Biochemical Society Transactions (2020) **48** 1819–1828
https://doi.org/10.1042/BST20200556

PORTLAND PRESS

Review Article

DNA copy number and structural variation (CNV) contributions to adult and childhood obesity

Megan Phillips[1], Jeganathan Ramesh Babu[1,2,3], Xu Wang[4,5,3] and Thangiah Geetha[1,2,3]

[1]Department of Nutrition, Dietetics, and Hospitality Management, Auburn University, Auburn, AL 36849, U.S.A.; [2]Boshell Metabolic Diseases and Diabetes Program, Auburn University, Auburn, AL 36849, U.S.A.; [3]Alabama Agricultural Experiment Station, Auburn University, Auburn, AL 36849, U.S.A.; [4]Department of Pathobiology, Auburn University, Auburn, AL 36849, U.S.A.; [5]HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, U.S.A.

Correspondence: Thangiah Geetha (thangge@auburn.edu)

In recent years, obesity has reached epidemic proportions globally and has become a major public health concern. The development of obesity is likely caused by several behavioral, environmental, and genetic factors. Genomic variability among individuals is largely due to copy number variations (CNVs). Recent genome-wide association studies (GWAS) have successfully identified many loci containing CNV related to obesity. These obesity-related CNVs are informative to the diagnosis and treatment of genomic diseases. A more comprehensive classification of CNVs may provide the basis for determining how genomic diversity impacts the mechanisms of expression for obesity in children and adults of a variety of genders and ethnicities. In this review, we summarize current knowledge on the relationship between obesity and the CNV of several genomic regions, with an emphasis on genes at the following loci: 11q11, 1p21.1, 10q11.22, 10q26.3, 16q12.2, 16p12.3, and 4q25.

Introduction

According to the National Health and Nutrition Examination Survey (NHANES), the prevalence of obesity in the United States is ~39.8% in adults and 18.5% in adolescents [1]. Obesity is a risk factor for some of the leading causes of mortality, including type 2 diabetes, heart disease, certain types of cancers, and COVID-19 [2]. What makes obesity different from other risk factors is that it is preventable [3]. The development of obesity is likely caused by several dietary, metabolic, and genetic factors [4]. Genetic variation plays a major role in determining the susceptibility or resistance an individual may have to the environment. Factors that contribute to this environment include reduced energy expenditure, access to fresh fruits and vegetables, increased high-calorie food intake, and socioeconomic status [5]. Previous analysis of the human genome points to single nucleotide polymorphisms (SNPs) to be the primary source of human genetic variation. However, current research shows copy number variations (CNVs) are responsible for a large extent of structural change in the genome of mammals [6]. A CNV is a DNA segment spanning 1 kb [7] or larger that has a variable number of copies among individuals in a population when compared with a reference genome [8]. Variations can be deletions or duplications [9]. Some CNVs do not appear to influence phenotype; however, several have been conclusively linked with disease. Recent studies have determined pathways involved in obesity pathogenesis by identifying the CNV and associated proteins increasing susceptibility to obesity. Also, there is evidence that interaction with other genetic or environmental factors may influence whether CNVs affect phenotypes [10]. For example, exposure to environmental mutagens, such as radiation, can lead to inherited genetic predispositions [11].

The mechanism of fork stalling and template switching (FoSTeS) and microhomology-mediated break-induced replication (MMBIR) model can also lead to genomic rearrangements of single exons, duplication or triplication of individual genes providing an innovative perspective for exploring gene and genome evolution [12]. The CNV represents a source of genetic diversity in human health,

Received: 14 May 2020
Revised: 5 July 2020
Accepted: 7 July 2020

Version of Record published: 29 July 2020

Biochemical Society Transactions (2020) 48 1819–1828
https://doi.org/10.1042/BST20200556

disease, and evolution [13,14]. The genome-wide association studies (GWAS) have identified several loci associated with obesity [15]. D'Angelo et al. described the CNVs of genes and genomic regions in obesity associated with developmental delay, intellectual disabilities (ID), and malformative features [16]. However, there is still more to be studied regarding the association of CNVs and obesity. In this article, we focus on the common CNVs, such as 11q11, 1p21.1, 10q11.22, 10q26.3, 16q12.2, 16p12.3 and 4q25 that has been found to associate with adult and childhood obesity.

The chromosome 11q11 locus

A potential link between olfactory receptors (ORs) and obesity has been proposed based on the observation of altered olfactory function in morbidly obese patients. Richardson et al. [17] reported that adult patients with a body mass index (BMI) greater than 45 were more likely to have olfactory dysfunction when compared with subjects with BMI of less than 45. The ability to detect chemicals in the environment is a sensory system dependent on ORs. A chemical compound has an odor when it is sufficiently volatile to be transported to the olfactory system [18]. OR neurons project their axon directly to the olfactory bulb in the brain. The central olfactory system is separated from the epithelium by the cribriform plate of the ethmoid bone. The cribriform plate, which supports the olfactory bulb, is perforated by a grouping of holes, known as the olfactory foramina, creating a channel for the olfactory nerves [19]. The first step in odor transduction begins when odorous ligands activate ORs on the olfactory epithelium [20]. The epithelium contains millions of ORs that are capable of binding with odor molecules. However, an odor molecule will not bind to just any OR. It must bind to the OR specifically designed to identify that molecule [21]. Since ORs belong to the G-protein-coupled receptor (GPCR) superfamily, the transduction of chemical information into electrical impulses involves signal amplification by G-protein-coupled activation of adenylyl cyclase. This facilitates a rise in cyclic AMP (cAMP), and subsequently, the opening of cyclic nucleotide-activated, nonselective cation channels. The cell membrane of the olfactory neuron becomes depolarized from the influx of cations through these channels. Ultimately, these results in an increase in the frequency of action potentials that travel down the axons to the glomeruli, which are spherical structures located in the outer part of the olfactory bulb [18]. OR genes have been identified on several chromosomes. With around 400 functional genes, they encompass one of the largest gene families in mammalian genomes [22].

The chromosome 11q11 locus, spanning ~80 kb, exclusively covers three OR genes, all belonging to the OLR family 4: OR4P4, OR4S2, and OR4C6 [23]. In addition to functional ORs, the gene family also contains a large number of pseudogenes, which no longer encode ORs due to loss of function mutations occurred during evolution [24]. Earlier studies have identified variation in the copy number of the 11q11 locus in association with obesity (Table 1, Figure 1). Copy number at this locus can range from 0 to 8 and appears to vary with gender, age, and ethnicity. León-Mimila et al. [25] reported a lower 11q11 copy number in Mexican children aged 6–12 years were significantly associated with lower obesity risk in children, but not in adult subjects. A hereditary factor has been observed in children of German descent. Family-based GWAS have demonstrated a lower 11q11 copy number in obese children compared with those of normal weight [23]. Zhang et al. observed a significant increase in the risk of obesity in Chinese children with deletions at the 11q11 locus. Interestingly, a cumulative effect was observed when these subjects also had deletions at two other loci related to obesity: 10q11.22 and 4q25. This suggests a robust collective association between increased CNV at-risk alleles and childhood obesity [26].

In addition to GWAS, 11q11 has been studied in individual case studies. Obesity is associated with a variety of liver abnormalities collectively known as nonalcoholic fatty liver disease (NAFLD) that is characterized by an accumulation of triglycerides in the liver (steatosis). This can range from simple steatosis to its inflammatory counterpart nonalcoholic steatohepatitis (NASH) [27]. NAFLD is associated with cardiometabolic syndrome (CMS), increasing the risk of insulin resistance, impaired glucose tolerance, hypertension, dyslipidemia, and central adiposity [28]. The 11q11 region has been shown to have one copy number less in adult patients with NAFLD when compared with controls, who had a copy number around 2 [27].

Salivary amylase 1 and obesity

The variation in the copy number appears to be influenced by several environmental factors, including stress levels, circadian rhythms, and diet [29]. Dietary habits have been shown to directly influence the copy number of certain obesity-related genes, Falchi et al. [30] was the first study to identify the genetic link between carbohydrate metabolism and obesity. Specifically, individuals consuming greater amounts of starch have been

shown to have higher levels of both salivary amylase protein and serum amylase compared with those consuming less starch. Recent research has indicated that the expression of the salivary amylase 1 (AMY1) gene on chromosome 1p21.1 is up-regulated by a high starch diet [31,32]. Salivary amylase, as well as pancreatic amylase, is secreted during the first steps of starch digestion, which continues after passing from the oral cavity into the stomach. These enzymes are specifically responsible for the hydrolysis of α-1,4 glycosidic bond, which yield oligosaccharides such as maltose and maltotriose during glycolysis [33]. Thus, high-AMY1 copy number and salivary amylase activity are both favorable for more efficient dietary starch digestion [34,35]. A significant negative association of AMY1 copy number with BMI and obesity was first reported by Falchi et al. [30] (Figure 1). AMY1 copy number is also found to be associated with serum salivary amylase enzyme levels and gene expression [30].

AMY1 is one of the most variable loci in copy number in the human genome, ranging from 2 to 20 copies [33]. Interestingly, the correlation of AMY1 copy number with obesity seems to be significant in the pediatric populations of various ethnicities. A recent study from our laboratory found a negative correlation between its copy number and BMI in pediatric population of both African American and European American ancestry in Alabama. Furthermore, the negative relationship of AMY1 copy number with obesity measurements was stronger in African American children than in European American children [36]. An Italian population known to rely on a high starch diet consisting of primarily complex carbohydrates demonstrated a significant correlation as well. A lower AMY1 copy number was associated with increased BMI in pediatric boys [37]. In a study done in Mexican children, having less than a specific number of copies increased the risk of obesity. An AMY1 copy number less than 6 correlated to a higher risk of obesity compared with those with at least six AMY1 copy numbers [25]. Similar studies have shown this correlation between being not only significant in pediatrics but also adult populations. A lower AMY1 copy number was associated with a higher BMI in both European and Asian adults [30].

In contrast, other studies observed differences in gender. In a Finnish study, Viljakainen et al. found no difference in AMY1 copy numbers between healthy individuals and individuals with a history of childhood-onset obesity. However, obese men had a higher copy number compared with obese females. Furthermore, AMY1 copy number correlated significantly with whole body fat percent and BMI only in obese females [35].

The genetic variation of the individual influences not only influences phenotype but also the microbiome composition. Previous studies have focused on the effect the gut microbiome may have on specific genotypes. Recent evidence supports the idea that gene copy number also varies. It has been established that AMY1 copy number is a genetic factor associated with microbiome composition and function. In a month-long diet intervention study, Poole et al. [38] not only showed that diet standardization was a catalyst for gut microbiome convergence, but also that AMY1 copy number correlated with the composition and function of oral and gut microbiomes. The microbiomes of individuals with low-AMY1 copy numbers had enhanced the capacity to metabolize complex carbohydrates. Interestingly, high-AMY1 copy number subjects had higher levels of salivary Porphyromonas. Gut microbiota of those individuals had a more considerable amount of resistant starch-degrading microbes and produced higher levels of short-chain fatty acids [38]. In a parallel study, León-Mimila et al. explored possible correlations between AMY1 copy number and several identified gut microbiota genera. The abundance of Prevotella was positively correlated to AMY1 copy number in the adult population. *Prevotella copri*, specifically, was two times higher in adults with at least 10 copies of AMY1 compared with those with less than four copies [25] (Table 1).

NPY4R and appetite control

In addition to AMY1, current research has uncovered another obesity-related gene involved in digestion. A CNV region located on chromosome 10q11.22 is known to cover the Neuropeptide Y Receptor type 4 (NPY4R) gene, also known as PPYR1 [23]. NPY4R encodes the NPY receptor that responds to the pancreatic polypeptide (PP), which has been shown to be an effective appetite inhibitor [39] (Figure 1). There are four genes in the NPY receptor family in humans, all of which are expressed in the hypothalamic region of the brain that is involved in appetite control and energy metabolism. Pancreatic PP-cells release PP postprandially in proportion to caloric intake to self-regulate pancreatic secretion activities [39]. Copy number gain of NPY in a 137 kb duplicated region in 7p15.3 was detected in early-onset obesity [40].

The relationship between CNV of the NPY4R gene and BMI, waist circumference, and dietary intake has been evaluated in recent studies. Zhang et al. identified 10q11.22 to be significantly associated with obesity in a pediatric population in China. The deletion of 10q11.22 was linked to a higher BMI and waist to height ratio.

Biochemical Society Transactions (2020) 48 1819–1828
https://doi.org/10.1042/BST20200556

The risk of obesity increased even further among the 10q11.22 deletion carriers who had meat-based diets, which could indicate a multiplicative interaction (MI) between deletions of 10q11.22 and a preference for meat-based meals [26]. It has been shown that protein, dietary fat, and to some degree, glucose stimulates PP production [41]. The intake of meat regularly may stimulate PP secretion, which would subsequently regulate the intake of food via PPYR1. However, in subjects with 10q11.22 deletions, PP cannot function normally. As a result, the 10q11.22 deletion carriers with meat-based diets might be more prone to developing obesity due to a pathologic inability to control food intake. In addition to a meat-based diet, there was a significant association between subjects with the 10q11.22 deletion and a preference for sweet foods. This could, in part, be due to glucose consumption stimulating the release of PP [26].

Another study in Chinese subjects found that deletions at NPY4R were not only significantly associated with higher BMI in pediatrics, but also in the elderly [42]. In contrast, Sun et al. [43] found no copy number changes at NPY4R in Chinese obese or healthy controls. Other studies have provided evidence of differences in both ethnicity and age group. Aerts et al. performed a mutation screen for variants in the NPY4R coding region in two groups: obese Belgian children and adolescents and lean White adults. The CNV analysis demonstrated a significantly higher frequency of NPY4R containing 10q11.22 CNV loss in the obese Belgian pediatric population. Furthermore, a CNV gain in this region was more prevalent in the lean White adult population [44]. Gender differences were observed in a community of Swedish adults. A positive correlation was found between NPY4R copy number and BMI, as well as waist circumference, in Swedish women. Each additional copy of NPY4R correlated with a 2.6 kg/m^2 increase in BMI and a 5.67 cm increase in waist circumference. The same findings were not seen in Swedish adult men [39]. Through a mutation screen in NPY and NPY2R genes of 436 obese children and adolescents, Aerts et al. [45] discovered missense mutations in both genes. No CNVs were detected in NPY2R in this population [45] (Table 1).

CYP2E1 and lipid metabolism

The ability of certain endogenous or exogenous factors to influence the activity level of genes has been shown to have a significant role in human health [46]. In addition to NPY4R, a region at the 10q26.3 locus comprises a superfamily of functional genes that are also inducible by several endogenous factors. The Cytochromes p450 superfamily (CYP) encodes enzymes and proteins that catalyze many reactions involving drug metabolism and synthesis of steroids, cholesterol, and lipids [47]. CYP proteins can be located in several different compartments of the cell [46]. In addition to their localization in the Golgi apparatus and plasma membrane, several different forms have been detected within microsomes present in the endoplasmic reticulum membrane, and in the mitochondria [48]. Different types of signals are required for targeting of P450 proteins to ER and mitochondria. Microsomal P450 is inserted into the rough-surfaced portion of the ER membrane through a non-cleavable hydrophobic signal sequence [49]. CYPs that play a major role in endogenous metabolisms are found mostly in mitochondria [50]. In contrast with microsomal P450s, mitochondrial P450s are comprised of cleavable amphipathic presequences that are important for targeting of the precursor protein (preprotein) to the mitochondria posttranslationally [51].

Of particular interest is Cytochrome P450 family 2 subfamily E member 1 (CYP2E1). CYP2E1 is important in signaling pathways related to diabetes and obesity, such as $\omega - 1$ hydroxylation of fatty acids like arachidonic acid. CYP2E1 also plays a role in the propanediol pathway, an alternative pathway of gluconeogenesis, which offers a way to generate glucose during starvation via hydroxylation of acetone, hydroxyacetone, and pyruvate [47]. One of the most distinguishing features that set CYP2E1 apart is its inducibility by a large variety of substrates [50]. Endogenous factors that induce CYP2E1 expression levels include obesity, diabetes, fasting, and chronic alcohol ingestion [52]. The fluctuation of CYP2E1 levels and expression might alter the balance, and subsequently, TG metabolism.

The expression of CYPs manifests in the liver, as well as other tissues, such as the brain, kidneys, and lungs. The first evidence of CYP2E1 localization within the liver was in 1997. Avadhani et al. [53] discovered the presence of the CYP2E1 protein located in the inner membrane of rat liver mitochondria. Hepatic CYP2E1 activity has been reported to be up-regulated in individuals with severe obesity, defined as a BMI higher than 35 kg/m^2 [54]. This finding can offer a potential explanation for why morbid obesity is often associated with fatty liver disease (steatosis) [55]. The degree of severity may range from steatosis alone to steatohepatitis (accompanying inflammation of the liver) with advanced fibrosis [56]. Emery et al. assessed the activity of CYP2E1 by rate of chlorzoxazone (CLZ) clearance, which has been used extensively to evaluate CYP2E1 activity. Both the total and unbound oral CLZ clearance (Cl_u/F) was approximately three times higher in morbidly obese subjects

compared with controls. One year after the morbidly obese subjects underwent gastroplasty, the total oral CLZ clearance and Cl_u/F declined by 46% and 35%, respectively. A positive association between the degree of steatosis and CYP2E1 activity before and after surgery suggests that increases in CYP2E1 activity may be related to the hepatic pathology of the liver resulting from morbid obesity (Figure 1). Furthermore, the up-regulation of CYP2E1 activity may accelerate liver injury when significant steatosis is present [55].

The CNV of CYP2E1 can influence activity by altering gene expression [57]. There is evidence showing that individuals have a copy number range of 1–4. Over 96% of individuals have two copies of the CYP2E1 gene. Three percent of people have three copies (tri-copy), and less than 2% of the population has either one copy (haplotype) or four copies (tetratype) [57]. Yang et al. reported association results in White and African American groups that indicated CNV at 10q26.3 might be a common variant for obesity across different ethnicities. The CNV found on CYP2E1 showed a strong positive association with both BMI and body fat mass in three independent populations: an unrelated sample of White subjects, a family-based sample of White subjects, and an unrelated sample of African American subjects [58] (Table 1).

One of the most severe lipid metabolism disorders related to obesity is Hypertriglyceridemia (HTG). HTG is characterized primarily by levels of plasma triglyceride higher than 1.7 mmol/l. When the level of plasma TGs is higher than 10 mmol/l, genetic factors are thought to play a dominant role in HTG [59]. Interestingly, individuals with CYP2E1 tri-copy and tetratype often show obesity and HTG phenotypes.

Fat mass and obesity associated (FTO) gene and shift from brown to white adipose cells

Another gene heavily involved with lipids is alpha-ketoglutarate dependent dioxygenase (FTO), which is located at the 16q12.2 chromosomal region. FTO was first identified through GWAS as the obesity-susceptibility gene [60]. FTO functions as an RNA and single-strand DNA demethylase that mediates oxidative demethylation of several different RNA species, including N6-methyladenosine (m6A) in mRNA transcripts. FTO is highly expressed in the adrenal glands and forebrain, specifically the hypothalamus and pituitary regions [61]. In doing so, FTO plays the role of regulator for adipogenesis and energy homeostasis (Figure 1). This contributes to the regulation of body size and body fat accumulation, specifically the differentiation into white or brown fat cells [62–64]. It is uncertain whether variations associated with obesity affect the function of the FTO gene directly, or modify the expression of adjacent genes, such as IRX3 [65]. A pathogenic intronic FTO variation (rs1421085) interrupts the sequence for ARID5B binding that has remained essentially unchanged throughout evolution. Without ARID5B binding, the two genes distal to FTO (IRX3 and IRX5) will be overexpressed, which alters the differentiation of pre-adipocytes and shifts from brown to white fat cells. This alteration results in a loss of mitochondrial thermogenesis and increased lipid storage [62]. This type of IRX-dependent shift from consuming energy to storing it in the form of adipocytes may offer the protection of body fat under conditions where energy supply is limited. This type of defense mechanism is pertinent for lean children in order to protect body weight [66].

Recent research suggests IRX3 may be the main mediator of obesity risk in children who are carriers of the FTO risk variant. González-Herrera et al. observed differences in CNV of the FTO gene that was associated with overweight status in Mexican-Mayan pediatric boys. Mexican Mayans from Yucatán have been reported to have a much higher prevalence of overweight and obesity than the national average. When comparing FTO risk-allele carriers to non-risk-allele carriers, adipocyte-specific expression of both IRX3 and IRX5 was increased only in lean children. Overweight boys showed higher average CNV than boys with normal weight [67] (Table 1). Interestingly, Mexican children from central and northern Mexico differ from those indicating a lack of FTO rs1421085 association with obesity [68].

Other obesity-related loci

In addition to the previously mentioned chromosomal regions, there are loci in which the biological function is not completely understood. Among these regions is 16p12.3, a 21 kb deletion located roughly 50 kb upstream of the gene GPCR, family C, group 5, member B (GPRC5b). The protein encoded by this gene is a member of the type-3 GPCR family. Research suggests the function of this protein may facilitate the cellular effects of retinoic acid on the G-protein signal transduction cascade. The direct correlation between GPRC5b and the molecular biology of obesity phenotypes is largely undiscovered [69]. Regarding the GPCR5 CNV, it has been previously suggested that the effect of the deletion of CNV is ethnic-specific, as it has been significantly

Biochemical Society Transactions (2020) **48** 1819–1828
https://doi.org/10.1042/BST20200556

Table 1. Summary of studies on the relation between DNA CNVs and obesity

Locus	Gene	CNV	Function(s)	Region of activity	Effect(s)
11q11	OR4P4, OR4S2, OR4C6 *Olfactory Receptor Family 4*	0–8	Specifically recognizes odorous molecules in the 1st step of odor transduction	Expressed in the olfactory bulb glomeruli of the brain	↓ CN =↓ obesity in Mexican children [25] ↓ CN = ↑ obesity in German children [23] ↓ CN = ↑ obesity in Chinese children [26] ↓ CN = ↑ obesity related to NASH in Malaysian adults [27]
1p21.1	AMY1 *Salivary Amylase 1*	2–20	Aids in hydrolysis of α-1,4 glycosidic linkages in starch metabolism	Expressed at high levels in salivary glands	↓ CN = ↑ obesity in EA and AA children [36] ↓ CN = ↑ BMI in Italian pediatric boys [37] CN < 6 = ↑ obesity in Mexican children [25] ↓ CN = ↑ obesity in European and Asian adults [30] ↑ CN = ↑ BMI and BF % in obese Finnish women [26] ↑ CN = ↑ obesity and *Porphyromonas* in Mexican adults [38]
10q11.22	NPY4R (PPYR1) *Neuropeptide Y Receptor Y4*	2–8	Y4 receptor responds to PP (appetite inhibitor)	Expressed in the hypothalamic region of the brain involved in appetite control and energy metabolism	↓ CN = ↑ BMI and WHtR in Chinese children [26] ↓ CN = ↑ BMI in Chinese elderly [42] ↓ CN = ↑ obesity in Belgian children [44] ↑ CN = ↓ obesity in White adults [44] ↑ CN = ↑ body weight in Swedish women [39]
10q26.3	CYP2E1 *Cytochrome p450 2E1*	1–4	Involved in cholesterol and lipid synthesis Induced by fasting, diabetes and obesity	Primarily expressed in the mitochondria of hepatic cells	↑ CN = ↑ BMI and body fat in EA adults [58] ↑ CN = ↑ BMI and body fat in AA adults [58]
16q12.2	FTO *Alpha-ketoglutarate Dependent Dioxygenase*	0–2	Involved in the regulation of thermogenesis, adipogenesis, fat mass, body weight and the control of adipocyte differentiation into brown or white fat cells.	Expressed in adrenal glands and brain, specifically the hypothalamus and pituitary gland.	↑ CN = ↑ obesity in Mexican-Mayan pediatric boys [67] ↓ CN = ↑ obesity in Chinese children [26]
16p12.3	GPRC5b *G-protein-coupled receptor, family C, group 5, member b*	0–2	Facilitates the cellular effects of retinoic acid on the G-protein signal transduction cascade	Not available	↓ CN = ↑ BMI and body fat in EA adults [69]
4q25	Intergenic region	NA	Modifier of rare ion channel mutation associated with familial atrial fibrillation	Cardiac ion channel	↓ CN = ↑ obesity in Chinese children [26]

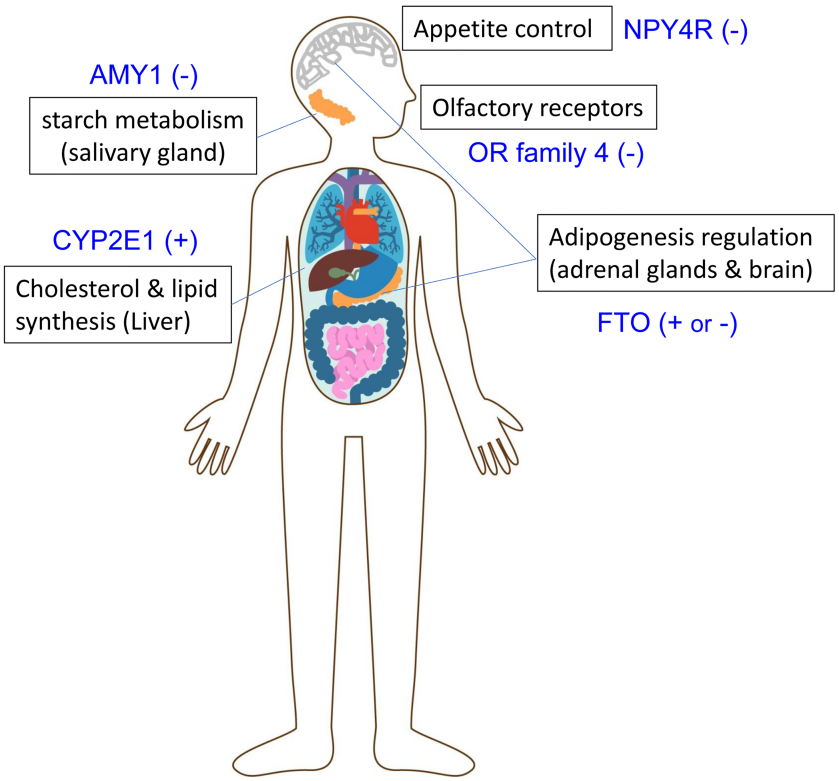

Figure 1. Copy number variation (CNV), physiological pathways and obesity.

The gene name and pathway information are labeled in the diagram. (+) positive correlation between CN and BMI; (–) negative correlation between CN and BMI.

associated with obesity in Europeans, who have a higher CNV deletion frequency. The same association was not found in Chinese populations who have a lower CNV deletion frequency. The GPCR5 CNV deletion frequency observed in the Mexican people by León-Mimila et al. was consistent with the previous findings in Chinese individuals. The copy number was not associated with obesity [25]. Other loci with an incomplete understanding contain CNV deletions that are correlated with a higher risk of childhood obesity. The 4q25 locus may be mediated by long-range regulators such as enhancers or repressors. CNVs may exert their effects on genes as far as 1 Mb away. Little is known about the biological function of this chromosomal region [26]. In obese-normal case-control studies, rare CNVs were found to be enriched in obese samples compared with control samples [7]. Wheeler et al. [7] performed a genome-wide CNV analysis in 1509 early-onset obesity patients and 5380 controls, and they discovered that significant rare CNV burden in severely obese cases. In another early-onset obesity study, rare CNVs were discovered in 19% of obese samples compared with 3% in control samples [70]. In a similar study of childhood obesity, clinically important rare CNVs were identified in 15% of the samples, suggesting that rare CNV events may explain a proportion of the missing heritability. Among these CNV loci, *16p11.2* microdeletions were observed in both studies and this region harbors an early-onset obesity risk gene *SH2B1* [71]. The discovery of rare obesity-related CNV would help to identify novel obesity candidate genes and elucidate the causal variants for obesity.

Conclusions and future directions

Human obesity has strong genetic predisposition with a relatively high heritability, but known GWAS hits only explain less than 10% of the phenotypic variability. Common and rare CNVs may contribute to the missing heritability in human obesity. This review summarized major obesity-related CNV loci, and understanding the genetic mechanisms will inform to the diagnosis and treatment of the diseases. A more comprehensive classification of CNVs may provide the basis for determining how genomic diversity affects the mechanisms of

Biochemical Society Transactions (2020) **48** 1819–1828
https://doi.org/10.1042/BST20200556

expression for obesity in children and adults of a variety of ethnicities. Compared with SNP association analysis, there is a paucity of GWAS analysis of obesity-related CNVs. The majority of the existing studies are targeted analysis or based on array CGH analysis, with limited resolution to discover novel CNVs. Whole-genome sequencing and *de novo* assembly approaches are needed to fully understand the role of CNVs in obesity. Another urgent need in obesity-related CNV research is the racial and ethnic health disparity in genomics.

Perspectives

- GWAS have successfully identified many loci containing CNV related to obesity.
- CNV determines how genomic diversity impacts the mechanisms of expression for obesity in children and adults of a variety of genders and ethnicities.
- Future research needed in obesity-related CNV research is the racial/ethnic health disparity in genomics.

Conflict of Interests

The authors declare that there are no competing interests associated with the manuscript.

Author Contributions

All authors contributed to the writing and editing of this manuscript.

Acknowledgements

This work was supported by the Alabama Agricultural Experiment Station and the Hatch Program of the National Institute of Food and Agriculture, U.S. Department of Agriculture to JRB, XW, and TG. XW is also supported by the National Science Foundation EPSCoR RII Track-4 Research Fellowship (NSF-OIA-1928770) and a generous laboratory start-up fund from the Auburn University College of Veterinary Medicine.

Abbreviations

AMY1, salivary amylase 1; CLZ, chlorzoxazone; CNV, copy number variations; CYP, Cytochromes p450 superfamily; FoSTeS, fork stalling and template switching; GPCR, G-protein-coupled receptor; GWAS, genome-wide association studies; MMBIR, microhomology-mediated break-induced replication; OR, olfactory receptor.

References

1 Hales, C.M., Carroll, M.D., Fryar, C.D. and Ogden, C.L. (2017) Prevalence of obesity among adults and youth: United States, 2015–2016. *NCHS Data Brief* **288**, 1–8
2 CDC (2020) Groups at Higher Risk for Severe Illness https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/groups-at-higher-risk.html. Accessed April 26 2020.
3 Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults–the evidence report. National Institutes of Health. (1998) *Obes. Res.* **6** Suppl 2, 51S–209S
4 Hammad, S.S. and Jones, P.J. (2017) Dietary fatty acid composition modulates obesity and interacts with obesity-related genes. *Lipids* **52**, 803–822 https://doi.org/10.1007/s11745-017-4291-9
5 Swinburn, B., Egger, G. and Raza, F. (1999) Dissecting obesogenic environments: the development and application of a framework for identifying and prioritizing environmental interventions for obesity. *Prev. Med.* **29**, 563–570 https://doi.org/10.1006/pmed.1999.0585
6 Hasin, Y., Olender, T., Khen, M., Gonzaga-Jauregui, C., Kim, P.M., Urban, A.E. et al. (2008) High-resolution copy-number variation map reflects human olfactory receptor diversity and evolution. *PLoS Genet.* **4**, e1000249 https://doi.org/10.1371/journal.pgen.1000249
7 Wheeler, E., Huang, N., Bochukova, E.G. Keogh, J.M., Lindsay, S., Garg, S. et al. (2013) Genome-wide SNP and CNV analysis identifies common and low-frequency variants associated with severe early-onset obesity. *Nat. Genet.* **45**, 513–517 https://doi.org/10.1038/ng.2607
8 Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D. et al. (2006) Global variation in copy number in the human genome. *Nature* **444**, 444–454 https://doi.org/10.1038/nature05329
9 Feuk, L., Carson, A.R. and Scherer, S.W. (2006) Structural variation in the human genome. *Nat. Rev. Genet.* **7**, 85–97 https://doi.org/10.1038/nrg1767
10 Clancy, S. (2008) Copy number variation. *Nat. Educ.* **1**, 95 https://www.nature.com/scitable/topicpage/copy-number-variation-445/

11 Arlt, M.F., Wilson, T.E. and Glover, T.W. (2012) Replication stress and mechanisms of CNV formation. *Curr. Opin. Genet. Dev.* **22**, 204–210 https://doi.org/10.1016/j.gde.2012.01.009
12 Zhang, F., Khajavi, M., Connolly, A.M., Towne, C.F., Batish, S.D. and Lupski, J.R. (2009) The DNA replication FoSTeS/MMBIR mechanism can generate genomic, genic and exonic complex rearrangements in humans. *Nat. Genet.* **41**, 849–853 https://doi.org/10.1038/ng.399
13 Zhang, F., Gu, W., Hurles, M.E. and Lupski, J.R. (2009) Copy number variation in human health, disease, and evolution. *Annu. Rev. Genom. Hum. Genet.* **10**, 451–481 https://doi.org/10.1146/annurev.genom.9.081307.164217
14 Kehrer-Sawatzki, H. (2007) What a difference copy number variation makes. *Bioessays* **29**, 311–313 https://doi.org/10.1002/bies.20554
15 Glessner, J.T., Bradfield, J.P., Wang, K., Takahashi, N., Zhang, H., Sleiman, P.M. et al. (2010) A genome-wide study reveals copy number variants exclusive to childhood obesity cases. *Am. J. Hum. Genet.* **87**, 661–666 https://doi.org/10.1016/j.ajhg.2010.09.014
16 D'Angelo, C.S. and Koiffmann, C.P. (2012) Copy number variants in obesity-related syndromes: review and perspectives on novel molecular approaches. *J. Obes.* **2012**, 845480 https://doi.org/10.1155/2012/845480
17 Richardson, B.E., Vander Woude, E.A., Sudan, R., Thompson, J.S. and Leopold, D.A. (2004) Altered olfactory acuity in the morbidly obese. *Obes. Surg.* **14**, 967–969 https://doi.org/10.1381/0960892041719617
18 Fahlbusch, K.G., Hammerschmidt, F.J., Panten, J., Pickenhagen, W., Schatkowski, D., Bauer, K. et al. (2003), Flavors and fragrances. In *Ullmann's Encyclopedia of Industrial Chemistry* https://doi.org/10.1002/14356007.a11_141
19 Bushak, L. (2015) How does your nose do what it does? The inner workings of our sense of smell. *Medical Daily*. https://www.medicaldaily.com/how-does-nose-smell-inner-workings-our-sense-smell-324566
20 Gaillard, I., Rouquier, S. and Giorgi, D. (2004) Olfactory receptors. *Cell. Mol. Life Sci.* **61**, 456–469 https://doi.org/10.1007/s00018-003-3273-7
21 Buck, L.B. (2004) Olfactory receptors and odor coding in mammals. *Nutr. Rev.* **62**, S184–S188 https://doi.org/10.1111/j.1753-4887.2004.tb00097.x
22 Young, J.M., Endicott, R.M., Parghi, S.S., Walker, M., Kidd, J.M. and Trask, B.J. (2008) Extensive copy-number variation of the human olfactory receptor gene family. *Am. J. Hum. Genet.* **83**, 228–242 https://doi.org/10.1016/j.ajhg.2008.07.005
23 Jarick, I., Vogel, C.I., 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**, 840–852 https://doi.org/10.1093/hmg/ddq518
24 O. Podlaha, J. and Zhang, (2010), Pseudogenes and their evolution, In *eLS* https://doi.org/10.1002/9780470015902.a0005118.pub2
25 Leon-Mimila, P., Villamil-Ramirez, H., Lopez-Contreras, B.E., Moran-Ramos, S., Macias-Kauffer, L.R., Acuna-Alonzo, V. et al. (2018) Low salivary amylase gene (AMY1) copy number is associated with obesity and gut prevotella abundance in Mexican children and adults. *Nutrients* **10**, 1607 https://doi.org/10.3390/nu10111607
26 Zhang, D., Li, Z., Wang, H., Yang, M., Liang, L., Fu, J. et al. (2015) Interactions between obesity-related copy number variants and dietary behaviors in childhood obesity. *Nutrients* **7**, 3054–3066 https://doi.org/10.3390/nu7043054
27 Zain, S.M., Mohamed, R., Cooper, D.N., Razali, R., Rampal, S., Mahadeva, S. et al. (2014) Genome-wide analysis of copy number variation identifies candidate gene loci associated with the progression of non-alcoholic fatty liver disease. *PLoS One* **9**, e95604 https://doi.org/10.1371/journal.pone.0095604
28 Fabbrini, E., Sullivan, S. and Klein, S. (2010) Obesity and nonalcoholic fatty liver disease: biochemical, metabolic, and clinical implications. *Hepatology* **51**, 679–689 https://doi.org/10.1002/hep.23280
29 Nater, U.M., Rohleder, N., Schlotz, W., Ehlert, U. and Kirschbaum, C. (2007) Determinants of the diurnal course of salivary alpha-amylase. *Psychoneuroendocrinology* **32**, 392–401 https://doi.org/10.1016/j.psyneuen.2007.02.007
30 Falchi, M., El-Sayed Moustafa, J.S., Takousis, P., Pesce, F., Bonnefond, A., Andersson-Assarsson, J.C. et al. (2014) Low copy number of the salivary amylase gene predisposes to obesity. *Nat. Genet.* **46**, 492–497 https://doi.org/10.1038/ng.2939
31 Bank, R.A., Hettema, E.H., Muijs, M.A., Pals, G., Arwert, F., Boomsma, D.I. et al. (1992) Variation in gene copy number and polymorphism of the human salivary amylase isoenzyme system in Caucasians. *Hum. Genet.* **89**, 213–222 https://doi.org/10.1007/BF00217126
32 Squires, B.T. (1953) Human salivary amylase secretion in relation to diet. *J. Physiol.* **119**, 153–156 https://doi.org/10.1113/jphysiol.1953.sp004835
33 Mandel, A.L., Peyrot des Gachons, C., Plank, K.L., Alarcon, S. and Breslin, P.A. (2010) Individual differences in AMY1 gene copy number, salivary alpha-amylase levels, and the perception of oral starch. *PLoS One* **5**, e13352 https://doi.org/10.1371/journal.pone.0013352
34 Perry, G.H., Dominy, N.J., Claw, K.G., Lee, A.S., Fiegler, H., Redon, R. et al. (2007) Diet and the evolution of human amylase gene copy number variation. *Nat. Genet.* **39**, 1256–1260 https://doi.org/10.1038/ng2123
35 Viljakainen, H., Andersson-Assarsson, J.C., Armenio, M., Pekkinen, M., Pettersson, M., Valta, H. et al. (2015) Low copy number of the AMY1 locus is associated with early-onset female obesity in Finland. *PLoS One* **10**, e0131883 https://doi.org/10.1371/journal.pone.0131883
36 Venkatapoorna, C.M.K., Ayine, P., Parra, E.P., Koenigs, T., Phillips, M., Babu, J.R. et al. (2019) Association of salivary amylase (AMY1) gene copy number with obesity in alabama elementary school children. *Nutrients* **11**, 1379 https://doi.org/10.3390/nu11061379
37 Marcovecchio, M.L., Florio, R., Verginelli, F., De Lellis, L., Capelli, C., Verzilli, D. et al. (2016) Low AMY1 gene copy number is associated with increased body mass index in prepubertal boys. *PLoS One* **11**, e0154961 https://doi.org/10.1371/journal.pone.0154961
38 Poole, A.C., Goodrich, J.K., Youngblut, N.D., Luque, G.G., Ruaud, A., Sutter, J.L. et al. (2019) Human salivary amylase gene copy number impacts oral and gut microbiomes. *Cell Host Microbe* **25**, 553–564.e557 https://doi.org/10.1016/j.chom.2019.03.001
39 Shebanits, K., Andersson-Assarsson, J.C., Larsson, I., Carlsson, L.M.S., Feuk, L. and Larhammar, D. (2018) Copy number of pancreatic polypeptide receptor gene NPY4R correlates with body mass index and waist circumference. *PLoS One* **13**, e0194668 https://doi.org/10.1371/journal.pone.0194668
40 Serra-Juhe, C., Martos-Moreno, G.A., Bou de Pieri, F., Flores, R., Gonzalez, J.R., Rodriguez-Santiago, B. et al. (2017) Novel genes involved in severe early-onset obesity revealed by rare copy number and sequence variants. *PLoS Genet* **13**, e1006657 https://doi.org/10.1371/journal.pgen.1006657
41 Karhunen, L.J., Juvonen, K.R., Huotari, A., Purhonen, A.K. and Herzig, K.H. (2008) Effect of protein, fat, carbohydrate and fibre on gastrointestinal peptide release in humans. *Regul. Pept.* **149**, 70–78 https://doi.org/10.1016/j.regpep.2007.10.008
42 Sha, B.Y., Yang, T.L., Zhao, L.J., Chen, X.D., 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 https://doi.org/10.1038/jhg.2009.10
43 Sun, C., Cao, M., Shi, J., Li, L., Miao, L., Hong, J. et al. (2013) Copy number variations of obesity relevant loci associated with body mass index in young Chinese. *Gene* **516**, 198–203 https://doi.org/10.1016/j.gene.2012.12.081

Biochemical Society Transactions (2020) **48** 1819–1828
https://doi.org/10.1042/BST20200556

44 Aerts, E., Beckers, S., Zegers, D., Van Hoorenbeeck, K., Massa, G., Verrijken, A. et al. (2016) CNV analysis and mutation screening indicate an important role for the NPY4R gene in human obesity. *Obesity (Silver Spring)* **24**, 970–976 https://doi.org/10.1002/oby.21435

45 Aerts, E., Geets, E., Sorber, L., Beckers, S., Verrijken, A., Massa, G. et al. (2018) Evaluation of a role for NPY and NPY2R in the pathogenesis of obesity by mutation and copy number variation analysis in obese children and adolescents. *Ann. Hum. Genet.* **82**, 1–10 https://doi.org/10.1111/ahg.12211

46 Robin, M.A., Anandatheerthavarada, H.K., Fang, J.K., Cudic, M., Otvos, L. and Avadhani, N.G. (2001) Mitochondrial targeted cytochrome P450 2E1 (P450 MT5) contains an intact N terminus and requires mitochondrial specific electron transfer proteins for activity. *J. Biol. Chem.* **276**, 24680–24689 https://doi.org/10.1074/jbc.M100363200

47 Porubsky, P.R., Meneely, K.M. and Scott, E.E. (2008) Structures of human cytochrome P-450 2E1. Insights into the binding of inhibitors and both small molecular weight and fatty acid substrates. *J. Biol. Chem.* **283**, 33698–33707 https://doi.org/10.1074/jbc.M805999200

48 Alberts, B., Johnson, A. and Lewis, J. (2002), The endoplasmic reticulum, In *Molecular Biology of the Cell* (Alberts, B., Johnson, A., Lewis, J., eds), pp. 669–694, Garland Science, New York, NY

49 Omura, T. and Ito, A. (1991) Biosynthesis and intracellular sorting of mitochondrial forms of cytochrome P450. *Methods Enzymol.* **206**, 75–81 https://doi.org/10.1016/0076-6879(91)06078-H

50 Knockaert, L., Fromenty, B. and Robin, M.A. (2011) Mechanisms of mitochondrial targeting of cytochrome P450 2E1: physiopathological role in liver injury and obesity. *FEBS J.* **278**, 4252–4260 https://doi.org/10.1111/j.1742-4658.2011.08357.x

51 Pfanner, N. (2000) Protein sorting: recognizing mitochondrial presequences. *Curr. Biol.* **10**, R412–R415 https://doi.org/10.1016/S0960-9822(00)00507-8

52 Lieber, C.S. (1997) Cytochrome P-4502E1: its physiological and pathological role. *Physiol. Rev.* **77**, 517–544 https://doi.org/10.1152/physrev.1997.77.2.517

53 Avadhani, N.G., Sangar, M.C., Bansal, S. and Bajpai, P. (2011) Bimodal targeting of cytochrome P450s to endoplasmic reticulum and mitochondria: the concept of chimeric signals. *FEBS J.* **278**, 4218–4229 https://doi.org/10.1111/j.1742-4658.2011.08356.x

54 Dixon, J.B., Bhathal, P.S. and O'Brien, P.E. (2001) Nonalcoholic fatty liver disease: predictors of nonalcoholic steatohepatitis and liver fibrosis in the severely obese. *Gastroenterology* **121**, 91–100 https://doi.org/10.1053/gast.2001.25540

55 Emery, M.G., Fisher, J.M., Chien, J.Y., Kharasch, E.D., Dellinger, E.P., Kowdley, K.V. et al. (2003) CYP2E1 activity before and after weight loss in morbidly obese subjects with nonalcoholic fatty liver disease. *Hepatology* **38**, 428–435 https://doi.org/10.1053/jhep.2003.50342

56 Angulo, P. (2002) Nonalcoholic fatty liver disease. *N. Engl. J. Med.* **346**, 1221–1231 https://doi.org/10.1056/NEJMra011775

57 Martis, S., Mei, H., Vijzelaar, R., Edelmann, L., Desnick, R.J. and Scott, S.A. (2013) Multi-ethnic cytochrome-P450 copy number profiling: novel pharmacogenetic alleles and mechanism of copy number variation formation. *Pharmacogenomics J.* **13**, 558–566 https://doi.org/10.1038/tpj.2012.48

58 Yang, T.L., Guo, Y., Shen, H., Li, J., Glessner, J.T., Qiu, C. et al. (2013) Copy number variation on chromosome 10q26.3 for obesity identified by a genome-wide study. *J. Clin. Endocrinol. Metab.* **98**, E191–E195 https://doi.org/10.1210/jc.2012-2751

59 Stahel, P., Xiao, C., Hegele, R.A. and Lewis, G.F. (2017) Polygenic risk for hypertriglyceridemia can mimic a major monogenic mutation. *Ann. Intern. Med.* **167**, 360–361 https://doi.org/10.7326/L17-0182

60 Scuteri, A., Sanna, S., Chen, W.M., Uda, M., Albai, G., Strait, J. et al. (2007) Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet.* **3**, e115 https://doi.org/10.1371/journal.pgen.0030115

61 Frayling, T.M., Timpson, N.J., Weedon, M.N., Zeggini, E., Freathy, R.M., Lindgren, C.M. et al. (2007) A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* **316**, 889–894 https://doi.org/10.1126/science.1141634

62 Claussnitzer, M., Hui, C.C. and Kellis, M. (2016) FTO obesity variant and adipocyte browning in humans. *N. Engl. J. Med.* **374**, 192–193 https://doi.org/10.1056/NEJMc1513316

63 Han, Z., Niu, T., Chang, J., Lei, X., Zhao, M., Wang, Q. et al. (2010) Crystal structure of the FTO protein reveals basis for its substrate specificity. *Nature* **464**, 1205–1209 https://doi.org/10.1038/nature08921

64 Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y. et al. (2011) N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* **7**, 885–887 https://doi.org/10.1038/nchembio.687

65 Smemo, S., Tena, J.J., Kim, K.H., Gamazon, E.R., Sakabe, N.J., Gomez-Marin, C. et al. (2014) Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature* **507**, 371–375 https://doi.org/10.1038/nature13138

66 Landgraf, K., Scholz, M., Kovacs, P., Kiess, W. and Korner, A. (2016) FTO obesity risk variants are linked to adipocyte IRX3 expression and BMI of children - relevance of FTO variants to defend body weight in lean children? *PLoS One* **11**, e0161739 https://doi.org/10.1371/journal.pone.0161739

67 Gonzalez-Herrera, L., Zavala-Castro, J., Ayala-Caceres, C., Perez-Mendoza, G., Lopez-Gonzalez, M.J., Pinto-Escalante, D. et al. (2019), Genetic variation of FTO: rs1421085T > C, rs8057044 G > A, rs9939609T > A, and copy number (CNV) in Mexican Mayan school-aged children with obesity/overweight and with normal weight. *Am. J. Hum. Biol.* **31**, e23192 https://doi.org/10.1002/ajhb.23192

68 Mejia-Benitez, A., Klunder-Klunder, M., Yengo, L., Meyre, D., Aradillas, C., Cruz, E. et al. (2013) Analysis of the contribution of FTO, NPC1, ENPP1, NEGR1, GNPDA2 and MC4R genes to obesity in Mexican children. *BMC Med. Genet.* **14**, 21 https://doi.org/10.1186/1471-2350-14-21

69 Yang, T.L., Guo, Y., Li, S.M., Li, S.K., Tian, Q., Liu, Y.J. et al. (2013) Ethnic differentiation of copy number variation on chromosome 16p12.3 for association with obesity phenotypes in European and Chinese populations. *Int. J. Obes. (Lond.)* **37**, 188–190 https://doi.org/10.1038/ijo.2012.31

70 Pettersson, M., Viljakainen, H., Loid, P., Mustila, T., Pekkinen, M., Armenio, M. et al. (2017) Copy number variants are enriched in individuals with early-onset obesity and highlight novel pathogenic pathways. *J. Clin. Endocrinol. Metab.* **102**, 3029–3039 https://doi.org/10.1210/jc.2017-00565

71 Selvanayagam, T., Walker, S., Gazzellone, M.J., Kellam, B., Cytrynbaum, C., Stavropoulos, D.J. (2018) Genome-wide copy number variation analysis identifies novel candidate loci associated with pediatric obesity. *Eur. J. Hum. Genet.* **26**, 1588–1596 https://doi.org/10.1038/s41431-018-0189-0