, however, been suggested to explain more than half of the genetically based BMI variance [79]. It thus seems highly likely that indeed non-additive effects remain to be detected. The joint analyses of many different European samples that stem from different populations in the meta-analyses conducted by Speliotes et al. [1] may readily obscure their detection. Similarly, if gene-environment interactions prove to be more or less specific for specific populations, the respective genes would potentially escape detection in meta-analyses encompassing samples of different populations.

Apart from the considerations mentioned above, additional aspects to explain the ‘missing heritability’ had been proposed [62]: (a) the parental contribution should be taken into account, as for instance certain risk alleles for cancer or type 2 diabetes mellitus were shown to confer risk only when inherited from a specific parent; (b) epigenetic effects beyond imprinting that are sequence-independent and that might be environmentally induced but can be transmitted for one or more generations; (c) phase-dependent interactions between variants that are not in linkage disequilibrium and that are difficult to detect without long-range haplotypes; (d) genetic variation influencing the expression of non-coding RNAs; (e) evidence for transgenerational genetic effects in which phenotypic variation in the present generation results from genetic variants in previous generations.

**Future Directions**

Elucidation of some of the mechanisms mentioned above will presumably lead to the identification of additional genetic mechanisms underlying body weight regulation. Current international projects will aid the understanding of the human genetic variation (e.g. 1000 Genomes Project; www.1000genomes.org) and alterations of the genetic architecture (e.g. genome-wide methylation analyses by chip technology; epigenetics). As the first confirmed obesity (pol)genes are now known, the analysis of gene-gene and gene-environment interactions can be undertaken.

**Disclosure**

The authors declared no conflict of interest.

### References


New Obesity Genes Explain Part of the Missing Heritability


