Chapter 4
Copy Number Variants and Their Contribution to the Risk of Obesity

Julia Sarah El-Sayed Moustafa and Philippe Froguel

Abstract Obesity is becoming an increasingly serious health concern, given its associated health risks and the growing number of people affected. Understanding the genetic factors underlying body weight regulation and obesity susceptibility has thus become an issue of paramount importance. Obesity has a high estimated heritability, yet much of this remains unexplained. Copy number variants (CNVs) represent a relatively understudied class of genetic variants which may account for some of this unexplained heritability. This chapter explores how copy number variation contributes to body weight regulation and obesity susceptibility.

Common CNVs associated with body mass index (BMI) and obesity have recently been identified, including variants on chromosomes 1p31.3, 8p21.2, 10q11.22, 11q11, and 16p12.3 at the *NEGR1*, *DOCK5*, *PPYR1*, *OR4P4*, *OR4S2*, and *OR4C6*, and *GPRC5B* loci. A number of rare CNVs have also recently been associated with extreme forms of obesity, including two on chromosome 16p11.2, consisting of a 593 kb deletion whose reciprocal duplication has been associated with increased risk of underweight, as well as a 220 kb deletion encompassing the *SH2B1* gene, which has been associated with overweight and obesity. Several studies have also reported enrichment in the global burden of large, rare CNVs among obese subjects, as well as the presence of several rare CNVs uniquely among obese cases.

In the case of large CNVs encompassing multiple genes, functional studies will be required to establish which gene or genes within each CNV are causative for the

J.S.El-Sayed Moustafa (✉)
Department of Genomics of Common Disease, Imperial College London, Hammersmith Hospital, 3rd Floor Burlington Danes Building, Du Cane Road, London W12 0NN, UK
e-mail: jelsayed@imperial.ac.uk

P. Froguel
Department of Genomics of Common Disease, Imperial College London, Hammersmith Hospital, 3rd Floor Burlington Danes Building, Du Cane Road, London W12 0NN, UK

Centre National de la Recherche Scientifique (CNRS)-Unité mixte de recherche de l'Université Lille 2 (UMR) 8199, Lille Pasteur Institute, Lille 59000, France
e-mail: p.froguel@imperial.ac.uk

observed phenotype. Improved methodologies, both laboratory and statistical, are also required to enable systematic investigation of complex structural variants such as VNTRs and multi-allelic loci. While developments in this field are promising, analysis of CNVs, both common and rare, has proven challenging, and reported associations require extensive follow-up in large replication samples in order to confirm their role in obesity susceptibility. It is hoped that such in-depth investigations will provide increased understanding of the role of CNVs in body weight regulation and risk of obesity.

Introduction

The study of the genetic factors underlying obesity susceptibility is a subject which has captured the attention of many within the scientific community, particularly due to the serious health risks faced by affected individuals, and the increased risk of obesity in their relatives. This chapter explores the contribution of copy number variants (CNVs) to body weight regulation and risk of obesity.

The Missing Heritability of Obesity

The current obesogenic environment, characterized by an increased consumption of widely available calorie-dense foods among many other factors, has no doubt driven the recent rise in obesity rates [1]. A question of extreme interest in the study of obesity, however, is why individual risk of obesity differs even between subjects exposed to the same environmental risk factors [1]. The answer to this question lies in the fact that obesity is a complex disease arising from a complex interplay of environmental risk factors, affecting all individuals within any given population, and individual genetic predisposition, which renders certain individuals more susceptible to obesity in the face of these environmental risk factors [1].

Despite this complex interaction, numerous studies have shown obesity to be a highly heritable trait. Several twin, adoption, and family studies examining the heritability of adiposity have reported heritability estimates for obesity ranging from approximately 40–70 %, with increased concordance levels between monozygotic twins, even those reared apart, compared to dizygotic twins [2–7].

Conversely, genetic variants associated with adiposity and obesity identified to date explain only approximately 2–4 % of the heritability of these traits [8, 9], with the vast majority of studies having focussed on the analysis of common single nucleotide polymorphisms (SNPs). This discrepancy between the estimated heritability of corpulence and the proportion of which has been explained to date has raised the important question of whether the heritability of obesity has been overestimated, or whether this "missing heritability" [10] could in fact be accounted for by forms of genetic variation not captured by genome-wide association studies (GWAS) of common SNP variants. One such class of variants that have received increased attention in recent years are CNVs.

Introduction to Copy Number Variation

A CNV is defined as a segment of DNA differing in the number of diploid copies carried by individuals within the population [11–14]. CNVs include simple bi-allelic deletions and duplications, as well as more complex, multi-allelic variants showing highly polymorphic patterns of copy number distribution at the population level (Fig. 4.1).

CNV discovery studies to date all concur that CNVs are widespread throughout the human genome, and are also observed in phenotypically healthy individuals [11–16]. While precise estimates of CNV frequencies and their average size have differed between studies, in the highest resolution genome-wide CNV discovery study carried out to date [14], a total of 8,599 CNVs above 443 bp, covering approximately 3.7 % of the genome, were independently validated, with a median CNV size of 2.7 kb and a median of 1,117 and 1,488 CNVs in European (CEU) and Yoruban African (YRI) subjects, respectively [14]. Of the approximately 5,000 validated CNVs which were subject to further investigation, 77 % were deletions, 16 % were duplications and 7 % were multi-allelic variants, although it is essential to consider that these frequencies may also be influenced in part by the respective ease of detection of these three forms of structural variation [14].

As shown in Fig. 4.2, CNVs were found to overlap 13.4 % of RefSeq genes, with a smaller proportion of deletions than duplications and multi-allelic variants overlapping genes [14]. CNVs were detected genome-wide, with CNVs shown to result in loss of function mutations at over 260 genes [14]. Any two subjects were found to differ in copy number at an average of approximately 0.78 % of the genome, affecting structure of approximately 2.7 % of gene transcripts [14]. Multiple studies have concurred that common bi-allelic CNVs are well-tagged by surrounding SNPs [13, 14, 17], while significantly less linkage disequilibrium has been detected between duplications and multi-allelic variants and their surrounding SNPs [14, 17]. In addition to tandem duplications, numerous dispersed duplications have also been detected, indicating that this may be an overlooked class of CNV [14] (Fig. 4.3).

Population genetic analyses of genomic structural variation thus suggest that CNVs are widely distributed in the human genome, with the majority of CNVs being of small size, with significant overlap between CNVs detected in different subjects [11–16]. Furthermore, CNV hotspots prone to recurrent recombination

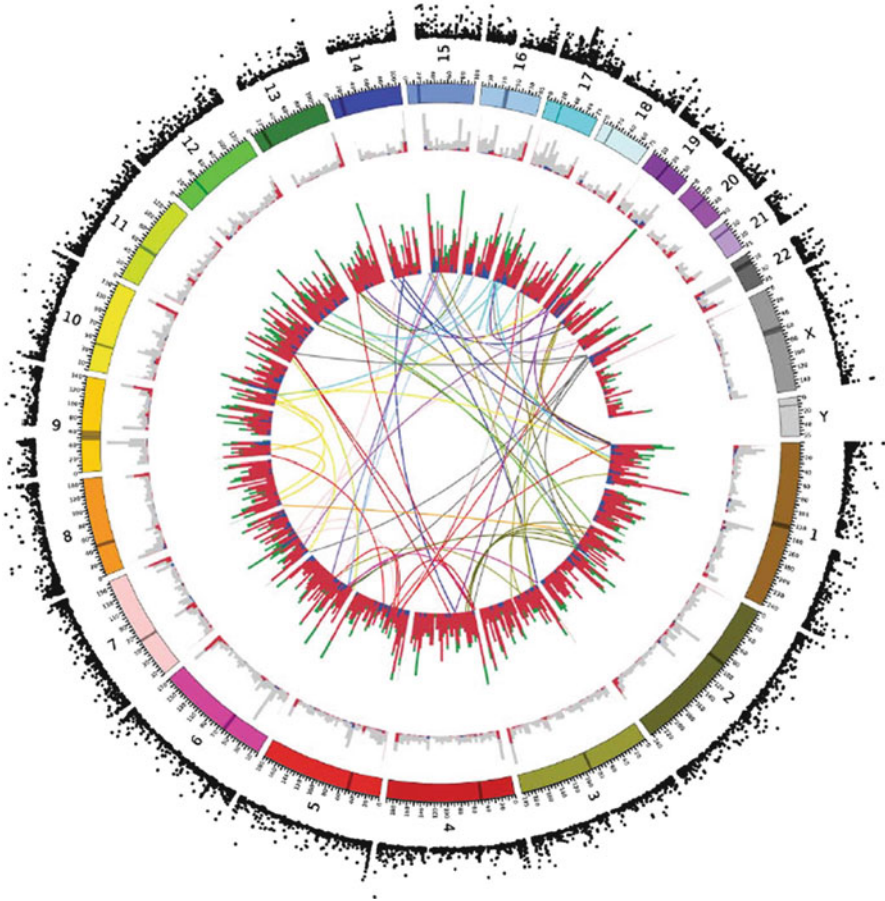**Deletion**   **Duplication**   **Multi-allelic CNV**
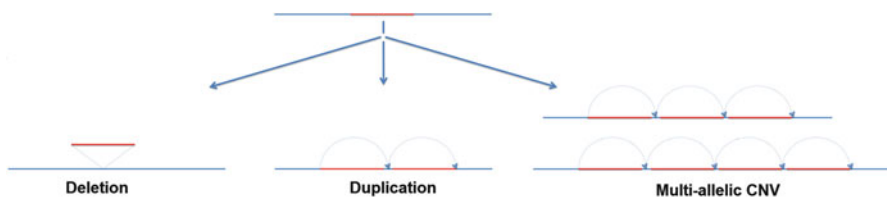
Fig. 4.1 Copy number variant (CNV) classes. CNVs may consist of simple deletions or duplications, or more complex rearrangements such as multi-allelic CNVs, where several allelic configurations exist for the same locus, varying in the number of copies of the duplicated region
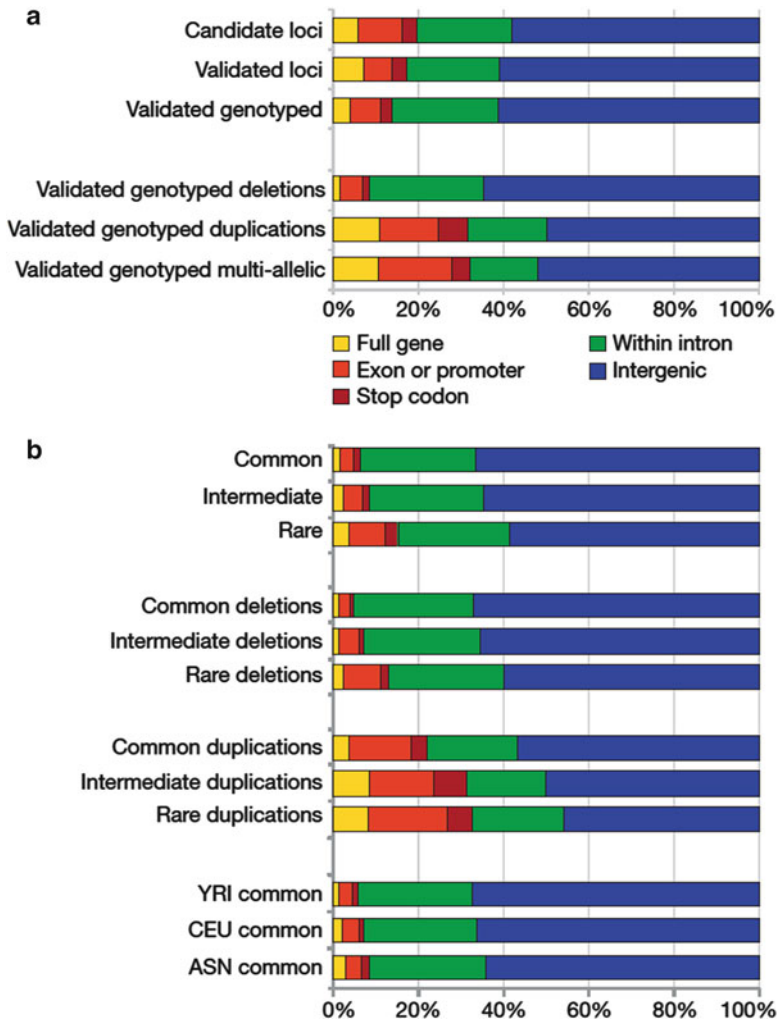
Fig. 4.2 Functional impact of CNVs in the genome. (**a**) Overall functional consequences of CNVs, stratified by level of validation and CNV type. (**b**) Functional impact of CNVs, stratified by CNV type, frequency, and sample geographic origin. YRI: Yoruba in Ibadan, Nigeria; CEU: CEPH (Utah residents with ancestry from northern and western Europe); ASN: Japanese in Tokyo, Japan + Han Chinese in Beijing, China. Figure reproduced with permission from Conrad et al. (2010) [14]

exist in the genome, particularly in the vicinity of segmental duplications [18] and sequence motifs such as *Alu* repeats [14, 19, 20]. In addition to common CNVs with identical breakpoints shared by multiple individuals, a multitude of rare and recurrent CNVs exist, a higher proportion of which overlap genes than do common structural variants, and might thus contribute significantly to interindividual phenotypic differences [14]. Similarly, a higher degree of overlap exists between genes and

Fig. 4.3 Circular plot of genome-wide CNV distribution reported by Conrad and colleagues [14]. The concentric circles depict, from inside to outside, stacked histograms of the numbers of deletions, duplication, and multi-allelic CNVs in *red*, *green*, and *blue*, respectively, the number of CNVs by mechanism of formation (NAHR, VNTR, and other shown in *blue*, *red*, and *grey*, respectively), and the degree of population differentiation between the Yoruban and European study samples of detected CNVs in the outermost circle, with the innermost circle depicting the origin and new location of dispersed duplications in the genome. Figure reproduced with permission from Conrad et al. (2010) [14]

complex structural variants such as multi-allelic CNVs and VNTRs, implicating these complex and understudied variants in phenotypic variability and disease susceptibility [14].

CNVs may influence gene expression levels either directly or indirectly through a number of different mechanisms, including deletion or duplication of entire genes, gene-disrupting CNVs, or through long-range effects mediated through disruption or insertion of regulatory elements such as enhancers or repressors [21, 22]. In the case of multi-allelic CNVs encompassing dosage sensitive genes, expression levels

Fig. 4.4 Dosage-sensitive genes. Dosage-sensitive genes are those at which changes in gene copy number result in changes in the quantity of mRNA produced

may be directly correlated with gene copy number [22] (Fig. 4.4). The phenotypic effects of CNVs and their potential contribution to disease susceptibility have thus become a topic of considerable interest.

Copy Number Variation in Adiposity and Obesity Susceptibility

The Contribution of Common Copy Number Variants to Body Weight Regulation

Given the previously noted potential functional influences of CNVs, a natural progression from CNV discovery studies was the investigation of their potential contribution to human disease susceptibility and the so-called "missing heritability" [10] of common diseases.

A large number of SNP association analyses have been conducted to date in both case–control samples and population cohorts for numerous common diseases [23], and the development of CNV prediction algorithms has enabled CNV prediction using these genome-wide SNP array data [24–27]. This has permitted the reuse of these data for CNV association studies. Similar to SNP GWAS, genome-wide CNV association studies have often focussed on common CNVs, usually defined as those having a population frequency above 5 %, with several associations between common CNVs and complex diseases, including obesity, having been reported in recent years.

Marginal association of a common CNV on chr10q11.22 encompassing the pancreatic polypeptide receptor 1 (*PPYR1*) gene with BMI has been reported in a Chinese population sample, with low copy number associated with increased BMI [28]. *PPYR1* ligands have previously been linked to the regulation of food intake in both human and animal studies [29–31], lending support to a potential role for CNVs encompassing this gene in body weight regulation. Furthermore, a common CNV at 11q11 encompassing the olfactory receptor genes *OR4P4*, *OR4S2*, and *OR4C6* has also been reported to show association with early-onset extreme obesity [32].

As well as CNVs consisting of di-allelic variants, more complex structurally variable regions may also contribute to increased risk of disorders such as obesity [33]. We have recently shown a complex copy number variable region on chromosome

8p21.2 to be significantly associated with susceptibility to severe obesity [33]. The region encompasses two variable number tandem repeats (VNTRs) flanking a 3,975 bp common deletion. Two of these three variants are located within the dedicator of cytokinesis gene (*DOCK5*), and all three structural variants were shown to be significantly associated with *DOCK5* gene expression levels [33]. The *DOCK5* gene is a member of the DOCK family of guanine nucleotide exchange factors (GEFs) [34], which are thought to be involved in a variety of cellular functions such as growth, differentiation, regulation of the actin cytoskeleton, vesicle transport, cell signalling, cell movement, phagocytosis, and apoptosis [35] through their role in the activation of members of the Rho/Rac-family GTPases [34]. Further investigation is required in order to establish the precise mechanism by which CNVs within the *DOCK5* region contribute to obesity susceptibility.

In addition to studies directly measuring copy number, some studies have also identified common CNVs potentially contributing to disease susceptibility through linkage disequilibrium with nearby SNPs. Using this approach, two common CNVs, one upstream of *NEGR1* and another near *GPRC5B*, have been linked to body weight through association of tagging SNPs with BMI in two large meta-analyses [8, 36]. Although the effect sizes observed at each of these loci were small, given the LD between these structural variants and their tagging SNPs, it has been suggested that these variants could potentially be causal variants [8, 36].

In spite of these reports, the role of common CNVs in disease susceptibility remains an issue of contention, with little replication of reported associations. A large study conducted by the Wellcome Trust Case Control Consortium (WTCCC) reported association of common CNVs at *IRGM* and *TSPAN8* with Crohn's disease and type 2 diabetes, respectively, as well as association of copy number at the HLA locus with each of Crohn's disease, rheumatoid arthritis, and type 1 diabetes [37]. However, apart from these reported hits, the authors found little evidence of association between common CNVs included in their analyses and any of the eight complex diseases in their study. The authors did however highlight the complexity of CNV prediction and association studies, reporting the confounding effects of several sources of systematic bias such as DNA source and quality, as well as batch effects, on CNV analyses [37]. Moreover, the authors also acknowledged that due to the extensive challenges in assaying more complex structural variants such as multi-allelic CNVs and VNTRs, their study was largely limited to common bi-allelic CNVs [37]. These observations highlight the need for additional investigation of the role of CNVs in complex disease susceptibility, focussing in particular on complex structural variants.

The Role of Rare Genomic Structural Variants in Adiposity and Risk of Obesity

Given the large proportion of the estimated heritability of obesity which remains unexplained, it has been suggested that some of this "missing heritability" [10] may
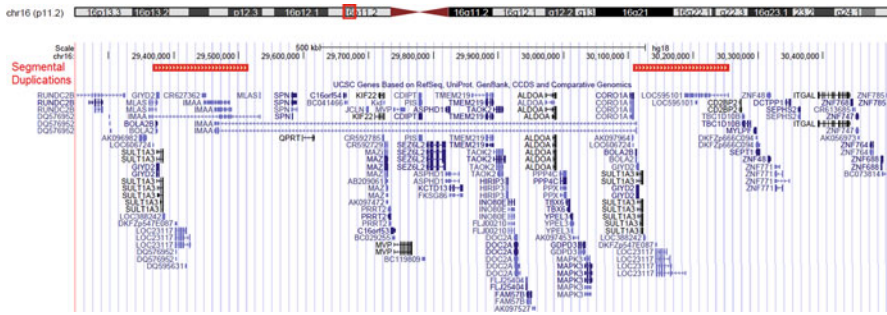
Fig. 4.5 UCSC genome browser view of the proximal 16p11.2 CNV region. The presence of two segmental duplications with high sequence similarity (depicted in *red*) results in the recurrent occurrence of deletions and duplications of the intervening 593 kb segment of unique DNA sequence in this region. Plot generated using the UCSC genome browser [56]

be accounted for by the collective effect of a large number of individually rare variants, each of large effect size [38]. Consistent with what has been observed in the case of SNPs, an increasing body of evidence is supporting the potential contribution of rare CNVs to susceptibility to complex diseases such as obesity, which will be the focus of this section.

Rare CNVs are generally defined as those with frequencies below 1 % in the general population [39, 40]. The rarity of these CNVs generally means that they are not well-tagged by surrounding common SNPs genotyped on GWAS panels. Given the inherent difficulties in accurately genotyping CNVs, analysis of rare CNVs has also principally focussed on variants of large size, often above 200–500 kb [39, 40]. Several large, rare CNVs have thus been reported to show association with body weight and risk of obesity.

Structural Variants Within the 16p11.2 Region

Several CNVs have been identified within the 16p11.2 region, with CNVs at two loci in this region showing association with either underweight, or increased risk of overweight or obesity [38, 39, 41, 42].

Copy Number Variation at the Proximal 16p11.2 Locus

In 2010, we reported association of a heterozygous deletion on chromosome 16p11.2 (chr16: 29,514,353–30,107,356) with highly increased risk of obesity [38]. The deletion encompasses 593 kb of unique sequence and contains 29 genes (Fig. 4.5), including multiple candidates for the obesity phenotype. The presence of two segmental duplications with high sequence similarity renders this locus prone to de novo structural rearrangements (Fig. 4.5), resulting in the occurrence of both deletions and duplications of the intervening DNA sequence [38].

This deletion was initially identified in our study at a frequency of approximately 2.9 % in a study sample of patients suffering from obesity-plus syndromes, whereby patients presented with obesity coupled with additional clinical features such as developmental delay and/or congenital abnormalities [38]. Further investigation revealed an additional 22 deletion carriers among subjects referred to clinical services for cognitive impairment, and 19 subjects among obesity case–control and population GWAS samples [38]. This variant was also reported concurrently in a study by Bochukova et al. [39].

Deletions at this locus resulted in a 30-fold increase in risk of obesity and 43-fold increased risk of morbid obesity, and were identified in 0.7 % of morbidly obese subjects included in our analysis [38]. The obesity phenotype observed among deletion carriers was frequently coupled with hyperphagia, suggesting it to be of potentially neurological origin. While no gender bias was detected in our analysis, an age-dependent effect for this CNV was observed, where penetrance of the obesity phenotype in deletion carriers was positively correlated with subject age [38]. A 0.4–0.7-fold reduction in gene expression levels was also observed for transcripts of genes located within the deleted segment, suggesting that haploinsufficiency for one or more of these genes may be causative for the obesity phenotype observed in deletion carriers [38]. In addition to increased risk of obesity, the deletion was also associated with increased head circumference [38].

A recent study also confirmed the association of this deletion with macrocephaly, but also reported significantly reduced cognitive functioning and an increased frequency of gross motor delay among deletion carriers [43]. Psychiatric comorbidities were reported in greater than 80 % of deletion carriers, while penetrance of obesity was over 70 % of carriers of the deletion in this study sample [43].

Apart from its association with obesity, copy number variation at this 16p11.2 locus has previously been associated with neurodevelopmental and psychiatric conditions, implicating this locus in a number of phenotypes. Both microdeletions and microduplications of the same locus at 29.5 Mb in the 16p11.2 region were shown to be associated with increased risk of autism spectrum disorders (ASD), accounting for approximately 1 % of ASD cases in one study [44]. On the other hand, duplications, but not deletions, at this locus were also linked to increased susceptibility to schizophrenia, with duplication carriers showing a 14.5-fold increased risk of schizophrenia [45]. These findings have since been replicated in a number of studies [46–48], confirming the contribution of these loci to increased risk of these disorders, and raising the interesting question of the interrelationship between the obesity and neurodevelopmental and psychiatric phenotypes associated with copy number variation at this locus.

A retrospective analysis of 16p11.2 deletions in a clinical sample of approximately 7,000 subjects—the majority of whom had presented with phenotypes such as developmental delay, autism spectrum disorder (ASD) or dysmorphism—identified 28 deletion carriers among this sample [49]. The age-dependence and juvenile onset of the obesity phenotype was confirmed, with obesity generally developing within the first decade of life. Furthermore, a gender-dependence for the 16p11.2 was reported, with male deletion carriers exhibiting a more severe phenotype than female carriers [49]. The incidence of obesity among deletion carriers diagnosed

with ASD was also noted to be higher than among autistic subjects not carrying the deletion, providing further support for the independent association between deletions at this locus and increased risk of obesity [49].

In a second study in 2011, we investigated the impact of the reciprocal 16p11.2 duplication on body mass and head circumference [41]. In a fascinating example of a mirror effect of gene dosage at this locus on phenotype, the reciprocal 16p11.2 duplication was associated with strongly increased risk of being underweight, with carriers of this duplication showed significantly reduced postnatal weight and BMI compared to non-duplication carriers [41]. For the purpose of this study, underweight was defined as a BMI ≤ 18.5 kg/m^2 in adults and BMI z-score ≤ 2 standard deviations from the mean for age and sex in children [41]. Underweight can have serious health repercussions, and is frequently associated with failure to thrive during childhood, eating and feeding disorders, as well as anorexia nervosa. Despite the potentially serious nature of this condition, little is known of the factors underlying its genetic susceptibility [50]. In this analysis, 50 % of the male duplication carriers under the age of 5 were diagnosed with a failure to thrive, while adult carriers of this duplication showed an 8.3-fold increased risk of being clinically underweight [41]. A gender effect was also observed, with males showing a trend towards increased severity. In addition to its observed effect on weight, the duplication was also associated with an increased frequency of restrictive and selective eating behaviors, mirroring the hyperphagic phenotype observed in carriers of the reciprocal deletion [41]. Similarly, duplication carriers were noted to show significant reduction in head circumference, which mirrored the macrocephaly associated with the reciprocal deletion [38].

The 16p11.2 duplication was also observed at a higher frequency among medically ascertained patients, recruited on the basis of developmental and cognitive delay or psychiatric phenotypes, than in non-medically ascertained population cohorts in this study, supporting the previously reported association of this duplication with cognitive, neurodevelopmental, and psychiatric phenotypes [44, 45].

A 220 kb Deletion on Chromosome 16p11.2 Encompassing the SH2B1 Gene

In addition to the previously described proximal 16p11.2 deletion and duplication shown to be associated with body weight regulation, additional CNVs within the 16p11.2 region have also been associated with obesity susceptibility.

A 220 kb deletion in this region encompassing nine genes, including the SH2B adaptor protein 1 (*SH2B1*) gene, has been reported to be associated with severe, hyperphagic, early-onset obesity [39]. Although carriers of this deletion have been reported to exhibit elevated fasting plasma insulin levels, conflicting observations have also been reported [39, 51].

SH2B1 is known to enhance leptin and insulin signalling, and animal studies have shown mice harboring homozygous null mutations in the *SH2B1* gene to
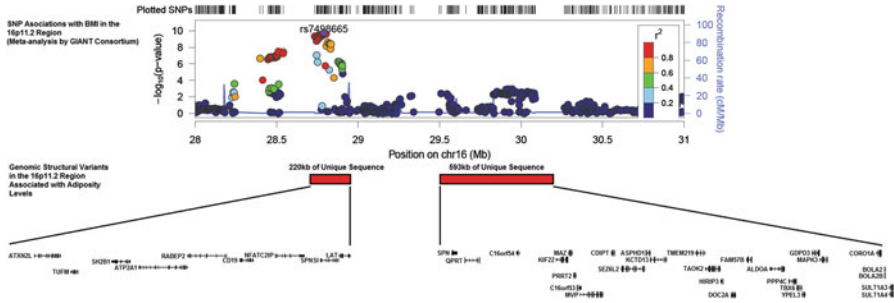
Fig. 4.6 The chromosome 16p11.2 region. (**a**) Association results for SNPs in the 16p11.2 region with BMI in a recent meta-analysis carried out by the GIANT consortium [8] Chromosome 16 genomic coordinates are plotted on the *x* axis, with minus $\log_{10}$(*P*-value) plotted on the *y* axis. An association peak can be seen at approximately 28.8 Mb. Plot generated using LocusZoom [57]. (**b**) The positions of two genomic structural variants associated with adiposity levels are depicted. A 220 kb deletion at chr16: 28.73–28.95 Mb and a 593 kb deletion at chr16: 29.51–30.11 Mb have been associated with obesity [38, 39] while a duplication of the latter 593 kb of unique sequence has also been associated with risk of being underweight [41]. The genes falling within each of the two CNVs are also shown. *CNV* copy number variant, *GIANT* Genetic Investigation of ANthropometric Traits, *SNP* single nucleotide polymorphism. First published in *Nature Reviews Endocrinology*, 2013, doi: 10.1038/nrendo.2013.57 by Nature Publishing Group

exhibit signs of metabolic syndrome, with a phenotype including obesity, hyperphagia and insulin resistance [52]. SNPs within *SH2B1* have also shown association with BMI in several meta-analyses [8, 36, 53], making it a strong candidate for the obesity phenotype observed in carriers of this 220 kb deletion. Figure 4.6 depicts association results for SNPs within the 16p11.2 region from a recent BMI meta-analysis, as well as the positions and gene content of both CNVs within the 16p11.2 region described in this chapter.

In addition to its association with severe obesity, this deletion encompassing *SH2B1* has also been linked to developmental delay. In an analysis of a clinical sample of approximately 23,000 patients referred for array comparative genome hybridization (aCGH) for phenotypic abnormalities including developmental delay and cognitive deficits, this medically ascertained sample was found to be enriched for this deletion, and assessment of additional anthropometric data available for a subset of the deletion carriers supported its association with early-onset obesity [42].

Global Burden of Rare Copy Number Variants in Obesity

In addition to the analysis of individual CNVs and their contribution to obesity susceptibility, another area of particular interest is whether the global burden of large, rare CNVs may be higher among subjects suffering from obesity. This is assessed by comparing the total number of rare CNVs above a defined size threshold observed in obese cases versus normal-weight control subjects.

Large, rare deletions have been reported to be enriched among obese cases compared to normal-weight controls in case–control analyses of global CNV burden [39, 40]. In these analyses, large CNVs were found to be overrepresented among obese cases, with this enrichment driven largely by deletions [39, 40]. Furthermore, a larger effect was observed when the analysis was limited to those CNVs which disrupt genes [40], highlighting the potential significance of genes located within these variants to obesity susceptibility.

Rare CNVs Present Exclusively in Cases

Another method of identifying CNVs which might be relevant to obesity susceptibility is to identify CNVs observed exclusively in obese cases and not in normal-weight controls. One study identified 17 CNVs present exclusively in three or more Caucasian obese subjects, eight of which were also observed only among African American obese subjects and no normal-weight controls [54]. Their presence solely in obese cases might suggest a potential role for these variants in obesity susceptibility, and replication of these observations in study samples of different ethnicities provides further support for their relevance to the pathogenesis of obesity [54].

While several studies have provided intriguing evidence for the involvement of rare CNVs on obesity susceptibility, it is essential to note that in the analysis of rare variants, wider replication in larger study samples will be necessary to firmly establish their contribution to disorders such as obesity.

From Genetic Variants to Their Physiological Impact: The Importance of Follow-up Studies in CNV Analyses

The identification of structural variants, both common and rare, associated with obesity susceptibility is providing insight into its pathogenetic origins and helping explain some of the missing heritability of this disorder. However, similar to what is observed in the case of common SNPs, there is often difficulty in translating these genetic findings into clear understanding of the underlying biological pathways and mechanisms responsible for this disease. In the case of CNVs, this problem is compounded by the fact that CNVs are often large and may encompass several genes, making it difficult to decipher which gene or genes are responsible for the phenotypic effects observed. Furthermore, CNVs may also have long-range effects, with variants shown to influence expression levels of genes up to several megabases away.

For this reason, it is important to follow up genetic associations with functional studies which attempt to understand how specific variants affect phenotype. As previously discussed, deletion and duplication of a 593 kb region on chromosome 16p11.2 has been associated with a mirror effect on various phenotypes, one of which is head circumference. Duplication of this region has been associated with microcephaly [41], while the reciprocal deletion has been associated with increased

head circumference. Through systematic over-expression and knockdown of each of the orthologous genes within the CNV region in zebrafish, the gene responsible for the variation in head circumference associated with copy number in this region was shown to be the potassium channel tetramerization domain containing 13 (*KCTD13*) gene [55]. Further studies of this type should be undertaken to identify the causal gene or genes for the obesity phenotype associated with the proximal 16p11.2 CNV. Similarly, functional exploration of other CNVs reported to be associated with obesity susceptibility would help in better delineating their physiological effects and in identifying the causative genes located within them.

Future Directions in the Study of CNVs in Obesity

The contribution of CNVs—both common and rare—to obesity susceptibility is becoming increasingly recognized, with progressively more reports of CNVs associated with adiposity levels. However, in spite of this mounting body of evidence, our understanding of the contribution of structural variants to complex diseases such as obesity remains rudimentary, particularly in the case of rare CNVs.

Extensive replication studies including larger numbers of subjects are now required in order to study reported structural variants more comprehensively, verify their reported associations with obesity susceptibility and provide better estimates of their effect sizes. As previously discussed, functional studies will also be necessary to uncover the underlying mechanisms by which such variants may contribute to body weight regulation.

Furthermore, novel methodologies, both technical and statistical, will be required to enable more systematic investigation of complex CNVs such as multi-allelic CNVs. It is hoped that such further in-depth analyses of structural variation may improve our understanding of the genetics factors underlying susceptibility to complex diseases such as obesity.

References

1. Friedman JM (2003) A war on obesity, not the obese. Science 299(5608):856–858
2. Stunkard AJ, Foch TT, Hrubec Z (1986) A twin study of human obesity. JAMA 256(1):51–54
3. Stunkard AJ, Harris JR, Pedersen NL, McClearn GE (1990) The body-mass index of twins who have been reared apart. N Engl J Med 322(21):1483–1487
4. Stunkard AJ, Sorensen TI, Hanis C, Teasdale TW, Chakraborty R, Schull WJ et al (1986) An adoption study of human obesity. N Engl J Med 314(4):193–198
5. Turula M, Kaprio J, Rissanen A, Koskenvuo M (1990) Body weight in the Finnish Twin Cohort. Diabetes Res Clin Pract 10(Suppl 1):S33–S36
6. Maes HH, Neale MC, Eaves LJ (1997) Genetic and environmental factors in relative body weight and human adiposity. Behav Genet 27(4):325–351
7. Wardle J, Carnell S, Haworth CM, Plomin R (2008) Evidence for a strong genetic influence on childhood adiposity despite the force of the obesogenic environment. Am J Clin Nutr 87(2):398–404

8. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU et al (2010) Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet 42(11):937–948

9. Day FR, Loos RJ (2011) Developments in obesity genetics in the era of genome-wide association studies. J Nutrigenet Nutrigenomics 4(4):222–238

10. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ et al (2009) Finding the missing heritability of complex diseases. Nature 461(7265):747–753

11. de Smith AJ, Tsalenko A, Sampas N, Scheffer A, Yamada NA, Tsang P et al (2007) Array CGH analysis of copy number variation identifies 1284 new genes variant in healthy white males: implications for association studies of complex diseases. Hum Mol Genet 16(23): 2783–2794

12. Korbel JO, Urban AE, Affourtit JP, Godwin B, Grubert F, Simons JF et al (2007) Paired-end mapping reveals extensive structural variation in the human genome. Science 318(5849): 420–426

13. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD et al (2006) Global variation in copy number in the human genome. Nature 444(7118):444–454

14. Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y et al (2010) Origins and functional impact of copy number variation in the human genome. Nature 464(7289):704–712

15. Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P et al (2004) Large-scale copy number polymorphism in the human genome. Science 305(5683):525–528

16. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y et al (2004) Detection of large-scale variation in the human genome. Nat Genet 36(9):949–951

17. Locke DP, Sharp AJ, McCarroll SA, McGrath SD, Newman TL, Cheng Z et al (2006) Linkage disequilibrium and heritability of copy-number polymorphisms within duplicated regions of the human genome. Am J Hum Genet 79(2):275–290

18. Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU et al (2005) Segmental duplications and copy-number variation in the human genome. Am J Hum Genet 77(1): 78–88

19. de Smith AJ, Walters RG, Coin LJ, Steinfeld I, Yakhini Z, Sladek R et al (2008) Small deletion variants have stable breakpoints commonly associated with alu elements. PLoS One 3(8):e3104

20. Conrad DF, Bird C, Blackburne B, Lindsay S, Mamanova L, Lee C et al (2010) Mutation spectrum revealed by breakpoint sequencing of human germline CNVs. Nat Genet 42(5):385–391

21. Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N et al (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science 315(5813):848–853

22. Pollex RL, Hegele RA (2007) Copy number variation in the human genome and its implications for cardiovascular disease. Circulation 115(24):3130–3138

23. Hindorff LA, MacArthur J (European Bioinformatics Institute), Wise A, Junkins HA, Hall PN, Klemm AK, Manolio TA. A Catalog of Published Genome-Wide Association Studies. Available at www.genome.gov/gwastudies. Accessed 11 Jun 2012

24. Coin LJ, Asher JE, Walters RG, Moustafa JS, de Smith AJ, Sladek R et al (2010) cnvHap: an integrative population and haplotype-based multiplatform model of SNPs and CNVs. Nat Methods 7(7):541–546

25. Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SF et al (2007) PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res 17(11):1665–1674

26. Colella S, Yau C, Taylor JM, Mirza G, Butler H, Clouston P et al (2007) QuantiSNP: an Objective Bayes Hidden-Markov Model to detect and accurately map copy number variation using SNP genotyping data. Nucleic Acids Res 35(6):2013–2025

27. Peiffer DA, Le JM, Steemers FJ, Chang W, Jenniges T, Garcia F et al (2006) High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. Genome Res 16(9):1136–1148

28. Sha BY, Yang TL, Zhao LJ, Chen XD, Guo Y, Chen Y et al (2009) Genome-wide association study suggested copy number variation may be associated with body mass index in the Chinese population. J Hum Genet 54:199–202

29. Batterham RL, Le Roux CW, Cohen MA, Park AJ, Ellis SM, Patterson M et al (2003) Pancreatic polypeptide reduces appetite and food intake in humans. J Clin Endocrinol Metab 88(8):3989–3992

30. Schmidt PT, Naslund E, Gryback P, Jacobsson H, Holst JJ, Hilsted L et al (2005) A role for pancreatic polypeptide in the regulation of gastric emptying and short-term metabolic control. J Clin Endocrinol Metab 90(9):5241–5246

31. Kamiji MM, Inui A (2007) Neuropeptide y receptor selective ligands in the treatment of obesity. Endocr Rev 28(6):664–684

32. Jarick I, Vogel CI, Scherag S, Schafer H, Hebebrand J, Hinney A et al (2011) Novel common copy number variation for early onset extreme obesity on chromosome 11q11 identified by a genome-wide analysis. Hum Mol Genet 20(4):840–852

33. El-Sayed Moustafa JS, Eleftherohorinou H, de Smith AJ, Andersson-Assarsson JC, Couto Alves A, Hadjigeorgiou E et al (2012) Novel association approach for variable number tandem repeats (VNTRs) identifies DOCK5 as a susceptibility gene for severe obesity. Hum Mol Genet 21(16):3727–3738

34. Cote JF, Vuori K (2002) Identification of an evolutionarily conserved superfamily of DOCK180-related proteins with guanine nucleotide exchange activity. J Cell Sci 115(Pt 24):4901–4913

35. Bustelo XR, Sauzeau V, Berenjeno IM (2007) GTP-binding proteins of the Rho/Rac family: regulation, effectors and functions in vivo. Bioessays 29(4):356–370

36. Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, Heid IM et al (2009) Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat Genet 41(1):25–34

37. Craddock N, Hurles ME, Cardin N, Pearson RD, Plagnol V, Robson S et al (2010) Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. Nature 464(7289):713–720

38. Walters RG, Jacquemont S, Valsesia A, de Smith AJ, Martinet D, Andersson J et al (2010) A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. Nature 463(7281):671–675

39. Bochukova EG, Huang N, Keogh J, Henning E, Purmann C, Blaszczyk K et al (2010) Large, rare chromosomal deletions associated with severe early-onset obesity. Nature 463(7281): 666–670

40. Wang K, Li WD, Glessner JT, Grant SF, Hakonarson H, Price RA (2010) Large copy-number variations are enriched in cases with moderate to extreme obesity. Diabetes 59(10): 2690–2694

41. Jacquemont S, Reymond A, Zufferey F, Harewood L, Walters RG, Kutalik Z et al (2011) Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. Nature 478(7367):97–102

42. Bachmann-Gagescu R, Mefford HC, Cowan C, Glew GM, Hing AV, Wallace S et al (2010) Recurrent 200-kb deletions of 16p11.2 that include the SH2B1 gene are associated with developmental delay and obesity. Genet Med 12(10):641–647

43. Zufferey F, Sherr EH, Beckmann ND, Hanson E, Maillard AM, Hippolyte L et al (2012) A 600 kb deletion syndrome at 16p11.2 leads to energy imbalance and neuropsychiatric disorders. J Med Genet 49(10):660–668

44. Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R et al (2008) Association between microdeletion and microduplication at 16p11.2 and autism. N Engl J Med 358(7):667–675

45. McCarthy SE, Makarov V, Kirov G, Addington AM, McClellan J, Yoon S et al (2009) Microduplications of 16p11.2 are associated with schizophrenia. Nat Genet 41(11):1223–1227

46. Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, Skaug J et al (2008) Structural variation of chromosomes in autism spectrum disorder. Am J Hum Genet 82(2):477–488

47. Kumar RA, KaraMohamed S, Sudi J, Conrad DF, Brune C, Badner JA et al (2008) Recurrent 16p11.2 microdeletions in autism. Hum Mol Genet 17(4):628–638

48. Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T et al (2007) Strong association of de novo copy number mutations with autism. Science 316(5823):445–449

49. Yu Y, Zhu H, Miller DT, Gusella JF, Platt OS, Wu BL et al (2011) Age- and gender-dependent obesity in individuals with 16p11.2 deletion. J Genet Genomics 38(9):403–409
50. Bulik CM, Slof-Op't Landt MC, van Furth EF, Sullivan PF (2007) The genetics of anorexia nervosa. Annu Rev Nutr 27:263–275
51. Walters R, Coin LJM, Ruokonen A, de Smith AJ, El-Sayed Moustafa JS, Jacquemont S, Elliott P, Esko T, Hartikainen AL, Laitinen J, Männik K, Martinet D, Meyre D, Nauck M, Schurmann C, Sladek R, Thorleifsson G, Thorsteinsdóttir U, Valsesia A, Waeber G, Zufferey F, Balkau B, Pattou F, Metspalu A, Völzke H, Vollenweider P, Stefansson K, Jarvelin MR, Beckmann JS, Froguel P, Blakemore AIF (2013) Rare genomic structural variants in complex disease: lessons from the replication of associations with obesity. PLoS One 8:e58048
52. Ren D, Li M, Duan C, Rui L (2005) Identification of SH2-B as a key regulator of leptin sensitivity, energy balance, and body weight in mice. Cell Metab 2(2):95–104
53. Doche ME, Bochukova EG, Su HW, Pearce LR, Keogh JM, Henning E et al (2012) Human SH2B1 mutations are associated with maladaptive behaviors and obesity. J Clin Invest 122(12):4732–4736
54. Glessner JT, Bradfield JP, Wang K, Takahashi N, Zhang H, Sleiman PM et al (2010) A genome-wide study reveals copy number variants exclusive to childhood obesity cases. Am J Hum Genet 87(5):661–666
55. Golzio C, Willer J, Talkowski ME, Oh EC, Taniguchi Y, Jacquemont S et al (2012) KCTD13 is a major driver of mirrored neuroanatomical phenotypes of the 16p11.2 copy number variant. Nature 485(7398):363–367
56. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM et al (2002) The human genome browser at UCSC. Genome Res 12(6):996–1006
57. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP et al (2010) LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 26(18): 2336–2337