Review article

Genetics of non-syndromic childhood obesity and the use of high-throughput DNA sequencing technologies

Ana Carolina Proença da Fonseca a, Claudio Mastronardi b, Angad Johar c, Mauricio Arcos-Burgos b, Gilberto Paz-Filho c,⁎

a Human Genetic Laboratory, Oswaldo Cruz Institute/FIOCRUZ, Rio de Janeiro, Brazil
b Institute of Translational Medicine, Universidad del Rosario, Bogota, Colombia
c Department of Genome Sciences, John Curtin School of Medical Research, The Australian National University, Australia

ARTICLE INFO

Article history:
Received 23 January 2017
Received in revised form 18 April 2017
Accepted 19 April 2017
Available online xxxx

Keywords:
DNA sequencing
Gene
Genetics
Leptin
Melanocortin
Obesity

ABSTRACT

Background: Childhood obesity is a serious public health problem associated with the development of several chronic diseases, such as type 2 diabetes mellitus, dyslipidemia, and hypertension. The elevated prevalence of obesity is mostly due to inadequate diet and lifestyle, but it is also influenced by genetic factors.

Objectives: To review recent advances in the field of the genetics of obesity. We summarize the list of genes associated with the rare non-syndromic forms of obesity, and explain their function. Furthermore, we discuss the technologies that are available for the genetic diagnosis of obesity.

Results: Several studies reported that single gene variants cause Mendelian forms of obesity, determined by mutations of major effect in single genes. Rare, non-syndromic forms of obesity are a result of loss-of-function mutations in genes that act on the development and function of the hypothalamus or the leptin-melanocortin pathway. These variants disrupt enzymes and receptors that play a role in energy homeostasis, resulting in severe early-onset obesity and endocrine dysfunctions. Different approaches and technologies have been used to understand the genetic background of obesity. Currently, whole genome and whole exome sequencing are important diagnostic tools to identify new genes and variants associated with severe obesity, but other approaches are also useful at individual or population levels, such as linkage analysis, candidate gene sequencing, chromosomal microarray analysis, and genome-wide association studies.

Conclusions: The understanding of the genetic causes of obesity and the usefulness and limitations of the genetic diagnostic approaches can contribute to the development of new personalized therapeutic targets against obesity.

© 2017 The Authors. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Childhood obesity has been recognized as one of the most serious public health problems of the 21st century,¹ and is defined by the World Health Organization as an abnormal or excessive accumulation of body fat, sufficient to cause adverse health effects. The body mass index (BMI, calculated by dividing weight [kilograms] by the height squared [meters]) is a clinical standard measure of overweight and obesity and provides an estimate of adiposity in children and adults. In the measurement of children and adolescents, BMI requires to be adjusted for both age and gender.¹²

The prevalence of childhood obesity has increased at alarming rates, affecting populations from developed and developing countries.³⁻⁵ The problem is global and tends to increase with the income level of these countries.⁶ In 2013, the prevalence of overweight and obesity in developed countries was 23.8% for boys and 22.6% for girls; in developing countries, it was 12.9% for boys and 13.4% for girls.⁷ The prevalence of childhood obesity in some European countries and the United States (US) has apparently reached a plateau; nevertheless, the prevalence continues to increase in developing countries.⁸ Recent data estimate that there are more than 41 million children worldwide under 5 years of age who are overweight or obese. Among those overweight children, almost half of them live in Asia and one quarter in Africa.¹⁻⁵

Unfortunately, obese children have more than a 50% probability of carrying their adiposity into adulthood, while nonobese children have 10%.⁹ Furthermore, childhood obesity is a risk factor for the development of different comorbidities at a younger age which used to be considered as diseases seen only in the adult, including type 2 diabetes mellitus, dyslipidemia, hypertension, obstructive sleep apnea...
and steatohepatitis. They also have psychosocial and emotional well-being consequences, since those children are more likely to have anxiety, depression, poor self-esteem and become victims of bullying. The increased risk for developing obesity-related disease is influenced by environmental, behavioral, genetic and ecological factors. Common forms of obesity, which are present in most cases, are caused by a combination of environmental factors with many gene variants of minor effect. However, at least ~7% of cases, are caused by a combination of environmental factors with many gene variants of major effect (Table 1). In the large universe of obesity, that small proportion of non-syndromic excess body weight, several genomic and...
2. Genetic basis of obesity

Many different hypotheses have been proposed to explain the origin of the obesity epidemic. It is broadly accepted that the two major putative contributors are specific food marketing and consumption habits, and decrease in physical activity. However, there are others contributors, such as infections, maternal age, sleep debt, endocrine disruptors (such as industrial chemicals), reduction in ambient temperature variations, intrauterine and intergenerational endocrine disruptors (such as industrial chemicals), reduction in temperature, and drug-induced weight gain. Despite the importance of all these contributors, it is known that genetic factors have a key role in the risk of one becoming obese.

The percentage of obesity that can be attributed to genetics varies from 6 to 85%. Despite high estimates of heritability and decades of molecular genetic investigation, a very small fraction of genes associated with obesity have been identified. Twin studies have been used to understand the genetic component of obesity. Bouchard et al. through experiments of overfeeding and regular exercises, reported that monozygotic twins had more similar changes of body weight, body composition and energy expenditure when compared to dizygotic twins. Additionally, adoption studies showed a strong relationship between the BMI of biological parents and their adopted children; however no similarity was found with adoptive parents. These studies suggest that obesity and adiposity are inheritable factors.

2.1. Polygenic obesity

It is well-characterized that the obesity epidemic is not caused by single gene disorders, but actually has a complex genetic background. The patient with common obesity has many genetic susceptibility variants, which have a minor effect on body weight. The effect of genes on the pathogenesis of obesity is more frequently stronger when combined with other genes and the environment. The association of these genes with obesity has been evaluated through several approaches, such as linkage analysis, candidate gene sequencing and GWAS.

Over the past years, genetic epidemiological studies have identified several genes which play important biological roles in human obesity. Recently, Locke et al. carried out GWAS and Metaobachip meta-analysis for BMI. Data was analyzed from 339,224 individuals of 125 different studies. A total of 97 BMI-associated loci were identified, which of 56 were novel. Many of these loci were in/near genes which play a role in different biological processes, such as, for example, the neuronal development (FAIM2, PTBP2 and ST5), hypothalamic expression and regulatory function (TMEM18, SGC3 and GFP), limb development (RAR8, TFAP2B and LMX1B), lipid biosynthesis and metabolism (CYP27A1, CYP11A1 and NR1H3). Furthermore, some genes have previously been associated with severe early onset obesity (BDNF, MC4R, SH2B1, TUB and POMC). These 97 loci account for 2.7% of BMI variation and it was suggested that their common variants influence nearly 21% cases of obesity. Despite this robust study, most of the genetic variability in human obesity remains unknown.

2.2. Mendelian or monogenic forms of non-syndromic obesity

Over the last 15–20 years, it has been characterized that single gene mutations can cause severe early-onset obesity in humans. Assessment of such cases has already provided seminal information on genes involved in critical pathophysiological pathways that can lead to obesity. These genes encode enzymes and receptors that have a physiologic role in the development of the hypothalamus and the leptin-melanocortin system.

2.2.1. The hypothalamic development and the appetite control

The hypothalamus is the region of the brain involved in the regulation of the body’s temperature, fertility, thirst and appetite. The role of the hypothalamus in severe obesity has been postulated since the early twentieth century. Initially, chemical and electrical experiments that damaged the hypothalamus of cats and mice were correlated with an increase or loss of weight, depending on the size and location of the lesion. These experiments suggested the existence of a specific circuit in the hypothalamus that promoted or suppressed food intake. In humans, patients with tumors in the hypothalamic region are obese, suggesting an involvement of the hypothalamus in the regulation of body weight. Nowadays, it is known that the hypothalamus acts to maintain energy homeostasis, by receiving information about the body’s nutritional status and energy stores, through neural and hormonal signals.

The neuronal differentiation of the hypothalamus is controlled by several proteins, including brain-derived neurotrophic factor (BDNF), transcription factor single-minded 1 (SIM1) and the tyrosine kinase receptor (TRKB). Interestingly, pathogenic mutations which disrupt these protein functions can lead to severe onset-obesity in humans.

The development of obesity results from imbalance between energy intake and expenditure over a prolonged period. One of the most important pathways responsible for controlling food intake and body weight is the leptin-melanocortin system, located in the hypothalamus. This system is an important pathway to energy balance that integrates information from energy stores and nutritional status in peripheral organs with a neural circuit.

This pathway is activated when leptin, produced by adipocytes, crosses the blood-brain barrier and binds to its receptors, localized in two different groups of neurons in the arcuate nucleus of the hypothalamus. Both groups of neurons act on the regulation of food ingestion and energy expenditure, which are important for energy homeostasis. One of these groups expresses neuropeptide Y (NPY) and agouti-related protein (AgRP) that stimulates food intake (orexigenic pathway), while the other expresses proopiocortin (POMC) and cocaine and amphetamine-regulated transcript (CART) which inhibits food consumption (anorexigenic pathway).

The brain can control energy homeostasis since it receives information about the body’s nutritional status and energy stores by peripheral hormones. When the body has food restrictions, it is fasting and/or has a decreased energy stock, the stomach secretes ghrelin that binds to the ghrelin receptor, localized in the orexigenic neurons, increasing the production of NPY and AgRP. The orexigenic neurons project to the paraventricular nucleus (PVN) of the hypothalamus and their neuropeptides act as antagonists of melanocortin receptors, which stimulate food intake and decrease the basal energy expenditure. However, when the body has adiposity and/or energy replacement, the white adipose tissue secretes leptin that binds to leptin receptor (LEPR), localized in the orexigenic neurons, suppressing this pathway. Furthermore, leptin binds to its receptor on anorexigenic neurons, increasing the production of POMC and CART. The POMC hormone is a protein precursor which is processed and cleaved into α melanocyte stimulating hormone (α-MSH) by proprotein convertase 1/3 (PC1/3) and binds to the melanocortin-4.
receptor (MC4R), located in the paraventricular nucleus (PVN). Activation of MC4R by α-MSH and CART increases a satiety signal, suppressing appetite and basal energy expenditure.²⁷–³³ Interestingly, mutations in genes involved in the leptin–melanocortin pathway demonstrate the critical role of energy balance and have been associated with development of severe early-onset obesity.²⁵,⁷⁹,⁸⁰ Fig. 1 summarizes the hypothalamic mechanisms involved in appetite control.

### 2.2.2. Recessive forms of non-syndromic monogenic obesity

The first evidence of monogenic obesity was found in rodents in the early 1950s. It was reported that ob/ob mutant mice have four times higher weight than normal animals.⁶⁷ Through further studies, Coleman⁷⁷ discovered that ob/ob mice have a deficiency of the signaling pathway from adipose tissue, which regulates food intake and metabolism. After positional cloning and sequencing, the leptin gene (LEP) was identified and characterized.⁷⁸ In humans, pathogenic mutations in the LEP gene are associated with low levels of leptin, hyperphagia and severe early-onset obesity.²⁵,⁷⁹,⁸⁰ For over 15 years, researchers have been evaluating phenotypic findings and the effects of leptin replacement in less than a dozen of the ≤ 10 leptin-deficient individuals identified in the world. Our group has been following one child and three adults from a consanguineous extended Turkish family (summarized in ²⁴,⁸¹,⁸²). These individuals have a Mendelian recessive mutation in the LEP gene, consisting of a C → T substitution in codon 105 of this gene, resulting in an Arg → Trp replacement in the mature protein. Three of them are the only leptin-deficient patients identified during adulthood in the world to date. Physiological doses of recombinant methionyl human leptin (r-metHuLeptin, metreleptin, Aegerion Pharmaceuticals, Inc., USA, 0.02–0.04 mg/kg/day) were started at ages 5 (boy), 27 (man), 30 and 40 (women), leading to significant improvements in body weight, endocrine function and behavior.⁸³ Leptin replacement was lifesaving, as eight members of this family with NEOSO, whom we presume to have been leptin-deficient, died of immune deficits in childhood.

The study of recessive forms of monogenic extreme obesity has been useful in delineating the role of genes from the leptin–melanocortin pathway. However, these recessive forms of obesity due to functional mutations in the LEP are very rare, and mostly identified in families with high levels of consanguinity. To date, ≤ 40 individuals have been found to be homozygous for each of these mutations. However, a systematic effort to sequence LEP in 73 individuals with NEOSO revealed that 22% were homozygous for LEP mutations, resulting in impaired leptin signaling.⁸⁰ This study showed that a systematic genetic examination of early-onset severe obesity, even by focusing on a few genes, could reveal a meaningful percentage of monogenic cases of obesity. It is important to note that, although most patients with the leptin-deficient phenotype have low or undetectable serum leptin levels, one patient had detectable, but biologically inactive serum leptin.⁸⁴

Besides LEP, homozygous loss of function mutation of its receptor LEPR is associated with severe hyperphagia and early-onset extreme obesity in humans. Initially, homozygous mutations in LEPR were identified in patients characterized by hyperphagia, early-onset morbid obesity, failure of pubertal development and impairment in secretion of growth hormone and thyrotropin. These patients had a base substitution in the splice donor site of exon 16, which results from skipping this exon and generates a truncated receptor that lacks both the transmembrane and intracellular domains.⁶⁰ Afterwards, Farooqi et al.²⁸ reported eight severely obese subjects that had nonsense or missense LEPR mutations. The nonsense mutations were predicted to result in a lack of all receptor isoforms, and the missense variants impaired receptor signaling. Recently, two novel frameshift mutations were identified in severely obese children by whole-exome sequencing. Both mutations generate an abnormal protein that lacks the leptin binding receptor domain (CRH2), affecting all isoforms.⁸⁵ On the whole, congenital deficiency of LEP or LEPR disrupts leptin signaling and is characterized by similar clinical data.²⁸

POMC is a complex propeptide expressed in the hypothalamus, pituitary gland and brainstem that is activated by tissue-specific post-translational processing.⁶² POMC generates a family of melanocortin peptides, for example the adrenocorticotropic hormone (ACTH) and α-, β- and γ-melanocyte-stimulating hormones (MSH). These peptides have an important role on skin pigmentation, the control of...
adrenal growth and energy balance. Different mutations that inactivate the POMC gene have been reported in children, leading to severe early-onset obesity, adrenal insufficiency and pigmentation abnormalities. Krude et al. reported the first human description of congenital deficiency of POMC. They identified two unrelated children who had severe childhood obesity, adrenal hypoplasia, red hair and pale skin. One patient had a compound heterozygous variant for two nonsense mutations and the second had a homozygous variant for the new initiation site of translation, localized in the 5-untranslated region. Although POMC mutations could result in a pigment abnormality, this characteristic was absent in different case reports. For example, Mendiratta et al. identified an 18-month-old female with congenital adrenal insufficiency caused by a homozygous C → A substitution in codon 231 of POMC gene, which was predicted to result in a premature stop codon. The clinical phenotype of the patient was similar to previous reports, however the child had dark colored hair and dark roots. More recently, another interesting variation in the classical phenotype of POMC deficiency was described by Samuels et al. They reported a 4-year-old girl and a 4-month-old boy who had high circulation levels of ACTH caused by variants that generate non-functional peptides, but immunologically detectable.

In the last 20 years, researchers have identified variants in POMC that cause congenital deficiency; however, only the adrenal insufficiency phenotype is treatable, by using hydrocortisone. This year, Kühnen et al. reported two obese POMC-deficient patients that were treated with daily subcutaneous synthetic melanocortin (setmelanotide), resulting in body fat loss, decreased appetite and darkening of the skin and hair. Setmelanotide, also known as RM-493, binds to the MC4R and activates the anorexigenic pathway, substituting the lack of MSH in those patients. Although this study shows a limitation in the number of patients, setmelanotide may offer a potential treatment for POMC deficiency in the future. Furthermore, they speculated that synthetic melanocortin might be useful for treating LEPR and PCSK1 deficiency patients.

In addition, null mutations in proprotein convertase subtilisin/kexin type 1 gene (PCSK1) cause a rare non-syndromic form of obesity. PCSK1 encodes the PC1/3 that is expressed in neural and endocrine tissues. In the brain, PC1/3 is highly expressed in the hypothalamus where it acts on cleavage processing of POMC to α-MSH. Several cases of PC1/3 deficiency were reported with hyperphagia, central diabetes insipidus, severe malabsorptive diarrhea and other endocrinopathies. This deficiency could be explained by mutations that affect protein maturation and generate impairment in the catalytic activity, resulting in a truncated protein or lack of protein. Although PC1/3 deficiency is described as an autosomal recessive disorder, the severe obesity phenotype has been associated with loss of one wild-allele. Philippe et al. identified a heterozygous non-synonymous mutation that is co-segregating with obesity in three generations of a French family. This variant, located in the second exon of PCSK1, affects both mutant and wild-type proteins activity in vitro. This autosomal dominant inheritance of obesity may be explained by the truncated propeptide which also interferes with the function of the wild-type protein.

Recently, Borman et al. identified a homozygous frameshift mutation in tubby bipartite transcription factor gene (TUB) which was associated with early-onset obesity as well as retinal dystrophy in humans. This variant is found in a conserved region of TUB family resulting in a truncated protein, which has an incorrect localization in the cell. Moreover, loss of function mutation in the TUB gene also induces obesity in mice. The role of TUB in energy balance has not been totally elucidated; however it is expressed in different regions of the hypothalamus, including the arcuate nuclei and paraventricular nuclei. Additionally, experimental data suggests that the TUB protein acts on modulating anorexigenic neuropeptides, and low levels of TUB expression result in increased food intake and adiposity. 2-5

2.2.3. Dominant forms of non-syndromic monogenic obesity

The progress in understanding the obesity background is mostly due to the identification of rare genetic forms. Although several gene variants have been identified to cause recessive forms of obesity, heterozygosity for deleterious mutations in BDNF, NTRK2, SIM1, MC4R, SH2B1, MRA2 and LRP2 have also been associated with severe early-onset obesity.

BDNF, NTRK2 and SIM1 are responsible for coding important proteins to the development of the hypothalamus. BDNF and its receptor TrkB (encoded by NTRK2) are involved in the proliferation, survival, and differentiation of neurons in the central nervous system, especially the hypothalamic neurons. SIM1 is a transcription factor that plays a major role in the development and function of the paraventricular nucleus, a critical region for food intake regulation. In humans, pathogenic mutations in those genes are associated with hyperphagia and severe obesity phenotype. A de novo heterozygous missense mutation (Y722C) in NTRK2 was found in an eight-year-old male patient with severe obesity, nociception, impaired short-term memory and learning. Functional analysis demonstrated that Y722C leads to TrkB dysfunction. Furthermore, a de novo paracentric inversion in chromosome 11 (46, XX, inv (11) (p13p15.3)), which disrupts one copy of the BDNF gene, was identified in a child with severe obesity, impaired cognitive function and hyperactivity.

Obesity also has been associated with SIM1 haploinsufficiency in humans. Initially, a female patient was identified with severe early-onset obesity that had a balanced translocation (46, XX, t(1:6)(p22.1;q16.2)), leading a disruption of SIM1 in one allele. In addition, rare missense mutations in SIM1 have been associated with obesity. Recently, the coding region of SIM1 was sequenced in 2100 severe early-onset obesity probands and in 1680 controls, identifying 13 heterozygous variants in 28 patients, in which none of these mutations would reduce the activity of SIM1 in vitro. Among these patients with pathogenic variants, 11 individuals had evidence of cognitive deficit including impaired concentration, emotional lability and memory deficit, in addition to severe early-onset obesity. A similar analysis was performed in 561 overweight/obese subjects, and identified four heterozygous rare non-synonymous variants in SIM1 gene, in which two of these mutations caused impairment in the SIM1 function. Therefore, BDNF, NTRK2 and SIM1 haploinsufficiency are associated with severe early-onset obesity as well as cognitive impairment in humans.

MC4R is the receptor for α-MSH and, as such, plays a key role in energy homeostasis, food intake and body weight, since it regulates satiety. Consistent with this role, loss of function MC4R mutations are associated with increased appetite that leads to severe early-onset obesity in humans. Nowadays, more than 150 variants are known, and the prevalence of deleterious MC4R variants ranges from 0.5% up to 5% in obese children. MC4R mutations are the most common genetic cause of monogenic obesity and also contribute to polygenic forms. The rare cases of MC4R deficiency are inherited in a co-dominant form. Several cases of MC4R deficiency described patients with early-onset obesity, hyperphagia, severe hyperinsulinemia and increased linear growth, although the endocrine and anthropometric phenotypes were not confirmed in other reports. Recently, the MC4R gene was sequenced in a large cohort of obese children and adolescents from Germany. MC4R mutations were detected in 22 probands, in which fourteen patients had variants causing loss of function or impaired the receptor signaling. Since pharmacological drugs are being developed to treat MC4R-deficient patients, it is extremely important to identify pathogenic mutations.

It has been recognized that accessory proteins can regulate melanocortin receptors such as MC4R. Melanocortin 2 receptor accessory protein 2 (MRAP2) is expressed in regions involved in energy balance, such as the hypothalamus and brainstem.
Furthermore, MRAP2 has been shown to interact with MC4R; its mRNAs is coexpressed in the brain, suggesting an involvement of MRAP2 in the energy homeostasis. To explore if MRAP2 mutations have an important role in human obesity, the coding region of this gene was screened in 483 children and adolescents with extreme obesity and 630 individuals with normal weight. Three missense mutations were detected among obese subjects, all in heterozygosis. In vitro functional analyses revealed that one of nonsynonymous variant (p.Gln174Arg) causes a partial MC4R loss of function. Additionally, four heterozygous mutations were identified in patients with severe early-onset obesity, suggesting that MRAP2 variants may contribute to rare cases of monogenic obesity.

Src homology 2 B adapter protein 1 (SH2B1) encodes a cytoplasmatic adaptor protein, which modulates leptin-melanocortin signaling. Heterozygous frameshift and missense variants in SH2B1 were associated with obesity, insulin resistance and behavioral abnormalities, including a tendency for social isolation and aggressive behavior. Furthermore, severe early-onset obesity and severe insulin resistance phenotypes were reported in a patient with a 220-kb segment deletion of chromosome 16p11.2, where several genes are located, including SH2B1. Several SH2B1 variants in heterozygosis have been reported and one in homozygosis. All variants except one co-segregated with obesity in families in a classical Mendelian manner. Additionally, kinase suppressor of ras 2 (KSR2), which acts as a scaffolding protein, was previously associated with obesity. KSR2 plays a role in energy homeostasis, insulin sensitivity and cellular fuel oxidation. Early studies demonstrated that deletion of Ksr2 in mice is associated with obesity and insulin resistance. Furthermore, KSR2 is located in a chromosome region linked to obesity and type 2 diabetes. To examine whether KSR2 is involved in human obesity, the coding region and intron/exon junctions of KSR2 were screened in 2101 patients with severe early-onset obesity and 1536 normal weight controls. A total of 31 different rare frameshift, nonsense or missense mutations were identified, 27 in severely obese subjects and 7 in controls. Most of the variants impaired the KSR2 function and were found in a heterozygous manner. Curiously, some loss of function mutations in KSR2 were not co-segregated with severe obesity, suggesting that other factors may modulate the phenotype caused by some variants.

More recently, we added LRP2 to the list of genes associated with dominant monogenic obesity. By submitting three Brazilian patient-family trios to WEC and parallel sequencing studies, we identified one male teenager with compound heterozygous mutations in the LRP2 gene (confirmed by Sanger sequencing). LRP2 binds to the long-form leptin receptor (Leprb), forming a complex that is co-localized and subjected to endocytosis in the hypothalamic neurons. Subsequently, the endocytosis of this co-localized complex leads to the activation of signal transducer and activator of transcription 3 (STAT3) signaling in hypothalamic neurons. As a consequence, food intake and body weight are decreased. In the absence of functional LRP2, STAT3 signaling is decreased. Therefore, hunger is stimulated and satiety decreased, leading to obesity.

3. Approaches for the genetic diagnosis of obesity

3.1. Linkage analysis

Linkage analysis is the most used method for the genetic mapping of Mendelian traits with familial aggregation. This genetic approach accounts for the transmission pattern from parents to offsprings, as well as the respective affection status of the family members, along with their allele frequencies for variants/loci under investigation. This quantification process also considers instances of recombination among loci being tested. All of this information is used to generate a likelihood ratio under the alternate hypothesis (i.e. evidence of linkage or co-segregation between variants and the disease). This is conditional on the premise that the recombination fraction is <0.5, and rejection of the null hypothesis (no linkage/co-segregation). Therefore, two linked loci yield a recombination frequency that is less than 0.5 according to the laws of segregation and independent assortment hypotheses, represented by a logarithm of the odds (LOD) score. It is used to identify loci which co-segregate with a disease or specific phenotype within related individuals. It is noteworthy that the discovery of the leptin gene by positional cloning used the concept of linkage in its methodological approach.

The availability of nuclear families (including parents and offspring) or extended families displaying the obesity phenotype would constitute a suitable paradigm to perform linkage analyses. It is also noteworthy to remark, that in this type of approach, genetic homogeneity is maximised, which also renders these analyses very powerful. This is the case even for small numbers of families, provided that cohorts are selected from genetically homogeneous populations. On the other hand, association studies are only able to detect allele-frequency differences between cases and controls of lower phenotypic effect and that requires extremely large sample sizes. Thus, the fact that the LOD score calculation accounts for many of these factors, and not just allele frequency differences between cases and controls, it renders a much more powerful outcome than GWAS and/or other association analyses. This would be especially the case if patients are carrying extreme phenotypic forms of the disease, when trying to identify rare variants of major effect. GWAS, on the other hand often relies on common-disease common-variant hypothesis, so at best they can only detect variants of small effect. In order to detect rare-variants of major effect in association studies, very large sample sizes are needed. The latter can be addressed with algorithms that use rare-variant collapsing/grouping strategies at the gene level (see below), but they do not have the advantages of mathematically incorporating information of familial segregation patterns, as mentioned earlier for linkage analyses.

Finally, new linkage algorithms have emerged that are capable of identifying de novo mutations as candidate disease variants. Thus, with all of these advantages, there have been new strategies developed to combine the strengths of linkage and association analyses. While this has the advantages of evaluating familial co-segregation, it will also ensure that the variants under investigation are not overrepresented in sets of unaffected controls, which may be unrelated to the families tested for linkage. In consideration of all of this, one has to state then, that linkage analyses are at least as powerful as association studies.

3.2. Candidate gene sequencing

This approach aims to study genetic markers in candidate genes for a specific characteristic. These genes are previously selected by evidence of involvement in relevant pathways to the phenotype studied as well as earlier results that showed a possible association of the gene and the disease. The candidate variants of this gene are then submitted to Sanger sequencing for validation, and their sequences are compared against the reference human genome. Sanger sequencing has been the gold-standard for validating genomic findings for over 2 decades. In well-selected cases, sequencing can lead to a genetic diagnosis in up to 46% of patients. Currently, next generation sequencing is a very cost-effective first line approach, even when there is high suspicion for a specific gene. Sanger sequencing is then employed to validate the next generation sequencing findings.
3.3. Chromosomal microarray analysis

Different studies are focused on identifying copy number variations (duplications or deletions or chromosomal segments; CNV) to increase the knowledge on obesity susceptibility. Previous studies reported that individuals with severe early-onset obesity have an enrichment of CNVs, suggesting that variation may represent a risk for this disease.\textsuperscript{127} Considering that fact, the chromosomal microarray analysis has been a useful tool to search for CNVs in the genome.\textsuperscript{128}

Chromosomal microarray is particularly useful for identifying microdeletions, microduplications, and most of abnormalities regarding chromosomal number. It can also, although with less power, identify most of the non-balanced rearrangements, imprinting, triploidy, and punctual mutations. However, it cannot identify larger deletions and mutations, such as fragile X syndrome, and balanced rearrangements, such as inversions and balanced translocations.

3.4. GWAS

The new age of technology has been allowing the identification of common variants associated with complex multigenic disorders. In GWAS, a huge number of DNA markers across the genome is scanned, and through statistical analysis, it is possible to identify the loci associated with disease or specific phenotype.\textsuperscript{69} Until now, GWA methodologies have been found in 119 common gene variants as being linked to obesity susceptibility or related traits.\textsuperscript{66} The fat mass and obesity-associated gene (FTO) was considered the first locus found in GWAS of obesity. Furthermore, this result was replicated in different studies including cohorts of European and African descent. These researchers identified a cluster of common variants in the first intron of FTO that is associated with increased BMI. Moreover, these common polymorphisms were associated with body weight and body fat distribution.\textsuperscript{129–131}

More recently, a GWAS meta-analysis study identified 97 loci linked to obesity explaining 2.7% of BMI variation. Although GWAS have been an efficient and robust method to identify regions that are predisposed to obesity, the results are not strong enough to explain the estimated heritability of this disease.\textsuperscript{22} Furthermore, GWAS are a useful approach for identifying common variants and loci associated with polygenic obesity at a population level, but not applicable in the diagnosis of Mendelian form at the individual level. Thus, more molecular studies and new methodologies are required to understand the genetics of obesity.

3.5. Next generation sequencing

The comprehension of the genetic background of obesity has required different molecular strategies, since it is not completely elucidated yet. Recently, a more precise and fast screening approach has been developed and applied successfully in the diagnosis of monogenic forms of obesity.\textsuperscript{59,132} Next generation sequencing (whole-genome and whole-exome) allows the identification of mutations that might lead to personalized medicine resulting in a better life for the patient.\textsuperscript{99} One of the most common examples is the severely obese patients carrying loss of function mutations in the LEP gene, who can be treated with recombinant leptin.\textsuperscript{11,133}

Next generation sequencing (NGS), particularly WEC, has been used in the clinics for the diagnosis of diseases such as neurodevelopmental and autoimmune disorders.\textsuperscript{134,135} It can provide a diagnosis in 22%–46% of cases, and is indicated by the American College of Medical Genetics and Genomics (ACMG) in case the phenotype and family history strongly suggest an unspecific genetic disease, when there is significant genetic heterogeneity, and when other specific tests are negative.\textsuperscript{136} Sequencing of trios (both biological parents, besides the proband) leads to higher chance of discovery of significant mutations (30%, vs 23% when just the proband is sequenced). Besides, it improves the detection of de novo mutations and of compound heterozygotes.\textsuperscript{137} When siblings are sequenced, the chance of discovery of recessive mutations is increased; singleton sequencing leads to higher diagnostic rates in children (23%, vs. 18% in adults).\textsuperscript{138}

Currently, whole-exome sequencing has become more commonly used to identify variants, since it is more affordable than the whole-genome sequencing, and it covers 95% of the protein-coding genome, comprised by ~20,000 genes. The application of whole-exome sequencing in the obesity field has recently begun, but it already helped to identify novel genes associated with severe early-onset obesity not detected from previous approaches (KSR2, SH2B1 and LRPR).\textsuperscript{32,40,46,118} Additionally, Saeed et al.\textsuperscript{80} analyzed 73 children with severe obesity from Pakistan and identified loss of function homozygous variants in LEP, LEPR, MC4R, NTRK2, PSCK1, POMC and SIM1. They found two probands with severe early-onset obesity and high leptin levels which have two novel LEPR mutations in homozygosity.\textsuperscript{27} Therefore, both whole-exome and whole-genome sequencing are important tools for genetic diagnosis of rare obesity forms.

One of the main challenges of WEC and WGS is the bioinformatics analysis of the massive amount of data that is generated. The employed analysis framework greatly varies among bioinformaticians, which can affect the ultimate result. Briefly, the resulting data is submitted for filtering using a database (usually, dbsNP), to remove common polymorphisms. Subsequently, variants are annotated by bioinformatics software, such as SIFT, PolyPhen, ConDEL, to highlight variants that affect coded proteins. Pathway analysis is performed, to assess the role of the variant in physiological pathways, and functional validation is conducted.

Technical limitations of WEC include the fact that 5% of the exome is not covered, and some mutations may be missed. Furthermore, WEC only covers protein coding regions which is less than 2% of the genome. Also, some areas composed by CG repeats are difficult to sequence, small insertions/deletions (indels) provide inaccurate data, and extended genes and pseudogenes are difficult to map. Extended structural variations are better detected by WGS, and CNVs are better assessed by chromosomal microarray. In NGS, often several potential pathogenic variants are identified, which clinical significance is uncertain – those are called “variants of uncertain significance”. It is up to the clinical geneticist to assess whether variants identified by NGS are associated with the phenotype.

Of clinical relevance in the USA, the ACMG determines that laboratories conducting NGS must analyze and report mutations in 56 genes that have been associated with 24 Mendelian disorders. Besides reporting mutations that are known to cause the disease that led the patient to seek genetic diagnosis, laboratories may also report incidental or secondary results, that are not related to the main reason of genetic testing. To minimize distress to the patient and their families, it is recommended that patients attend pre-test counseling, when they are informed about the possibility of incidental findings, and choose whether they want to be informed or not (“opt-out”).

Negative results from the NGS do not exclude disease. Since those approaches and tools are in constant evolution, data can be reanalyzed after 6–12 months, when new genes, bioinformatics algorithms, and case reports are made available. If the WEC provides negative results, another alternative is to perform WGS, which more uniformly covers coding regions not well covered by WEC, and is able to detect structural rearrangements and CNVs (unlike WEC).\textsuperscript{139}

3.5.1. RVAT (rare-variant association tests)

While analyses from GWAS have successfully identified disease-causing variants in the past, they rely on the common disease, common variant hypothesis. Therefore, such analyses are not suitable
for rare variant studies, which could lead to missing heritability, especially given the hypothesis which states that disease causing alleles of large effect are likely to be of low frequency. Hence, in recent years, novel methods to perform rare-variant association tests (RVATs) have been developed. The most notable of these is the KBAC or Kernel Based Adaptive Cluster. Each rare variant within a particular gene is collapsed into multi-site genotypes. Thus, in brief, multi-site genotype frequency differences between cases and controls are calculated, as opposed to quantifying single variants only. This gene-based association test ensures that rare variants of potentially major effect are not completely discounted due to small individual effect size.140

4. Network analysis of genes underpinning rare non-syndromic childhood forms of obesity

In order to better understand the potential mechanistic role of LRP2 in morbid obesity, as previously described by our group, we ran analyses of pathways and network building algorithms from the Metacore GeneGo Software Suite from Thomson Reuters (version 6.29, build 68,613). Further details of the statistical analysis implemented in the various network-building algorithms can be found in supplementary materials and elsewhere.135 These procedures allowed for the heuristic integration of maps and networks and corresponding ontologies for obesity related processes in order to quantitatively and qualitatively evaluate the biological role of corresponding ontologies for obesity related processes in order to quantitatively and qualitatively evaluate the biological role of LRP2 with and 5.524 × 10−70 respectively (Table 2). Therefore, the probability that these nodes within the network (including proteins encoded by LRP2 as well as LEPR and its ligand are simultaneously linked (shown by these algorithms) to regulation of lipid storage and carbohydrate metabolism, means that they may have some potential pathway and network interactions in the biological processes and mechanisms of morbid obesity (Fig. 2). The main processes in which LRP2 was found associated to a subset of the genes causing monogenic forms of obesity are nutrient sensing processes and carbohydrate and lipid metabolism (Table 2). In Table 2, it is shown that there is evidence for the relatedness of LRP2 with monogenic genes causing extreme forms of morbid obesity as quantified by the following statistical analyses described below. Firstly, the z-score of 143.67 of the constructed network yielded a P-value of 8.48 × 10−70. Thus, this indicates that there is a significant degree of enrichment of the network with genes from the candidate input list. Hence, this network is potentially a highly accurate representation of the potential biological links and interactions of a subset of these candidate genes out of all those analyzed. The P-values for these three particular processes are 1.011 × 10−12, 1.106 × 10−12 and 5.524 × 10−70 respectively (Table 2). Therefore, the probability that these nodes within the network (including proteins encoded by LRP2 as well as LEPR, LEPR and POMC) are associated with these processes simply due to chance (according to the hypergeometric distribution calculation under the null hypothesis) is considerably low. Thus, based on this evidence, one can postulate that a mutation in the LRP2 gene can alter those mentioned processes and pathways, which may have substantial implications in morbid obesity.

5. Genetic diagnosis applied for patient suspected of NEOSO

Monogenic obesity syndromes and nonsyndromes are rare and mainly result in severe early onset-obesity. These disorders may occur due to chromosomal abnormalities and/or highly penetrant genetic variants in critical genes for regulation of energy intake and expenditure.135 Recently, a clinical practice guideline was published suggesting how to proceed with patients suspected of Mendelian forms of obesity.142 Specialists have suggested that all assessment of

### Table 2

<table>
<thead>
<tr>
<th>Gene</th>
<th>Process ontology</th>
<th>Network nodes</th>
<th>Ontology P-value</th>
<th>Network z-score</th>
<th>Network P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRP2</td>
<td>Regulation of lipid metabolism</td>
<td>Leptin binding to lepton receptor (activation), Receptor binds and activates STAT5B</td>
<td>1.725e−7</td>
<td>143.67</td>
<td>8.48e−70</td>
</tr>
<tr>
<td>LEP</td>
<td>Response to nutrient levels</td>
<td>Leptin binding to lepton receptor (activation), Casein kinase II phosphorylate LRP2 (activation), LRP2 inhibits cytochrome c by binding</td>
<td>1.011e−12</td>
<td>143.67</td>
<td>8.48e−70</td>
</tr>
<tr>
<td>LRP2</td>
<td>Regulation of carbohydrate metabolism</td>
<td>Leptin binding to lepton receptor (activation), Casein kinase II phosphorylate LRP2 (activation), LRP2 inhibits cytochrome c by binding</td>
<td>1.106e−3</td>
<td>143.67</td>
<td>8.48e−70</td>
</tr>
<tr>
<td>LRP2</td>
<td>Regulation of protein metabolic process</td>
<td>Leptin binding to lepton receptor (activation), Casein kinase II phosphorylate LRP2 (activation), LRP2 inhibits cytochrome c by binding</td>
<td>2.330e−30</td>
<td>143.67</td>
<td>8.48e−70</td>
</tr>
<tr>
<td>LEP</td>
<td>Negative regulation of proteolysis</td>
<td>Leptin binding to lepton receptor (activation), LRP2 inhibits cytochrome c by binding</td>
<td>8.48e−70</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

P-values for process ontologies from the GeneGo database are given for the corresponding nodes in the network. Degree of enrichment for constructed network is calculated by z-score, based on number of genes from the input list. Significance of enrichment is given in Network P-value.
children and adults with severe obesity should be careful, since it may help to guide the disorder diagnosis.

Based on the practice guideline, physical examination and medical history about the patient is required to establish the age of obesity onset, presence of hyperphagia, developmental delay and/or dysmorphic features. In addition, family history should be obtained to identify the consanguineous relationships, other family cases with severe obesity/bariatric surgery and the ethnic and geographical origin of the family.

Abnormalities in the medical history and/or the physical examination may be a potential endocrine, genetic or hypothalamic obesity disorders. Patients with severe early onset obesity (before 5 years age), neurodevelopmental abnormalities and/or severe hyperphagia are susceptible to the Mendelian form of obesity and genetic testing should be performed. When these patients have a development delay, karyotype and DNA methylation studies are a useful strategy for identifying Prader–Willi Syndrome. Patients with negative results and additional features (retinal dystrophy, photophobia or nystagmus) could indicate Bardet–Biedl syndrome, Alström syndrome or Tub deficiency. However, patients with negative results and without these additional features could be tested for Albright Hereditary Osteodystrophy, BDNF, TrkB e SIM1 deficiency.

Additionally, severe obesity patients without development delays should have leptin, proinsulin and insulin levels assessed. When these subjects have decreased or undetectable leptin levels, it might indicate LEP deficiency. Nevertheless, elevated ratio of proinsulin/insulin (‘mature’) levels suggest PCSK1 deficiency. Finally, patients with normal levels of those hormones may be genetically tested to different deficiencies, such as LEP, LEPR, POMC, MC4R, SH2B1 and KSR2.

Importantly, consideration should be given to the ethical aspects of genetic testing and its implications. Pretest counseling is warranted, to establish realistic expectations for unveiling a causative gene variant, and to make patients and families aware that a positive result is unlikely to change the clinical management, given the lack of available targeted therapies. Also, the discovery of incidental findings unrelated to the reason for testing might create additional stress to the patient, and should be clearly disclosed. Furthermore, the implications of a positive finding to biological relatives should be discussed with the patient, as well as the fact that the interpretation of a positive result should be correlated with clinical findings. Finally, although genetic testing costs, particularly of NGS, are decreasing, widespread use of these diagnostic tools still adds financial burden to patients and the healthcare system.

6. Current clinical approach in non-syndromic monogenic obesity

Promising approaches based on high-throughput next-generation sequencing technologies have the potential to guide a robust genetic diagnosis and identify new genes and pathways in the future. Molecular diagnosis of monogenic obesity is important to the patients and their families, since the clinicians could manage the patient appropriately and provide the genetic counseling. The management of these patients based on lifestyle and behavioral modification appears to not be so efficient long-term, since they are relatively resistant to losing and/or maintaining the weight loss. Bariatric surgery is an effective therapy for severe obesity; however, the efficacy in patients with monogenic forms of obesity has extremely limited data and is still up for debate. Adult patients with heterozygous MC4R mutations who were submitted to Roux-en-Y gastric bypass had an effect of excess weight loss compared to a cohort of controls without MC4R mutations. Those patients were able to lose about 60% of excess weight after the surgery. In addition, four adolescents with heterozygous MC4R mutations were evaluated in a restrictive bariatric surgery, in which three underwent laparoscopic adjustable gastric banding and one was submitted to vertical sleeve gastrectomy. The four patients and their
Bonnefond et al.149 found that patients carrying gain of function MC4R mutations had a higher risk of reoperation. Therefore, more studies are necessary.

In most cases, the individual variability of becoming obese is affected by them. LEP: leptin; LEPR: leptin receptor; PCSK1: proprotein convertase subtilisin/kexin type 1; POMC: pro-opiomelanocortin; TUB: tubby bipartite transcription factor; SH2B1: Src homology 2 B adaptor protein 1; MC4R: melanocortin-4 receptor; MRAP2: melanocortin 2 receptor accessory proteins 2; KSR2: kinase suppressor of ras 2; SIM1: single-minded homolog 1; BDNF: brain-derived neurotrophic factor; NTRK2: neurotrophic tyrosine kinase receptor type 2; LRP2: low-density lipoprotein receptor 2; FFA: free fatty acids.

Fig. 3. Sites of involvement of the various genes associated with obesity and the physiological pathways affected by them. Schematic representation of the sites of action of the proteins expressed by the genes associated with obesity, and the pathways affected by them. LEP: leptin; LEPR: leptin receptor; PCSK1: proprotein convertase subtilisin/kexin type 1; POMC: pro-opiomelanocortin; TUB: tubby bipartite transcription factor; SH2B1: Src homology 2 B adaptor protein 1; MC4R: melanocortin-4 receptor; MRAP2: melanocortin 2 receptor accessory proteins 2; KSR2: kinase suppressor of ras 2; SIM1: single-minded homolog 1; BDNF: brain-derived neurotrophic factor; NTRK2: neurotrophic tyrosine kinase receptor type 2; LRAP2: low-density lipoprotein receptor 2; FFA: free fatty acids.

matched controls had a similar rate of absolute weight loss, BMI change and percentage of excess weight loss at 1 year, postbariatric surgery.148 In contrast to the above findings, Aslan et al.147 have reported an unsuccessful case of weight loss in an adolescent patient (18.7-year-old) who is compound heterozygous for two loss of function MC4R mutations. Those findings suggest the complete functional loss of MC4R may render the bariatric surgery an ineffective procedure to lose weight in the long-term. Unexpectedly, Bonnefond et al.146 found that patients carrying gain of function MC4R mutations had higher risk of reoperation. Therefore, more studies about the effect of bariatric surgery in patients with MC4R mutations are necessary.

Nowadays, specific treatment is available for patients with leptin deficiency. In these cases, adults and children are treated with subcutaneous injection of recombinant human leptin, reducing food intake and weight. Furthermore, the administration of metreleptin improves most of the metabolic and endocrine dysfunctions seen in these patients.33,131,135 Unfortunately, leptin treatment is inefficient in LEPR-deficient patients, since their receptors are non-functional. Additionally, there is no specific treatment for other human cases of non–syndromic monogenic obesity. The physiological systems that control appetite and regulate body weight may harbour numerous targets for anti-obesity drugs. Thus, researchers are developing new molecules for pharmacological obesity treatment which act on the leptin–melanocortin pathway; however it is still in clinical trials.131

7. Conclusion

Clearly, obesity is a complex disease influenced by different factors. In most cases, the individual variability of becoming obese is due to many gene variants of minor effect. However, several studies have reported patients with severe early-onset obesity and endocrine dysfunctions caused by single gene mutations that disrupt the development of the hypothalamus and leptin–melanocortin pathway (Fig. 3). In the past few years, several molecular approaches and technologies were developed and have provided a small understanding of the genetic background of obesity. New genetic tools, such as whole-genome/exome sequencing, have been identifying novel variants and genes associated with common and rare forms of obesity. Those high-throughput DNA sequencing technologies have increasingly been used in the clinics, and their contribution to the development of new therapeutic targets for Mendelian obesity, and to the improvement quality of life remains to be determined.

References


