Faculteit Farmaceutische, Biomedische en Diergeneeskundige Wetenschappen

Study of copy number and DNA sequence variations in candidate genes for obesity

Studie naar de rol van structurele en DNA sequentie variaties in kandidaatgenen voor obesitas

Proefschrift voorgelegd tot het behalen van de graad van
Doctor in de Wetenschappen: Biochemie en Biotechnologie aan de Universiteit Antwerpen
te verdedigen door

Evi Aerts

Promotoren: Prof. Dr. Wim Van Hul
Prof. Dr. Luc Van Gaal
Co-promotor: Dr. Sigri Beckers

Antwerpen, 2017

“The weigh of life”

Lora Van der Kuylen

20/09/2017

MEMBERS OF THE JURY

Prof. Dr. Wim Van Hul, *promotor*

Center of Medical Genetics, University of Antwerp and University Hospital of Antwerp, Belgium

Prof. Dr. Luc Van Gaal, *promotor*

Department of Endocrinology, Diabetology and Metabolic Diseases, Antwerp University Hospital, Belgium

Dr. Sigri Beckers, *co-promotor*

Center of Medical Genetics, Antwerp University Hospital, Belgium

Prof. Dr. Wim Vanden Berghe, *chairman*

Department of Biomedical Sciences, University of Antwerp, Belgium

Prof. Dr. Julie van der Zee

Department of Biomedical Sciences, University of Antwerp, Belgium

Prof. Dr. Inge Gies

Department of Pediatrics, Division of Pediatric Endocrinology, University Hospital Brussels, Belgium

Dr. Eveline Dirinck

Department of Endocrinology, Diabetology and Metabolic Diseases, Antwerp University Hospital, Belgium

Dr. Mieke van Haelst

Academic Medical Center, Amsterdam, Netherlands

TABLE OF CONTENTS

TABLE OF CONTENTS

I. SUMMARY/SAMENVATTING

SUMMARY

Obesity is a highly heritable complex and heterogeneous disorder, characterized by an excessive accumulation of adipose tissue that results from a persistent imbalance between energy intake and expenditure. The heritability of obesity is generally estimated between 40-70%, indicating that genetic factors play an important role in the aetiology of the disease. So far, 97 BMI-associated loci were identified that account for approximately 2.7% of BMI variation, leaving a substantial component of the heritability unexplained. Since 2009, genome-wide CNV surveys have associated a number of CNVs with early-onset extreme obesity with the hope of shedding some light on the heritability. Further investigation of the candidate genes in these regions might offer novel insights into the genetic architecture of obesity. In this thesis, we aimed to study copy number and DNA sequence variations in selected candidate genes for obesity.

In the first chapter, we focused on ***SH2B1*** and the distal *SH2B1*-encompassing chr.16p11.2 CNV region by performing an extensive mutation and CNV analysis. Based on the equal frequency of *SH2B1* variants that we found among obese and lean individuals, it seems unlikely that these variants contribute majorly to the development of obesity. Furthermore, we were not able to detect carriers of the distal chr.16p11.2 deletion in our population of severely overweight or obese children and adolescents.

In the second chapter, we investigated the role of **NPY** and its receptors (**NPY2R** and **NPY4R**) in the development of obesity. We performed a mutation analysis for *NPY, NPY2R* and *NPY4R*. CNV screening was only performed for *NPY2R* and *NPY4R*. Mutation screening of *NPY* did not result in the identification of pathogenic mutations. This underlines the strong evolutionary conservation of the gene. For *NPY2R*, 1 rare non-synonymous variant F87I was identified in an 8-year old obese carrier. For this variant a possible pathogenic effect cannot be ruled out at this moment and functional experiments are required to elucidate its role in the pathogenesis of obesity. In contrast to *NPY* and *NPY2R,* the results of our mutation and CNV screening of *NPY4R* support an essential role for genetic variation within this gene in the susceptibility of obesity. Mutation analysis resulted in the identification of fifteen rare non-synonymous heterozygous variants. For two variants that could only be identified in our patient population (M116T and V271M), we were able to demonstrate receptor dysfunction and thus a pathogenic effect by performing a luciferase reporter assay. These findings were in line with the results of our CNV analysis in which we demonstrated a significantly higher frequency of *NP4YR*-containing 10q11.22 CNV loss in the patient population. Taken together, the presence of structural (12

10q11.22 CNV loss carriers) and single-base variation (2 pathogenic mutations) within *NPY4R* at least partially explains the obese phenotype of 3.7% and 0.6% of our patient population, respectively. These results indicate that *NPY4R* is a very important risk factor for obesity with prevalences close to those reported for *MC4R* mutations.

In the third chapter, we investigated the link between obesity and genetic variation in ***FGF21*** and ***UCP1***, two key regulators of brown and beige adipogenesis. For *UCP1*, we designed an extensive mutation and CNV screen in our patient and control population. Mutation screening resulted in the identification of 3 rare non-synonymous variants (G57S, T157I and T227I). *In silico* analysis predicted a potential impact of these variants on the function of UCP1. Functional experiments are still ongoing. In our CNV screening, we could not identify CNV in the *UCP1* region. Structural variation in this gene is unlikely to majorly contribute to the heritability of obesity.

For *FGF21,* mutation screening resulted in the identification of two rare non-synonymous variants R47Q and L142V. By performing luciferase reporter assays, we were able to demonstrate FGF21 dysfunction for L142V. For R47Q, we could only show a trend towards decreased activity.

In conclusion, we identified new and possibly pathogenic mutations and we implicated one gene as a new important risk factor for obesity. These results help in understanding the involvement of these genes in the pathogenesis of obesity.

SAMENVATTING

Obesitas is een chronische aandoening die gekenmerkt wordt door een excessieve accumulatie van lichaamsvet. De heritabiliteit van obesitas wordt algemeen geschat op 40-70%, hetgeen aantoont dat genetische factoren een belangrijke rol spelen in de etiologie van de aandoening. Tot op heden werden er 97 BMI-geassocieerde loci geïdentificeerd, maar samen verklaren deze slechts 2.7% van de variantie in BMI, waardoor een groot aandeel van de heritabiliteit nog niet verklaard kan worden. Onlangs werden structurele variaties in verband gebracht met obesitas. Door verdere karakterisatie van de genen in deze CNV regio's, kan hun belang in de pathogenese van obesitas onderzocht worden. Met deze thesis willen we de bijdrage van structurele en DNA sequentie variaties in geselecteerde kandidaatgenen tot de ontwikkeling van obesitas nagaan.

In het eerste hoofdstuk werd er gefocust op ***SH2B1*** en de distale *SH2B1*-omvattende 16p11.2 CNV regio. Zo werd er een uitgebreide mutatie- en CNV screening uitgevoerd. De mutatiescreening resulteerde tot de identificatie van een gelijk aantal *SH2B1*-varianten in de patiënten- en controlepopulatie. Op basis hiervan werd er gesuggereerd dat *SH2B1* varianten vermoedelijk geen groot effect op de ontwikkeling van obesitas uitoefenen in de algemene populatie. Verder kon er geen drager van de distale 16p11.2 CNV geïdentificeerd worden in onze populatie van kinderen en adolescenten met ernstig overgewicht of obesitas.

In het tweede hoofdstuk van deze thesis werd de rol van **NPY** en zijn receptoren (**NPY2R** en **NPY4R**) in de ontwikkeling van obesitas verder onderzocht. Zo werd er voor *NPY, NPY2R* en *NPY4R* een mutatie analyse uitgevoerd. De twee receptorgenen werd eveneens gescreend op de aanwezigheid van CNVs. Bij de mutatiescreening van *NPY* werden geen pathogene mutaties gevonden. Voor *NPY2R* werd 1 zeldzame niet-synonieme variant gevonden bij een 8-jarig obees kind. Verder functioneel onderzoek is noodzakelijk om de impact ervan op de werking van NPY2R en de rol in de pathogenese van obesitas op te helderen. In tegenstelling tot de resultaten van *NPY* en *NPY2R,* werd er voor *NPY4R* een sterke link aangetoond met obesitas. De mutatie-analyse van *NPY4R* resulteerde in de identificatie van vijftien niet-synonieme heterozygote varianten. Voor de twee varianten (M116T en V271M) die exclusief in de patiëntenpopulatie werden teruggevonden, kon via een luciferase reporter assay een verminderde werking van de receptor aangetoond worden. Deze bevindingen sluiten aan bij de resultaten van onze CNV analyse waarbij we een significant hogere frequentie aantoonden van de *NPY4R*-omvattende deletie in onze patiëntenpopulatie in vergelijking met de frequentie in

de controlepopulatie. De data van de mutatie- en CNV analyse tonen aan dat genetische variatie in *NPY4R*, door de aanwezigheid van pathogene puntmutaties en deleties in het gen, ten minste gedeeltelijk verantwoordelijk is voor het obese fenotype van respectievelijk 3.7 en 0.6% van onze obese studiepopulatie.

In het derde hoofdstuk, werd er gefocust op twee genen die een sleutelrol spelen in de bruine en beige adipogenese: ***FGF21*** en ***UCP1***. Voor het *UCP1* gen werd er een uitgebreide mutatie en CNV screening uitgevoerd in onze studiepopulaties. De mutatie-analyse resulteerde tot de identificatie van 3 zeldzame niet-synonieme varianten (G57S, T157I en T227I). Via *in silico* analyse werd een potentieel effect van de varianten op de functie van UCP1 aangeduid. Functionele experimenten worden momenteel uitgevoerd om deze predictie te bevestigen. Bij de CNV screening werden geen CNVs in de *UCP1* regio gevonden. Structurele variatie in *UCP1* zal dus bijgevolg minimaal bijdragen tot de vatbaarheid voor obesitas in de algemene populatie. Voor *FGF21* resulteerde de mutatie-analyse in de identificatie van twee zeldzame niet-synonieme varianten R47Q en L142V. In een luciferase reporter assay werd er een verminderde FGF21 activiteit aangetoond voor L142V. Voor R47Q kon enkel een trend van verminderde activiteit van het proteïne bevestigd worden.

We kunnen besluiten dat in deze studie nieuwe en mogelijk pathogene mutaties geïdentificeerd werden. Verder werd er aangetoond dat *NPY4R* een belangrijke rol speelt bij de vatbaarheid voor obesitas. Deze resultaten zullen bijdragen tot een beter begrip van de betrokkenheid van deze genen bij obesitas.

II. INTRODUCTION

THE ROLE OF COPY NUMBER VARIATION IN THE PATHOGENESIS OF OBESITY

Evi Aerts[1] and Wim Van Hul[1]

[1]Center of Medical Genetics, University of Antwerp, Belgium

Submitted, Obesity (Silver Spring)

ABSTRACT

Obesity is an heritable and multifactorial disorder that is the result of an imbalance between energy intake and expenditure. The extensive genetic heterogeneity of obesity has been one of the major problems to identify the genetic factors underlying the heritable risk of obesity. Copy number variation (CNV) accounts for a major proportion of human genetic variation and has been predicted to have an important role in genetic susceptibility to obesity. Since 2009, genome-wide CNV surveys have associated a number of CNVs with early-onset extreme obesity with the hope of shedding some light on the heritability. Further investigation of the genes in the linked CNV regions might offer novel insights into the genetic architecture of obesity and might help unraveling new obesity pathways. The *NPY4R* gene, located in the associated 10q11.22 CNV region, is the most important example of a gene that was identified as a major risk factor for obesity by CNV screening. In this paper, we review the involvement of CNVs in the pathogenesis of obesity and its implications towards pathogenic mechanisms, genetic screening and treatment.

INTRODUCTION

Obesity is a highly heritable complex and heterogeneous disorder, characterized by an excessive accumulation of adipose tissue that results from a persistent imbalance between energy intake and expenditure. With its prevalence reaching epidemic proportions, it has become the most common health disorder worldwide, contributing to the increased risk of many chronic diseases such as type 2 diabetes, cardiovascular diseases and certain cancers [1-3]. Because of the burden this disease poses on an individual's health, research efforts searching for factors contributing to the disease's etiology have intensified in the past decades [4]. Numerous studies have shown that 40-70% of the interindividual variability in BMI is attributed to genetic factors [5, 6]. So far, only a limited number of single gene mutations, causing rare monogenic obesity, is known, many of which are associated with the leptin-melanocortin pathway (figure 1). In this pathway, leptin is secreted by adipocytes in proportion to the amount of fat mass. Binding of leptin with its receptors in the *nucleus arcuatus* leads to the activation of the melanocortin-4 receptors followed by appetite suppression and stimulation of energy expenditure. With a prevalence of 4% among severely obese patients, *Melanocortin-4 receptor* (*MC4R*) gene mutations are the most prevalent cause of monogenic obesity [7]. In contrast, only eleven different mutations in the *leptin* gene have been discovered, making leptin deficiency a rare form of monogenic obesity (< 1/1 000 000) [8-15]. Leptin deficiency is currently, still the only treatable type of monogenic obesity [10, 16, 17]. Leptin receptor deficiency is more common and has a prevalence of up to 1% among early-onset obese patients [18-20]. Despite the importance of the leptin-melanocortin pathway in the pathogenesis of obesity, mutations in the genes involved in this pathway account for not more than 5% of all childhood severe obesity cases. The genetics of obesity for most individuals are complex and involve the interaction of multiple genes and environmental factors. Therefore, extensive genome-wide association studies (GWASs) have been performed to identify common single-nucleotide polymorphisms (SNPs), underlying the heritable risk of common obesity. So far, 97 BMI-associated loci are identified that account for approximately 2.7% of BMI variation, leaving a substantial component of the heritability unexplained [21-23].

Copy Number Variation (CNV) accounts for a major proportion of human genetic variation and has been predicted to have an important role in genetic susceptibility to common disease. Over the years, CNVs have been linked to dozens of human diseases including mental retardation, autism, schizophrenia, Charcot-Marie-Tooth disease type 1A and Crohn's disease [24-27]. In this

paper, we review the involvement of CNV in the pathogenesis of obesity with the hope of shedding some light on the heritability.

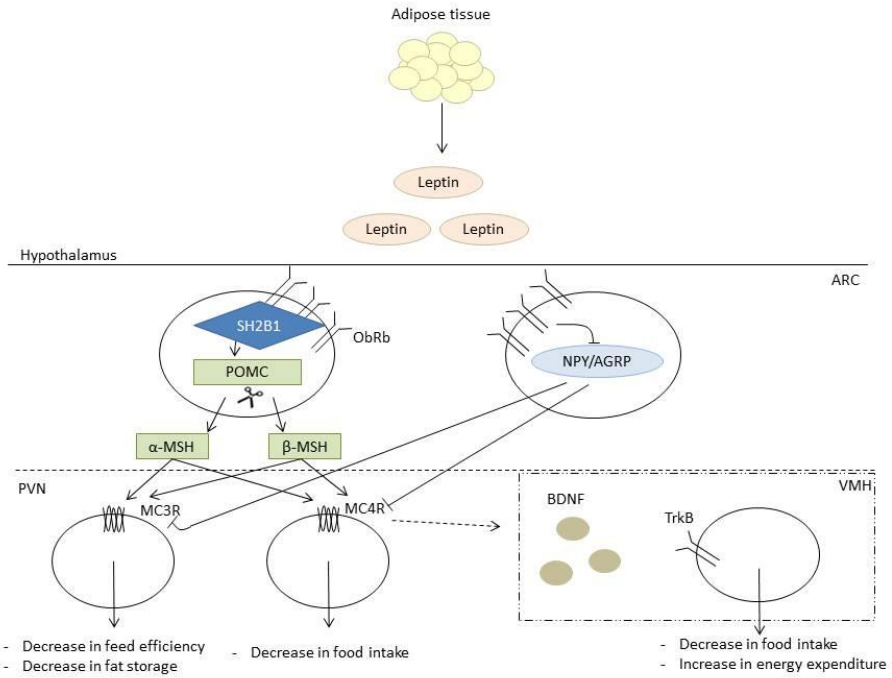

Figure 1: Schematic overview of the leptin-melanocortin pathway
Leptin is secreted from adipocytes in proportion to the amount of fat mass and travels to the brain via the bloodstream. In the hypothalamus, leptin binds to its receptor, ObRb. This leads to an increased expression of POMC in the anorexigenic part of the leptin-melanocortin pathway. POMC is cleaved by PC1 to its active forms α-MSH and β-MSH. These peptides bind to MC3R and MC4R, thereby inhibiting food intake, fat storage and feed efficiency. In the orexigenic part of the pathway, binding of leptin to its receptor, inhibits the expression of AgRP, an orexigenic protein that binds MC3R and MC4R. AgRp, agouti-related peptide; ARC, arcuate nucleus; BDNF, brain-derived neurotrophic factor; MC3R, melanocortin-3 receptor; MC4R, melanocortin-4 receptor; MSH, melanocyte-stimulating hormone; ObRb, long form of the leptin receptor; PC1, prohormone convertase 1; POMC, proopiomelanocortin; PVN, paraventricular nucleus; TrkB, tropomyosin-related kinase receptor B; VMH, ventromedial hypothalamus.

GENETIC VARIATION

Advances in the analysis of human genomes have revealed an unexpected amount of variability in human populations. The first and most common form of genetic variation involves the small-scale change of a single base pair, often called point mutations, or the deletion or insertion of a few bases. These changes may lead to one of four types of mutation: missense, nonsense, frameshift and splice site mutations and affect the function of a single gene. For monogenic obesity, these single gene mutations are directly responsible for the pathogenesis of obesity, while common obesity is caused by a complex interplay between multiple variants and environmental factors. The second form of genetic variation involves large-scale mutations in chromosomal structure including large copy number variants, chromosomal translocations and inversions. These changes can affect the functioning of numerous genes, resulting in major phenotypic consequences.

There are several models by which genetic variants affect gene regulation depending on the type, nature and position of the variant in the genome and functional genomic elements. These genetic variations may impact protein function and regulation in two ways: changes to transcript sequences by coding variants on one hand, and changes to transcript abundance by dosage or regulatory variants on the other hand. Furthermore, also epigenetic mechanisms are involved in the regulation of gene expression. Epigenetics is the study of heritable genome modifications that affect gene expression without actually changing the DNA sequence. The four major determinants of epigenetic regulation are tissue specific and include DNA methylation patterns, histone modification, chromatin confirmation and noncoding RNAs. Failures in these epigenetic marks or imprinting are not only known to cause extreme forms of obesity such as Prader-Willi syndrome, but are also associated with the general susceptibility of obesity.

However, in this review we focus on CNVs. Last few years CNVs have emerged as a potential factor that might explain individual differences in obesity risk. CNVs are defined as chromosomal regions with sizes of 1kb or larger being interindividual present in variable numbers. There are two major mechanisms for CNV generation, namely non-allelic homologous recombination (NAHR) and non-homologous end joining (NHEJ) (figure 2).

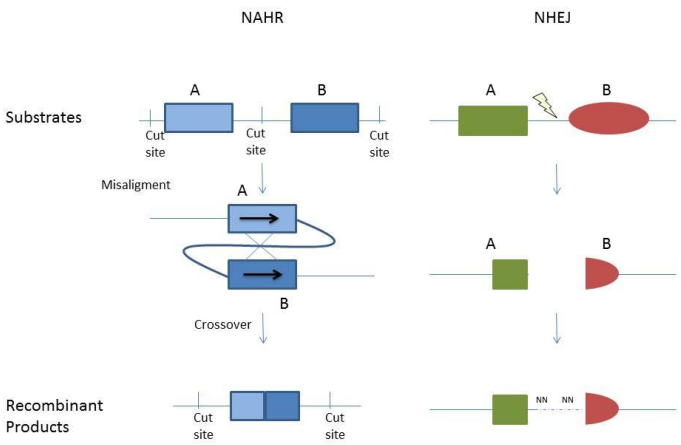

Figure 2: Schematic overview of NAHR and NHEJ
NAHR and NHEJ are two major mechanisms for CNV generation. NAHR occurs between two lengths of DNA that have high sequence similarity. NHEJ maintains the integrity of the DNA by repairing DNA double stranded breaks. The break ends are directly ligated without the need for a homologous template.

During the last decade, advances in whole-genome technologies, including array comparative genomic hybridization, SNP microarrays and genome sequencing have enhanced the ability to detect these alterations, revealing that they are widely spread along the human genome. So far, 552586 CNVs have been reported in the Database of Genomic variants with an estimated 75.6% of exons and 91.2% of transcripts that are overlapped by at least one CNV. Therefore, CNVs are a significant source of human genetic variation accounting for population diversity and human diseases. Pathogenic CNVs can cause Mendelian or sporadic traits, or can be associated with complex diseases by various molecular mechanisms including gene dosage, gene fusion, gene disruption and position effects [28-31]. The first mechanism, gene dosage, implies the number of copies of a particular gene that is present in the genome and is related to the amount of gene product that is expressed. Gene fusion is the process by which the complete or partial sequence of two or more genes are fused into a single transcript. This leads to the expression of hybrid proteins or to the misregulated transcription of one gene by the cis regulatory elements of the other. The mechanism of gene disruption implies the disruption of a target gene by the directed insertion of DNA within the gene, while position effect involves the deleterious change of gene expression due to the change in position of the gene relative to its normal chromosomal environment within the promotor region, silencers/enhancers and local chromatin environment of the gene locus. The transcription unit and minimal promotor region of the

rearranged gene remain intact. To explore the contribution of CNVs to obesity, genome-wide CNV surveys have been performed in patient populations of early-onset extremely obese individuals. These surveys reported a number of CNVs, especially large (>500 kilobases) and rare (<1%) deletions, conferring high risk to moderate and early-onset extreme obesity [32-37]. Few studies focused on the importance of common or complex multi-allelic CNV for early-onset extreme obesity. The genes that are impacted by CNVs might encompass known or novel candidate genes for follow-up studies (table 1).

Table 1: Overview of the most important BMI-associated CNVs

Locus	Position (Mb)	Size (kb)	Overlap Genes	**Regulatory pathways**	**References**
1p21.1	Chr1:103,900,000-104,200,000	N.A.	AMY1A, AMY2B, AMY1C, AMY2A, AMY2B	- Alpha-amylase activity - Carbohydrate metabolic process - Feeding behavior	[38]
1q31.1	Chr1:72,541,074-72,583,749	42,7	NEGR1	- Locomotory behavior - Neuron project development	[22, 23, 39]
4q31	Chr4:141,598,764-143,656,669	2057,9	ELMOD2, IL15, INPP4B, RNF150, TBD1D9, UCP1, ZNF330	- Oxidative phosphorylation uncoupler activity	[35]
10q11.22	Chr10:46,338,178-46,812-351	474,2	GLUDP2, NPY4R, GPRIN2, SYT15, BMS1P2, LOC642826, LOC643650, ANXA8L1, CTGLF7, LOC728643, LOC728657, LOC100132646, FAM25B, LOC100133189	- Pancreatic polypeptide receptor signaling - GPCR signaling	[32, 34, 40]
11q11	Chr11:55,130,596-55,210,165	79,6	OR4P1P, OR4V1P, OR4P4, OR4S2, OR4C6	- Olfactory signaling pathway - GPCR signaling	[34]
16p11.2	Chr16:28,731,428-28,951,376	220,0	ATXN2L, SH2B1, TUFM, RABEP2, CD19, SPNS1, LAT, ATP2A1, NFATC2IP GPRC5B	- Intracellular signal transduction	[33, 41]
16p12.3	Chr16:19,853,734-19.19,874,731	21,0		- GPCR signaling - Locomotory behavior - Glucose homeostasis - Regulation of protein kinase activity	[34]

RARE CNVS

Chromosome 4q31

In 2010, a 2.1 Mb deletion on chromosome 4q31 was identified as an important risk factor in the study of obesity. The deletion was found in a severely obese patient with a BMI of 46.2 kg/m^2 that inherited the CNV from his father. Genetic evidence coming from animal models and association studies, has identified the *Uncoupling protein 1 (UCP1)* gene as the strongest *a priori* candidate gene for obesity in this CNV region [35]. UCP1 plays an important role in the regulation of energy homeostasis, cold- and diet-induced thermogenesis and in the protection against oxidative stress. It is a mitochondrial carrier protein that is localized on the inner membrane of mitochondria in brown adipose tissue where it diminishes the proton gradient by uncoupling cellular respiration and mitochondrial ATP synthesis. By releasing this chemical energy as heat, activation of UCP1 increases the basal energy expenditure [42]. Numerous studies investigated the association of three polymorphisms in the promotor region (-3826A/G, -1766A/G and -112A/C) and A64T in exon 2 and M299L in exon 5 of the *UCP1* gene with obesity. Although the results of these studies were not always consistent, the -3826A/G polymorphism was reported to clearly contribute to the susceptibility of obesity and obesity related parameters [43-46]. In animal studies, *UCP1*-ablation in itself induced obesity in C57Bl6 mice that were kept under thermoneutral conditions [47].

To evaluate whether genetic and structural variation in *UCP1* might be implicated in human energy regulation, we performed an extensive mutation and CNV screen in our patient population of severely overweight or obese children and adolescents. As we could not identify CNV in the *UCP1* region, structural variation in this gene is unlikely to contribute majorly to the obese phenotype. However, our mutation analysis resulted in the identification of three rare non-synonymous heterozygous variants, two of which could only be identified in obese individuals and one that could only be found in a lean adult. Further functional experiments are necessary to determine the effect of these variants on the function of UCP1.

16p11.2 deletion

The short arm of human chromosome 16 is particularly enriched for large segmental duplications that serve as breakpoints for NAHR (figure 2), resulting in recurrent duplications and deletions [48, 49]. In 2010, Bochukova *et al.* investigated the contribution of copy number variation to the pathogenesis of obesity in 362 patients with severe early-onset obesity, 143 of whom also suffered from developmental delay. The most common CNV that was enriched among these severely obese patients, was identified on 16p11.2 with a minimal overlapping segment of 220kb [33]. This minimal deleted region contains genes involved in immunity (*CD19, LAT, NFATC2IP*), neurological diseases (*TUFM, ATP2A1*) and genes of unknown function (*ATXN2L, RABEP2, SPNS1*), as well as the *Src homology 2B adaptor protein 1* gene (*SH2B1)* that is involved in the leptin-melanocortin signaling pathway. Genetic evidence coming from genome-wide association and animal studies strengthened the case for a role of *SH2B1* in the development of human obesity. Genome-wide association studies identified the A484T variant in *SH2B1* (rs7498665) as one of the top-hits for obesity . The rs7498665 variant allele would increase obesity risk by about 11-26%, a finding that we could also replicate in our Belgian population [22, 39, 50]. In our research group, we also demonstrated that non-synonymous variations in the *SH2B1* gene are frequent in both lean and obese groups, with distinctive variations being present on either side of the weight spectrum. Although the equal variation frequency does not immediately support disease causality, it cannot be excluded that some variations are weight-increasing or –decreasing [41]. Furthermore, the generation of an *Sh2b1* null mouse also established the protein's involvement in obesity, as these animals were obese with severe hyperphagia, hyperleptinemia, hyperinsulinemia, hyperlipidemia, hepatic steatosis, hyperglycemia and glucose intolerance.

In the study of Bochukova *et al.,* the 220 kb *SH2B1*-encompassing 16p11.2 deletion was identified in 5 unrelated patients and it was found to be associated with highly penetrant familial early-onset obesity, hyperphagia and severe insulin resistance. The reported overall prevalence of the *SH2B1*-containing deletion in severely obese patients was 0.41% (5 out of 1219) compared to only 0.027% (2 out of 7366) among controls. Furthermore, two patients with developmental delay in addition to severe obesity were carriers of a longer 1.7 Mb deletion that not only encompassed the distal 16p11.2 deletion (220 kb), but also included a 593kb region that was already associated with autism spectrum disorder, epilepsy, developmental delay and obesity [51, 52]. Later on, this proximal 593kb deletion was recognized as a syndrome, "the

16p11.2 microdeletion syndrome", with phenotypic variety that ranges from asymptomatic to developmental delay, mild cognitive impairment, behavioral problems, dyslexia, speech-language impairments, motor delay and dysmorphologies [51-54]. This broader phenotype was originally not attributed to the distal 16p11.2 deletion. However, this was not in line with an overview that was published in 2011, reviewing all known carriers of the distal 16p11.2 deletion. Of all reported cases, about one third had developmental problems and almost 23% had an unusual facial morphology [55]. Other common traits were obesity, behavioral problems, seizures, autism and speech delay. In contrast to the study of Bochukova, this review reported that in addition to obesity, the distal 220kb deletion on itself predisposes for developmental delay and intellectual disability, similarly to the proximal deletion. These conflicting data prompted us to further investigate the prevalence of the *SH2B1*-encompassing deletion in a patient population of severely overweight or obese children and adolescents without developmental or behavioral problems [41]. As we were not able to detect CNV in the 16p11.2 region among 421 patients, we suggest that the prevalence of this deletion in individuals with no other traits besides obesity might indeed be very low. The other genes in the deleted region might be responsible for the developmental and behavioral problems in its carriers. However, more recently another explanation was hypothesized for this observation. Loviglio *et al.* demonstrated in 2016 that both 16p11.2 regions, whose CNVs are linked to overlapping phenotypes, are reciprocally engaged in complex chromatin looping. Rearrangement of the 16p11.2 region would affect the three-dimensional chromatin architecture by shifting regulatory elements between domains or by modifying domain boundaries. This might lead to the disruption of expression networks that involve multiple genes and pathways. Elucidation of chromatin contacts can be proposed as a new and effective tool to unravel the genes participating in the pathogenesis of obesity [56, 57].

Chromosome 16p12.3

In 2010, a genome-wide association study linked 18 additional loci to BMI, one of which included a CNV near *GPRC5B* that was tagged by the T-allele of rs12444979. This CNV encloses a 21kb deletion that lies 50kb upstream of the *G protein coupled receptor 5B (GPRC5B)* on chromosome 16p12.3[23]. This was successfully confirmed in an independent study that associated the CNV 16p12.3 with obesity-related phenotypes in the European population [58]. In 2012 follow-up studies were performed, investigating the potential involvement of *GPRC5B* in

disrupting energy homeostasis in mice. However, in contrast to the human phenotype, *Gprc5b* deficient mice were found to be protected against diet-induced obesity and insulin resistance because of reduced inflammation in their white adipose tissue [59]. GPRC5B is a lipid raft-associated transmembrane phosphoprotein in the plasma membrane that acts as a major node in the adipose signaling system. Further functional studies are essential to elucidate its detailed role in the pathogenesis of obesity.

COMMON CNVS

chromosome 1p21

Although the majority of common CNVs are simple and bi-allelic, an understudied set of common CNVs are complex and less characterized as they have resisted effective analysis by most molecular methods. These multi-allelic CNVs (mCNVs) involve genomic segments that seem to vary widely in copy number, in patterns that imply the existence of more than two segregating alleles. Compared to the bi-allelic CNVs, the mCNVs appear enriched for phenotypic associations [60]. For obesity, a strong correlation was reported between the copy number of mCNV *salivary α-amylase (AMY1)* and common obesity, with lower number of gene copies linked to increased BMI. Increased *AMY1* copy number was positively associated with both amylase gene expression and serum enzyme levels . The OR value of 1.19 per copy of *AMY1* translates into an eightfold difference in risk of obesity between subjects in the top and bottom 10% of the copy number distribution [38]. The *AMY1* locus is located in a tandem CNV region on human chromosome 1p21.1, with a copy number varying between two and twenty. In the original assembly of the locus sequence three *AMY1* gene repeats are reported: *AMY1A, AMY1B and AMY1C,* with *AMY1B* in inverted orientation (figure 3).

These three copies of *AMY1* are more than 99.9% identical in DNA sequence and are more than 93% identical to *AMY2A* and *B* genes, suggesting that these are all repeat units of *AMY1*. The salivary α-amylase is a major digestive enzyme in the saliva that catalyzes the hydrolysis of starch into sugars. All amylases are glycoside hydrolases that break down long-chain carbohydrates. During evolution, starch has become a prominent component of the human diet. Therefore, the average human has three times more *AMY1* copies than chimpanzees [61, 62]. This is in line with the lower number of *AMY1* repeats that was associated with an increased BMI.

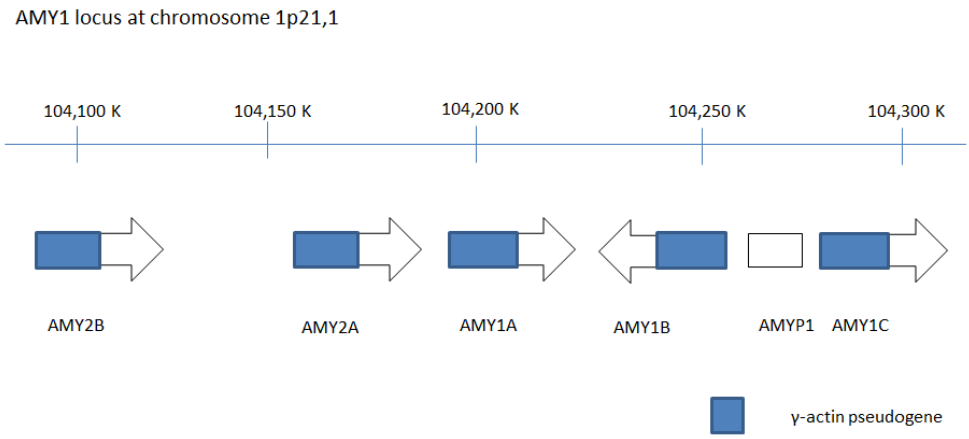

Figure 3: *AMY1* genes and α-amylase gene cluster on chromosome 1p21
The *AMY1* genes (*AMY1A, AMY1B* and *AMY1C*) that encode the salivary α-amylase are located in a gene cluster on 1p21 that also includes two pancreatic α-amylase genes (*AMY2A* and *AMY2B*) and a pseudogene *AMYP1*. The uppermost exon and the following intron and 5'-flanking region of the *AMY* genes are superimposed with a gamma actin pseudogene sequence.

Chromosome 1q31.1

Neuronal growth regulator 1 (NEGR1) belongs to the group of GPI-anchored adhesion molecules that controls cellular processes as neurite outgrowth and synapse formation. It is one of the first genes in the expanding list of common obesity loci that were identified via GWASs [63]. This was further confirmed in independent studies linking the *NEGR1* locus not only to BMI and obesity, but also to a variety of obesity-related traits such as birth weight, infancy weight gain and subcutaneous fat mass [64-67]. Two copy number polymorphisms (CNPs) upstream of *NEGR1* were associated with severe obesity that segregate on distinct haplotypes: a protective 8 kb deletion with inversed association and a larger 43 kb deletion with positive association to obesity (figure 4) [23, 39, 68]. Although the most strongly associated SNP rs2815752 represents the 43 kb deletion, conditional analysis revealed that the 8 kb deletion is the major driver of the association and that it has a strong sex bias in effect size [69, 70]. Both deletion polymorphisms do not overlap coding sequence but encompass several important conserved elements, including binding sites of transcription factors *NKX3.1* (43 kb deletion) and *NKX6.1* (8kb deletion). *NKX3.1* is a putative prostate tumor suppressor that is expressed in a largely prostate-specific and androgen-regulated manner [71]. So *NKX3.1,* given its male specificity, might have a

trans-regulatory role in the association between the deletion and obesity explaining the observed sex bias in effect size. As the 8kb deletion removes the binding site for *NKX6.1*, a potent transcriptional repressor, upregulation of *NEGR1* might protect against obesity while downregulation might contribute to the predisposition to obesity. This is in line with a recent study that showed an increased body weight and food intake in mice with a decreased expression of *Negr1* in periventricular hypothalamic areas. However, systemic *Negr1* KO mice show a decreased body weight, food intake and locomotor activity. Developmental effects could account for these incompatible outcomes [72, 73]. In either way, these results support a clear role for *NEGR1* in the control of body weight and also point to a key role of the central nervous system in body weight regulation.

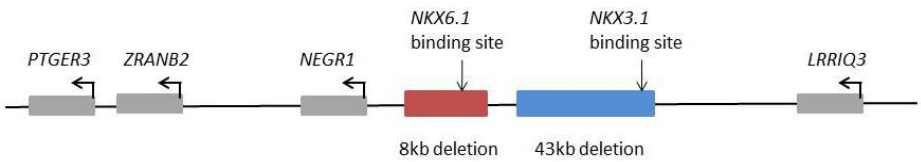

Figure 4: Overview of two CNPs upstream of *NEGR1*
These two CNPs are associated with obesity and segregate on distinct haplotypes: a protectective 8kb deletion with inversed association to obesity and a larger 43kb deletion with positive association to obesity. Conditional analysis revealed the small 8kb deletion to be the major driver of the association with obesity.

Chromosome 10q11.22

Another important CNV region that warrants further follow-up in the study of obesity, is the 10q11.22 CNV region. In 2009, a genomewide CNV analysis reported the strong association of a 193kb CNV region at this chromosome (from 46.36Mb to 46.56Mb) with early-onset extreme obesity. Subjects with 10q11.22 CNV loss were reported to have 12.4% higher BMI values, while subjects with CNV gain in this region had 5.4% lower BMI values when compared to normal diploid subjects [32]. In general, this CNV region would contribute to 1.6% of BMI variation. Later, independent studies confirmed the strong association between the 10q11.22 deletion and an increased risk of childhood obesity in a Chinese and European population [34, 35, 37]. The 193kb 10q11.22 CNV region includes four genes: *Synaptotagmin XV* (*SYT15*), G protein regulated inducer of neurite outgrowth 2 (GPRIN2), *heterogeneous nuclear ribonucleoprotein A1*

pseudogene (*LOC728643*) and *Pancreatic Polypeptide receptor 1* (*PPYR1*). One of these genes, *PPYR1*, also named as the *Neuropeptide Y4 receptor gene* (*NPY4R*), is a seven transmembrane domain G-protein-coupled receptor belonging to the Y receptor family that is highly implicated in the energy homeostasis. The other three genes that are impacted by this CNV are not known to be directly involved in the regulation of food intake. *NPY4R* is predominantly expressed in peripheral tissues such as the pancreas, colon, liver, heart and intestine [74-76]. Pancreatic polypeptide (PP) is the preferential agonist of the Y4 receptor and functions as a satiety signal via the Y4 receptors [77]. Intravenous or intraperitoneal administration of PP to mice increases metabolic rate and reduces the obese phenotype of *ob/ob* mice by decreasing hyperglycemia, insulin resistance and hyperlipidemia [78, 79]. Seemingly contradictory results for the role of NPY4R in the pathogenesis of obesity came from germline and conditional *Npy4r* knockout mouse models. These knockout models report hyperphagia in combination with a reduced body weight, reduced amount of white adipose tissue and increased plasma PP levels, while *NPY4R* deficiency in humans is associated with obesity [80, 81]. In contrast to mice that express a functional Y6 receptor, the *NPY6R* gene only exists as a pseudogene in humans. The presence of an extra functioning receptor might alter the ligand-binding patterns of the other Y receptors. Furthermore, the mice *Npy4r* amino acid sequence is only 76% identical to human *NPY4R*. So although mice are not the most appropriate model to investigate NPY4R, they clearly demonstrate the link between *NPY4R* and body weight regulation. In 2015 and 2016, we and others performed a CNV analysis, investigating genetic and structural variation among obese children and adolescents in a Chinese and Belgian population respectively. We demonstrated a significantly higher frequency of *NPY4R* containing 10q11.22 CNV loss among obese patients ($p=0.0003$), while CNV gain in this region was more prevalent among lean individuals. Mutation analysis in the Belgian population resulted in the identification of fifteen rare non-synonymous heterozygous variants. Two of these were only identified among obese patients and were found to result in receptor dysfunction. In general, the presence of pathogenic non-synonymous mutations and deletions within the *NPY4R* could al least explain 3.7% and 0.6% of this obese population. This makes this gene a very important risk factor for obesity with prevalences close to those reported for *MC4R* mutations [37, 40].

Chromosome 11q11

The human olfactory receptor (OR) genes form the largest multigene family with 906 genes and pseudogenes that are organized in genomic clusters on almost every chromosome [82]. Compared to other mammals, the genomic content of OR genes in humans is significantly reduced as it has undergone an accelerated process of pseudogenization. This has led to the functional inactivation of more than 50% of the human OR repertoire [83, 84]. The genomic OR clusters arose during evolution via numerous tandem duplications and interchromosomal duplications. Due to these dynamic evolutionary processes, the OR gene loci belong to one of the most genetically diverse regions in the human genome that are significantly enriched in CNV regions, suggesting that CNVs might play an important role in the evolution of ORs [84]. In 2010, Jarick *et al*. searched genome-wide for associations between common CNVs and early-onset extreme obesity and identified a common 80kb CNV region on chromosome 11q11 exclusively covering three olfactory receptor genes: *olfactory receptor family 4 subfamily P member 4 (OR4P4), subfamily S member 2 (OR4S2)* and *subfamily C member 6 (OR4C6)* [34]. The genes encompassing this CNV region might be important regulators of food intake. Olfactory receptors are G protein coupled receptors that play an essential role in the specific recognition of millions of odorous compounds. As the perception of smell acts on appetite or satiety signals, variations in human olfactory genes might indeed influence eating behaviors and lead to hyperphagia and overweight/obesity [85, 86]. This was confirmed by a study in 2012 that associated SNPs in a cluster of 7 OR genes on chromosome 19p13 with obesity-related parameters [87]. Furthermore a recent functional clustering study investigating the exome sequencing data of 30 extremely obese subjects showed that predicted damaging missense variants in olfactory receptor genes on chromosome 1q and rare variants in the protocadherin genes were found to colocalize in subjects with extreme obesity, suggesting a synergistic effect of these genes in the predisposition to obesity [88]. Future studies investigating the genetic and structural variation of the olfactory receptor genes might shed further light on the role of these genes in the pathogenesis of obesity.

DISCUSSION

Obesity is an heritable and multifactorial disorder that is the result of an imbalance between energy intake and expenditure, caused by the complex interaction of genetic and environmental factors. The extensive genetic heterogeneity of obesity has been one of the major problems to

identify the genetic factors underlying the heritable risk of obesity. To date, three main approaches have been used to identify and investigate the potential of previously unknown variants, starting with linkage, candidate gene association and GWA studies. With the advent of next generation sequencing (NGS) technology, large-scale investigation of whole genomes has been within reach with unprecedented throughput and speed. Despite the enormous volumes of data coming from these six GWA and NGS studies, the fraction of BMI variance that can be attributed to common SNPs stays minimal. These findings bring up the discussion on the common disease-common variant model versus common disease-rare variant model. The first model supports the proposition that common variants with small effect sizes are a major source of genetic variance for complex diseases. Numerous genetic factors would provide small independent and additive contributions to the complex phenotype. The rare allele model states that multiple highly penetrant rare variants are major contributors to genetic susceptibility of common diseases. Although GWAS have successfully identified 97 BMI-associated loci for obesity, altogether they only account for approximately 2.7% of BMI variation, leaving a substantial component of the heritability unexplained. One of the arguments that must be considered, is that the proportion of heritability that can be captured by GWAS, depends on how well the causal variants are tagged by tagSNPs. Ungenotyped causal variants that are in low linkage disequilibrium or have a lower allele frequency than the selected tagSNP might not reach genome-wide statistical significance, underestimating the percentage of heritability that can be explained [89]. However, in the case of height, for which the same methods have been applied, over 50% of the genetic variance can be explained by common variants. Taken together, it can be argued that common variants do not capture all the genetic variance for obesity and so the remaining heritability must be found in other sources of genetic variation. The last few years, there has been growing interest in the epigenetic regulation of gene expression and its role in the pathogenesis of obesity. Epigenetics, a rapidly developing field focusing on heritable changes, alter gene expression without affecting the underlying genomic sequences. These modifications are tissue specific and include DNA methylation and changes in chromatic organization by histone modifications, genomic imprinting, non-covalent mechanisms and non-coding micro RNAs [90]. Since the advances in microarray and sequencing technologies, major progress has been made in high-throughput profiling of genome-wide DNA methylation and hydroxy-methylation. A better understanding and incorporation of these epigenetic data may further increase our knowledge of underlying disease mechanisms such as epigenetic mediation and environmental risk factors. Additionally, this might eventually lead to

the clinical use of epigenetics as biomarkers for diagnosis, prognosis and response to specific treatment.

Like single nucleotide variants, also CNVs have been predicted to have an important role in the genetic susceptibility of obesity. Structural variation results from DNA recombination-, replication and repair- based processes. Genome-wide CNV surveys have associated a significant number of CNV regions with early-onset extreme obesity. Further investigation of the interesting candidate genes in these regions would offer novel insights into the genetic architecture of obesity. However, to date limited follow-up has been performed. The impact of a particular CNV on the expression of overlapping genes might go beyond the direct effect of the altered copy number. Large genomic rearrangements can lead to chromatin modifications that not only influence the expression of local genes, but extend to flanking regions, reflecting long-range *cis*-regulatory elements. With the recent burst of technological development in NGS and the use of composite pipelines for CNV analysis of NGS data, further comprehensive investigation of these data is an essential step in order to increase our knowledge about the distribution, formation and evolution of CNVs and their implication in the pathogenesis of obesity. Eventually this knowledge will help unraveling new pathways and new insights in existing pathways.

Based on this knowledge, possibilities to develop genetic tests for the prediction of obesity susceptibility and the option for personalized medicine may arise in the future. Major genetic determinants of obesity risk involve mutations in *MC4R, LEP, LEPR , FTO* and copy number variation in *NPY4R*. However, for general complex obesity, all current risk alleles together are not sufficient to have any clinical value and additional genes or different types of genetic variants remain to be discovered.

REFERENCES

1. Lew, E.A. and L. Garfinkel, *Variations in mortality by weight among 750,000 men and women.* J Chronic Dis, 1979. **32**(8): p. 563-76.
2. Lavie, C.J., R.V. Milani, and H.O. Ventura, *Obesity and cardiovascular disease: risk factor, paradox, and impact of weight loss.* J Am Coll Cardiol, 2009. **53**(21): p. 1925-32.
3. Hunter, D.J. and W.C. Willett, *Diet, body size, and breast cancer.* Epidemiol Rev, 1993. **15**(1): p. 110-32.
4. Wang, Y.C., et al., *Health and economic burden of the projected obesity trends in the USA and the UK.* Lancet, 2011. **378**(9793): p. 815-25.
5. Stunkard, A.J., T.T. Foch, and Z. Hrubec, *A Twin Study of Human Obesity.* Jama-Journal of the American Medical Association, 1986. **256**(1): p. 51-54.
6. Bouchard, C., et al., *The Response to Long-Term Overfeeding in Identical-Twins.* New England Journal of Medicine, 1990. **322**(21): p. 1477-1482.
7. Hinney, A., A.L. Volckmar, and N. Knoll, *Melanocortin-4 receptor in energy homeostasis and obesity pathogenesis.* Prog Mol Biol Transl Sci, 2013. **114**: p. 147-91.
8. Montague, C.T., et al., *Congenital leptin deficiency is associated with severe early-onset obesity in humans.* Nature, 1997. **387**(6636): p. 903-8.
9. Strobel, A., et al., *A leptin missense mutation associated with hypogonadism and morbid obesity.* Nat Genet, 1998. **18**(3): p. 213-5.
10. Gibson, W.T., et al., *Congenital leptin deficiency due to homozygosity for the Delta133G mutation: report of another case and evaluation of response to four years of leptin therapy.* J Clin Endocrinol Metab, 2004. **89**(10): p. 4821-6.
11. Mazen, I., et al., *A novel homozygous missense mutation of the leptin gene (N103K) in an obese Egyptian patient.* Mol Genet Metab, 2009. **97**(4): p. 305-8.
12. Fischer-Posovszky, P., et al., *A new missense mutation in the leptin gene causes mild obesity and hypogonadism without affecting T cell responsiveness.* J Clin Endocrinol Metab, 2010. **95**(6): p. 2836-40.
13. Fatima, W., et al., *Leptin deficiency and leptin gene mutations in obese children from Pakistan.* Int J Pediatr Obes, 2011. **6**(5-6): p. 419-27.
14. Saeed, S., et al., *High prevalence of leptin and melanocortin-4 receptor gene mutations in children with severe obesity from Pakistani consanguineous families.* Mol Genet Metab, 2012. **106**(1): p. 121-6.

15. Wasim, M., et al., *Role of Leptin Deficiency, Inefficiency, and Leptin Receptors in Obesity.* Biochem Genet, 2016. **54**(5): p. 565-72.
16. Farooqi, I.S., et al., *Effects of recombinant leptin therapy in a child with congenital leptin deficiency.* N Engl J Med, 1999. **341**(12): p. 879-84.
17. Farooqi, I.S., et al., *Beneficial effects of leptin on obesity, T cell hyporesponsiveness, and neuroendocrine/metabolic dysfunction of human congenital leptin deficiency.* J Clin Invest, 2002. **110**(8): p. 1093-103.
18. Farooqi, I.S., et al., *Clinical and molecular genetic spectrum of congenital deficiency of the leptin receptor.* N Engl J Med, 2007. **356**(3): p. 237-47.
19. Mazen, I., et al., *Homozygosity for a novel missense mutation in the leptin receptor gene (P316T) in two Egyptian cousins with severe early onset obesity.* Mol Genet Metab, 2011. **102**(4): p. 461-4.
20. Andiran, N., N. Celik, and F. Andiran, *Homozygosity for two missense mutations in the leptin receptor gene (P316:W646C) in a Turkmenian girl with severe early-onset obesity.* J Pediatr Endocrinol Metab, 2011. **24**(11-12): p. 1043-5.
21. Locke, A.E., et al., *Genetic studies of body mass index yield new insights for obesity biology.* Nature, 2015. **518**(7538): p. 197-U401.
22. Willer, C.J., et al., *Six new loci associated with body mass index highlight a neuronal influence on body weight regulation.* Nat Genet, 2009. **41**(1): p. 25-34.
23. Speliotes, E.K., et al., *Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index.* Nature Genetics, 2010. **42**(11): p. 937-U53.
24. Moser, H.W., *Genetic causes of mental retardation.* Ann N Y Acad Sci, 2004. **1038**: p. 44-8.
25. Chelly, J., et al., *Genetics and pathophysiology of mental retardation.* Eur J Hum Genet, 2006. **14**(6): p. 701-13.
26. Schaschl, H., T.J. Aitman, and T.J. Vyse, *Copy number variation in the human genome and its implication in autoimmunity.* Clin Exp Immunol, 2009. **156**(1): p. 12-6.
27. Boerkoel, C.F., et al., *Molecular mechanisms for CMT1A duplication and HNPP deletion.* Ann N Y Acad Sci, 1999. **883**: p. 22-35.
28. Olson, H., et al., *Copy number variation plays an important role in clinical epilepsy.* Ann Neurol, 2014. **75**(6): p. 943-58.
29. Zhang, F., et al., *Copy number variation in human health, disease, and evolution.* Annu Rev Genomics Hum Genet, 2009. **10**: p. 451-81.

30. Pinto, D., et al., *Functional impact of global rare copy number variation in autism spectrum disorders.* Nature, 2010. **466**(7304): p. 368-72.
31. Castellani, C.A., et al., *Copy number variation distribution in six monozygotic twin pairs discordant for schizophrenia.* Twin Res Hum Genet, 2014. **17**(2): p. 108-20.
32. Sha, B.Y., et al., *Genome-wide association study suggested copy number variation may be associated with body mass index in the Chinese population.* J Hum Genet, 2009. **54**(4): p. 199-202.
33. Bochukova, E.G., et al., *Large, rare chromosomal deletions associated with severe early-onset obesity.* Nature, 2010. **463**(7281): p. 666-70.
34. Jarick, I., et al., *Novel common copy number variation for early onset extreme obesity on chromosome 11q11 identified by a genome-wide analysis.* Human Molecular Genetics, 2011. **20**(4): p. 840-852.
35. Wang, K., et al., *Large copy-number variations are enriched in cases with moderate to extreme obesity.* Diabetes, 2010. **59**(10): p. 2690-4.
36. Biesecker, L.G., et al., *The ClinSeq Project: piloting large-scale genome sequencing for research in genomic medicine.* Genome Res, 2009. **19**(9): p. 1665-74.
37. Zhang, D., et al., *Interactions between obesity-related copy number variants and dietary behaviors in childhood obesity.* Nutrients, 2015. **7**(4): p. 3054-66.
38. Falchi, M., et al., *Low copy number of the salivary amylase gene predisposes to obesity.* Nat Genet, 2014. **46**(5): p. 492-7.
39. Thorleifsson, G., et al., *Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity.* Nat Genet, 2009. **41**(1): p. 18-24.
40. Aerts, E., et al., *CNV analysis and mutation screening indicate an important role for the NPY4R gene in human obesity.* Obesity (Silver Spring), 2016. **24**(4): p. 970-6.
41. Aerts, E., et al., *Genetic and structural variation in the SH2B1 gene in the Belgian population.* Mol Genet Metab, 2015. **115**(4): p. 193-8.
42. Kozak, L.P. and R. Anunciado-Koza, *UCP1: its involvement and utility in obesity.* Int J Obes (Lond), 2008. **32 Suppl 7**: p. S32-8.
43. Brondani, L.A., et al., *The role of the uncoupling protein 1 (UCP1) on the development of obesity and type 2 diabetes mellitus.* Arq Bras Endocrinol Metabol, 2012. **56**(4): p. 215-25.

44. Hayakawa, T., et al., *Phenotypic characterization of the beta3-adrenergic receptor mutation and the uncoupling protein 1 polymorphism in Japanese men.* Metabolism, 1999. **48**(5): p. 636-40.
45. Heilbronn, L.K., et al., *Association of -3826 G variant in uncoupling protein-1 with increased BMI in overweight Australian women.* Diabetologia, 2000. **43**(2): p. 242-4.
46. Nakano, T., et al., *A/G heterozygote of the A-3826G polymorphism in the UCP-1 gene has higher BMI than A/A and G/G homozygote in young Japanese males.* J Med Invest, 2006. **53**(3-4): p. 218-22.
47. Feldmann, H.M., et al., *UCP1 ablation induces obesity and abolishes diet-induced thermogenesis in mice exempt from thermal stress by living at thermoneutrality.* Cell Metab, 2009. **9**(2): p. 203-9.
48. Rees, E., et al., *CNV analysis in a large schizophrenia sample implicates deletions at 16p12.1 and SLC1A1 and duplications at 1p36.33 and CGNL1.* Hum Mol Genet, 2014. **23**(6): p. 1669-76.
49. D'Angelo, C.S. and C.P. Koiffmann, *Copy number variants in obesity-related syndromes: review and perspectives on novel molecular approaches.* J Obes, 2012. **2012**: p. 845480.
50. Beckers, S., et al., *Replication of the SH2B1 rs7498665 association with obesity in a Belgian study population.* Obes Facts, 2011. **4**(6): p. 473-7.
51. Kumar, R.A., et al., *Recurrent 16p11.2 microdeletions in autism.* Hum Mol Genet, 2008. **17**(4): p. 628-38.
52. Bijlsma, E.K., et al., *Extending the phenotype of recurrent rearrangements of 16p11.2: deletions in mentally retarded patients without autism and in normal individuals.* Eur J Med Genet, 2009. **52**(2-3): p. 77-87.
53. Ghebranious, N., et al., *A novel microdeletion at 16p11.2 harbors candidate genes for aortic valve development, seizure disorder, and mild mental retardation.* Am J Med Genet A, 2007. **143A**(13): p. 1462-71.
54. Weiss, L.A., et al., *Association between microdeletion and microduplication at 16p11.2 and autism.* N Engl J Med, 2008. **358**(7): p. 667-75.
55. Barge-Schaapveld, D.Q., et al., *The atypical 16p11.2 deletion: a not so atypical microdeletion syndrome?* Am J Med Genet A, 2011. **155A**(5): p. 1066-72.
56. Loviglio, M.N., et al., *Chromosomal contacts connect loci associated with autism, BMI and head circumference phenotypes.* Mol Psychiatry, 2016.

57. Blumenthal, I., et al., *Transcriptional consequences of 16p11.2 deletion and duplication in mouse cortex and multiplex autism families.* Am J Hum Genet, 2014. **94**(6): p. 870-83.
58. Yang, G., et al., *Ethnic differences in oro-facial somatosensory profiles-quantitative sensory testing in Chinese and Danes.* J Oral Rehabil, 2013. **40**(11): p. 844-53.
59. Kim, Y.J., et al., *GPRC5B activates obesity-associated inflammatory signaling in adipocytes.* Sci Signal, 2012. **5**(251): p. ra85.
60. Usher, C.L. and S.A. McCarroll, *Complex and multi-allelic copy number variation in human disease.* Brief Funct Genomics, 2015. **14**(5): p. 329-38.
61. Fortna, A., et al., *Lineage-specific gene duplication and loss in human and great ape evolution.* PLoS Biol, 2004. **2**(7): p. E207.
62. Wilson, G.M., et al., *Identification by full-coverage array CGH of human DNA copy number increases relative to chimpanzee and gorilla.* Genome Res, 2006. **16**(2): p. 173-81.
63. Hashimoto, T., S. Maekawa, and S. Miyata, *IgLON cell adhesion molecules regulate synaptogenesis in hippocampal neurons.* Cell Biochem Funct, 2009. **27**(7): p. 496-8.
64. Renstrom, F., et al., *Replication and extension of genome-wide association study results for obesity in 4923 adults from northern Sweden.* Hum Mol Genet, 2009. **18**(8): p. 1489-96.
65. Zhao, J., et al., *The role of obesity-associated loci identified in genome-wide association studies in the determination of pediatric BMI.* Obesity (Silver Spring), 2009. **17**(12): p. 2254-7.
66. Lee, K.T., et al., *Neuronal genes for subcutaneous fat thickness in human and pig are identified by local genomic sequencing and combined SNP association study.* PLoS One, 2011. **6**(2): p. e16356.
67. Bauer, F., et al., *Obesity genes identified in genome-wide association studies are associated with adiposity measures and potentially with nutrient-specific food preference.* Am J Clin Nutr, 2009. **90**(4): p. 951-9.
68. Willer, C.J., et al., *Six new loci associated with body mass index highlight a neuronal influence on body weight regulation.* Nat Genet, 2009. **41**(1): p. 25-34.
69. Wheeler, E., et al., *Genome-wide SNP and CNV analysis identifies common and low-frequency variants associated with severe early-onset obesity.* Nat Genet, 2013. **45**(5): p. 513-7.

70. Magi, R., et al., *Contribution of 32 GWAS-identified common variants to severe obesity in European adults referred for bariatric surgery.* PLoS One, 2013. **8**(8): p. e70735.

71. Gurel, B., et al., *NKX3.1 as a marker of prostatic origin in metastatic tumors.* Am J Surg Pathol, 2010. **34**(8): p. 1097-105.

72. Boender, A.J., et al., *The obesity-associated gene Negr1 regulates aspects of energy balance in rat hypothalamic areas.* Physiol Rep, 2014. **2**(7).

73. Lee, A.W., et al., *Functional inactivation of the genome-wide association study obesity gene neuronal growth regulator 1 in mice causes a body mass phenotype.* PLoS One, 2012. **7**(7): p. e41537.

74. Bard, J.A., et al., *Cloning and functional expression of a human Y4 subtype receptor for pancreatic polypeptide, neuropeptide Y, and peptide YY.* J Biol Chem, 1995. **270**(45): p. 26762-5.

75. Larsen, P.J. and P. Kristensen, *The neuropeptide Y (Y4) receptor is highly expressed in neurones of the rat dorsal vagal complex.* Brain Res Mol Brain Res, 1997. **48**(1): p. 1-6.

76. Parker, R.M. and H. Herzog, *Regional distribution of Y-receptor subtype mRNAs in rat brain.* Eur J Neurosci, 1999. **11**(4): p. 1431-48.

77. Katsuura, G., A. Asakawa, and A. Inui, *Roles of pancreatic polypeptide in regulation of food intake.* Peptides, 2002. **23**(2): p. 323-9.

78. Asakawa, A., et al., *Characterization of the effects of pancreatic polypeptide in the regulation of energy balance.* Gastroenterology, 2003. **124**(5): p. 1325-36.

79. Gettys, T.W., et al., *Insulin-sparing effects of pancreatic polypeptide in congenitally obese rodents.* Pancreas, 1991. **6**(1): p. 46-53.

80. Sainsbury, A., et al., *Synergistic effects of Y2 and Y4 receptors on adiposity and bone mass revealed in double knockout mice.* Mol Cell Biol, 2003. **23**(15): p. 5225-33.

81. Sainsbury, A., et al., *Important role of hypothalamic Y2 receptors in body weight regulation revealed in conditional knockout mice.* Proc Natl Acad Sci U S A, 2002. **99**(13): p. 8938-43.

82. Glusman, G., et al., *The complete human olfactory subgenome.* Genome Res, 2001. **11**(5): p. 685-702.

83. Hasin, Y., et al., *High-resolution copy-number variation map reflects human olfactory receptor diversity and evolution.* PLoS Genet, 2008. **4**(11): p. e1000249.

84. Kambere, M.B. and R.P. Lane, *Co-regulation of a large and rapidly evolving repertoire of odorant receptor genes.* BMC Neurosci, 2007. **8 Suppl 3**: p. S2.

85. Yeomans, M.R., *Taste, palatability and the control of appetite.* Proc Nutr Soc, 1998. **57**(4): p. 609-15.

86. Flier, J.S., *Obesity wars: molecular progress confronts an expanding epidemic.* Cell, 2004. **116**(2): p. 337-50.

87. Choquette, A.C., et al., *Association between olfactory receptor genes, eating behavior traits and adiposity: results from the Quebec Family Study.* Physiol Behav, 2012. **105**(3): p. 772-6.

88. Mariman, E.C., et al., *Olfactory receptor genes cooperate with protocadherin genes in human extreme obesity.* Genes Nutr, 2015. **10**(4): p. 465.

89. Yang, J., et al., *Genome partitioning of genetic variation for complex traits using common SNPs.* Nat Genet, 2011. **43**(6): p. 519-25.

90. Albuquerque, D., et al., *Current review of genetics of human obesity: from molecular mechanisms to an evolutionary perspective.* Mol Genet Genomics, 2015. **290**(4): p. 1191-221.

III. OBJECTIVES

Obesity is a highly heritable complex disorder that has reached epidemic proportions worldwide. Its increasing prevalence and associated morbidity and mortality impose an enormous burden on human health.

Contribute to the unravelling of the genetic basis of this complex disorder is essential and as such the main objective of this thesis.

To address this general aim, the following specific objectives were set out:

- **Clarify the role of selected structural variation in the development of obesity.** This was performed by screening severely overweight or obese children and adolescents for the presence of four BMI-associated CNVs.

- **Genetic follow-up of genes that are impacted by BMI-associated CNVs.** These CNVs might encompass known or novel candidate genes for obesity. We screened for mutations in four genes (*SH2B1, NPY2R, UCP1* and *NPY4R)* that were impacted by the selected CNVs. Furthermore, we decided to screen two additional genes (*NPY* and *FGF21)* that are functionally linked to these candidate genes.

- **Functional characterization of newly identified variants with a possible pathogenic effect by *in vitro* techniques.**

Eventually this study will improve the understanding of the etiology of obesity and may allow the development of early preventive and therapeutic strategies for people at high risk.

IV. RESULTS

PART 1

SH2B1 GENE

GENETIC AND STRUCTURAL VARIATION IN THE SH2B1 GENE IN THE BELGIAN POPULATION.

Evi Aerts[1*], Sigri Beckers[1*], Doreen Zegers[1], Jasmijn K Van Camp[1], Kim Van Hoorenbeeck[2], Guy Massa[3], An Verrijken[4], Ilse L Mertens[4], Stijn L Verhulst[2], Raoul R Rooman[2], Luc F Van Gaal[4] & Wim Van Hul[1]

[1] Centre of Medical Genetics, University of Antwerp, Antwerp, Belgium

[2] Department of Pediatrics, Antwerp University Hospital, Antwerp, Belgium

[3] Department of Pediatrics, Jessa Hospital, Hasselt, Belgium

[4] Department of Endocrinology, Diabetology and Metabolic Diseases, Antwerp University Hospital, Antwerp, Belgium

* These authors contributed equally.

Mol Genet Metab (2015) 115 (4): 193-19

ABSTRACT

Objective Animal studies, genome-wide association and genomic structural variation studies have identified the *SH2B1* gene as a candidate gene for obesity. Therefore, we have designed an extensive mutation and copy number variation (CNV) analysis investigating the prevalence of genetic and structural variation in *SH2B1* in the Belgian population.

Design and Methods In the first part of this study, we performed a mutation screen for variants in the *SH2B1* coding region in 581 obese children and adolescents and 433 healthy, lean individuals with high-resolution melting curve analysis followed by direct sequencing.

In the second part of this study, Multiplex Amplicon Quantification (MAQ) analysis was used to identify CNVs in the distal *SH2B1*-containing chr.16p11.2 region in 421 obese children and adolescents with no developmental delay or behavioral phenotype.

Results Mutation analysis resulted in the identification of fifteen rare non-synonymous heterozygous variants. Several of these were found both in lean and obese subjects, suggesting these are neutral polymorphisms. However, six private, heterozygous, non-synonymous variations were present in obese children only. Furthermore, we also identified six missense variants solely in lean individuals.

CNV analysis could not identify carriers of the distal 16p11.2 deletion in our population.

Conclusion Our mutation analysis has demonstrated that variation in the *SH2B1* gene is frequent in both lean and obese groups, with distinctive variations being present on either side of the weight spectrum. Although the equal variation frequency does not immediately support disease causality, it cannot be excluded that some variations are weight-increasing or –decreasing. Further functional testing of the variants will be necessary to fully understand the impact of these variants on SH2B1 .

We were not able to detect carriers of the distal 16p11.2 deletion in our study population. As we excluded patients with developmental or behavioral problems, we suggest that in addition to obesity, the distal deletion might predispose for these traits. Further characterization of the phenotype is therefore necessary to clearly identify the phenotype of the distal 16p11.2 microdeletion syndrome.

INTRODUCTION

Maintaining a balance between energy intake and expenditure is vital in the process of energy homeostasis. Once the balance shifts to either end of the spectrum, health problems arise. The most prevalent health issue caused by a disturbance in energy balance is obesity, with an estimated 500 million obese adults worldwide [1]. Because of the burden this disease poses on an individual's health and its ever increasing costs for national health budgets, research efforts searching for factors contributing to the disease's etiology have intensified in the past decades. This has led to the recognition of the leptin-melanocortin signaling pathway as a key regulator of food intake and energy expenditure [2]. Mutations in several genes coding for proteins involved in this pathway have proven to be responsible for early-onset, severe obesity in populations worldwide. Melanocortin-4 receptor (*MC4R*) gene mutations have shown to be the most prevalent genetic cause of this so-called monogenic obesity known to date, with over 150 distinct mutations reported so far [3]. In contrast, only six different mutations in the leptin gene have been discovered, making leptin deficiency a rare form of monogenic obesity [4-10]. Leptin deficiency is currently, however, still the only treatable type of monogenic obesity [6, 11, 12].

A relatively new candidate gene involved in the leptin-melanocortin signaling pathway is the *SH2B1* gene encoding the Src homology 2B adaptor protein 1, a protein involved in a variety of signalling pathways. SH2B1 was initially identified as a JAK2-interacting protein through a yeast two-hybrid screen [13]. Further investigations demonstrated that SH2B1 also enhanced JAK2 activation after binding of leptin to its receptor [14, 15], thus demonstrating that the protein is a positive regulator of leptin signalling and suggesting that it might be implicated in obesity.

The generation of an *Sh2b1* null mouse clearly established the protein's involvement in obesity, as these animals were obese with severe hyperphagia, hyperleptinemia, hyperinsulinemia, hyperlipidemia, hepatic steatosis, hyperglycemia and glucose intolerance. They also exhibited an increased expression of the orexigenic proteins NPY and AgRP [16, 17]. Hypothalamic restoration of *Sh2b1* rescues the hyperphagia, obesity, hyperglycemia and glucose intolerance seen in the knock-out mice. These $Sh2b1^{TgKO}$ mice also showed improved leptin signalling and protection against high fat diet induced obesity [18]. Recently, the SH2B1 expression pattern was investigated in normal and obese mice. Compared with normal mice of the same age, SH2B1 mRNA expression in the obese mice was decreased with increased levels of serum leptin and fasting insulin levels.

Further genetic evidence has also emerged, strengthening the case for a role of SH2B1 in the development of human obesity. The first additional evidence came from genome-wide

association studies, which identified the A484T variant in *SH2B1* (rs7498665) as one of the top-hits for obesity [19, 20]. The rs7498665 variant allele increases obesity risk by about 11-26% in these studies, a finding which we could also replicate in our Belgian population [21].

For the first part of this study we hypothesized that loss-of-function mutations in SH2B1 could also play a role in the pathogenesis of early-onset obesity. Two recent reports estimate that about 1% of obese children carry a non-synonymous SH2B1 variation (excluding the known polymorphism T484A) [22, 23]. Furthermore, it also seems likely that gain-of-function mutations could lead to optimized weight regulation through an improved response to leptin signalling. Therefore we designed an elaborate mutation analysis, including not only a set of obese children and adolescents but also a cohort of lean adults. This set-up will thus allow us to study the effect of genetic variation in SH2B1 at both sides of the weight spectrum.

In the second part of this study, we investigated the importance of copy number variation (CNV) in the *SH2B1*-containing 16p11.2 region. Apart from the proximal 16p11.2 deletion (593kb) [24] which is already recognized as a microdeletion syndrome associated with developmental delay, autism spectrum disorder, epilepsy and obesity, less is known about the smaller distal 220 kb deletion. In 2010, CNV-analysis in a cohort of Caucasian patients with severe early-onset obesity identified the distal *SH2B1*-containing deletion on 16p11.2 and found it to be associated with highly penetrant familial early-onset obesity, hyperphagia and severe insulin resistance. Developmental delay was only attributed to the extended deletions that also included the proximal deletion. These findings are partly in contrast to other studies that identified the distal SH2B1-containing deletion in patients with a variety of traits often in combination with obesity. The latter studies suggested that in addition to obesity, the distal 220kb deletion strongly prediposes to developmental delay and intellectual disability [25, 26]. The conflicting data regarding the phenotype associated with the distal deletion prompted us to further investigate the prevalence of the distal deletion in a Belgian population of obese children and adolescents with no developmental or behavioral problems.

METHODS AND PROCEDURES

Study population

Mutation screening in the coding region of the *SH2B1* gene was performed in a population of 581 severely overweight and obese children and adolescents (269 boys, 312 girls), and 433

healthy lean adults. For CNV-analysis we screened 421 overweight and obese children and adolescents (169 boys, 252 girls). Overweight/obese children and adolescents (adolescent if age ≥ 12 yrs) are unrelated and were recruited at the Child Obesity Clinics from the Antwerp University Hospital (Antwerp) and Jessa Hospital (Hasselt), both in Belgium. The Flemish Growth Charts 2004 were used to identify overweight and obese individuals. The percentile lines that cross BMI 25 and 30 kg/m^2 at 18 years of age on the Flemish age- and sex-specific BMI growth curves were used as cut-off for the diagnosis of overweight and obesity respectively [27, 28]. Patients with mutations in *Leptin* and the *Melanocortin-4* receptor causing known monogenic obesity have been excluded from the screening sample. Population characteristics are summarized in Table 1 en 2.

Our lean population, used for mutation screening, consisted of 433 healthy, lean adults (18.5 kg/m^2 ≤ BMI < 25 kg/m^2; 146 men, 287 women; Table 1) recruited among employees from the Antwerp University Hospital and the University of Antwerp, and among couples seeking prenatal counselling at the Centre for Medical Genetics (due to increased triple test or high maternal age).

All participants gave their written informed consent. The study protocol was approved by the local ethics committee.

Table 1: Characteristics of the population used for mutation analysis

	Obese subjects		Lean adults
	Children	Adolescents	
N	272	309	433
Male/female	139/133	130/179	146/287
Age (years)	8.6 ± 2.4	15.2 ± 2.2	35.5 ± 7.4
Weight (kg)	52.3 ± 17.5	95.2 ± 21.6	65.2 ± 8.7
Height (m)	1.36 ± 0.20	1.65 ± 0.20	1.71 ± 0.10
BMI (kg/m^2)	27.2 ± 4.1	34.2 ± 5.0	22.2 ± 1.7
BMI Z-score	2.7 ± 0.6	2.5 ± 0.4	N.A.

Mean value ± standard deviation is shown for all parameters, except N and gender distribution (absolute numbers). N.A., not applicable

Table 2: Characteristics of the population used for CNV analysis

	Obese subjects	
	Children	Adolescents
N	177	244
Male/female	69/108	100/144
Age (years)	8.7 ± 2.4	15.6 ± 2.3
Weight (kg)	53.3 ± 17.2	97.9 ± 23.5
Height (m)	1.31 ± 0.36	1.63 ± 0.30
BMI (kg/m^2)	26.0 ± 6.8	34.6 ± 6.8
BMI Z-score	2.4 ± 0.8	2.5 ± 0.5

Mean value ± standard deviation is shown for all parameters, except N and gender distribution (absolute numbers).

Anthropometry

Height was measured to the nearest 0.1 cm. Weight was measured on a digital scale to the nearest 0.1 kg. Body mass index (BMI) was calculated for all individuals as weight (in kg) over height (in m) squared. For children and adolescents, BMI Z-scores were calculated using the data from the Flemish Growth Curves 2004 [27].

Mutation analysis of SH2B1

Mutation analysis was performed on genomic DNA isolated from whole blood by standard procedures [29]. The *SH2B1* gene has 5 different transcripts (GenBank NM_001145795.1, NM_001145796.1, NM_001145797.1, NM_001145812.1 and NM_015503.2), encoding 3 protein isoforms. We designed 11 PCR amplicons to screen all *SH2B1* coding exons (including intron-exon boundaries) for mutations with high-resolution melting curve analysis (HRM) using the Lightcycler LC480 Real-Time PCR System (Roche, Penzberg, Germany). Real-time PCR was performed in a reaction volume of 10 µl. The amplification mixture included 10 ng of template DNA, 1.5 mM $MgCl_2$, 0.15 mM dNTP's, 500 nM primers, 0.015 U/µl GoTaq (Promega Corporation, Madison, WI, USA), 1x GoTaq buffer and the saturating dye LCGreen+ (Idaho Technology, Salt Lake City, UT, USA) in a 0.5x concentration. Following amplification, samples were denaturated at 95°C, renaturated at 40°C and then melted between 60°C and 90°C while constantly measuring fluorescence. Samples with resulting melting curves deviating from wild type were sequenced. Sequencing was performed with ABI BigDye Terminator v1.1 Cycle Sequencing kits on an ABI Prism Genetic Analyzer 3130xl (Applied Biosystems Inc, Foster City,

CA, USA). Identified mutations were confirmed by resequencing. Primer sequences for HRM and sequencing are available upon request.

Prediction programs

In silico analysis to evaluate possible pathogenic effects of the identified non-synonymous variants on protein function was performed using four different prediction programs: PolyPhen-2 [30], SIFT [31], SNPs&Go [32], and MutPred [33].

CNV-analysis

a) MAQ

Genomic DNA was isolated from whole blood by standard procedures [29]. To detect and analyse CNVs, we used Multiplex Amplicon Quantification assays (MAQ), consisting of a multiplex PCR based amplification of fluorescently labeled target and reference amplicons, followed by capillary electrophoresis and fragment analysis. MAQ-assays were performed on 20 ng of genomic DNA. Target primers were designed with the Flagged-You MAQ primer design tool and its concentrations were optimized according to the manufacturer's protocol. Final primer concentrations varied between 0.04 and 0.79 µM in the Flagged-you MAQ PCR mix which was prepared according to the manufacturer's protocol [34]. Target region was set at chr 16:28858666-28895886 (NCBI Build GRCh37.p13) and comprises both *SH2B1* and *ATPase 2A1 (ATP2A1).* Two negative control samples from lean individuals were included in each experiment for accurate normalisation. One positive control sample with known 16p11.2 deletion was used for optimization and quality control. Standard PCR cycle conditions were followed. MAQ primer sequences are available upon request.

b) MLPA

Multiplex Ligation-dependent Probe Amplification analysis was optimized as an alternative technique to confirm samples with deviating MAQ dosage quotient (DQ). For MLPA analysis, the SALSA reagent Kit was used (MRC-Holland). Target and reference primers were designed by H-MAPD, a probe design suite for MLPA-assays [35]. In each experiment two control samples were included, allowing accurate normalisation. One positive control sample with known deletion was used for optimization. MLPA primer sequences are available on request. MLPA was performed according to the manufacturer's protocol [36].

Capillary electrophoresis and fragment analysis

Fragment analysis of the resulting MAQ- and MLPA-PCR products was performed on an ABI Prism Genetic Analyzer 3130xl (Applied Biosystems Inc, Foster City, CA, USA). The generated raw MAQ data were analysed using the MAQ-S software package (Multiplicom) [34], designed to calculate and visualise the dosage quotient which reflects the copy number of each target amplicon. A dosage quotient of 1.25 -1.75 was considered indicative of a duplication, a DQ of 0.25-075 was indicative of a deletion.

For MLPA, the Genemapper software (Applied Biosystems) and the Excell tool Fluodosage P034/35v.2 were used to calculate the DQ and to visualise the electropherograms.

Statistical analysis

For rs7498665, Hardy Weinberg equilibrium (HWE) was calculated using the HWE tool from LINKUTIL [37]. Comparison of genotype and allele frequencies among cases and controls was performed using Pearson's chi square analysis. Significance level was set at p=0.05. All statistical analyses were performed using SPSS version 20.0 (SPSS, Chicago, IL, USA).

RESULTS

Detection of novel mutations in SH2B1

Screening of all coding exons and intron-exon boundaries in the different transcripts of *SH2B1* identified twenty-four different heterozygous coding variants in our populations: eight synonymous variants, one common polymorphism (T484A) and fifteen non-synonymous variations (Table 3). Of the latter, three variants were identified both in lean and obese subjects, six were found exclusively in obese cases and another six were seen only in lean adults. In total, we found non-synonymous *SH2B1* variations (excluding T484A) at a frequency of 2.8% in obese children/adolescents (= 16/581), and 3.9% in lean adults (= 17/433), showing that the mutation burden is high in both groups. The αA663V, αP689L and αE713D variations were all rare variants found both in obese and lean individuals (Table 3), suggesting that these are neutral polymorphisms. Six obese children each carried a private, heterozygous, non-synonymous variation: P19L, R43S, T175N, S188L, αA667V and αA672T (Table 3).

Table 3: Coding variants identified in obese and lean individuals.

Variation		Obese	Lean	SNP
gene	protein			Rs7498665
c.46c>t	P16S	-	1	AG
c.56c>t	P19L	1	-	AA
c.127c>a	R43S	1	-	AA
c.199c>t	R67C	-	1	AA
c.524c>a	T175N	1	-	GG
c.563c>t	S188L	1	-	AA
c.613g>a	G205R	-	1	AA
c.1298g>a	R433H	-	1	AG
c.1846t>c	S616P	-	1	AG
βc.1970c>t / γc.2020c>t	βT656I/γP674S	-	1	AG
αc.1988c>t	αA663V	8	9	(*)
αc.2000c>t	αA667V	1	-	AA
αc.2014g>c	αA672T	1	-	GG
αc.2066c>t	αP689L	1	2	AA/AG,AG
αc.2139g>c	αE713D	1	1	AG/AG

Position of variants on gene and protein levels is shown, as well as their frequency in the obese and lean groups (in absolute numbers). Numbering on gene level is based on cDNA sequence (GenBank NM_001145795.1, NM_001145796.1, NM_001145797.1, NM_001145812.1 and NM_015503.2), following the recommendations by the Human Genome Variation Society [38]. The protein isoforms in which the variant is present are also indicated. Genotype of the rs7498665 SNP is specified.
(*) Among A663V-carriers, the AA- genotype was identified in 5 control and 5 patient samples, the AG-genotype was identifed in 4 control and 3 patient samples.

We identified six other heterozygous, non-synonymous variants solely in lean adults: P16S, R67C, G205R, R433H, S616P and βT656I/γP674S. [19, 20]. An overview of the results is given in table 3. The presence of rs7498665 among variant carriers was specified as the SNP allele increases obesity risk by about 11%. This association was already confirmed in our study population [21]. Developmental delay or intellectual disability was not reported for any of the screened samples. All were proven to be unrelated.

In silico analysis

Prediction programs indicate that most of the variants found in either obese or lean individuals are suspected not to have a functional effect, except for R67C. For this variation, 3 out of 4 programs indicate that an effect on SH2B1's function is to be expected (Table 4).

Table 4: *In silico* analysis for all non-synonymous variants.

Variant	PolyPhen-2	SIFT	SNPs&Go	MutPred
P16S	Benign	Tolerated	Neutral	Neutral
P19L	Benign	Tolerated	Neutral	Neutral
R43S	Probably damaging	Tolerated	Neutral	Possibly deleterious
R67C	Probably damaging	Not tolerated	Neutral	Possibly deleterious
T175N	Benign	Not tolerated	Neutral	Neutral
S188L	Benign	Tolerated	Neutral	Neutral
G205R	Possibly damaging	Tolerated	Neutral	Neutral
R433H	Benign	Tolerated	Neutral	Neutral
S616P	Benign	Not tolerated	Disease	Neutral
αA667V	Benign	Tolerated	Neutral	Neutral
αA672T	Benign	Not tolerated	Neutral	Neutral
βT656I	Probably damaging	Tolerated	N.D.	Neutral
γP674S	Possibly damaging	Not tolerated	N.D.	Neutral

N.D., not determined

Detection of CNVs in SH2B1

421 severely overweight or obese children and adolescents were screened for structural variation in *SH2B1* with MAQ. We could not detect any copy number gain or loss in the target region. 12 samples with a slightly deviating DQ (1-1.25 and 0.75-1) were re-analyzed by MLPA analysis which confirmed that no deletions or duplications were present.

DISCUSSION

Genetic evidence coming from animal studies, genome-wide association studies and studies looking for structural variation has clearly implicated the *SH2B1* gene in the etiology of obesity. Moreover, the SH2B1 protein has been shown to act as a positive regulator of the leptin-melanocortin pathway, one of the most important regulators of food intake. Although the *SH2B1* gene has been identified as a novel candidate gene for obesity, no data were available on the frequency of *SH2B1* mutations in obese and lean individuals and on the prevalence of the distal *SH2B1*-containing deletion in an obese population with no developmental or behavioral problems. In contrast to the proximal 16p11.2 deletion syndrome (593kb), there are only a few previous reports which investigated the distal deletion and characterized the phenotype of its

carriers. The first reported cases were a patient and his father who both had developmental problems and minor dysmorphisms. Additional growth parameters were not reported [39].

In 2009, Bochukova *et al.* investigated the contribution of the distal deletion in 1219 severely obese patients and reported a frequency of 0.41% (5 out of 1219) in this population. Two of these patients were carriers of a larger deletion that included both the distal and the proximal deletion and had mild developmental delay in addition to severe obesity. The developmental delay was attributed to the proximal deletion as the other three patients were only carriers of the distal deletion and were not known to have behavioral problems. Limited data on their phenotype were published. The prevalence among controls was reported to be significantly lower 0.027% (2 out of 7366) when compared to patients.

In 2011, an overview of all published cases carrying the distal deletion was presented. Of all reported patients, about one third had developmental problems and almost 23% had an unusual facial morphology. Other common traits were obesity, behavioral problems, seizures, autism and speech delay. Few cases were reported since then with a phenotypical spectrum that ranges from developmental delay, behavioral problems and obesity to a normal phenotype.

In the first part of this study we performed an elaborate mutation analysis on both obese children/adolescents and lean adults to determine the frequency of *SH2B1* variation in these groups. We identified fifteen rare non-synonymous variants: three of these were found in both obese and lean individuals, six variations were private to the obese population and another six were solely found in lean adults. These results clearly demonstrate that variation in SH2B1 is frequent in both the obese and lean population, and thus would suggest that these variants are unlikely to be causative for the obese phenotype . Furthermore, we would also expect to find nonsense mutations in the obese population since deletion of SH2B1 has been reported to contribute to obesity [24]. However, no proven loss-of-function mutations were identified here. In silico analysis for the twelve variants exclusively present in either of the two weight-groups also indicates that these variants do not have a major impact on the protein's function, except for R67C (Table 3). However, we cannot exclude that the mutations found solely in obese or lean subjects, can have different effects on the protein, for example by distinctive localization in specific functional domains of the protein.

The different genotypes of rs7498665 were specified for all rare variant carriers as previous studies in our research group confirmed that the rs7498665 variant allele increases obesity risk by about 11-26%. [19-21]. A combination of a rare variant (T175 and αA672T) with the

homozygote rs749665 genotype might have an aggregated effect on the SH2B1 protein function and obesity.

When looking at the localization in the SH2B1 protein (Figure 1), it is remarkable that mutations mostly occur outside of the known functional domains of the protein, with the exception of R43S, R67C and S616P. These latter mutations are expected to disturb the dimerization domain (R43S and R67C) and SH2 domain (S616P), respectively. However, no clear distinction between mutations found in the discrete weight groups is possible on the basis of their localization.

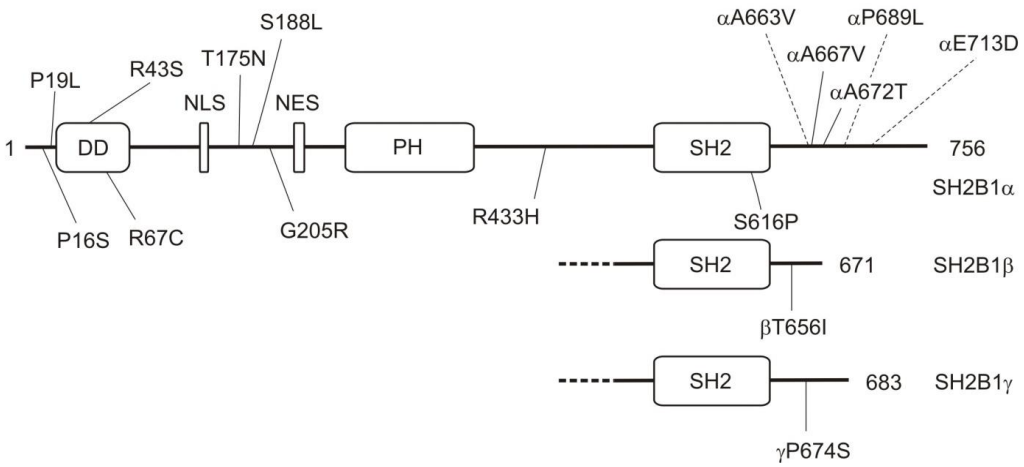

Figure 1: Structure of the SH2B1 protein and location of the missense variations.
A schematic representation of the SH2B1 protein structure is shown for all three isoforms. Functional domains are indicated by boxes. Variants found in obese subjects are shown above the protein structure, while variations indicated below the protein structure were found in lean adults. The three variants found in both groups are indicated by dotted lines. DD, dimerization domain; NLS, nuclear localization signal; NES, nuclear export signal; PH, pleckstrin homology domain; SH2, Src homology 2 domain.

Combining these observations with the *in silico* results (Table 3), would identify R67C as the most promising mutation identified here. Since this mutation was found in a lean woman, we speculate that the presence of an extra cysteine in this region could provide additional stability to the SH2B1 dimer, resulting in better translation of the leptin signal to the downstream effector proteins. However, functional testing of this hypothesis is essential in understanding the true effect of this mutation on SH2B1 dimerization and functioning. Furthermore, we also see that the three variants that are present in both weight categories (αA663V, αP689L and αE713D) are the only variations that are exclusively expressed in the α-isoform of the protein. This would

suggest that this isoform is less crucial for SH2B1's function in weight regulation. Previous research has already indicated that the β-isoform seems most crucial for the weight-regulating function of SH2B1, since neuron-specific restoration of SH2B1β corrected the obesity seen in *Sh2b1* KO mice [18].Most of the variants reported here have not been described previously. Only four variants (T175N, R433H, S616P and αA663V) have been reported by the 1000 Genomes project [40], the NCBI Exome Sequencing Project [41], and/or the ClinSeq project [42] at a minor allele frequency of 0.001-0.003. This would again suggest that these are neutral polymorphisms although we cannot exclude that the variation carriers identified in these projects were overweight/obese since no phenotypical data is available. For αA663V, our own data also indicate that this variant is unlinked to obesity since it was found at similar frequency in both obese and lean subjects. However, a recent study investigated the functional effect of the variation and reported it to impair the ability of *SH2B1* to enhance growth hormone induced cell motility although the variation did not cosegregate with obesity in the families [43].

Recently, two other publications describe mutation screening of *SH2B1* in obese subjects [44, 45] and identify potential pathogenic *SH2B1* variants in obese patients at a frequency of about 1%, which is in line with our results. Of the variants reported here, only T175N and βT656I/γP674S were also reported in these earlier studies. T175N was previously identified in both obese and lean subjects [44, 45]. βT656I/γP674S, however, was found in a lean adult in our study, while Volckmar *et al.* reported this variant in an obese subject only [45]. This would thus suggest that these variations are polymorphisms unlinked to weight regulation.

Remarkably, we report a similar variation frequency in lean adults, while the two previous studies found no private variants in lean subjects. In the study by Volckmar *et al.*, this is explained by their study set-up. They screened the entire *SH2B1* gene region in 95 obese subjects and subsequently only genotyped the variants found in this pilot project in their large populations of both obese and lean subjects [45]. Doche and colleagues, however, also screened a group of 500 control subjects for mutations in *SH2B1* and could not identify any non-synonymous variants in this group, although we would expect to find about five mutation carriers based on our current results [44]. This could possibly be explained by a difference in recruitment and/or ethnicity, but since no details on the inclusion/exclusion criteria for their control population was reported, no clear explanation can be given at the moment.

In the second part of this study, we wanted to investigate the prevalence of the distal 16p11.2 deletion in a population of children and adolescents with severe early-onset obesity, but with no developmental or behavioral problems.

In contrast to the study of Bochukova *et al.* which reported a prevalence of 0.41% in a population of patients with early-onset obesity, we were not able to detect any CNV. As we specifically excluded patients with developmental delay, we suggest that this is an important factor of difference in comparison with the previously reported study populations. As the prevalence of this deletion in patients with no other traits beside obesity might be very low, we suggest that in addition to obesity the distal deletion might strongly predispose for other traits in combination with obesity.

Although it seems unlikely that *SH2B1* variants contribute to the development of obesity, *SH2B1* deficiency might have an important role in the development of obesity in patients carrying the distal deletion. Additional genes in the 220 kb region might be responsible for the other phenotypic characteristics: *ATXN2L, TUFM, ATP2A1, RABEP2, CD19, SPNS1, NFATC2IP* and *LAT*. Some of these genes are already reported to be involved in neurological diseases (*TUFM*, *ATP2A1*) and immunity (*CD19, NFATC2IP, LAT)*.

In conclusion, we report here the presence of potentially functional missense mutations in SH2B1 in about 1% of both obese and lean individuals. Based on the equal frequency of mutations in both weight groups and their scattered localization in the protein, it seems unlikely that SH2B1 variants contribute to the development of obesity. However we cannot exclude distinct effect of the variations found exclusively in the obese or lean cohorts. Functional experiments determining the effect of the variants reported here on SH2B1 expression, localization, dimerization and activity are thus essential.

CNV analysis was not able to detect structural *SH2B1* variation in a our obese population with no other traits beside obesity. We suggest that the other genes in the 220kb deleted region might be responsible for the developmental and behavioral problems in its carriers. In the future it will be important to investigate the prevalence and phenotype of the distal deletion in an extended study population in which patients present a variability of abnormal phenotypes including developmental delay, behavioral problems, mild facial dysmorphology and intellectual disability. Detailed clinical information for all of these patients by a clinical geneticist might be indispensable to define a specific recognizable phenotype.

ACKNOWLEDGEMENTS

This study was supported by an Interuniversity Attraction Pole Project (Phase VII project 43, BELSPO) and a TOP project from the University of Antwerp. SB is a postdoctoral researcher of the "Fund for Scientific Research Flanders" (FWO-Vlaanderen). EA holds a PhD Grant for

Strategic Basic Research from the "Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen)".

SB designed the study, carried out experiments and wrote the manuscript. EA performed experiments and wrote the manuscript. DZ, JKVC and LVG contributed to the discussion. KVH, GM, AV, ILM, SLV, RRR, and LVG recruited and phenotyped subjects. WVH designed the study and wrote the manuscript. All authors read and approved the final version of the manuscript.

REFERENCES

1. WHO. *Fact Sheet N°311*. 2013 22 July 2013]; Available from: http://www.who.int/mediacentre/factsheets/fs311/en/.
2. Beckers, S., et al., *The role of the leptin-melanocortin signalling pathway in the control of food intake*. Crit Rev Eukaryot Gene Expr, 2009. **19**(4): p. 267-87.
3. Tao, Y.X., *The melanocortin-4 receptor: physiology, pharmacology, and pathophysiology*. Endocr Rev, 2010. **31**(4): p. 506-43.
4. Montague, C.T., et al., *Congenital leptin deficiency is associated with severe early-onset obesity in humans*. Nature, 1997. **387**(6636): p. 903-8.
5. Strobel, A., et al., *A leptin missense mutation associated with hypogonadism and morbid obesity*. Nat Genet, 1998. **18**(3): p. 213-5.
6. Gibson, W.T., et al., *Congenital leptin deficiency due to homozygosity for the Delta133G mutation: report of another case and evaluation of response to four years of leptin therapy*. J Clin Endocrinol Metab, 2004. **89**(10): p. 4821-6.
7. Mazen, I., et al., *A novel homozygous missense mutation of the leptin gene (N103K) in an obese Egyptian patient*. Mol Genet Metab, 2009. **97**(4): p. 305-8.
8. Fischer-Posovszky, P., et al., *A new missense mutation in the leptin gene causes mild obesity and hypogonadism without affecting T cell responsiveness*. J Clin Endocrinol Metab, 2010. **95**(6): p. 2836-40.
9. Fatima, W., et al., *Leptin deficiency and leptin gene mutations in obese children from Pakistan*. Int J Pediatr Obes, 2011. **6**(5-6): p. 419-27.
10. Saeed, S., et al., *High prevalence of leptin and melanocortin-4 receptor gene mutations in children with severe obesity from Pakistani consanguineous families*. Mol Genet Metab, 2012. **106**(1): p. 121-6.
11. Farooqi, I.S., et al., *Effects of recombinant leptin therapy in a child with congenital leptin deficiency*. N Engl J Med, 1999. **341**(12): p. 879-84.
12. Farooqi, I.S., et al., *Beneficial effects of leptin on obesity, T cell hyporesponsiveness, and neuroendocrine/metabolic dysfunction of human congenital leptin deficiency*. J Clin Invest, 2002. **110**(8): p. 1093-103.
13. Rui, L., et al., *Identification of SH2-Bbeta as a substrate of the tyrosine kinase JAK2 involved in growth hormone signaling*. Mol Cell Biol, 1997. **17**(11): p. 6633-44.

14. Duan, C., M. Li, and L. Rui, *SH2-B promotes insulin receptor substrate 1 (IRS1)- and IRS2-mediated activation of the phosphatidylinositol 3-kinase pathway in response to leptin.* J Biol Chem, 2004. **279**(42): p. 43684-91.
15. Li, Z., et al., *SH2B1 enhances leptin signaling by both Janus kinase 2 Tyr813 phosphorylation-dependent and -independent mechanisms.* Mol Endocrinol, 2007. **21**(9): p. 2270-81.
16. Duan, C., et al., *Disruption of the SH2-B gene causes age-dependent insulin resistance and glucose intolerance.* Mol Cell Biol, 2004. **24**(17): p. 7435-43.
17. Ren, D., et al., *Identification of SH2-B as a key regulator of leptin sensitivity, energy balance, and body weight in mice.* Cell Metab, 2005. **2**(2): p. 95-104.
18. Ren, D., et al., *Neuronal SH2B1 is essential for controlling energy and glucose homeostasis.* J Clin Invest, 2007. **117**(2): p. 397-406.
19. Willer, C.J., et al., *Six new loci associated with body mass index highlight a neuronal influence on body weight regulation.* Nat Genet, 2009. **41**(1): p. 25-34.
20. Thorleifsson, G., et al., *Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity.* Nat Genet, 2009. **41**(1): p. 18-24.
21. Beckers, S., et al., *Replication of the SH2B1 rs7498665 association with obesity in a Belgian study population.* Obes Facts, 2011. **4**(6): p. 473-7.
22. Doche, M.E., et al., *Human SH2B1 mutations are associated with maladaptive behaviors and obesity.* J Clin Invest, 2012. **122**(12): p. 4732-6.
23. Volckmar, A.L., et al., *Mutation screen in the GWAS derived obesity gene SH2B1 including functional analyses of detected variants.* BMC Med Genomics, 2012. **5**: p. 65.
24. Bochukova, E.G., et al., *Large, rare chromosomal deletions associated with severe early-onset obesity.* Nature, 2010. **463**(7281): p. 666-70.
25. Su, T., et al., *[Expression change of SH2B1, SOCS3, PTP1B and NPY in mice hypothalamus and its relation with obesity].* Zhong Nan Da Xue Xue Bao Yi Xue Ban, 2014. **39**(1): p. 43-8.
26. Barge-Schaapveld, D.Q., et al., *The atypical 16p11.2 deletion: a not so atypical microdeletion syndrome?* Am J Med Genet A, 2011. **155A**(5): p. 1066-72.
27. Roelants, M., R. Hauspie, and K. Hoppenbrouwers, *References for growth and pubertal development from birth to 21 years in Flanders, Belgium.* Ann Hum Biol, 2009. **36**(6): p. 680-94.

28. Hauspie, R. and M. Roelants. *Flemish Growth Charts 2004*. 2004 22 July 2013]; Available from: http://www.vub.ac.be/groeicurven.
29. Miller, S.A., D.D. Dykes, and H.F. Polesky, *A simple salting out procedure for extracting DNA from human nucleated cells.* Nucleic Acids Res, 1988. **16**(3): p. 1215.
30. Adzhubei, I.A., et al., *A method and server for predicting damaging missense mutations.* Nat Methods, 2010. **7**(4): p. 248-9.
31. Ng, P.C. and S. Henikoff, *Predicting deleterious amino acid substitutions.* Genome Res, 2001. **11**(5): p. 863-74.
32. Calabrese, R., et al., *Functional annotations improve the predictive score of human disease-related mutations in proteins.* Hum Mutat, 2009. **30**(8): p. 1237-44.
33. Li, B., et al., *Automated inference of molecular mechanisms of disease from amino acid substitutions.* Bioinformatics, 2009. **25**(21): p. 2744-50.
34. Multiplicom. Multiplicom the multiplex experts.; Available from: http://www.multiplicom.com.
35. Zhi, J., *MAPD: a probe design suite for multiplex ligation-dependent probe amplification assays.* BMC Res Notes, 2010. **3**: p. 137.
36. MRC-Holland. Available from: http://www.mlpa.com.
37. J, O., *Documentation to linkage utility programs.* . 2013.
38. den Dunnen, J.T. and S.E. Antonarakis, *Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion.* Hum Mutat, 2000. **15**(1): p. 7-12.
39. Bijlsma, E.K., et al., *Extending the phenotype of recurrent rearrangements of 16p11.2: deletions in mentally retarded patients without autism and in normal individuals.* Eur J Med Genet, 2009. **52**(2-3): p. 77-87.
40. Abecasis, G.R., et al., *A map of human genome variation from population-scale sequencing.* Nature, 2010. **467**(7319): p. 1061-73.
41. NHLBI. *NHLBI GO Exome Sequencing Project - Exome Variant Server*. 2013 14 March 2013]; Available from: http://evs.gs.washington.edu/EVS/.
42. Biesecker, L.G., et al., *The ClinSeq Project: piloting large-scale genome sequencing for research in genomic medicine.* Genome Res, 2009. **19**(9): p. 1665-74.
43. Pearce, L.R., et al., *Functional characterisation of obesity-associated variants involving the alpha and beta isoforms of human SH2B1.* Endocrinology, 2014: p. en20141264.

44. Doche, M.E., et al., *Human SH2B1 mutations are associated with maladaptive behaviors and obesity.* J Clin Invest, 2012. **122**(12): p. 4732-6.

45. Volckmar, A.L., et al., *Mutation screen in the GWAS derived obesity gene SH2B1 including functional analyses of detected variants.* BMC Med Genomics, 2012. **5**(1): p. 65.

PART 2

NPY AND ITS RECEPTORS

COMBINATION OF CNV ANALYSIS AND MUTATION SCREENING INDICATES AN IMPORTANT ROLE FOR THE NPY4R GENE IN HUMAN OBESITY.

Evi Aerts[1], Sigri Beckers[1], Doreen Zegers[1], Kim Van Hoorenbeeck[2], Guy Massa[3], An Verrijken[4], Stijn L Verhulst[2], Luc F Van Gaal[4] & Wim Van Hul[1]

[1] Centre of Medical Genetics, University of Antwerp, Antwerp, Belgium

[2] Department of Pediatrics, Antwerp University Hospital, Antwerp, Belgium

[3] Department of Pediatrics, Jessa Hospital, Hasselt, Belgium

[4] Department of Endocrinology, Diabetology and Metabolism, Antwerp University Hospital, Antwerp, Belgium

Obesity (Silver Spring) (2016) 24 (4): 970-976

ABSTRACT

Objective Genomewide copy number variation (CNV) analyses have associated the 10q11.22 CNV with obesity. As the *NPY4R* gene is the most interesting candidate gene in this region, we hypothesized that both genetic and structural variation in *NPY4R* might be implicated in the pathogenesis of obesity.

Methods In the first part of this study, we screened 326 children and adolescents with obesity, and 298 healthy lean individuals for CNV in the *NPY4R*-containing chr.10q11.22 region. In the second part of this study, we performed a mutation screen for variants in the *NPY4R* coding region in 356 children and adolescents with obesity, and 337 healthy lean adults.

Results In our CNV analysis we demonstrated a significantly higher frequency of *NPY4R* containing 10q11.22 CNV loss in the patient population (p=0.0003), while CNV gain in this region was more prevalent in our control population (p=0.031). Mutation analysis resulted in the identification of fifteen rare non-synonymous heterozygous variants. For two variants that could only be identified in our patient population, we were able to demonstrate receptor dysfunction and thus a pathogenic effect

Conclusion In conclusion, these data support an essential role for genetic and structural variation within the *NPY4R* gene in the pathogenesis of obesity.

INTRODUCTION

Obesity is a highly heritable complex disorder that has reached epidemic proportions worldwide. Its increasing prevalence and associated morbidity imposes an enormous burden on human health [1]. Although an obesogenic environment and a passive lifestyle are main contributors to obesity prevalence, numerous studies have shown that 40-70% of the interindividual variability in BMI is attributed to genetic factors [2, 3]. Given the estimated heritability, extensive genome-wide association studies (GWASs) have been performed to identify common single-nucleotide polymorphisms (SNPs) underlying the heritable risk for human obesity [4, 5]. Recently, genome-wide CNV surveys have associated a number of CNVs with early-onset extreme obesity [6-8]. The genes that are impacted by these CNVs might encompass known or novel interesting candidate genes for follow-up studies. The *Neuronal growth regulator 1 (NEGR1)* gene was one of the first common variants that has been associated with BMI due to the existence of two distinct deletion alleles upstream of *Negr1*. Although its function in the etiology of obesity still has to be determined, altered expression of the gene contributes to the regulation of BMI [9]. In 2014, Falchi et al. identified significant associations of a multi-allelic CNV encompassing the *salivary amylase* gene (*AMY1)* with BMI and obesity, providing the first link between carbohydrate metabolism and BMI [10]. Another important CNV region that warrants further follow-up, is the 10q11.22 CNV region. In 2009, Sha *et al.* reported that CNV at 10q11.22 (from 46.36 Mb to 46.56Mb) was associated with BMI in the Chinese population. Copy number loss would lead to a 12.4% higher BMI value and copy number gain to a 5.4% lower BMI value when compared to normal diploid subjects. Regression analysis showed that CNV at 10q11.22 contributed to 1.6% of BMI variation [6]. Later, two independent studies confirmed this association in the European population [7, 8]. Additional evidence came from a CNV study by Zhang *et al.*, in which they demonstrated a strong association of the 10q11.22 deletion with an increased risk of childhood obesity [11].

The 193kb 10q11.22 CNV includes four genes: *Synaptotagmin XV (SYT15), G protein regulated inducer of neurite outgrowth 2 (GRPIN2), heterogeneous nuclear ribonucleoprotein A1 pseudogene (LOC728643)* and *Pancreatic Polypeptide receptor 1 (PPYR1).* [12]. One of these genes, *PPYR1,* also named as the *Neuropeptide Y4 receptor* gene (*NPY4R*), is a seven transmembrane domain G-protein-coupled receptor belonging to the Y receptor family which is highly implicated in the energy homeostasis. The other three genes that are impacted by this CNV are not known to be directly involved with the regulation of food intake. The NPY4R is predominantly expressed in peripheral tissues such as the pancreas, colon, liver, heart and intestine [13-15]. Pancreatic

polypeptide (PP) is the preferential agonist of the Y4 receptor and functions as a satiety signal via the Y4 receptors [16]. Seemingly contradictory results for the role of NPY4R in the pathogenesis of obesity came from germline and conditional *NPY4R* knockout mouse models. These knockout models report hyperphagia in combination with a reduced body weight, reduced amount of white adipose tissue and increased plasma PP levels, while *NPY4R* deficiency in humans is associated with obesity [17, 18]. In contrast to mice that express a functional Y6 receptor, the *NPY6R* gene only exists as a pseudogene in humans. The presence of an extra functioning receptor might alter the ligand-binding patterns of the other Y receptors. Based on the above data we hypothesized that both genetic and structural variation in the *NPY4R* gene might be implicated in the pathogenesis of obesity. Therefore we designed a CNV screen to confirm the association of the 10q11.22 CNV with obesity followed by an extensive mutation screen to investigate the prevalence of genetic variation in both our patient population of children with a severely overweight or obese phenotype and our control population of lean adults.

METHODS AND PROCEDURES

Study population

For CNV-analysis we screened 324 children and adolescents with an obese or severely overweight phenotype (171 boys and 153 girls), and 298 healthy lean adults. Mutation screening in the coding region of *NPY4R* was performed in a population of 356 (324 individuals of the CNV study population) children and adolescents with obese or severely overweight characteristics (187 boys, 169 girls), and 337 healthy lean adults. Children and adolescents with obesity (adolescent if age ≥ 12 years) are unrelated and were recruited at the Child Obesity Clinics from the Antwerp University Hospital (Antwerp) and Jessa Hospital (Hasselt), both in Belgium. The Flemish Growth Charts 2004 were used to identify individuals with obesity. Patients with mutations in *Leptin* and the *Melanocortin-4* receptor (*MC4R)* causing known monogenic obesity and carriers of the 16p11.2 CNV have been excluded from the screening sample. Our white lean population consists of healthy, lean adults (18.5 kg/m^2 ≤ BMI < 25 kg/m^2) recruited among employees from the Antwerp University Hospital and the University of Antwerp, and among couples seeking prenatal counselling at the Centre for Medical Genetics (due to increased triple test or high maternal age). All participants gave their written informed consent. For children participating in this study, parental permission was given. The study protocol was

approved by the local ethics committee. Population characteristics are summarized in Table 1-2. The current study design and study population allows to detect an odds ratio of 5.69 with a power of 80% at a significance level of 5%. Power calculations were performed using R (version 3.2.2).

Table 1: Characteristics of the population used for CNV analysis of *NPY4R*

	Subjects with an obese or overweight phenotype		Lean adults
	Children	Adolescents	
N	181	143	298
Male/female	71/110	100/43	168/124
Age (years)	9,6 ± 0,2	13,8 ± 0,2	34,2 ± 0,4
Weight (kg)	54.2 ± 1,6	99.4 ± 2,5	66,2 ± 0,5
Height (m)	1,29 ± 0,04	1,63 ± 0,03	1,73 ± 0,01
BMI (kg/m^2)	27.3 ± 0,5	33,9 ± 0,5	22,0 ± 0,1
BMI Z-score	2,48 ± 0,05	2,6 ± 0,04	N.A.

Mean value ± standard error of the mean is shown for all parameters, except N and gender distribution (absolute numbers). N.A., not applicable

Table 2: Characteristics of the population used for mutation analysis of *NPY4R*

	Subjects with an obese or overweight phenotype		Lean adults
	Children	Adolescents	
N	195	161	337
Male/female	77/118	110/51	192/145
Age (years)	9,4 ± 0,2	13,9 ± 0,2	34,6 ± 0,2
Weight (kg)	53.1 ± 1,6	98.4 ± 2,5	65,9 ± 0,5
Height (m)	1,29 ± 0,04	1,63 ± 0,03	1,73 ± 0,01
BMI (kg/m^2)	27.3 ± 0,5	33,8 ± 0,5	22,0 ± 0,1
BMI Z-score	2,50 ± 0,05	2,6 ± 0,04	N.A.

Mean value ± standard error of the mean is shown for all parameters, except N and gender distribution (absolute numbers). N.A., not applicable

Anthropometry

Height was measured to the nearest 0.5 cm. Weight was measured on a digital scale to the nearest 0.1 kg. BMI was calculated for all individuals as weight (in kg) over height (in m) squared. For children and adolescents, BMI Z-scores were calculated using the data from the "Flemish Growth Charts 2004". Using these charts, the cutoff values for overweight and obesity are defined as the percentile line on the chart that crosses BMI 25 kg/m^2 and, respectively, BMI 30 kg/m^2 at 18 years of age [12, 19].

MAQ

Genomic DNA was isolated from whole blood by standard procedures [20]. To detect and analyse CNVs, we used Multiplex Amplicon Quantification (MAQ) assays, consisting of a multiplex PCR based amplification of fluorescently labeled target and reference amplicons, followed by capillary electrophoresis and fragment analysis. MAQ assays were performed on 20ng of genomic DNA. Target primers were designed with the Flagged-You MAQ primer design tool and its concentrations were optimized according to the manufacturer's protocol. Final primer concentrations varied between 0.07 and 0.95 µM in the Flagged-you MAQ PCR mix [21]. Target region was set at chr 10:47075071-47099291 (NCBI Build GRCh37.p13) and comprises *NPY4R*. Two negative control samples from lean individuals were included in each experiment for accurate normalisation. Standard PCR cycle conditions were followed. MAQ primer sequences are available upon request.

MLPA

Multiplex Ligation-dependent Probe Amplification analysis was optimized as an independent technique to confirm samples with deviating MAQ dosage quotient (DQ). For MLPA analysis, the SALSA reagent Kit was used (MRC-Holland). Target and reference primers were designed by H-MAPD, a probe design suite for MLPA-assays [22]. In each experiment two control samples were included, allowing accurate normalisation. MLPA primer sequences are available on request. MLPA was performed according to the manufacturer's protocol [23].

Capillary electrophoresis and fragment analysis

Fragment analysis of the resulting MAQ and MLPA products was performed on an ABI Prism Genetic Analyzer 3130xl (Applied Biosystems Inc, Foster City, CA, USA). The MAQ-S software package was used (Multiplicom) to calculate and visualise the dosage quotient which reflects the

copy number of each target amplicon [21]. A dosage quotient of 1.25 -1.75 was considered indicative of a duplication, a DQ of 0.25-075 was indicative of a deletion.

Statistical analysis

Comparison of CNV frequencies among cases and controls was performed using Pearson's chi square analysis. Significance level was set at p=0.05 and statistical analyses were performed using SPSS version 22.0 (SPSS, Chicago, IL, USA).

Mutation analysis of NPY4R

The *NPY4R* gene has 2 different transcripts (GenBank accession number NM_001278794.1, NM_005972.5), encoding 1 protein isoform of 375 amino acids. Mutation screening of the entire coding region of the *NPY4R* gene, including intron-exon boundaries was performed by high-resolution melting curve analysis (HRM) in combination with direct sequencing. HRM, which is performed using the LightCycler 480 Real-Time PCR System (Roche Applied Science, Mannheim, Germany), is a highly sensitive real-time PCR technique to detect novel mutations in a large sample set. Samples with melting curves deviating from wild type were sequenced. Sequencing was performed with ABI BigDye Terminator v1.1 Cycle Sequencing kits on an ABI Prism Genetic Analyser 3130xl (Applied Biosystems Inc, Foster City, CA, USA). Identified mutations were confirmed by resequencing.

Prediction programs

In silico analysis to evaluate possible pathogenic effects of the identified non-synonymous variants on protein function was performed using four different prediction programs: PolyPhen-2 [24], SIFT [25], SNAP [26], Conseq [27], Mutation Taster [28] and MutPred [29].

Expression constructs and mutagenesis

Full-length human wild-type *NPY4R* (NM_005972.5) was cloned in a pCMV6-Entry expression vector obtained from Origene Technologies. Four different missense mutations (L79F, M116T, C168Y and V271M) were individually introduced into the WT *NPY4R* expression vector, using the QuickChange II Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) according to the manufacturer's protocol. The WT sequence and the presence of mutations were checked and confirmed by direct sequencing using an ABI 3130xl Genetic Analyzer (Applied Biosystems).

Luciferase assay for cAMP measurement

For monitoring the cAMP pathway in cells transiently expressing WT or mutant *NPY4R*, a luciferase assay was performed. Human embryonic kidney (HEK) 293T cells were cultured in Dulbecco's modified Eagle medium (DMEM) (Invitrogen) supplemented with fetal bovine serum (10% v/v, Invitrogen), penicillin/streptomycin and L-glutamine. Twenty-four hours prior to transfection, cells were plated at a density of 0.3 x 10^5 cells/well in 96-well plates. Cells were transiently transfected using Fugene6 (Roche Applied Science) with 25 ng pCRE-luc (Stratagene, La Jolla, CA, USA), 2.5 ng pRL-TK and 25 ng WT/mutant NPY4R pCMV6-Entry or empty pcDNA3.1 construct per well. Each transfection was carried out in triplicate and repeated independently in three separate experiments. Twenty-four hours later, cells were washed in PBS and then treated with various concentrations of Pancreatic Polypeptide (Bachem, Germany; 0-10^{-8}-10^{-6} M PP) in the presence of 1µM Forskolin and incubated for 3 hours. Cells were then lysed and assayed for luciferase activity using the Dual-Glo™ Luciferase Assay System (Promega, Madison, WI, USA) on a Glomax-multi Detection System (Turner Biosystems, Sunnyvale, CA) according to the manufacturer's protocol. The ratio of the firefly and pRL-TK Renilla luciferase measurement was calculated. Relative ratios of Firefly luciferase over Renilla luciferase readings were normalized to the relative ratio calculated for the transfection results with the empty pcDNA3.1 vector. Comparison between two measurements for a single experiment was performed using a Student's t-test. Values of $p < 0.05$ were considered significant. Statistical tests were provided by the SPSS 22.0 software package (SPSS Inc.).

RESULTS

CNV-analysis

324 children and adolescents with obesity and 298 lean adults were screened for structural variation in the *NPY4R* region with MAQ. In our patient population we identified 12 cases with 10q11.22 CNV loss and 1 case with CNV gain (being homozygous for a duplication resulting in 4 copies of *NPY4R*). No deletions of the 10q11.22 region could be detected in our control population. Furthermore, CNV gain was determined among 7 lean adults, two of which were homozygous for the duplication. The samples with deletions or duplications in the *NPY4R* region were re-analyzed by MLPA analysis to confirm the presence of structural variation. An overview of the characteristics of the CNV carriers is shown at table 3. By performing Pearson's

chi square analysis, we confirmed a significant difference in the frequency of 10q11.22 CNV loss and CNV gain among subjects with an obese or overweight phenotype and individuals with an healthy weight. CNV loss was more prevalent in our patient population with obese or severely overweight characteristics (3.68%) when compared to our control population (0%)(p=0.001). Furthermore, a higher frequency of 10q11.22 gain was observed in our control population (2.68%) when compared to its frequency among subjects with obesity (0.31%)(p=0.031). In conclusion, the estimated percentage of CNV loss in the patient population is 3.7%, with the 95% confidence interval ranging from 5.8% to 1.6% (normal approximation).

Table 3: Characteristics of the *NPY4R* CNV carriers

	Subjects with an obese or overweight phenotype		Lean adults	
	CNV Loss	CNV gain	CNV loss	CNV gain
N	12	1	0	7
Male/female	5/7	1/0	N.A.	2/5
Age (years)	12,0 ± 3,4	10,08	N.A.	36,0 ± 6,8
BMI (kg/m^2)	32,0 ± 6,1	26,31	N.A.	21,8 ± 1,8
BMI Z-score	2,66 ± 0,5	2,25	N.A.	N.A.

Mean value ± standard deviation of the mean is shown for all parameters, except N and gender distribution (absolute numbers). N.A., not applicable

Detection of rare variants in NPY4R

Screening the coding region of the *NPY4R* gene in our populations identified twenty-four rare different heterozygous coding variants: nine synonymous variants and fifteen non-synonymous variants (figure 1). Of the latter, five variants were identified among both lean individuals and subjects with obesity, six were found exclusively among patients with obesity and four were only seen in lean adults. As the V40M, R239W, R239Q, R246C and L287P variations were all variants found at a similar frequency both in patients with a severely overweight or obese phenotype and lean subjects, we suggest these are rather neutral polymorphisms. Furthermore, four heterozygous variants were exclusively found among lean adults: L79F, C168Y, A169V and R246H. Six children with unhealthy weight each carried a private, heterozygous, non-synonymous variant: V54M, A81T, M116T, K190E, V247M and V271M. In total, the frequency of non-synonymous variants in our patient population (4.49%) is similar to the frequency found in lean adults (5.34%). An overview of the results is shown in table 4.

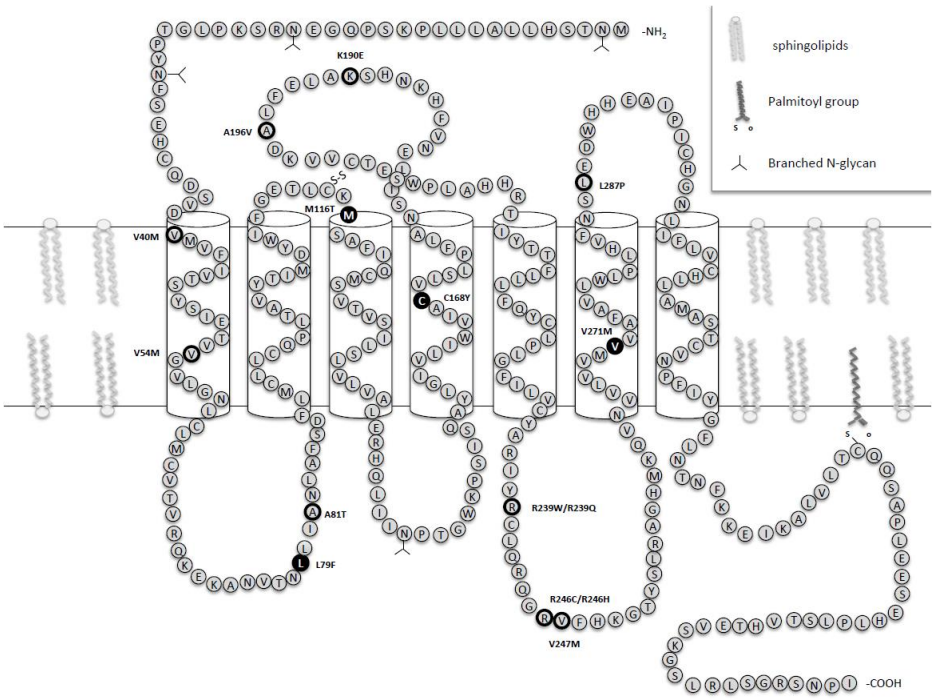

Figure 1: Schematic representation of the NPY4R
The non-synonymous variants of the NPY4R that were identified in this study are highlighted with a black border; the most interesting variants that were selected by in silico analysis for further functional research are colored in black.

Table 4: Coding non-synonymous variants identified in subjects with obesity and lean individuals

Variation		Obese	Lean	MAF
gene	protein			
c.118g>a	V40M	1	1	<0,001
c.160g>a	V54M	1	-	<0,001
c.235c>t	L79F	-	1	<0,001
c.241g>a	A81T	1	-	<0,001
c.347t>c	M116T	1	-	<0,001
c.503g>a	C168Y	-	1	<0,001
c.568g>a	K190E	1	-	<0,001
c.587c>t	A196V	-	1	<0,001
c.715c>t	R239W	3	5	0,004
c.716g>a	R239Q	4	6	0,014
c.736g>a	R246C	1	1	<0,001
c.7137g>a	R246H	-	1	<0,001
c.739g>a	V247M	1	-	<0,001
c.811g>a	V271M	1	-	<0,001
c.860t>c	L287P	1	1	<0,001

Position of variants on gene and protein levels is shown, as well as their frequency among individuals with obesity and among subjects with an healthy weight (in absolute numbers) and the reported MAF in a European population (ExAC browser) [30]. Numbering on gene level is based on cDNA sequence (GenBank NM_001278794.1, NM_005972.5), following the recommendations by the Human Genome Variation Society [31].

In silico analysis was used to select the most interesting variants for further functional research (table 5). Prediction programs indicate that most of the variants found in either subjects with obesity or in lean individuals are suspected not to have a functional effect, except for L79F, M116T, C168Y, and V271M (figure 1). Two of these could only be identified among lean subjects (L79F and C168Y), the other two are exclusively found in our cohort with obesity. For these variations a minimum of 3 out of 6 prediction programs indicate that an effect on NPY4R's function is to be expected (table 5).

Table 5: *In silico* analysis for the non-synonymous variants, identified exlusively among patients or controls

Variant	PolyPhen-2	SIFT	Conseq	MutPred	SNAP	Mutation Taster
V54M	Benign	Tolerated	Moderate	Tolerated	Neutral	Neutral
L79F	Probably damaging	Tolerated	Conserved	Not tolerated	Neutral	Not neutral
A81T	Benign	Tolerated	Moderate	Tolerated	Neutral	Neutral
M116T	Benign	Tolerated	Moderate	Not tolerated	Neutral	Not neutral
C168Y	Probably damaging	Not tolerated	Moderate	Tolerated	Neutral	Not neutral
K190E	Benign	Tolerated	Moderate	Tolerated	Neutral	Neutral
A196V	Benign	Tolerated	Moderate	Tolerated	Neutral	Neutral
R246H	Benign	Tolerated	Moderate	Tolerated	Neutral	Neutral
V247M	Benign	Tolerated	Variable	Tolerated	Neutral	Neutral
V271M	Probably damaging	Not tolerated	Conserved	Not tolerated	Not neutral	Not neutral

Functional analysis of mutant Y4 receptors

In vitro signaling experiments were carried out to determine the influence of the selected mutations on the cAMP signaling function of the receptor. We performed a luciferase assay to monitor the PP-induced inhibition of Forskolin-stimulated cAMP production. First of all, we assayed the intracellular cAMP signaling pathway in HEK293T cells transiently expressing the WT *NPY4R* to confirm intracellular cAMP is decreased in response to increasing stimulation (0M-$1E^{-4}$M PP) of its preferential ligand PP in a dose-dependent manner. Stimulation with a minimal amount of approximately $1E^{-6}$M PP induces the maximum response of the WT receptor, while stimulation with $1E^{-8}$ M PP induces a response halfway the baseline after 3 hours of exposure (figure 2). In addition, we stimulated the HEK293T cells transiently expressing WT, mutant or Empty pcDNA3.1 vector with these specific amounts of PP (0M, $1E^{-8}$M and $1E^{-6}$ M) in order to compare the properties of the WT receptor to those of the four mutant receptors. Cells transfected with the two mutated constructs found only in lean individuals (L79F and C168Y), didn't demonstrate a significantly different response. However, for the two mutants (M116T and V271M) found in individuals with obesity, a significant difference was seen when compared to WT (stimulation with $1E^{-8}$M PP). This effect remains significant for V271M even after stronger

stimulation (stimulation with $1E^{-6}$M PP). In conclusion, the estimated percentage of pathogenic mutations in the patient population is 0.56%, with the 95% confidence interval ranging from 1.34% to 0.21% (normal approximation).

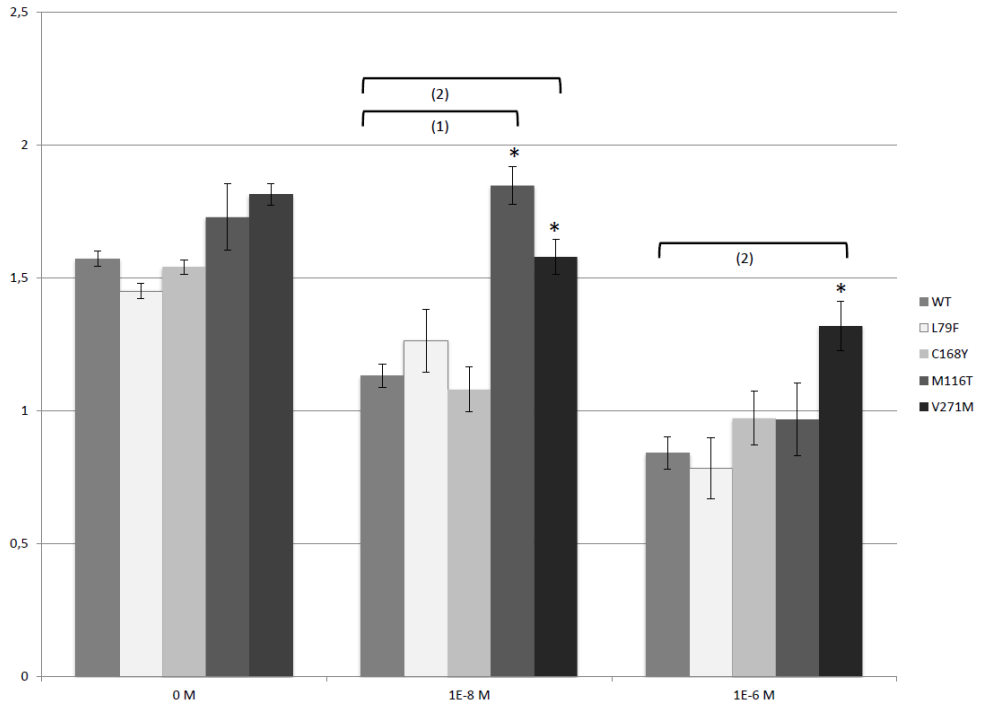

Figure 2: PP-induced cAMP signaling in HEK293T cells expressing WT or mutant NPY4R
PP induced inhibition of the cAMP signaling activity of the NPY4R in HEK293T cells expressing WT or mutant NPY4R. The ratio of the firefly and pRL-TK Renilla luciferase measurement was calculated. Relative ratios of Firefly luciferase over Renilla luciferase readings were normalized to the relative ratio calculated for the transfection results with the empty pcDNA3.1 vector (y-axis). Bars represent average values ± SD. Each transfection was carried out in triplicate and repeated independently in three separate experiments. Intracellular cAMP is decreased in response to stimulation with increasing concentration PP in a dose-dependent manner. The asterisk indicates significance at $p < 0,05$. (1) Cells transfected with mutated M116T NPY4R demonstrated a significantly lower reponse after stimulation with 1E-8M PP, but reached maximal response after stimulation with 1E-6M PP. (2) Stimulation of the mutated V271M NPY4R demonstrated a significant loss in signalling activity when compared to WT at both 1E-6 and 1E-8M PP concentration.

DISCUSSION

In the first part of this study we investigated the prevalence of copy number variations in our patient and control population. The prevalence of *NPY4R* containing 10q11.22 CNV loss was significantly higher in our population with obesity when compared to its frequency in our control population. Among those carriers that were called positive for CNV loss, none were homozygous for the deletion. These findings are in line with previous studies that also associated CNV loss at 10q11.22 with obesity and an increasing BMI [6-8]. Additionally, for two of the patients we were able to test a sibling with obesity. In both cases, the sibling was confirmed to carry the deletion as well. Further family data are necessary to definitely draw conclusions about the segregation of the CNVs with obesity in these families. Furthermore, we also confirmed the association of CNV 10q11.22 with a significantly higher frequency of CNV gain in lean individuals [6, 8]. Only one observation is not completely in line with the assumption that CNV gain has a BMI decreasing effect, the homozygous presence of a *NPY4R* duplication in a 10-year-old child with severely overweight characteristics. Additional genetic or environmental factors might be responsible for the overweight phenotype in this patient.

In the second part of this study we investigated the presence of non-synonymous point mutations in our patient and control population. This mutation analysis resulted in the identification of fifteen rare non-synonymous variants, of which six were found exclusively among patients and four were only seen in lean adults. Four of these private mutations (L79F, M116T, C168Y and V271M) were predicted to have a potential effect on the receptor-activity and were selected for functional studies. In these studies we investigated the effect of the selected variants on cAMP signaling by monitoring the PP-induced inhibition of Forskolin-stimulated cAMP increase. L79F and C168Y are those variants that were privately identified in lean adults and that were expected to have a rather gain of function effect on the NPY4R. For L79F, the leucine residue at codon 79 is located at the first intracellular loop of the receptor, known for its function in G-protein coupling [32]. We could only demonstrate a minor increase of the receptor activity while stimulating with 0M and 1E-6M PP ($p > 0.05$). At 0M and $1E^{-8}$M PP stimulation the same trend ($p > 0.05$) was observed for the C168Y mutant of which the highly conserved cysteine is localized at the fourth transmembrane helix. The substitution of a sulfhydryl to an hydroxyl group at this codon might have a positive effect on the conformational structure of the receptor and ligand binding pocket.

For mutations M116T and V271M, we were able to demonstrate receptor dysfunction as both mutant receptors showed markedly decreased receptor activity in respone to PP stimulation. The M116T and V271M mutations localize to the first extracellular loop and to the sixth transmembrane helix respectively, two regions known for its function in ligand-binding and receptor activation [32-34]. For these mutations, the substitution might lead to a conformational change in the binding pocket prohibiting the proper binding of receptor ligands. Although the M116T mutant receptor shows markedly decreased signaling properties in response to stimulation with $1E^{-8}$M PP, its activity after maximal stimulation approaches the maximal performance of the WT receptor. This is in contrast with the V271M receptor, of which its maximal performance reaches to less than 40% of the maximal activity of the WT receptor. Previous studies reported that the top of transmembrane helix two, six and seven form the core of the peptide binding pocket. Therefore, this mutation is suggested to cause a more important sterical hindrance than the M116T mutant, leading to the strongest NPY4R dysfunction. Taken together, these data led to the conclusion that M116T and V271M are pathogenic mutants, while no significant functional effect could be demonstrated for L79F and C168Y. However, it has also been reported that the NPY receptors are linked to calcium signaling pathways and to phospholipase C [35, 36]. Further functional analyses regarding the effect of the mutations on these signaling pathways might be necessary to draw definite conclusions about the disease-causality of these variants. One of the other limitations of our study is that we were not able to functionally validate our CNV variants. This could be done by expression analysis. However, as we don't have acces to patient tissue samples expressing *NPY4R* (predominantly pancreas, liver, heart and intestine), we were not able to functionally validate our CNV variants. Furthermore, our approach of using a patient population of children and adolescents with severe early-onset obesity and a control population of never-overweight adults is one of the strengths of our study but it also has its weaknesses. As we have no information about confounding factors such as diet and alcohol consumption, we are not able to adjust our analysis for these factors.

CONCLUSION

We were able to demonstrate receptor dysfunction and thus a pathogenic effect for two variants M116T and V271M that could only be identified in our patient population. These findings were in line with the results of our CNV-analysis in which we demonstrated a significantly higher frequency of *NPY4R* containing 10q11.22 CNV loss in the patient population. Taken together, the presence of structural (12 10q11.22 CNV loss carriers) and genetic variation (2 pathogenic

mutations) within the *NPY4R* gene at least partially explains the obese phenotype of 3.7% and 0.6% of our patient population, respectively. This makes this gene a very important risk factor for obesity with prevalences close to those reported for *MC4R* mutations. In conclusion, these data support an essential role for genetic and structural variation within the *NPY4R* gene in the pathogenesis of obesity.

ACKNOWLEDGEMENTS

This study was supported by an Interuniversity Attraction Pole Project (Phase VII project 43, BELSPO) and a TOP project from the University of Antwerp. EA holds a PhD Grant for Strategic Basic Research from the "Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen)". SB is a postdoctoral researcher of the "Fund for Scientific Research Flanders" (FWO-Vlaanderen).

Study design: EA, SB and WVH; Data collection and Analysis: EA; Recruiting and phenotyping the subjects: DZ, AV, GM, KVH, SV, LVG; Manuscript preparation: EA.

REFERENCES

1. Wang, Y.C., et al., *Health and economic burden of the projected obesity trends in the USA and the UK*. Lancet, 2011. **378**(9793): p. 815-25.
2. Stunkard, A.J., T.T. Foch, and Z. Hrubec, *A Twin Study of Human Obesity*. Jama-Journal of the American Medical Association, 1986. **256**(1): p. 51-54.
3. Bouchard, C., et al., *The Response to Long-Term Overfeeding in Identical-Twins*. New England Journal of Medicine, 1990. **322**(21): p. 1477-1482.
4. Shungin, D., et al., *New genetic loci link adipose and insulin biology to body fat distribution*. Nature, 2015. **518**(7538): p. 187-U378.
5. Locke, A.E., et al., *Genetic studies of body mass index yield new insights for obesity biology*. Nature, 2015. **518**(7538): p. 197-U401.
6. Sha, B.Y., et al., *Genome-wide association study suggested copy number variation may be associated with body mass index in the Chinese population*. Journal of Human Genetics, 2009. **54**(4): p. 199-202.
7. Wang, K., et al., *Large copy-number variations are enriched in cases with moderate to extreme obesity*. Diabetes, 2010. **59**(10): p. 2690-4.
8. Jarick, I., et al., *Novel common copy number variation for early onset extreme obesity on chromosome 11q11 identified by a genome-wide analysis*. Human Molecular Genetics, 2011. **20**(4): p. 840-852.
9. Wheeler, E., et al., *Genome-wide SNP and CNV analysis identifies common and low-frequency variants associated with severe early-onset obesity*. Nat Genet, 2013. **45**(5): p. 513-7.
10. Falchi, M., et al., *Low copy number of the salivary amylase gene predisposes to obesity*. Nat Genet, 2014. **46**(5): p. 492-7.
11. Zhang, D., et al., *Interactions between obesity-related copy number variants and dietary behaviors in childhood obesity*. Nutrients, 2015. **7**(4): p. 3054-66.
12. Roelants, M., R. Hauspie, and K. Hoppenbrouwers, *References for growth and pubertal development from birth to 21 years in Flanders, Belgium*. Ann Hum Biol, 2009. **36**(6): p. 680-94.
13. Bard, J.A., et al., *Cloning and functional expression of a human Y4 subtype receptor for pancreatic polypeptide, neuropeptide Y, and peptide YY*. J Biol Chem, 1995. **270**(45): p. 26762-5.

14. Larsen, P.J. and P. Kristensen, *The neuropeptide Y (Y4) receptor is highly expressed in neurones of the rat dorsal vagal complex.* Brain Res Mol Brain Res, 1997. **48**(1): p. 1-6.
15. Parker, R.M. and H. Herzog, *Regional distribution of Y-receptor subtype mRNAs in rat brain.* Eur J Neurosci, 1999. **11**(4): p. 1431-48.
16. Katsuura, G., A. Asakawa, and A. Inui, *Roles of pancreatic polypeptide in regulation of food intake.* Peptides, 2002. **23**(2): p. 323-9.
17. Sainsbury, A., et al., *Synergistic effects of Y2 and Y4 receptors on adiposity and bone mass revealed in double knockout mice.* Mol Cell Biol, 2003. **23**(15): p. 5225-33.
18. Sainsbury, A., et al., *Y4 receptor knockout rescues fertility in ob/ob mice.* Genes Dev, 2002. **16**(9): p. 1077-88.
19. Hauspie, R. and M. Roelants. *Flemish Growth Charts 2004*. 2004 22 July 2013]; Available from: http://www.vub.ac.be/groeicurven.
20. Miller, S.A., D.D. Dykes, and H.F. Polesky, *A simple salting out procedure for extracting DNA from human nucleated cells.* Nucleic Acids Res, 1988. **16**(3): p. 1215.
21. Multiplicom. Multiplicom the multiplex experts.; Available from: http://www.multiplicom.com.
22. Zhi, J., *MAPD: a probe design suite for multiplex ligation-dependent probe amplification assays.* BMC Res Notes, 2010. **3**: p. 137.
23. MRC-Holland. Available from: http://www.mlpa.com.
24. Adzhubei, I.A., et al., *A method and server for predicting damaging missense mutations.* Nat Methods, 2010. **7**(4): p. 248-9.
25. Ng, P.C. and S. Henikoff, *Predicting deleterious amino acid substitutions.* Genome Res, 2001. **11**(5): p. 863-74.
26. Bromberg, Y., G. Yachdav, and B. Rost, *SNAP predicts effect of mutations on protein function.* Bioinformatics, 2008. **24**(20): p. 2397-8.
27. Berezin, C., et al., *ConSeq: the identification of functionally and structurally important residues in protein sequences.* Bioinformatics, 2004. **20**(8): p. 1322-4.
28. Schwarz, J.M., et al., *MutationTaster evaluates disease-causing potential of sequence alterations.* Nat Methods, 2010. **7**(8): p. 575-6.
29. Li, B., et al., *Automated inference of molecular mechanisms of disease from amino acid substitutions.* Bioinformatics, 2009. **25**(21): p. 2744-50.
30. Exome Aggregation Consortium (ExAC), C., MA.

31. den Dunnen, J.T. and S.E. Antonarakis, *Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion.* Hum Mutat, 2000. **15**(1): p. 7-12.

32. Ritter, S.L. and R.A. Hall, *Fine-tuning of GPCR activity by receptor-interacting proteins.* Nat Rev Mol Cell Biol, 2009. **10**(12): p. 819-30.

33. Peeters, M.C., et al., *GPCR structure and activation: an essential role for the first extracellular loop in activating the adenosine A2B receptor.* FASEB J, 2011. **25**(2): p. 632-43.

34. Wheatley, M., et al., *Lifting the lid on GPCRs: the role of extracellular loops.* Br J Pharmacol, 2012. **165**(6): p. 1688-703.

35. Motulsky, H.J. and M.C. Michel, *Neuropeptide Y mobilizes Ca2+ and inhibits adenylate cyclase in human erythroleukemia cells.* Am J Physiol, 1988. **255**(6 Pt 1): p. E880-5.

36. Reynolds, E.E. and S. Yokota, *Neuropeptide Y receptor-effector coupling mechanisms in cultured vascular smooth muscle cells.* Biochem Biophys Res Commun, 1988. **151**(2): p. 919-25.

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

Evi Aerts[1], Ellen Geets[1], Laure Sorber[1], Dr. Sigri Beckers[1], Dr. An Verrijken[2,] Prof. Dr. Guy Massa[3], Dr. Kim Van Hoorenbeeck[4], Prof. Dr. Stijn L Verhulst[4], Prof. Dr. Luc F Van Gaal[2] & Prof. Dr. Wim Van Hul[1]

[1] Department of Medical Genetics, University of Antwerp, Antwerp, Belgium

[2] Department of Endocrinology, Diabetology and Metabolic Diseases, Antwerp University Hospital, Antwerp, Belgium

[3] Department of Pediatrics, Jessa Hospital, Hasselt, Belgium

[4] Department of Pediatrics, Antwerp University Hospital, Antwerp, Belgium

Annals of human genetics (2017) Accepted

ABSTRACT

Background Neuropeptide Y (NPY) and its G-protein coupled NPY Y2 Receptor (NPY2R) are highly expressed in orexigenic NPY/Agouti-related peptide neurons within the arcuate nucleus, a major integrator of appetite control in the hypothalamus. As *NPY* and *NPY2R* are interesting candidate genes for obesity, we hypothesized that genetic variation in these genes might be implicated in the pathogenesis of obesity.

Methods In the first part of this study we performed a mutation analysis of the coding region of *NPY* and *NPY2R* with high-resolution melting curve analysis. For the highly conserved *NPY* gene, an extended population of 436 obese children and adolescents was screened, while for *NPY2R* a smaller subset of 306 patients was used. A control population of 300 healthy individuals was screened for *NPY2R* to determine the general prevalence of the variants found among patients. Direct sequencing was performed for samples with melting patterns deviating from wild-type.

In the second part of this study, Multiplex Amplicon Quantification (MAQ) analysis was performed in 308 obese children and adolescents to detect CNV in the *NPY2R* region.

Results Mutation analysis of the *NPY* gene led to the identification of 1 common missense variant (L7P; MAF 0.04), while the screening of the *NPY2R* gene resulted in the identification of 1 rare missense variant F87I in the patient population.

In our CNV analysis, we could not identify copy number variation in the *NPY2R* region among obese children and adolescents.

Conclusion In summary, this study clearly indicates that genetic variation in *NPY* and *NPY2R* is at low frequency and thus does not make a major contribution to the obese phenotype in the general population.

INTRODUCTION

Obesity is a highly heritable complex disorder, characterized by an excessive accumulation of adipose tissue that results from a persistent imbalance between energy intake and expenditure. With a prevalence reaching epidemic proportions, obesity imposes an enormous burden on human health. Numerous studies showed that 40-70% of the interindividual variability in BMI can be attributed to genetic factors [1, 2]. So far, only a limited number of rare monogenic causes of obesity is known, many of which are associated with the leptin-melanocortin pathway. In this pathway, leptin is secreted by adipocytes in proportion to the amount of fat mass. Binding of leptin with its receptors in the *nucleus arcuatus* leads to the activation of the melanocortin-4 receptors followed by appetite suppression and stimulation of energy expenditure. *Melanocortin-4 receptor* (*MC4R*) gene mutations are the most prevalent cause of monogenic obesity, with over 160 distinct mutations reported so far [3]. In contrast, only eleven different mutations in the *leptin* gene have been discovered, making leptin deficiency a rare form of monogenic obesity [4-11]. Leptin deficiency is currently, still the only treatable type of monogenic obesity [6, 12, 13]. Despite the importance of the leptin-melanocortin pathway in the pathogenesis of obesity, mutations in the genes involved in this pathway account for not more than 5% of all childhood severe obesity cases. The genetics of obesity for most individuals is complex and involves the interaction of multiple genes and environmental factors. Therefore, extensive genome-wide association studies (GWASs) were performed to identify common single-nucleotide polymorphisms (SNPs), underlying the heritable risk of common obesity. So far, 97 BMI-associated loci were identified that account for approximately 2.7% of BMI variation, leaving a substantial component of the heritability unexplained [14-16].

Recently, there is considerable interest in the role of gut hormones and their receptors in the pathogenesis of obesity. One of these gut hormones, neuropeptide Y (NPY), belongs to the pancreatic polypeptide (PP) family together with peptide-yy (PYY) and PP, and is reported to be one of the most potent orexigenic peptides in the brain. This 36-amino acid neuromodulator is released by neurons in the central and peripheral nervous system and stimulates food intake by decreasing the latency to eat, increasing the motivation to eat and by delaying satiety [17]. In the leptin-melanocortin pathway, binding of leptin with its receptors in the *nucleus arcuatus* leads to the inhibition of orexigenic *NPY* expression and results in a decrease of food intake [18-20].

Additional indications confirmed the importance of NPY in the regulation of energy homeostasis and food intake. The common non-synonymous SNP rs16139 in the *NPY* gene was associated to

obesity-related phenotypes. Subjects with this L7P substitution had higher BMI (+0.58 kg/m^2) and higher risk of being obese (p=0.002) [21-23]. Further insights into the role of NPY in energy homeostasis were gained by animal studies. In 1991, Billington and colleagues showed that central administration of NPY in rats leads to weight gain and obesity due to enhanced feeding, increased white fat lipid storage and decreased energy expenditure [24]. This was confirmed in *NPY*-overexpressing mice which developed increased body weight gain, hyperglycemia and hyperinsulinemia after a 50% sucrose-loaded diet [25]. In contrast, 24-h food intake in *NPY* knockout mice did not differ from their wild-type littermates. As the levels of orexigenic agouti-related peptide (AgRP) were significantly increased, the seemingly normal phenotype of *NPY* knockout mice could result from developmental compensation by other neuropeptides [26]. Furthermore, *NPY* deficiency in the *ob/ob* mice led to a less obese phenotype because of reduced food intake and increased energy expenditure. These results suggest that NPY and leptin synergistically act on the regulation of food intake and weight regulation [27].

In mammals, NPY targets a population of at least four different NPY receptors. NPY Y1 and Y5 receptors are known to play an important role in the stimulation of feeding, whereas NPY Y2 and Y4 receptors (NPY2R-NPY4R) seem to have roles in the inhibition of appetite [28-30]. Recently, we were able to demonstrate an essential role for the *NPY4R* gene in the pathogenesis of obesity. In our study, the presence of copy number (*NPY4R* containing 10q11.22 deletion) and sequence variation (pathogenic mutations) within the *NPY4R* gene at least partially explained the obese phenotype of 3.7% and 0.6% of our patient population, respectively [31]. This makes the *NPY4R* gene an important risk factor for obesity with prevalences close to those reported for *Melanocortin-4 receptor* (*MC4R*) mutations, the most prevalent cause of monogenic obesity. The NPY4R seems to work synergistically with the NPY2R, which is highly expressed in orexigenic NPY/AgRP neurons within the arcuate nucleus [32-34]. In 1999, *NPY2R* knockout mice were reported to develop increased body weight, food intake and fat deposition. They also showed an attenuated response to leptin administration but a normal response to NPY-induced food intake [29]. In contrast with the obese phenotype they observed in germline knockout mice, a lean phenotype was reported in hypothalamus-specific *NPY2R*-deleted mice. Although this is in contrast with the expected phenotype in humans, this might be explained by compensatory mechanisms present in mice, who have a functional Y6 receptor [35]. In 2004, two common single nucleotide polymorphisms (SNPs) in exon 2 (rs2880415 and rs1047214) of the *NPY2R* were identified in a white British population and were found to be in strong linkage disequilibrium. Men who were homozygous for the rarer variant, had significantly lower BMI

(p=0.017), waist-to-hip ratio (p=0.013) and higher non-esterified fatty acid levels (p=0.01) [36]. Similar to the British result, the association of these SNPs with severe obesity was confirmed in Pima Indian males [37]. Further evidence that strengthens the case for a role of *NPY2R* in the pathogenesis of obesity, came from a study that reported concordance of a small 500bp deletion, located in intron 1 of the *NPY2R* gene, in severely obese members of a specific pedigree. No additional carriers of this deletion could be found in other pedigrees and therefore the study of the CNV was not pursued at that point. However, because of this deletion, they genotyped two SNPs on either side of the deletion and one SNP within the deleted region (rs10461238) in a control and patient population and showed that these SNPs were significantly associated with BMI [38].

Based on the above data, we hypothesize that sequence variation in the *NPY* and *NPY2R* genes may be implicated in the pathogenesis of obesity. Therefore, we designed an extensive mutation screen for these genes in a patient population of severely overweight and obese children.

In the second part of this study we performed CNV-analysis on the *NPY2R* gene region to study whether the previously reported *NPY2R* deletion contributes to obesity in the general population or whether it is a private mutation in the specific pedigree.

METHODS AND PROCEDURES

Study population

As the *NPY* gene is highly conserved, we screened the coding region in an extended patient population of 436 severely overweight and obese children and adolescents (197 boys, 239 girls; Table 1). For *NPY2R*, 306 patients (148 boys, 158 girls) were screened for mutation analysis (Table 2) and 311 patients were screened for CNVs (144 boys and 167 girls) (Table 3).

A control population of 300 healthy individuals was screened for mutations in the *NPY2R* gene to determine the general prevalence of the variants that were found among patients

Our obese patients (adolescent if age ≥ 12 years) are unrelated and were recruited at the Child Obesity Clinics from the Antwerp University Hospital (Antwerp) and Jessa Hospital (Hasselt) in Belgium. The percentile lines of the Flemish Growth Charts 2004 that cross BMI 25 and 30 kg/m^2 at 18 years of age on the Flemish age- and sex-specific BMI growth curves were used as cut-off for the diagnosis of overweight and obesity respectively [39, 40]. Patients with mutations in *Leptin* or *MC4R* that cause known monogenic obesity have been excluded from our study population. Among our patients, a minor proportion is of non-Caucasian ethnicity (<10%).

Our lean control population, used for mutation screening, consisted of 300 healthy, lean adults (18.5 kg/m^2 ≤ BMI < 25 kg/m^2; 124 men, 176 women; Table 2) recruited among employees from the Antwerp University Hospital and the University of Antwerp, and among couples seeking prenatal counselling at the Centre for Medical Genetics (due to increased triple test or high maternal age).

All participants gave their written informed consent. The study protocol was approved by the local ethics committee.

Table 1: Characteristics of the population used for mutation analysis of *NPY*

	Obese subjects	
	Children	Adolescents
N	241	195
Male/female	134/107	63/132
Age (years)	9,3 ± 0,2	14.1 ± 0,2
Weight (kg)	52,7 ± 1,4	98,7 ± 2,0
Height (m)	1,26 ± 0,04	1,45 ± 0,01
BMI (kg/m^2)	27,7 ± 0,4	34,0 ± 0,5
BMI Z-score	2,52 ± 0,05	2,6 ± 0,04

Mean value ± standard error of the mean is shown for all parameters, except N and gender distribution (absolute numbers).

Table 2: Characteristics of the population used for mutation analysis of *NPY2R*

	Obese subjects		Lean adults
	Children	Adolescents	
N	171	135	300
Male/female	109/62	39/96	124/176
Age (years)	9,5 ± 0,4	13,6 ± 0,2	34,4 ± 0,3
Weight (kg)	52,4 ± 1,7	95,7 ± 2,5	65,8 ± 0,5
Height (m)	1,37 ± 0,02	1,65 ± 0,01	1,73 ± 0,01
BMI (kg/m^2)	28,3 ± 0,4	33,5 ± 0,5	21,9 ± 0,1
BMI Z-score	2,58 ± 0,04	2,6 ± 0,04	N.A.

Mean value ± standard error of the mean is shown for all parameters, except N and gender distribution (absolute numbers). N.A., not applicable

Table 3: Characteristics of the population used for CNV analysis

	Obese subjects	
	Children	Adolescents
N	168	143
Male/female	101/67	43/100
Age (years)	9,5 ± 0,2	13,7 ± 0,2
Weight (kg)	53,2 ± 1,2	95,8 ± 2,3
Height (m)	1,36 ± 0,02	1,65 ± 0,01
BMI (kg/m^2)	28,0 ± 0,4	33,7 ± 0,5
BMI Z-score	2,56 ± 0,04	2,61 ± 0,04

Mean value ± error of the mean is shown for all parameters, except N and gender distribution (absolute numbers).

Anthropometry

Height was measured to the nearest 0.1 cm. Weight was measured on a digital scale to the nearest 0.1 kg. Body mass index (BMI) was calculated for all individuals as weight (in kg) over height (in m) squared. For children and adolescents, BMI Z-scores were calculated based on the data from the Flemish Growth Curves 2004 [39, 41].

Mutation analysis of NPY and NPY2R

Blood samples from all patients and control subjects were obtained for extraction of genomic DNA by standard procedures [42]. Mutation screening of the entire coding region of the *NPY* and *NPY2R* gene (Genbank accession number NM_000905.3 and NM_000910.3), including intron-exon boundaries was performed by high-resolution melting curve analysis (HRM) in combination with direct sequencing. The *NPY* gene consists of 3 coding exons, the *NPY2R* gene consists of 2 exons, 1 of which is coding. This coding region was divided in amplicons of approximately 300 base pairs (bp) for HRM and 700 bp for sequencing. HRM, which is performed using the LightCycler 480 Real-Time PCR System (Roche Applied Science, Mannheim, Germany), is a highly sensitive real-time PCR technique to detect novel mutations in a large sample set. HRM was performed according to the standard procedure that was previously described [31, 43]. Primer sequences for HRM and Sanger sequencing are listed in table 4.

Table 4: Primer sequences used for mutation and CNV analysis of *NPY* and *NPY2R*

Gene	Application	Forward Primer	Reverse Primer
NPY	SEQ	gtgtggtagcaggaggagg	agtgcctgatactgtcctgc
		ggcaggaatttgactaggaattt	agcaccaaataacctgagctaaa
		tcacgatctgatattccacatgg	cagaattcagcacagtggct
	HRM	gttgagccttctgtgcctg	ggatctcctggtgtgcagg
		gcttcatacacctagcttgct	gtctctgacttccctcccct
		acccttgctcatacctcagg	ggaaaaggccagagagcaag
NPY2R	SEQ	tgttttcctcgttccattggt	aatactcccggaagatggcc
		gaaaatgggtcctgtcctgtg	ccagccagttttccaggaat
	HRM	gccaagtggacctgtactga	tggatcaccaaggagttgc
		tccatcatcttgcttggggt	ccggtccagggcaattact
		tcaccttgacagtaattgccc	cccagaggcaaaacatacaaga
		agagcatctatggcactgtct	gtcaatgtcaacggcaagct
		cctacactcgcatttggagt	tggaacactgtgaagatgagtt
		Agcttgccgttgacattgac	Acaccacagcttccttagaca
	MAQ	Agcggataacaatttcacacaggaaagaagatggtatcccaatgc	gtttctttgtatgggaagaattagagaaatgc
		agcggataacaatttcacacaggacccactggcaggagctt	gtttcttgaagggaagtttcaagaatgctac
		agcggataacaatttcacacaggctggaattcaaagataaggcaac	gtttcttgaactgcacacctcctcaatc
		agcggataacaatttcacacaggaatggaaaggtaaagggcaga	gtttcttacttactgaatctaggcaaatggaa

SEQ, sequencing; HRM, High-resolution melting curve analysis; MAQ, Multiplex Amplicon Quantification analysis

Prediction programs

In silico analysis to evaluate possible pathogenic effects of the identified non-synonymous variants on protein function was performed using four different prediction programs: PolyPhen-2 [44], SIFT [45], SNPs&Go [46], and MutPred [47].

Statistical analysis

For rs1047214 and rs2880415, Hardy Weinberg equilibrium (HWE) was calculated using the HWE tool from LINKUTIL [48]. Comparison of genotype and allele frequencies among cases and controls was performed using Pearson's chi square analysis. Significance level was set at p=0.05. All statistical analyses were performed using SPSS version 20.0 (SPSS, Chicago, IL, USA).

CNV-analysis

MAQ

To detect and analyse CNVs, we used Multiplex Amplicon Quantification assays (MAQ) according to the standard procedure that was previously described [31, 43, 49]. Target region was set at chr 4:156123313-156143207 (NCBI Build GRCh37.p13) and comprises *NPY2R* so the previously reported deletion can be detected. Two negative control samples from lean individuals were included in each experiment for accurate normalisation. With a sensitivity and specificity approaching 100%, MAQ can be used as a valuable diagnostic tool for reliable detection of copy number changes [50]. Standard PCR cycle conditions were followed. MAQ primer sequences are listed in table 4.

Capillary electrophoresis and fragment analysis

Fragment analysis of the resulting MAQ-PCR products was performed on an ABI Prism Genetic Analyzer 3130xl (Applied Biosystems Inc, Foster City, CA, USA). The MAQ-S software package was used (Multiplicom) to calculate and visualise the dosage quotient (DQ) which reflects the copy number of each target amplicon [49]. A DQ of 1.25 -1.75 was considered indicative of a duplication, a DQ of 0.25-075 was indicative of a deletion.

RESULTS

Mutation analysis of NPY

Mutation analysis of the entire coding region of the *NPY* gene led to the identification of 1 rather common missense variation rs16139 (L7P). The prevalence of this variant in our patient population was similar to the one that was reported for the European population (MAF 0.04). Furthermore, two common synonymous variants (S50S and S68S) and one rare synonymous variant (A28A) were found. As we were not able to identify a rare non-synonymous mutation in our obese population, we decided to not screen our population of lean adults. An overview of the results was given in table 5.

Mutation analysis of NPY2R

Mutation analysis of the coding region of the *NPY2R* gene in the patient population resulted in the identification of 3 rare synonymous variants (L53L, G326G and G370G), 2 common SNPs (I195I and I312I) and 1 rare non-synonymous heterozygous variant F87I. The patient bearing the F87I mutation is an 11-year-old boy with severe obesity. His BMI reaches a score of 32.1 kg/m^2 with a BMI z-score of 2.49 and a percentage BMI of 188%. As no family data were available, it was not possible to determine whether the mutation segregates with obesity in his family. To evaluate the pathogenicity of F87I, *in silico* analysis using a set of four different prediction programs (PolyPhen-2 [44], SIFT [45], SNPs&Go [46] and MutPred [47]) was performed. 3 out of 4 prediction programs indicated that it has a possible impact on the structure and function of the NPY2R. To determine whether F87I is a neutral polymorphism or a pathogenic mutation, 300 lean individuals were screened for genetic variation in *NPY2R*. Screening our control population identified two additional synonymous variants P105P and A278A. None of the other rare variants were identified in control subjects. The allele frequencies of SNPs rs1047214 (I195I) and rs2880415 (I312I) were not significantly different among cases and controls. An overview of the results was given in table 5.

Table 5: Coding variants identified in obese and lean individuals.

Gene	SNP	Variation Gene	Variation Protein	MAF obese cases/ lean adults	MAF ExAC EU
NPY	Rs16139	c.20T>C	L7P	0.04/ND	0.04
	Rs5572	c.84G>A	A28A	0.002/ND	$7.5*10^{-5}$
	Rs5573	c.150G>A	S50S	0.50/ND	0.51
	Rs5574	c.204C>T	S68S	0.46/ND	0.47
NPY2R	Rs2342674	c.159C>T	L53L	0.005/-	0.0002
	Rs144160377	c.259T>A	F87I	0.002/-	0.0003
	Rs148709959	c.315G>A	P105P	-/0.002	0.0003
	Rs1047214	c.585C>T	I195I	0.56/0.51	0.54
	Rs550095329	c.834G>A	A278A	-/0.002	-
	Rs2880415	c.936C>T	I312I	0.56/0.51	0.54
	Rs138080356	c.978C>T	G326G	0.002/-	0.002
	Rs138652181	c.1110C>A	G370G	0.002/-	$1.5*10^{-5}$

Position of variants on gene and protein levels is shown, as well as their frequency in the obese and lean groups and the reported MAF in a European non-Finnish population (Exome Aggregation Consortium browser). Numbering on gene level is based on cDNA sequence (GenBank NM_001145795.1, NM_001145796.1, NM_001145797.1, NM_001145812.1 and NM_015503.2), following the recommendations by the Human Genome Variation Society [51]. ND, not determined

CNV-analysis of NPY2R

We screened 308 obese children and adolescents for copy number variation in *NPY2R* with MAQ. We could not detect any copy number gain or loss in the target regio. The previously reported *NPY2R* deletion could not be identified in the patient population.

DISCUSSION

In the first part of this study, we screened the coding region of *NPY* in 436 severely overweight/obese children and adolescents. This mutation analysis resulted in the identification of 2 common synonymous variants (S50S and S68S), 1 rare synonymous variant (A28A) and 1 rather common missense variation (L7P). The rs16139 variant (L7P) is the only well documented non-synonymous missense variant that was reported for the *NPY* gene. This SNP is located in the signal peptide part of the preproNPY and leads to an increase in NPY synthesis and secretion [52]. It was originally discovered in Dutch and Finnish populations where it was linked with higher serum levels of total and LDL cholesterol in obese subjects. Since then, the L7P substitution was further associated with a variety of metabolic, cardiovascular and behavioral

parameters in combination with obesity, increased BMI, hyperlipidemia, atherosclerosis, major depressive disorder and alcohol use [23, 53-56]. As we did not screen our lean control population for the rs16139 variant, we cannot exclude association with obesity.

Furthermore, we identified two carriers of the rs5572 variant (A28A) of which one also suffered from *pubertas praecox*. Previous studies already reported a possible role for NPY in the initiation of puberty [57]. Further research is necessary to evaluate a possible association between rs5572 and puberty. However, although synonymous substitutions can affect transcription, splicing, mRNA transport and translation, these variations do not have a high priority for further research. That also counts for the synonymous variants rs5573 and rs5574 that were found in our patient population at a frequency that is similar to their corresponding MAF in the European population (ExAC database).

As we were not able to identify a rare non-synonymous variation in the *NPY* gene among 436 severely overweight and obese children, the prevalence of missense variants among obese individuals in *NPY* must be low. Due to its important role in a variety of physiological processes, the *NPY* gene must be under strong negative selection against mutations. This is in line with the strong evolutionary conservation across diverse species that was reported for the *NPY* gene [58].

In the second part of our study, we screened the coding region of the *NPY2R* gene for genetic and structural variation in our study population. The NPY2R has previously been reported as a crucial receptor that works synergistically with the NPY4R as major regulators of energy homeostasis. Recently, we were able to demonstrate an essential role for genetic (pathogenic mutations) and structural variation within the *NPY4R* gene in the pathogenesis of obesity [31]. However, for the *NPY2R* gene only a limited number of data were available on the prevalence of point mutations and structural variation in this gene. To the best of our knowledge, only one research group reported the presence of three missense mutations (L40F, F87I, A172T) in a rather small study population of 100 unrelated early-onset obese individuals and only one 500 bp deletion in the *NPY2R* gene was associated with obesity [36, 38]. In this study, we aimed to further unravel the function of NPY2R by performing a mutation screen and by analysing CNV in this gene.

Our mutation analysis resulted in the identification of 1 rare missense variation (F87I) in a white obese patient. After finding the F87I variation in a population of severely overweight and obese children, we screened our control population. No lean carrier of this allele could be identified among the 300 healthy controls. Unfortunately, neither the patient nor the family was willing to cooperate in further research, so we were unable to perform segregation analysis. The mutated

residue localizes to the transmembrane segment II, known for its function in ligand-dependent and ligand-independent activation of 7TM receptors [59]. As *in silico* analysis predicts a probably pathogenic impact for F87I, the variant was initially thought to impair ligand binding and signaling. However, the variation had previously been reported not to segregate with obesity in a family, of which an 8-year-old white boy with a BMI SDS of 3.8 was found to carry the F87I allele in heterozygous form. The mutant allele was inherited from the lean mother [36]. Therefore, we hypothesize that F87I is a rare variant which is not directly associated with the pathogenesis of obesity. However, as they were also not able to identify an additional lean carrier among their lean controls, we cannot rule out other possible genetic determinants causing the lean phenotype of the mother. Functional studies will be necessary to definitely clarify the role of F87I variant in the energy homeostasis.

Furthermore, we identified 5 different rare synonymous variants of which three (L53L, G326G and G370G) could solely be detected among obese children and two (P105P and A278A) were only present in lean individuals. As all three patients who harbored the L53L variant were of non-Caucasian origin and no lean white carriers could be detected, this variant is suggested to be ethnic group specific. These observations are in line with a previous study where also no normal-weight white and three non-Caucasian L53L carriers were reported: two were South African Zulu and one was of Afro-Caribbean origin [36]. For the two common SNPs that were associated with severe obesity in the British and Pima Indian population, we could not detect significantly different allele frequencies among cases and controls in our Caucasian population [36, 37]. To evaluate this further, a larger study population is required to achieve adequate statistical power.

In our CNV analysis, we screened 308 obese children and adolescents for structural variation in *NPY2R* with MAQ. Previously a study reported the concordance of a small 500bp deletion, located in intron 1 of the *NPY2R* gene, in severely obese members of a specific pedigree. Based on this deletion, they genotyped two SNPs on either side of the deletion (rs10461239 3′ of the deletion; rs17376826 5′ of the deletion) and one SNP within the deleted region (rs10461238). They showed highly significant association for rs1737826 with BMI that are additive to *FTO* and *MC4R* gene effects, suggesting that genetic variation in the *NPY2R* gene might increase the risk of early-onset obesity [38]. As we could not detect any copy number gain or loss in the target regio, we suggest that the deletion was a very rare or private mutation in the pedigree.

In summary, these results clearly demonstrate that sequence variation in *NPY* and *NPY2R* and CNV in *NPY2R* is rare, and thus would suggest that variants in these genes are unlikely to majorly contribute to the obese phenotype among the general population.

REFERENCES

1. Stunkard, A.J., T.T. Foch, and Z. Hrubec, *A Twin Study of Human Obesity*. Jama-Journal of the American Medical Association, 1986. **256**(1): p. 51-54.
2. Bouchard, C., et al., *The Response to Long-Term Overfeeding in Identical-Twins*. New England Journal of Medicine, 1990. **322**(21): p. 1477-1482.
3. Hinney, A., A.L. Volckmar, and N. Knoll, *Melanocortin-4 receptor in energy homeostasis and obesity pathogenesis*. Prog Mol Biol Transl Sci, 2013. **114**: p. 147-91.
4. Montague, C.T., et al., *Congenital leptin deficiency is associated with severe early-onset obesity in humans*. Nature, 1997. **387**(6636): p. 903-8.
5. Strobel, A., et al., *A leptin missense mutation associated with hypogonadism and morbid obesity*. Nat Genet, 1998. **18**(3): p. 213-5.
6. Gibson, W.T., et al., *Congenital leptin deficiency due to homozygosity for the Delta133G mutation: report of another case and evaluation of response to four years of leptin therapy*. J Clin Endocrinol Metab, 2004. **89**(10): p. 4821-6.
7. Mazen, I., et al., *A novel homozygous missense mutation of the leptin gene (N103K) in an obese Egyptian patient*. Mol Genet Metab, 2009. **97**(4): p. 305-8.
8. Fischer-Posovszky, P., et al., *A new missense mutation in the leptin gene causes mild obesity and hypogonadism without affecting T cell responsiveness*. J Clin Endocrinol Metab, 2010. **95**(6): p. 2836-40.
9. Fatima, W., et al., *Leptin deficiency and leptin gene mutations in obese children from Pakistan*. Int J Pediatr Obes, 2011. **6**(5-6): p. 419-27.
10. Saeed, S., et al., *High prevalence of leptin and melanocortin-4 receptor gene mutations in children with severe obesity from Pakistani consanguineous families*. Mol Genet Metab, 2012. **106**(1): p. 121-6.
11. Wasim, M., et al., *Role of Leptin Deficiency, Inefficiency, and Leptin Receptors in Obesity*. Biochem Genet, 2016. **54**(5): p. 565-72.
12. Farooqi, I.S., et al., *Effects of recombinant leptin therapy in a child with congenital leptin deficiency*. N Engl J Med, 1999. **341**(12): p. 879-84.
13. Farooqi, I.S., et al., *Beneficial effects of leptin on obesity, T cell hyporesponsiveness, and neuroendocrine/metabolic dysfunction of human congenital leptin deficiency*. J Clin Invest, 2002. **110**(8): p. 1093-103.
14. Locke, A.E., et al., *Genetic studies of body mass index yield new insights for obesity biology*. Nature, 2015. **518**(7538): p. 197-U401.

15. Willer, C.J., et al., *Six new loci associated with body mass index highlight a neuronal influence on body weight regulation.* Nat Genet, 2009. **41**(1): p. 25-34.
16. Speliotes, E.K., et al., *Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index.* Nature Genetics, 2010. **42**(11): p. 937-U53.
17. Beck, B., *Neuropeptide Y in normal eating and in genetic and dietary-induced obesity.* Philos Trans R Soc Lond B Biol Sci, 2006. **361**(1471): p. 1159-85.
18. Maffei, M., et al., *Leptin levels in human and rodent: measurement of plasma leptin and ob RNA in obese and weight-reduced subjects.* Nat Med, 1995. **1**(11): p. 1155-61.
19. Considine, R.V., et al., *Serum immunoreactive-leptin concentrations in normal-weight and obese humans.* N Engl J Med, 1996. **334**(5): p. 292-5.
20. Ding, B., et al., *Human neuropeptide Y signal peptide gain-of-function polymorphism is associated with increased body mass index: possible mode of function.* Regul Pept, 2005. **127**(1-3): p. 45-53.
21. van Rossum, C.T., et al., *Polymorphisms in the NPY and AGRP genes and body fatness in Dutch adults.* Int J Obes (Lond), 2006. **30**(10): p. 1522-8.
22. Pesonen, U., *NPY L7P polymorphism and metabolic diseases.* Regul Pept, 2008. **149**(1-3): p. 51-5.
23. Yeung, E.H., et al., *Polymorphisms in the neuropeptide Y gene and the risk of obesity: findings from two prospective cohorts.* J Clin Endocrinol Metab, 2011. **96**(12): p. E2055-62.
24. Billington, C.J., et al., *Effects of intracerebroventricular injection of neuropeptide Y on energy metabolism.* Am J Physiol, 1991. **260**(2 Pt 2): p. R321-7.
25. Kaga, T., et al., *Modest overexpression of neuropeptide Y in the brain leads to obesity after high-sucrose feeding.* Diabetes, 2001. **50**(5): p. 1206-10.
26. Patel, H.R., et al., *Neuropeptide Y deficiency attenuates responses to fasting and high-fat diet in obesity-prone mice.* Diabetes, 2006. **55**(11): p. 3091-8.
27. Erickson, J.C., G. Hollopeter, and R.D. Palmiter, *Attenuation of the obesity syndrome of ob/ob mice by the loss of neuropeptide Y.* Science, 1996. **274**(5293): p. 1704-7.
28. Nguyen, A.D., et al., *Y1 and Y5 receptors are both required for the regulation of food intake and energy homeostasis in mice.* PLoS One, 2012. **7**(6): p. e40191.
29. Naveilhan, P., et al., *Normal feeding behavior, body weight and leptin response require the neuropeptide Y Y2 receptor.* Nat Med, 1999. **5**(10): p. 1188-93.

30. Katsuura, G., A. Asakawa, and A. Inui, *Roles of pancreatic polypeptide in regulation of food intake.* Peptides, 2002. **23**(2): p. 323-9.
31. Aerts, E., et al., *CNV analysis and mutation screening indicate an important role for the NPY4R gene in human obesity.* Obesity (Silver Spring), 2016. **24**(4): p. 970-6.
32. Kassis, S., et al., *Neuropeptide Y inhibits cardiac adenylate cyclase through a pertussis toxin-sensitive G protein.* J Biol Chem, 1987. **262**(8): p. 3429-31.
33. Westlind-Danielsson, A., et al., *Neuropeptide Y receptors and the inhibition of adenylate cyclase in the human frontal and temporal cortex.* Neurosci Lett, 1987. **74**(2): p. 237-42.
34. Sainsbury, A., et al., *Synergistic effects of Y2 and Y4 receptors on adiposity and bone mass revealed in double knockout mice.* Mol Cell Biol, 2003. **23**(15): p. 5225-33.
35. Sainsbury, A., et al., *Important role of hypothalamic Y2 receptors in body weight regulation revealed in conditional knockout mice.* Proc Natl Acad Sci U S A, 2002. **99**(13): p. 8938-43.
36. Hung, C.C., et al., *Studies of the peptide YY and neuropeptide Y2 receptor genes in relation to human obesity and obesity-related traits.* Diabetes, 2004. **53**(9): p. 2461-6.
37. Ma, L., et al., *Variations in peptide YY and Y2 receptor genes are associated with severe obesity in Pima Indian men.* Diabetes, 2005. **54**(5): p. 1598-602.
38. Hunt, S.C., et al., *Polymorphisms in the NPY2R gene show significant associations with BMI that are additive to FTO, MC4R, and NPFFR2 gene effects.* Obesity (Silver Spring), 2011. **19**(11): p. 2241-7.
39. Roelants, M., R. Hauspie, and K. Hoppenbrouwers, *References for growth and pubertal development from birth to 21 years in Flanders, Belgium.* Ann Hum Biol, 2009. **36**(6): p. 680-94.
40. Hauspie, R. and M. Roelants. *Flemish Growth Charts 2004*. 2004 22 July 2013]; Available from: http://www.vub.ac.be/groeicurven.
41. Hauspie, R. and M. Roelants, *Flemish growth charts 2004*. . 2004.
42. Miller, S.A., D.D. Dykes, and H.F. Polesky, *A simple salting out procedure for extracting DNA from human nucleated cells.* Nucleic Acids Res, 1988. **16**(3): p. 1215.
43. Aerts, E., et al., *Genetic and structural variation in the SH2B1 gene in the Belgian population.* Mol Genet Metab, 2015. **115**(4): p. 193-8.
44. Adzhubei, I.A., et al., *A method and server for predicting damaging missense mutations.* Nat Methods, 2010. **7**(4): p. 248-9.

45. Ng, P.C. and S. Henikoff, *Predicting deleterious amino acid substitutions.* Genome Res, 2001. **11**(5): p. 863-74.

46. Calabrese, R., et al., *Functional annotations improve the predictive score of human disease-related mutations in proteins.* Hum Mutat, 2009. **30**(8): p. 1237-44.

47. Li, B., et al., *Automated inference of molecular mechanisms of disease from amino acid substitutions.* Bioinformatics, 2009. **25**(21): p. 2744-50.

48. J, O., *Documentation to linkage utility programs.* . 2013.

49. Multiplicom. Multiplicom the multiplex experts.; Available from: http://www.multiplicom.com.

50. Kumps, C., et al., *Multiplex Amplicon Quantification (MAQ), a fast and efficient method for the simultaneous detection of copy number alterations in neuroblastoma.* BMC Genomics, 2010. **11**: p. 298.

51. den Dunnen, J.T. and S.E. Antonarakis, *Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion.* Hum Mutat, 2000. **15**(1): p. 7-12.

52. Mitchell, G.C., et al., *A common single nucleotide polymorphism alters the synthesis and secretion of neuropeptide Y.* J Neurosci, 2008. **28**(53): p. 14428-34.

53. Niskanen, L., et al., *Leucine 7 to proline 7 polymorphism in the neuropeptide Y gene is associated with enhanced carotid atherosclerosis in elderly patients with type 2 diabetes and control subjects.* J Clin Endocrinol Metab, 2000. **85**(6): p. 2266-9.

54. Wang, Y., et al., *A neuropeptide Y variant (rs16139) associated with major depressive disorder in replicate samples from Chinese Han population.* PLoS One, 2013. **8**(2): p. e57042.

55. Mottagui-Tabar, S., et al., *A novel single nucleotide polymorphism of the neuropeptide Y (NPY) gene associated with alcohol dependence.* Alcohol Clin Exp Res, 2005. **29**(5): p. 702-7.

56. Karvonen, M.K., et al., *Association of a leucine(7)-to-proline(7) polymorphism in the signal peptide of neuropeptide Y with high serum cholesterol and LDL cholesterol levels.* Nat Med, 1998. **4**(12): p. 1434-7.

57. Sutton, S.W., et al., *Neuropeptide Y (NPY): a possible role in the initiation of puberty.* Endocrinology, 1988. **123**(4): p. 2152-4.

58. Blomqvist, A.G., et al., *Strong evolutionary conservation of neuropeptide Y: sequences of chicken, goldfish, and Torpedo marmorata DNA clones.* Proc Natl Acad Sci U S A, 1992. **89**(6): p. 2350-4.

59. Benned-Jensen, T. and M.M. Rosenkilde, *The role of transmembrane segment II in 7TM receptor activation.* Curr Mol Pharmacol, 2009. **2**(2): p. 140-8.

PART 3

GENES INVOLVED IN BROWN FAT

BROWN ADIPOSE TISSUE INDUCTION AND OBESITY: GENETIC ASPECTS OF UCP1.

Evi Aerts[1], Kim Van Hoorenbeeck[2], Guy Massa[3], An Verrijken[4], Ilse L Mertens[4], Stijn L Verhulst[2], Luc F Van Gaal[4] & Wim Van Hul[1]

[1] Centre of Medical Genetics, University of Antwerp, Antwerp, Belgium

[2] Department of Pediatrics, Antwerp University Hospital, Antwerp, Belgium

[3] Department of Pediatrics, Jessa Hospital, Hasselt, Belgium

[4] Department of Endocrinology, Diabetology and Metabolic Diseases, Antwerp University Hospital, Antwerp, Belgium

Work in progress

ABSTRACT

Objective Obesity is a highly heritable complex and heterogeneous disorder, characterized by an excessive amount of adipose tissue. Numerous studies have shown that 40-70% of the interindividual variability in BMI is attributed to genetic factors, of which approximately 2,7% are currently explained by associated polymorphisms. Copy number variants (CNVs) account for a major proportion of human genetic variation and have been predicted to have an important role in genetic susceptibility to common disease. Genomewide copy number variation (CNV) analyses have associated the 4q31 CNV, containing the *UCP1* gene, with obesity. *UCP1*, encoding for the thermogenic protein in brown and beige adipose tissue, is the most interesting candidate gene in this region, we hypothesized that both genetic and structural variation in *UCP1* may be implicated in the pathogenesis of obesity.

Design and methods In the first part of this study, Multiplex Amplicon Quantification (MAQ) analysis was used to identify CNVs in the *UCP1*-containing chr.4q31 region in 306 obese children and adolescents. In the second part of this study, we performed a mutation screen for variants in the *UCP1* coding region in 643 obese children and adolescents, and 445 healthy lean adults. To functionally identify potential pathogenic effects, plate-based respirometry (Seahorse technology) will be performed on overexpressed UCP1 variants in intact cells.

Results and conclusion In our CNV analysis we could not identify CNV in the *UCP1* region in our population. Mutation analysis resulted in the identification of three rare non-synonymous heterozygous variants, 2 of which could only be found among obese individuals (G57S and T227I) and 1 that was only identified in a lean subject (T157I). By performing *in silico* analysis, we determined that these 3 variants are probably damaging to the protein structure and might have a disease causing effect. Further functional testing are ongoing to fully understand the impact on UCP1. As we could not identify any CNV in the *UCP1* region, structural variation in this gene is unlikely to majorly contribute to the obese phenotype.

INTRODUCTION

Copy number variation (CNV) accounts for a major proportion of human genetic variation, underlying population diversity and human diseases ranging from neuropsychiatric diseases to cancer. With the advent of genome-wide analysis tools, including array comparative genomic hybridization, SNP microarrays and genome sequencing, the ability to detect these alterations has enhanced, revealing that they are widely spread along the human genome. So far, the database of Genomic variants reports a presence of 552586 CNVs in the human genome with an estimated 75.6% of exons and 91.2% of transcripts that are overlapped by at least one CNV. In order to shed some light on the missing heritability of obesity, the contribution of CNVs was explored by genome-wide CNV surveys in patient populations of early-onset obese individuals. The genes that are impacted by these CNVS might encompass known or novel candidate genes for follow-up studies. One of these surveys identified a 2.1Mb deletion on chromosome 4q31, encompassing *Uncoupling protein 1 (UCP1)* and Interleukin-15 (*IL15),* as an important risk factor in the study of obesity. The deletion was found in a severely obese patient with a BMI of 46.2 kg/m^2 that inherited the CNV from his father. Based on additional genetic evidence, the *UCP1* gene was identified as the strongest a priori candidate gene for obesity in this CNV region [1]. Numerous studies investigated the association of three polymorphisms in the promotor region (-3826A/G, -1766A/G and -112A/C) and A64T in exon 2 and M299L in exon 5 of the *UCP1* gene with obesity. Although the results of these studies were not always consistent, the -3826A/G polymorphism was reported to clearly contribute to the susceptibility of obesity and obesity-related parameters [2-5]. In animal studies, *UCP1*-ablation on itself induced obesity in C57Bl6 mice that were kept under thermoneutral conditions [6]. UCP1 is a thermogenic protein that plays a key role in the brown and beige adipogenesis. While white adipose tissue (WAT) is the main tissue of energy storage, brown adipose tissue (BAT) significantly contributes to the control of body temperature and energy expenditure. More recently, a third type of adipocytes was identified, termed beige adipocytes, with inducible brown-like thermogenic activities. However, while brown adipocytes express high levels of *UCP1* and other thermogenic genes under basal conditions, beige adipocytes require activators such as agonists of the β-adrenergic receptor or peroxisome proliferator-activated receptor-γ. UCP1 is a mitochondrial carrier protein that generates heat and diminishes the proton gradient by uncoupling cellular respiration and mitochondrial ATP synthesis (figure 1) [7]. This process, termed non-shivering thermogenesis, has the capacity to increase energy expenditure and thus may offer new therapeutic strategies to counteract obesity. In 2011, a small-scale mutation screen was performed among 100 obese

patients and 100 control subjects. In this study, 7 variations in the promotor region, 4 in the intronic region and 4 in the exonic region (two coding variants A64T and M229L) were identified. Although they concluded that *UCP1* variants could represent thrifty factors that promote energy storage, they advised to screen the *UCP1* gene in larger populations [8]. Therefore, we designed an extensive mutation and CNV screen in our study population of severely overweight or obese children and adolescents to study whether genetic variation might be implicated in the pathogenesis of obesity.

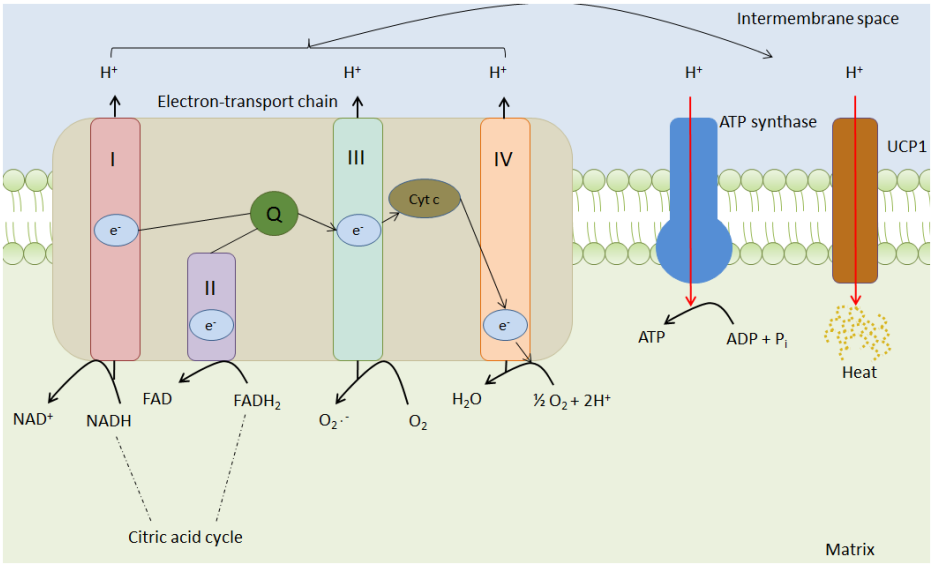

Figure 1: UCP1 location and function in the mitochondrial respiratory chain (MRC).
During respiration, protons are pumped through the MRC complexes (I-IV), generating a proton gradient. This proton gradient drives the synthesis of ATP by the ATP-synthase complex. UCP1 catalyzes a regulated proton leakage and consequently increases basal energy expenditure.

METHODS AND PROCEDURES

Study population

For CNV-analysis we screened 343 severely overweight and obese children and adolescents (136 boys and 165 girls). Mutation screening in the coding region of *UCP1* was performed in a population of 644 severely overweight and obese children and adolescents (277 boys, 367 girls), and 445 healthy lean adults. Severely overweight/obese children and adolescents (adolescent if age ≥ 12 years) are unrelated and were recruited at the Child Obesity Clinics from the Antwerp

University Hospital (Antwerp) and Jessa Hospital (Hasselt), both in Belgium. The Flemish Growth Charts 2004 were used to identify overweight and obese individuals [9, 10]. Patients with mutations in *Leptin* and the *Melanocortin-4* receptor (*MC4R)* causing known monogenic obesity and carriers of the 16p11.2 CNV have been excluded from the screening sample. Our white lean population consists of healthy, lean adults (18.5 kg/m^2 ≤ BMI < 25 kg/m^2) recruited among employees from the Antwerp University Hospital and the University of Antwerp, and among couples seeking prenatal counselling at the Centre for Medical Genetics (due to increased triple test or high maternal age). All participants gave their written informed consent. For children participating in this study, parental permission was given. The study protocol was approved by the local ethics committee. Population characteristics are summarized in table 1.

Anthropometry

Height was measured to the nearest 0.1 cm. Weight was measured on a digital scale to the nearest 0.1 kg. Body mass index (BMI) was calculated for all individuals as weight (in kg) over height (in m) squared. For children and adolescents, BMI Z-scores were calculated using the data from the Flemish Growth Curves 2004. The cutoff values for overweight and obesity are defined as the percentile line on the chart that crosses BMI 25 kg/m^2 and, respectively, BMI 30 kg/ m^2 at 18 years of age [9, 10].

Table 1: Characteristics of the population used for mutation and CNV analysis of *UCP1*

	Mutation analysis		CNV analysis
	Obese patients	Lean Adults	Obese patients
N	644	445	343
Male/female	277/367	253/192	136/165
Age (years)	10,8 ± 0,4	34,4 ± 0,3	10,6 ± 0,4
Weight (kg)	75,4 ± 1,7	65,8 ± 0,5	75,1 ± 1,6
Height (m)	1,46 ± 0,02	1,73 ± 0,01	1,45 ± 0,02
BMI (kg/m^2)	30,6 ± 0,4	21,9 ± 0,1	30,3 ± 0,3
BMI Z-score	2,58 ± 0,04	N.A.	2,58 ± 0,04

Mean value ± standard error of the mean is shown for all parameters, except N and gender distribution (absolute numbers).

MAQ

Extraction of genomic DNA from blood samples from all patients and control subjects were performed using standard procedures [11]. To detect and analyse CNVs, we used Multiplex Amplicon Quantification assays (MAQ), consisting of a multiplex PCR based amplification of fluorescently labelled target and reference amplicons, followed by capillary electrophoresis and fragment analysis. MAQ-assays were performed on 20 ng of genomic DNA. Target primers were designed with the Flagged-You MAQ primer design tool and its concentrations were optimized according to the manufacturer's protocol. Final primer concentrations varied between 0.05 and 0.95 µM in the Flagged-you MAQ PCR mix which was prepared according to the manufacturer's protocol [12]. Target region was set at chr 4:141480800-141491000 (NCBI Build GRCh37.p13) and comprises *UCP1*. Two negative control samples from lean individuals were included in each experiment for accurate normalisation. Standard PCR cycle conditions were followed. MAQ primer sequences are available upon request.

Capillary electrophoresis and fragment analysis

Fragment analysis of the resulting MAQ products was performed on an ABI Prism Genetic Analyzer 3130xl (Applied Biosystems Inc, Foster City, CA, USA). The MAQ-S software package was used (Multiplicom) to calculate and visualise the dosage quotient which reflects the copy number of each target amplicon [12]. A dosage quotient of 1.25 -1.75 was considered indicative of a duplication, a DQ of 0.25-075 was indicative of a deletion.

Mutation analysis of UCP1

Blood samples from all patients and control subjects were obtained for extraction of genomic DNA by standard procedures [11]. Mutation screening of the entire coding region of the *UCP1* gene (Genbank accession number NM_021833.4), including intron-exon boundaries was performed by high-resolution melting curve analysis (HRM) in combination with direct sequencing. The *UCP1* gene consists of 6 coding exons (<300 bp). HRM, which is performed using the LightCycler 480 Real-Time PCR System (Roche Applied Science, Mannheim, Germany), is a highly sensitive real-time PCR technique to detect novel mutations in a large sample set. This real-time PCR was performed in a reaction volume of 10 µl. The amplification mixture included 10 ng of template DNA, 1.5 mM $MgCl_2$, 0.15 mM dNTP's, 500 nM primers, 0.015 U/µl GoTaq (Promega Corporation, Madison, WI, USA), 1x GoTaq buffer and the saturating dye LCGreen+ (Idaho Technology, Salt Lake City, UT, USA) in a 0.5x concentration. Following

amplification, samples were denaturated at 95°C, subsequently renaturated at 40°C and then melted between 60°C and 90°C while constantly measuring fluorescence. Samples with melting curves deviating from wild type were sequenced. Sequencing was performed with ABI BigDye Terminator v1.1 Cycle Sequencing kits on an ABI Prism Genetic Analyzer 3130xl (Applied Biosystems Inc, Foster City, CA, USA). Primer sequences for HRM and sequencing are available upon request.

Prediction programs

In silico analysis to score the deleteriousness of the identified non-synonymous variants on protein function was performed using the Combined Annotation Dependent Depletion (CADD) score. This broadly applicable metric objectively weighs and prioritizes functional, deleterious and disease causal variants [13]. A CADD score of 20 means that a variant is amongst the top 1% of deleterious variants in the human genome, while a variant with a CADD score of 30 is on the 0.1% of human variants.

Statistical analysis

For rs45539933 and rs2270565, Hardy Weinberg equilibrium (HWE) was calculated using the HWE tool from LINKUTIL [14]. Comparison of genotype and allele frequencies among cases and controls was performed using Pearson's chi square analysis. Significance level was set at p=0.05. All statistical analyses were performed using SPSS version 20.0 (SPSS, Chicago, IL, USA).

Expression constructs and mutagenesis

Full-length human wild-type *UCP1* (NM.021833.4), expressed in a mammalian expression vector with Neomycin selection marker (EX-C0618-M02) obtained from Labomics. Three different missense mutations (G57S, T157I and T227I) were individually introduced into the WT *UCP1* expression vector, using the QuickChange II Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) according to the manufacturer's protocol. The WT sequence and the presence of mutations were checked and confirmed by direct sequencing using an ABI 3130xl Genetic Analyzer (Applied Biosystems).

Extracellular flux analysis

Extracellular flux analysis is used to evaluate UCP1 function in a quantitative manner. For this functional assay, we collaborate with the Mitochondrial Biology Unit of the Helmholtz Zentrum München, led by Prof. Dr. Jastroch. These experiments are still ongoing.

RESULTS

CNV-analysis of UCP1

We screened 343 obese children and adolescents for copy number variation in *UCP1* with MAQ. We could not detect any copy number gain or loss in the target regio. The previously reported 2.1Mb deletion on chromosome 4q31, encompassing *Uncoupling protein 1 (UCP1)* and *Interleukin-15 (IL15)* could not be identified in the patient population.

Mutation analysis of UCP1

Mutation analysis of the coding region of the *UCP1* gene resulted in the identification of 11 different heterozygous coding variants in our populations: 5 rare synonymous variants (A80A, S88S, S184S, C225C and V233V), 2 common polymorphisms (A64T and M229L) and 4 rare non-synonymous variants (R47W, G57S, T157I and T227I). Of the latter, R47W was identified in both, an obese patient and a lean individual. Two variants were exclusively found among obese patients (G57S and T227I) and one was only identified in a lean adult (T157I). For the two common SNPs there were no significant differences in genotype frequencies between obese and control subjects. This is in contrast with the slightly higher mean BMI value that was reported for patients carrying the polymorphic allele of the M229L and A64T substitution in a Japanese and Korean population respectively [15, 16]. *In silico* analysis was used to select the most interesting variants for further functional research (CADD score >20): R47W, G57S, A64T, T157I and T227I. As the prevalences of R47W and A64T in the obese and lean population were not significantly different, functional experiments were not performed for these variants. An overview of the results is found in table 2.

Table 2: Coding variants identified in obese and lean individuals.

SNP	Variation		MAF obese cases/lean adults	MAF ExAC EU	CADD score
	gene	protein			
Rs150067245	c.118C>T	R47W	0.002/0.002	0.005	25
Rs141520915	c.169G>A	G57S	0.002/-	<0.0001	34
Rs45539933	c.190G>A	A64T	0.12/0.11	0.08	28.7
Rs79610439	c.240G>C	A80A	0.003/-	0.004	12.5
Rs138241273	c.264C>T	S88S	-/0.002	0.0002	13.61
-	c.470C>T	T157I	-/0.002	0.08	22.3
Rs202215906	c.552T>G	S184S	0.002/-	<0.0001	13.61
Rs142052143	c672C>T	C225C	0.002/-	0.006	6.978
Rs148598275	c.680C>T	T227I	0.006/-	<0.0001	24.6
Rs2270565	c.682A>T	M229L	0.11/0.11	0.08	10.28
-	c.696T>A	V233V	0.002/-	-	13.04

Position of variants on gene and protein levels is shown, as well as their frequency in the obese and lean groups and the reported MAF in a European non-Finnish population (Exome Aggregation Consortium browser). Numbering on gene level is based on cDNA sequence (GenBank NM_021833.4), following the recommendations by the Human Genome Variation Society [17].

Functional analysis of mutant UCP1

Work in progress.

DISCUSSION

In the first part of this study we performed CNV-analysis on the *UCP1* gene region to study whether the previously reported 2.1Mb deletion contributes to obesity in the general population. In 2010, this deletion, encompassing *UCP1* and *IL15,* was identified as an important risk factor for the study of obesity. The deletion was found in a severely obese patient with a BMI of 46.2 kg/m^2 that inherited the CNV from his father. No further clinical details or segregation data were reported.

As we could not detect any copy number gain or loss in the target regio, we suggest that the deletion is a very rare or private mutation in the pedigree that does not majorly contribute to the obese phenotype among the general population.

In the second part of this study, we screened the coding region of *UCP1* in 644 severely overweight/obese children and adolescents and 445 healthy lean adults. This mutation analysis resulted in the identification of 4 rare non-synonymous variants with high CADD scores of which two (G57S and T227I) were found exclusively among patients, while T157I was only seen in a lean adult. With CADD scores above 20, these variants were predicted to have potential effect on the activity of UCP1 and were selected for functional studies. These studies are still ongoing.

As T157I was privately identified in a lean adult, we expect the variant to have a rather gain of function effect on UCP1 and brown adipogenesis. This is in contrast to G57S and T227I that were predicted to have a loss of function effect on UCP1 leading to decreased energy expenditure and increasing obesity risk.

REFERENCES

1. Wang, K., et al., *Large copy-number variations are enriched in cases with moderate to extreme obesity.* Diabetes, 2010. **59**(10): p. 2690-4.
2. Brondani, L.A., et al., *The role of the uncoupling protein 1 (UCP1) on the development of obesity and type 2 diabetes mellitus.* Arq Bras Endocrinol Metabol, 2012. **56**(4): p. 215-25.
3. Hayakawa, T., et al., *Phenotypic characterization of the beta3-adrenergic receptor mutation and the uncoupling protein 1 polymorphism in Japanese men.* Metabolism, 1999. **48**(5): p. 636-40.
4. Heilbronn, L.K., et al., *Association of -3826 G variant in uncoupling protein-1 with increased BMI in overweight Australian women.* Diabetologia, 2000. **43**(2): p. 242-4.
5. Nakano, T., et al., *A/G heterozygote of the A-3826G polymorphism in the UCP-1 gene has higher BMI than A/A and G/G homozygote in young Japanese males.* J Med Invest, 2006. **53**(3-4): p. 218-22.
6. Feldmann, H.M., et al., *UCP1 ablation induces obesity and abolishes diet-induced thermogenesis in mice exempt from thermal stress by living at thermoneutrality.* Cell Metab, 2009. **9**(2): p. 203-9.
7. Kozak, L.P. and R. Anunciado-Koza, *UCP1: its involvement and utility in obesity.* Int J Obes (Lond), 2008. **32 Suppl 7**: p. S32-8.
8. Labruna, G., et al., *Sequence Analysis of the UCP1 Gene in a Severe Obese Population from Southern Italy.* J Obes, 2011. **2011**: p. 269043.
9. Roelants, M., R. Hauspie, and K. Hoppenbrouwers, *References for growth and pubertal development from birth to 21 years in Flanders, Belgium.* Ann Hum Biol, 2009. **36**(6): p. 680-94.
10. Hauspie, R. and M. Roelants, *Flemish growth charts 2004.* . 2004.
11. Miller, S.A., D.D. Dykes, and H.F. Polesky, *A simple salting out procedure for extracting DNA from human nucleated cells.* Nucleic Acids Res, 1988. **16**(3): p. 1215.
12. Multiplicom. Multiplicom the multiplex experts.; Available from: http://www.multiplicom.com.
13. Kircher, M., et al., *A general framework for estimating the relative pathogenicity of human genetic variants.* Nat Genet, 2014. **46**(3): p. 310-5.
14. J, O., *Documentation to linkage utility programs.* . 2013.

15. Cha, M.H., et al., *Association of UCP1 genetic polymorphisms with blood pressure among Korean female subjects.* J Korean Med Sci, 2008. **23**(5): p. 776-80.

16. Mori, H., et al., *A polymorphism in the 5' untranslated region and a Met229-->Leu variant in exon 5 of the human UCP1 gene are associated with susceptibility to type II diabetes mellitus.* Diabetologia, 2001. **44**(3): p. 373-6.

17. den Dunnen, J.T. and S.E. Antonarakis, *Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion.* Hum Mutat, 2000. **15**(1): p. 7-12.

IDENTIFICATION AND FUNCTIONAL CHARACTERIZATION OF FGF21 MUTATIONS IN OBESE INDIVIDUALS.

Evi Aerts[1], Kim Van Hoorenbeeck[2], Guy Massa[3], An Verrijken[4], Ilse L Mertens[4], Stijn L Verhulst[2], Luc F Van Gaal[4] & Wim Van Hul[1]

[1] Centre of Medical Genetics, University of Antwerp, Antwerp, Belgium

[2] Department of Pediatrics, Antwerp University Hospital, Antwerp, Belgium

[3] Department of Pediatrics, Jessa Hospital, Hasselt, Belgium

[4] Department of Endocrinology, Diabetology and Metabolic Diseases, Antwerp University Hospital, Antwerp, Belgium

Submitted

ABSTRACT

Objective Brown adipose tissue is a key site of heat production that has the potential to increase energy expenditure. After the discovery of active brown fat and inducible "brown-like" or beige adipocytes in human adults, browning of white adipose tissue has been considered an attractive obesity target. The fibroblast growth factor 21 (FGF21) is one of the most essential inducers of beige adipocytes and so dysregulation of this endocrine factor may be implicated in energy regulation. Therefore, we hypothesized that pathogenic *FGF21* mutations might be involved in the pathogenesis of obesity.

Methods We screened 406 severely overweight or obese children and adolescents for mutations in the coding region of the *FGF21* gene with high-resolution melting curve analysis. Direct sequencing was performed for samples with melting patterns deviating from wild-type. To determine the effect of the mutations on the FGF21-induced activity of the FGFR1, luciferase reporter assays were performed to measure cAMP levels.

Results Our mutation analysis resulted in the identification of two non-synonymous coding variants in two unrelated obese individuals: p.R47Q and p.L142V. For p.L142V we were able to demonstrate a significant decrease of the FGF21-induced cAMP signaling activity.

Conclusion With this study, we have demonstrated the presence of a rare variant in a severely obese 6-year-old girl (BMI z-score 3.01) and confirmed its pathogenic effect by performing a luciferase reporter assay. In conclusion, these data confirm the importance of genetic variation within the *FGF21* gene in energy regulation.

INTRODUCTION

Obesity is an important health problem with a massive impact on human health. It is a complex and chronic disorder with a high and ever-increasing prevalence. Apart from an obesogenic environment and a passive lifestyle, numerous studies have shown that genetic factors play an important role in the pathogenesis of obesity, as the estimated heritability is about 40-70%. Because of the burden this disease poses on an individual's health, research efforts searching for factors contributing to the disease's etiology have intensified in the past decades. This led to the recognition of the leptin-melanocortin pathway as a key regulator of energy homeostasis. Mutations in the genes of this pathway have proven to be responsible for severe early-onset monogenic obesity. With a prevalence of 4% among severely obese patients and over 160 distinct mutations reported so far, melanocortin-4 receptor (*MC4R*) gene mutations have shown to be the most prevalent genetic cause of monogenic obesity known to date [1]. In contrast, only six different mutations in the leptin gene have been discovered, making leptin deficiency a rare form of monogenic obesity (< 1/1 000 000) [2-9]. Leptin deficiency is currently, however, still the only treatable type of monogenic obesity [4, 10, 11]. [4, 10, 11]. Leptin receptor deficiency is more common and has a prevalence of 1% among early-onset obese patients [12]. Furthermore, genome-wide association studies have been performed to identify common single-nucleotide polymorphisms (SNPs) underlying the heritable risk for human obesity. So far, 97 BMI-associated loci were identified that account for approximately 2.7% of BMI variation, leaving a substantial component of the heritability unexplained [13-15].

Recently, there has been considerable interest in the role of brown and beige adipogenesis in energy regulation. Fibroblast growth factor 21 (FGF21) is a predominantly liver derived hormone that has emerged as a key regulator of glucose and lipid homeostasis and belongs to the FGF19 subfamily of endocrine growth factors. FGF21 binds multiple FGF receptors including FGFR1, -2, -3 and -4, of which FGFR1 with the highest affinity, in the presence of co-receptor *β-Klotho* (*KLB)*. It binds KLB with its C terminus, while its N terminus interacts with the D2 and D3 immunoglobulin-like domains of the receptor. The FGF receptors are tyrosine kinase receptors that are expressed at the plasma membrane and ligand binding induces several known signalling cascades including the mitogen-activated protein kinase (MAPK)/extracellular signal-regulated kinase (ERK) pathway [16]. FGF21 is an important metabolic regulator that induces *UCP1* expression and plays a key role in the brown and beige adipogenesis. It is induced upon thermogenic stimulation and regulates metabolic parameters such as body weight, glucose homeostasis and plasma triglycerides [17-19].

Association studies confirm the importance of FGF21 in energy regulation and food intake. The common SNP rs838133 has been associated with obesity related phenotypes, including macronutrient and sugar intake [20, 21]. In 2009, human *FGF21* overexpression in mice was reported to result in decreased body weight and increased resistance to diet-induced obesity (DIO). Consistent with these findings, pharmacological treatment with human FGF21 for 2 weeks in DIO and *ob/ob* mice lowered the body weight with 20%, predominantly by a reduction in adiposity. As no direct decrease of caloric intake was observed, the FGF21-induced weight loss resulted from increased energy expenditure. *FGF21* KO mice were reported to have mild weight gain and slightly impaired glucose homeostasis, underlining its important role in long-term energy homeostasis. In both obese mice and humans, FGF21 levels are greatly increased, leading to the suggestion that obesity may give rise to an *FGF21*-resistant state.

Based on the above data, we hypothesized that sequence variation in the *FGF21* gene might be implicated in the pathogenesis of obesity. Therefore, we performed an extensive mutation screen in our patient population and designed a luciferase assay for monitoring the signaling capacity of FGFR1 after stimulation with mutant FGF21.

METHODS AND PROCEDURES

Study population

We screened the coding region of *FGF21* in 406 severely overweight and obese children and adolescents (179 boys and 227 girls). Severely overweight/obese children and adolescents (adolescent if age ≥ 12 years) are unrelated and were recruited at the Child Obesity Clinics from the Antwerp University Hospital (Antwerp) and Jessa Hospital (Hasselt), both in Belgium. The Flemish Growth Charts 2004 were used to identify overweight and obese individuals. The percentile lines that cross BMI 25 and 30 kg/m^2 at 18 years of age on the Flemish age- and sex-specific BMI growth curves were used as cut-off for the diagnosis of overweight and obesity respectively [22, 23]. Patients with mutations in *Leptin* and the *Melanocortin-4* receptor causing known monogenic obesity have been excluded from the screening sample. All participants gave their written informed consent. For children participating in this study, parental permission was given. The study protocol was approved by the local ethics committee. Population characteristics are summarized in table 1.

Table 1: Characteristics of the population used for mutation analysis of *FGF21*

	Obese subjects	
	Children	Adolescents
N	208	198
Male/female	116/92	63/135
Age (years)	9,3 ± 0,2	13,9 ± 0,2
Weight (kg)	53,9 ± 1,5	95,2 ± 1,7
Height (m)	1,19 ± 0,05	1,39 ± 0,06
BMI (kg/m^2)	28,2 ± 0,4	33,4 ± 0,6
BMI Z-score	2,51 ± 0,07	2,6 ± 0,05

Mean value ± standard error of the mean is shown for all parameters, except N and gender distribution (absolute numbers).

Anthropometry

Height was measured to the nearest 0.1 cm. Weight was measured on a digital scale to the nearest 0.1 kg. Body mass index (BMI) was calculated for all individuals as weight (in kg) over height (in m) squared. For children and adolescents, BMI Z-scores were calculated using the data from the Flemish Growth Curves 2004 [22].

Mutation analysis

Genomic DNA was isolated from whole blood by standard procedures [24]. The *FGF21* gene (GenBank accession number NM_019113), encoding 1 protein isoform of 209 amino acids. Mutation screening of the entire coding region of the *FGF21* gene, including intron-exon boundaries was performed by high-resolution melting curve analysis (HRM) in combination with direct sequencing. The coding region was divided in amplicons of approximately 300 base pairs (bp) for HRM and 700 bp for sequencing. HRM, which is performed using the LightCycler 480 Real-Time PCR System (Roche Applied Science, Mannheim, Germany), is a highly sensitive real-time PCR technique to detect novel mutations in a large sample set. HRM was performed in a reaction volume of 10 µl. The amplification mixture included 10 ng of template DNA, 1.5 mM $MgCl_2$, 0.15 mM dNTP's, 500 nM primers, 0.015 U/µl GoTaq (Promega Corporation, Madison, WI, USA), 1x GoTaq buffer and the saturating dye LCGreen+ (Idaho Technology, Salt Lake City, UT, USA) in a 0.5x concentration. Following amplification, samples were denaturated at 95°C, subsequently renaturated at 40°C and then melted between 60°C and 90°C while constantly measuring fluorescence. Samples with melting curves deviating from wild type were sequenced.

Sequencing was performed with ABI BigDye Terminator v1.1 Cycle Sequencing kits on an ABI Prism Genetic Analyser 3130xl (Applied Biosystems Inc, Foster City, CA, USA). Identified mutations were confirmed by resequencing. Primer sequences for HRM and sequencing are available upon request.

Prediction programs

In silico analysis to score the deleteriousness of the identified non-synonymous variants on protein function was performed using the Combined Annotation Dependent Depletion (CADD) score. This broadly applicable metric objectively weighs and prioritizes functional, deleterious and disease causal variants [25].

Expression constructs and mutagenesis

Full-length human wild-type *FGF21* (NM_005972.5), *FGFR1* (NM_023110.2) and *KLB* (NM_175737.3) were cloned in a pReceiver-M02 expression vector obtained from GeneCopoeia. Two different missense mutations (R47Q and L142V) were individually introduced into the WT *FGF21* expression vector, using the QuickChange II XL Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) according to the manufacturer's protocol. The WT sequence and the presence of mutations were checked and confirmed by direct sequencing using an ABI 3130xl Genetic Analyzer (Applied Biosystems).

Luciferase assay for cAMP measurement

For monitoring the FGF21-induced cAMP signaling activity of FGFR1 in cells transiently expressing WT or mutant *FGF21*, a luciferase assay was performed. Human embryonic kidney (HEK) 293T cells were cultured in Dulbecco's modified Eagle medium (DMEM) (Invitrogen) supplemented with fetal bovine serum (10% v/v, Invitrogen), penicillin/streptomycin and L-glutamine. Twenty-four hours prior to transfection, cells were plated at a density of 0.3 x 10^5 cells/well in 96-well plates. Cells were transiently transfected using Fugene6 (Roche Applied Science) with 25 ng pCRE-luc (Stratagene, La Jolla, CA, USA), 2.5 ng pRL-TK, 5ng KLB, 15ng FGFR1 and 5 ng WT/mutant FGF21 pReceiver-M02 or empty pcDNA3.1 construct per well. Each transfection was carried out in triplicate and repeated independently in three separate experiments. Forty-eight hours later, cells were washed in PBS, lysed and assayed for luciferase activity using the Dual-Glo™ Luciferase Assay System (Promega, Madison, WI, USA) on a Glomax-multi Detection System (Turner Biosystems, Sunnyvale, CA) according to the

manufacturer's protocol. The ratio of the firefly and pRL-TK Renilla luciferase measurement was calculated. Relative ratios of Firefly luciferase over Renilla luciferase readings were normalized to the relative ratio calculated for the transfection results with the empty pcDNA3.1 vector. Comparison between two measurements for a single experiment was performed using a Student's t-test. Values of $p < 0.05$ were considered significant. Statistical tests were provided by the SPSS 22.0 software package (SPSS Inc.).

RESULTS

Mutation analysis

Screening of the coding region of the *FGF21* gene in our population identified 7 heterozygous variants; two common SNPs (G12G and L174P), 3 rare synonymous variants (T6T, H140H and Y207Y) and 2 non-synonymous variants (R47Q and L142V). The allele frequencies of the two common variants rs1047214 (I195I) and rs2880415 (I312I) in our patient population were not significantly different from the reported minor allele frequency in the European non-Finnish population in the ExAC database. CADD was used to score the deleteriousness of the identified variants on protein function. The two rare non-synonymous variants were predicted to have the most functional effect and were selected for *in vitro* signaling experiments. An overview of the results was given in table 2.

Table 2: Coding variants identified in patient population

SNP	Variation Gene	Variation Protein	MAF obese cases	MAF ExAC EU	CADD score
Rs200052141	c.18C>A	T6T	0.001	0.0008	11.76
Rs838133	c.36A>G	G12G	0.59	0.56	3.6
Rs36123953	c.140G>A	R47Q	0.001	-	23.0
Rs3745710	c.420C>T	H140H	0.08	0.09	7.0
Rs144978172	c.424C>G	L142V	0.001	<0.0001	9.6
Rs739320	c.521T>C	L174P	0.65	0.63	0.3
Rs838130	c.621C>T	Y207Y	0.002	0.0005	6.0

Position of variants on gene and protein levels is shown, as well as their frequency in the obese and lean groups and the reported MAF in a European non-Finnish population (Exome Aggregation Consortium browser). Numbering on gene level is based on cDNA sequence (GenBank NM_021833.4), following the recommendations by the Human Genome Variation Society [26].

Functional analysis of mutant FGF21

In vitro signaling experiments were carried out to determine the influence of mutant FGF21 on the cAMP signaling function of FGFR1. We performed a luciferase assay to monitor the FGF21-induced activation of cAMP production. First of all, we assayed the intracellular cAMP signaling pathway in HEK293 cells transiently expressing increasing amounts of *FGF21* to confirm intracellular cAMP is increased in response to increasing amounts of FGF21 (figure 1). Furthermore, HEK293 cells were transfected with mutant FGF21 in order to compare the properties of WT and mutant FGF21. Cells transfected with the two mutated constructs demonstrated a decreased response on cAMP activity. However, only for L142V this difference was significantly different from WT (figure 2). In conclusion, we were able to demonstrate FGF21 dysfunction on cAMP signaling and thus pathogenic effect for L142V.

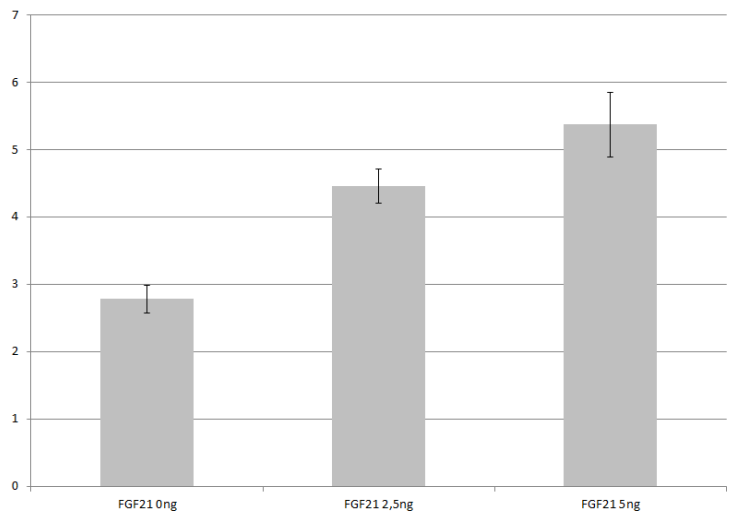

Figure 1: FGF21-induced signaling activity of FGFR1 in HEK293T cells expressing 0-2.5-5 ng WT FGF21.
The ratio of the firefly and pRL-TK Renilla luciferase measurement was calculated. Relative ratios of Firefly luciferase over Renilla luciferase readings were normalized to the relative ratio calculated for the transfection results with the empty pcDNA3.1 vector (y-axis). Bars represent average values ± SD. Each transfection was carried out in triplicate and repeated independently in three separate experiments. Intracellular cAMP signaling of FGFR1 is increased in response to stimulation with increasing concentration *FGF21*.

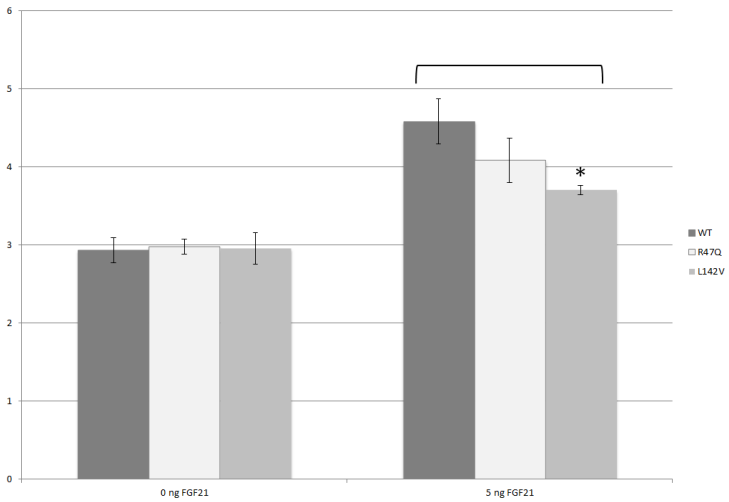

Figure 2: FGF21-induced signaling activity of FGFR1 in HEK293T cells expressing WT or mutant FGF21. Cells transfected with mutated L142V FGF21 demonstrated a significantly lower cAMP signaling activity after stimulation with 5ng mutant FGF21 when compared to WT FGF21.

DISCUSSION

In this study we investigated whether sequence variation in the *FGF21* gene might be implicated in the pathogenesis of obesity. Therefore, we designed an extensive mutation screen of *FGF21* in our patient population of severely overweight or obese children and adolescents. This resulted in the identification of 2 common SNPs, 3 synonymous variants and two rare non-synonymous variants R47Q and L142V. As the prevalence of both common SNPs in our patient population was similar to the one that was reported for the European population (ExAC database), we assume that they don't contribute majorly to the obese phenotype of our obese patients. For the identified non-synonymous variants R47Q and L142V, a functional effect was predicted. When looking at the localization of the variants and the known functional domains of the protein, L142V occurs outside a known functional domain, while R47Q is estimated to disturb an important receptor interaction site. With a luciferase reporter assay using cAMP-responsive constructs, we investigated the signaling activity of mutant FGF21 on FGFR1. For L142V, a significant decrease in cAMP signaling activity was measured when compared to WT FGF21, while for R47Q only a trend towards decreased signaling activity was demonstrated. This is in contrast to the CADD scores that predicted the most severe pathogenic effect for R47Q. L142V was identified in a 6-year-old severely obese girl (BMI z-score 3.01). Unfortunately, neither the

patient nor the family was willing to cooperate in further research, so we were unable to perform segregation analysis. Taken together, these data led to the conclusion that L142V is a pathogenic mutant while a trend towards decreased activity was determined for R47Q. Based on the findings in *FGF21* KO mice, the observed decrease of FGF21 activity is in line with the decreased induction of beige adipogenesis that was expected in obese subjects.

In conclusion, we were able to demonstrate FGF21 dysfunction and thus a pathogenic effect for L142V. Further functional testing regarding the effect on other signaling pathways of the FGF receptor might be necessary to draw definite conclusions regarding the definite disease-causality of these variants. These data support an essential role for genetic variation within the *FGF21* gene in the pathogenesis of obesity.

REFERENCES

1. Hinney, A., A.L. Volckmar, and N. Knoll, *Melanocortin-4 receptor in energy homeostasis and obesity pathogenesis.* Prog Mol Biol Transl Sci, 2013. **114**: p. 147-91.
2. Montague, C.T., et al., *Congenital leptin deficiency is associated with severe early-onset obesity in humans.* Nature, 1997. **387**(6636): p. 903-8.
3. Strobel, A., et al., *A leptin missense mutation associated with hypogonadism and morbid obesity.* Nat Genet, 1998. **18**(3): p. 213-5.
4. Gibson, W.T., et al., *Congenital leptin deficiency due to homozygosity for the Delta133G mutation: report of another case and evaluation of response to four years of leptin therapy.* J Clin Endocrinol Metab, 2004. **89**(10): p. 4821-6.
5. Mazen, I., et al., *A novel homozygous missense mutation of the leptin gene (N103K) in an obese Egyptian patient.* Mol Genet Metab, 2009. **97**(4): p. 305-8.
6. Fischer-Posovszky, P., et al., *A new missense mutation in the leptin gene causes mild obesity and hypogonadism without affecting T cell responsiveness.* J Clin Endocrinol Metab, 2010. **95**(6): p. 2836-40.
7. Fatima, W., et al., *Leptin deficiency and leptin gene mutations in obese children from Pakistan.* Int J Pediatr Obes, 2011. **6**(5-6): p. 419-27.
8. Thakur, S., et al., *A novel mutation of the leptin gene in an Indian patient.* Clin Genet, 2014. **86**(4): p. 391-3.
9. Funcke, *Monogenic forms of childhood obesity due to mutations in the leptin gene.* Molecular and Cellular Pediatrics, 2014. **1**(3).
10. Farooqi, I.S., et al., *Effects of recombinant leptin therapy in a child with congenital leptin deficiency.* N Engl J Med, 1999. **341**(12): p. 879-84.
11. Farooqi, I.S., et al., *Beneficial effects of leptin on obesity, T cell hyporesponsiveness, and neuroendocrine/metabolic dysfunction of human congenital leptin deficiency.* J Clin Invest, 2002. **110**(8): p. 1093-103.
12. Farooqi, I.S., et al., *Clinical and molecular genetic spectrum of congenital deficiency of the leptin receptor.* N Engl J Med, 2007. **356**(3): p. 237-47.
13. Locke, A.E., et al., *Genetic studies of body mass index yield new insights for obesity biology.* Nature, 2015. **518**(7538): p. 197-206.
14. Speliotes, E.K., et al., *Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index.* Nature Genetics, 2010. **42**(11): p. 937-U53.

15. Willer, C.J., et al., *Six new loci associated with body mass index highlight a neuronal influence on body weight regulation.* Nat Genet, 2009. **41**(1): p. 25-34.
16. Ornitz, D.M. and N. Itoh, *The Fibroblast Growth Factor signaling pathway.* Wiley Interdiscip Rev Dev Biol, 2015. **4**(3): p. 215-66.
17. Berglund, E.D., et al., *Fibroblast growth factor 21 controls glycemia via regulation of hepatic glucose flux and insulin sensitivity.* Endocrinology, 2009. **150**(9): p. 4084-93.
18. Coskun, T., et al., *Fibroblast growth factor 21 corrects obesity in mice.* Endocrinology, 2008. **149**(12): p. 6018-27.
19. Adams, A.C., et al., *The breadth of FGF21's metabolic actions are governed by FGFR1 in adipose tissue.* Mol Metab, 2012. **2**(1): p. 31-7.
20. Chu, A.Y., et al., *Novel locus including FGF21 is associated with dietary macronutrient intake.* Hum Mol Genet, 2013. **22**(9): p. 1895-902.
21. Soberg, S., et al., *FGF21 Is a Sugar-Induced Hormone Associated with Sweet Intake and Preference in Humans.* Cell Metab, 2017. **25**(5): p. 1045-1053 e6.
22. Roelants, M., R. Hauspie, and K. Hoppenbrouwers, *References for growth and pubertal development from birth to 21 years in Flanders, Belgium.* Ann Hum Biol, 2009. **36**(6): p. 680-94.
23. Hauspie, R. and M. Roelants. *Flemish Growth Charts 2004*. 2004 22 July 2013]; Available from: http://www.vub.ac.be/groeicurven.
24. Miller, S.A., D.D. Dykes, and H.F. Polesky, *A simple salting out procedure for extracting DNA from human nucleated cells.* Nucleic Acids Res, 1988. **16**(3): p. 1215.
25. Kircher, M., et al., *A general framework for estimating the relative pathogenicity of human genetic variants.* Nat Genet, 2014. **46**(3): p. 310-5.
26. den Dunnen, J.T. and S.E. Antonarakis, *Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion.* Hum Mutat, 2000. **15**(1): p. 7-12.

V. GENERAL DISCUSSION AND CONCLUSION

ABSTRACT

The aim of this PhD research was to clarify the role of 6 selected candidate genes in the pathogenesis of obesity. Maintaining a balance between energy intake and expenditure is vital in the process of energy homeostasis. Once the balance shifts to either end of the spectrum, health problems arise. With its prevalence reaching epidemic proportions, obesity has become the most important health disorder worldwide. The heritability of obesity is generally estimated between 40-70%, indicating that genetic factors play an important role in the aetiology of obesity [1, 2]. So far, only a limited number of monogenic causes of obesity is known, many of which are associated with the leptin-melanocortin pathway. Despite the importance of this pathway, mutations in the genes involved in this pathway account for not more than 5% of all childhood severe obesity cases. For complex obesity, all current risk alleles together capture less than 3% of BMI variation, leaving a substantial proportion of the heritability unexplained. The remaining heritability must be found in other sources of genetic variation. Copy number variation accounts for a major proportion of human genetic variation and it is predicted to play an important role in the genetic susceptibility of common complex diseases. Since 2009, genome-wide CNV analyses have associated numerous CNV regions with early-onset extreme obesity [3-5]. Further investigation of the genes in these regions might offer novel insights into the genetic architecture of obesity. However, to date limited follow-up has been performed and this prompted us to further investigate the most interesting CNV regions in a population of severely overweight or obese children and adolescents. So far, 552586 CNVs have been reported in the Database of Genomic variants with an estimated 75,6% of exons and 91,2% of transcripts that are overlapped by at least one CNV . Therefore, CNVs are a significant source of human genetic variation accounting for population diversity and human diseases [6-8]. To explore the contribution of CNVs to obesity, genome-wide CNV surveys have been performed in patient populations of early-onset obese patients. These surveys reported a number of CNVs that confer high risk to moderate and extreme obesity.

DISCUSSION OF THE RESULTS OF MUTATION AND CNV SCREENING IN THIS THESIS

SH2B1

In this PhD project, we first investigated the prevalence of pathogenic mutations and CNV in the *SH2B1* gene in the Belgian population. SH2B1, a positive regulator of the leptin-melanocortin pathway, has been clearly implicated in the aetiology of obesity by animal, association and CNV studies. However, limited data were available on the presence of single-base and structural variation in the *SH2B1* gene in obese and lean individuals. Therefore, we started with a mutation analysis to screen the *SH2B1* coding region in 581 obese children and adolescents and 433 healthy lean individuals. This resulted in the identification of fifteen rare non-synonymous heterozygous variants (table 1). Several of these were found in both lean and obese subjects, suggesting these are neutral polymorphisms. However, six private non-synonymous variations were only present in obese children and six other variants could only be identified in lean individuals. Although this equal variation frequency does not immediately support disease causality, it cannot be excluded that some variations are weight-increasing or -decreasing based by a distinctive localization in specific functional domains of the protein. It was remarkable that the identified mutations mostly occur outside the known functional domains of the protein, with the exception of R43S and R67C that are located in the dimerization domain and S616P in the SH2 domain. To understand the specific impact on the function of SH2B1, functional experiments are necessary. In addition to the mutation screen, we investigated the presence of the distal *SH2B1*-encompassing 16p11.2 deletion in a population of 421 obese children and adolescents with no developmental or behavioural phenotype. In contrast to previously published studies that reported a prevalence of 0,41% in a population of patients with early-onset obesity, we were not able to detect any CNV. As we specifically excluded patients with developmental delay, we suggest that this is a main factor of difference. The prevalence of this deletion in patients with no other traits beside obesity might be very low. Therefore, we suggest that in addition to obesity the distal deletion might strongly predispose for other traits beside obesity. Additional genes in the 220kb region (*ATXN2L, TUFM, ATP2A1, RABEP2, CD19, SPNS1, NFATC2IP* and *LAT)* might be responsible for the other phenotypic characteristics. In the future, it will be important to investigate the prevalence and phenotype of the distal 16p11.2 deletion in an extended study population of patients that present a variability of abnormal phenotypes including developmental delay, behavioural problems, mild facial dysmorphology and intellectual disability. Detailed clinical investigation of these patients by a clinical geneticist

might be indispensable to define a specific and recognizable phenotype. Altogether, these data suggest that it is unlikely that genetic variation in the *SH2B1* gene contributes majorly to the pathogenesis of obesity.

NPY and its receptors NPY2R and NPY4R

In the second part of this thesis, we investigated the role of NPY and its receptors in energy regulation. NPY, one of the most potent orexigenic peptides, plays an essential role in the leptin-melanocortin pathway, the most important regulator of energy homeostasis known so far. Genetic evidence coming from association and animal studies confirmed the importance of this neuromodulator in energy regulation. However, little was known about the possible role of point mutations within the *NPY* gene. This offered us a chance to clarify its role by performing an extensive mutation analysis. Additionally, we decided to not only screen *NPY*, but also two of its receptors, *NPY2R* and *NPY4R*. These receptors work synergistically as major regulators of energy homeostasis. Because of the CNV in these genes that was additionally linked to obesity, the genes were particularly interesting in this thesis. Therefore, we screened the genes for both the presence of single nucleotide and structural variation.

For *NPY,* one common missense variation rs16139 was identified in our patient population. This SNP, located in the signal peptide part of the preproNPY, has been associated with a variety of metabolic, cardiovascular and behavioural parameters. However, as the prevalence of the SNP in our patient population was similar to the one that was reported for the European population, we assume that it does not contribute majorly to the obese phenotype of its carriers. We were not able to identify rare non-synonymous variants in the *NPY* gene. The absence of these variants in our obese population underlines the strong evolutionary conservation of *NPY*. Due to its important role in a variety of physiological processes, *NPY* must be under strong negative selection against mutations.

The mutation screening of *NPY2R* resulted in the identification of one missense variant F87I that was exclusively found in the obese population (table 1). However, previous studies reported that this variant didn't segregate with obesity in a family. The 8-year-old obese carrier inherited the variant from his lean mother. Furthermore, we were not able to find additional carriers of the previously reported *NPY2R*-encompassing deletion among our obese individuals. Therefore, we suggest that this deletion was a very rare or a private mutation in the specific pedigree. Altogether, genetic variation in *NPY* and *NPY2R* is unlikely to majorly contribute to the obese

phenotype in the general population. This is in contrast to the results of our mutation and CNV screening in the *NPY4R* gene. Mutation screening resulted in the identification of fifteen rare non-synonymous heterozygous variants. Six of these variants (V54M, A81T, M116T, K190E, V247M and V271M) were exclusively found among obese patients and four other variants (L79F, C168Y, A169V and R246H) could only be identified among lean adults. *In silico* analysis was used to select the most interesting variants for functional research: L79F, C168Y, M116T and V271M. Luciferase reporter assays were carried out to determine the influence of the selected variants on the cAMP signalling activity of the receptor. For the two variants (L79F and C168Y) that were exclusively found in the control population, no significant effect on receptor activity could be determined. This is in contrast to the two variants (M116T and V271M) that were identified among obese patients and for which significant receptor dysfunction could be reported. These findings were in line with the strong results of our CNV study. In this study, we demonstrated a significantly higher frequency of *NPY4R*-encompassing deletions in the patient population. Taken together, the presence of pathogenic structural and single nucleotide variation within the *NPY4R* gene explains, at least partially, the obese phenotype of 3,7% and 0,6% of our patient population respectively. This makes this gene an important risk factor for obesity with prevalences close to those reported for the most common form of monogenic obesity, pathogenic *MC4R* mutations. Altogether these data support an essential role for genetic variation within the *NPY4R* gene in the pathogenesis of obesity.

Two key regulators in brown and beige adipogenesis: UCP1 and FGF21

In the last part of this thesis, we investigated the importance of genetic variation in *UCP1* and *FGF21*. Both genes are two key regulators in brown and beige adipogenesis. The *UCP1* gene was initially selected based on the link between *UCP1*-encompassing CNV and the susceptibility of obesity. A 2,1 Mb deletion on chromosome 4q31, encompassing *UCP1* and *Interleukin-15 (IL-15),* was found in a severely obese patient with a BMI of 46,2 kg/m^2. Additional studies indicated an essential role for the *UCP1* gene in the pathogenesis of obesity. Association studies confirmed the link between three SNPs in the promotor region of *UCP1* and human obesity. Furthermore, in animal studies *UCP1* ablation on itself also induced obesity in C57BI6 mice that were kept under thermoneutral conditions, confirming the importance of *UCP1* in energy regulation. However, no information was available on the prevalence of single-nucleotide and structural variation in the gene. Therefore, we designed an extensive mutation and CNV screen in our patient and control population. These analyses resulted in the identification of 3 non-

synonymous rare variants (G57S, T157I and T227I) (table 1). G57S and T227I could only be identified in obese individuals, while T157I could only be found in a lean adult. The T227I was particularly interesting as it was identified in four obese individuals. *In silico* analysis predicted a potential impact of the found variants on the function of UCP1. Therefore, UCP1 function was assessed in a quantitative manner using flux analysis for intact HEK293 cells. Furthermore, the exact function in isolated mitochondria will be determined. Altogether these assays will provide insights into the mutant's potential to uncouple oxidative phosphorylation in isolated mitochondria. For these experiments we collaborate with the Mitochondrial Biology Unit of Prof. Jastroch in the German Research Center for Environmental Health Institute for Diabetes and Obesity that is specialized in the functional characterization of UCP1 in mammalian HEK293 cells. These experiments are still ongoing and its results will be included in a later version of my thesis.

Due to the therapeutic potential that UCP1 holds in energy expenditure, we decided to look further into the brown and beige adipogenesis. Apart from UCP1, FGF21 plays a key role into the browning of white adipose tissue in adaptive thermogenesis. As no information was available about variation in this gene, we performed mutation screening in our patient population of severely overweight or obese children and adolescents. This led to the identification of two rare non-synonymous variants: R47Q and L142V. *In silico* analysis predicted a possible functional impact of the variants. However, further functional characterization of the mutants was necessary to provide insights into their potential to activate the intracellular pathways of the FGF21 receptor: cAMP, AP-1, NFAT and NF-KB signalling pathway. Luciferase reporter assays were performed to identify the effect of the variant on FGF21. For L142V, we were able to demonstrate FGF21 dysfunction as the mutant showed significantly decreased cAMP and NFAT signalling properties. For R47Q only a trend towards decreased activation of these two pathways was shown. Taken together, we demonstrated an essential role for FGF21 in energy regulation. This confirms the therapeutic potential of FGF21 in energy regulation.

Conclusion

Table 1 summarizes the main findings from our mutation and CNV studies for all investigated candidate genes. We believe that these results contribute to the current knowledge on the aetiology of human obesity. However, as obesity is a common and complex disorder, interpretation of the results is not straightforward and segregation analysis is often complicated. So before the results of genetic screenings can add major value to genetic counselling as well as

to the development of pharmacotherapy for obesity, functional studies are necessary to clearly define the pathogenic effect of the genetic variation.

Table 1: Results from mutation, copy number variation and functional studies

Candidate gene	Mutation analysis	Copy number variation study	Functional Study
SH2B1	• 24 genetic variants found • 15 non-synonymous variants • 6 rare non-synonymous variants (P19L, R43S, T175N, S188L, αA667V and αA672T) exclusively found in obese individuals • 6 rare non-synonymous variants (P16S, R67C, G205R, R433H, S616P and βT656I/γP674S) exclusively found in lean individuals • High mutation burden in both groups: 2.8% in obese children/adolescents and 3.9% in lean adults	• No distal 16p11.2 deletion was found in our population of severely overweight or obese children and adolescents without developmental or behavioral problems	• Functional study of the variants is necessary to fully understand the impact of these variants on SH2B1
NPY	• 4 genetic variants found • 1 common non-synonymous variant (L7P) • Low prevalence of genetic variation in *NPY* highlights its important role in a variety of physioloical processes	• N.A.	• N.A.
NPY2R	• 4 genetic variants • 1 rare non-synonymous variant (F87I) • *In silico* analysis predicted a possible impact of the mutant F87I on NPY2R	• No CNVs were found in *NPY2R*-region	• Functional study of the variants is necessary to fully understand the impact of these variants on NPY2R
NPY4R	• 24 genetic variants • 15 non-synonymous variants • 6 rare non-synonymous variants (V54M, A81T, M116T, K190E, V247M and V271M) exclusively found in obese individuals • 4 rare non-synonymous variants (L79F, C168Y, A196V and R246H) exclusively found in lean individuals • *In silico* analysis predicted a possible impact of mutants L79F, M116T, C168Y	• Significantly higher frequency of *NPY4R*-encompassing CNV loss in obese individuals (3.68%) when compared to lean controls (0%) • Significantly higher frequency of *NPY4R*-encompassing CNV gain in lean controls (2.68%) when compared to obese individuals (0.31%) • *NPY4R*-encompassing CNV loss at least partially explains the obese phenotype of	• Luciferase reporter assay demonstrated a pathogenic effect for two mutants (M116T and V271M) • Pathogenic *NPY4R* mutations explain at least partially the obese phenotype in 0.6% of our patient population of severely overweight or obese children and adolescents

	and V271M on NPY4R	3.7% of our patient population of severely overweight or obese children and adolescents	
UCP1	• 11 genetic variants • 6 non-synonymous variants • 2 rare non-synonymous variants (G57S and T227I) exclusively found in obese individuals • 1 rare non-synonymous variant (T157I) exclusively found in lean individuals	• No CNVs were found in *UCP1*-region	• "ongoing"
FGF21	• 2 genetic variants • 2 rare non-synonymous (R47Q and L142V) • *In silico* analysis predicted a possible impact of these mutants on FGF21	• N.A.	• Luciferase reporter assay demonstrated a pathogenic effect for the L142V mutant

The main findings from the mutation, copy number variation and functional studies are summarized, for all investigated candidate genes.

MISSING HERITABILITY

Obesity results from a complex interaction between genetic and environmental factors. The search for genetic variants contributing to the 40-70% heritability of obesity, has been a challenging task. So far, only a limited number of causes of monogenic obesity is known, many of which are associated with the leptin-melanocortin pathway. *MC4R* gene mutations are the most prevalent cause of monogenic obesity, *Leptin* deficiency is the only treatable one. Despite the importance of this brain pathway in the pathogenesis of obesity, mutations in the genes of this pathway are not responsible for more than 5% of all childhood obesity cases. In addition, 97 BMI-associated loci were identified that account for 2,7% of BMI variation, leaving a substantial part of the heritability unexplained.

The main focus of this thesis was to investigate copy number variation as a potential factor to explain part of the missing heritability of obesity. Therefore, we made a selection of interesting candidate genes that are localized in linked CNV regions. Genetic variation in these genes might offer novel insights into the genetic architecture of obesity. In contrast with this candidate gene approach, there are alternative strategies to unravel the genetic contribution to complex diseases, some of which are explained below.

Genome-wide association studies (GWAS)

For obesity, extensive GWAS have been performed to identify genetic loci that underlie the genetic risk of obesity. So far, 97 BMI-associated loci were linked with BMI, waist-to-hip ratio, body fat percentage or extreme obesity. Together, these loci explain only 2,7% of BMI variation while genome-wide estimates suggest that common variation accounts for >20%. The *Fat mass and obesity associated* gene (*FTO*) region harbours the strongest genetic association with obesity [9]. This locus explains 0,34% of BMI variation in the white population and for each additional *FTO* risk allele, BMI increases by 0,39 kg/m^2 [10, 11]. Although these loci represent an important step in understanding the physiological mechanisms leading to obesity, altogether they are not sufficient to have any clinical value.

Next generation sequencing

Next generation sequencing (NGS) is a high-throughput technique whereby millions of sequence reads can be processed in parallel. With longer read lengths and decreased sequencing costs per mega base, the number and diversity of sequenced genomes increased.

This led to an enhanced understanding of how genome sequence variants underlie phenotype and disease. NGS projects such as the 1000 Genomes project and ExAC database harmonize exome sequencing data to make them available for the wider scientific community. Based on these data, whole genomes of healthy individuals and subjects with a particular disease can be compared to identify the genetic variants that are causative for the phenotype of interest. This strategy is particularly useful when the disease phenotype is well-defined and additional family members are available to distinguish causal variants from benign polymorphisms. However for complex, multifactorial and highly prevalent diseases such as obesity, this strategy is less straightforward. More specific obesity initiatives such as the UK10K project with an obesity sample set of 2000 exomes, provide genome-wide sequencing data of deeply phenotyped individuals with extreme and early-onset obesity. This specific focus increases the power of the study to identify causal variants at an unprecedented level and holds potential for further clarification of obesity genes.

Rare variants

The analysis of rare variants is more challenging than that of common variants. GWAS have typically focussed on very common SNPs (MAF>5%) in the human genome. If ungenotyped causal variants have a lower allele frequency than the SNPs in the GWAS, they will be in low LD with the common SNPs, and the resulting p-value will be proportionally attenuated. So the proportion of the heritability that can be explained by common SNPs depends on how well the causal variants are tagged by these SNPs. There are two genetic obesity models: common disease-common variant model and the common disease-rare variant model. The first model supports the proposition that common variants with small effect sizes are a major source of genetic variance for complex diseases. Numerous genetic factors would provide small independent and additive contributions to the complex phenotype. In contrast, the rare allele model states that multiple highly penetrant rare variants with important effect sizes contribute majorly to the susceptibility of obesity. This category of less common and even rare variants is not sufficiently frequent to be captured in genome wide association studies and their effect size was too small to be detected in screening for monogenic disease (figure 1). Taken together, it can be argued that common variants do not capture all the genetic variance for obesity and so the remaining heritability might be found in other sources of genetic variance.

Targeted follow-up of a deeply phenotyped cohort of extremely lean individuals on one hand and severely obese subjects on the other hand allows to study genetic variation at both ends of

the weight spectrum. These extreme phenotypes are more likely to be caused by genetic factors that contribute to energy homeostasis. Combining these well-defined and extreme cohorts might be a productive approach to unravel the genetics of complex obesity.

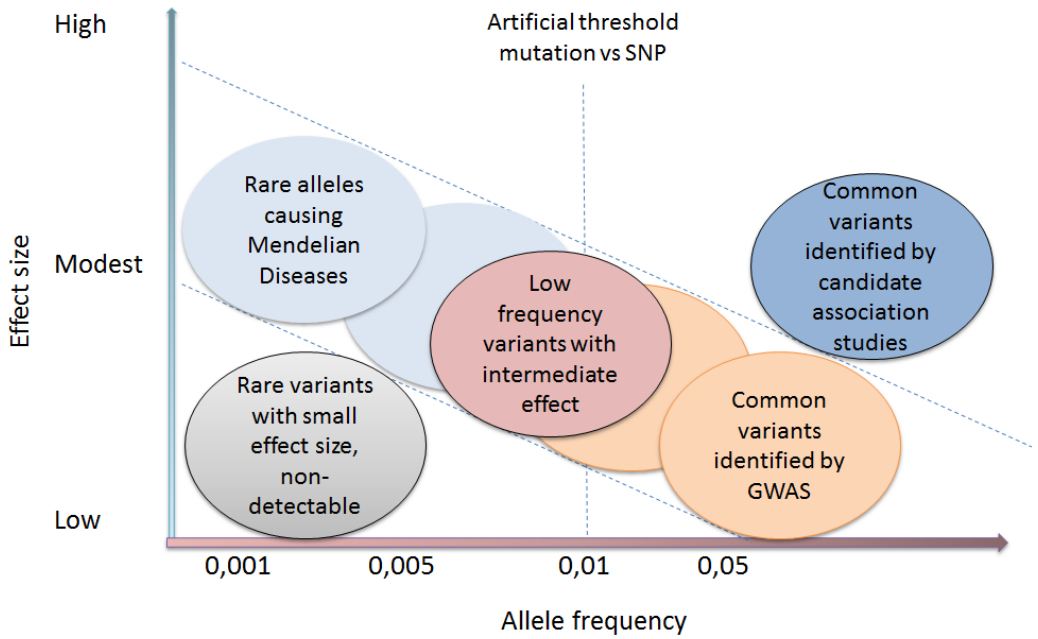

Figure 1: Schematic representation of genetic variants spectrum (McCarthy/Manolia model)(Adapted from [12]). Allele frequency and effect size are generally inversely related with common variants having small effects and rare variants large effects. Common variants with large effects are generally rare and subject to strong purifying selection, rare variants with small/moderate effects are difficult to detect.

Epigenetic studies

Epigenetics is the study of heritable genome modifications that affect gene expression without actually changing the DNA sequence. The four major determinants of epigenetic regulation are tissue specific and include DNA methylation patterns, histone modification, chromatin confirmation and noncoding RNAs. Failures in these epigenetic marks or imprinting are not only known to cause extreme forms of obesity such as Prader-Willi syndrome, but are also associated with the general susceptibility of obesity. Pre- and neonatal environmental exposures modify the epigenome and mediate the expression of genes that are associated with obesity or adiposity. Hence, this link between genetics, disease and environment might play a decisive role in the aetiology of obesity. Advances of high throughput profiling, such as the Infinium MethylationEPIC Beadchip (Illumina), CpG pyrosequencing, chromatin immunoprecipitation sequencing (ChIP-Seq) and reduced representation bisulfite sequencing (RRBS) have enabled

researchers to identify thousands of differentially expressed genes under epigenetic regulation. Projects such as the Human Epigenome Consortium aim to identify, catalogue and interpret genome-wide DNA methylation patterns of all human genes in all major tissues. Completion of this project will increase our understanding of the genetics and epigenetics that contribute to obesity susceptibility.

Therapeutic potential of genetic findings

Since the discovery of leptin in 1994, and the observation that its replacement reverses morbid obesity in leptin-deficient mice and humans, major research efforts have been undertaken with the hope that leptin replacement therapy would be a powerful tool to treat common obesity. However, due to leptin resistance, this was only the case for patients with congenital leptin deficiency (CLD). For these individuals, daily injection of the pharmaceutical form of recombinant methionyl human leptin (Metreleptin), normalized most of their phenotypes, including hyperphagia, metabolic and hormonal disturbances.

Apart from leptin, MC4R is also a well-recognized and important mediator of body weight homeostasis. However, unlike leptin, which is not effective in common obesity due to leptin resistance, the effect of MC4R agonism on energy regulation is more pronounced in obese than in lean rats. This makes MC4R an attractive target for a therapeutic agent as MC4R mediates the effect of leptin on energy homeostasis. Since then, a variety of pharmacologically modified MC4R agonists have been generated. Although these agonists were often able to reduce food intake and body weight, many have serious side effects, including increased heart rate and blood pressure. Currently no effective MC4R agonists are in use for clinical treatment of humans.

Given the importance of NPY and its receptors in the pathogenesis of obesity, the development of pharmacological agents targeting the NPY receptors is attractive for research and clinical settings. Especially the *NPY4R* gene, holds an important potential as pathogenic mutations occur at a prevalence of 4% among early-onset obese patients. A number of promising chemical compounds, agonists against the NPY receptors, show potent anti-obesity effects. However, most of these compounds show a lack of selectivity, low oral bioavailability, poor brain penetrability and interaction with other receptors. Possible toxicity or long-term effects still hamper the compounds into a clinical setting [13]. It appeared that the PP analogues for NPY receptors don't majorly improve the obese phenotype due to the short half-life of PP. The lack of sensitivity might be solved by targeting peripheral receptors with compounds that are

specifically designed not to cross the blood-brain barrier. However, it has also become clear that targeting more than 1 NPY receptor, may have synergistic effects in treating obesity.

Furthermore, possibilities to develop a genetic test for the prediction of obesity susceptibility and conferring the option for personalized medicine may arise in the future. However, for now all current risk alleles together are not sufficient to have any clinical value and additional genes or different types of genetic variants remain to be discovered.

In conclusion, we were able to investigate the involvement of single-nucleotide and structural variation in 6 candidate genes for obesity (table 1). We believe that the results of these studies contribute to knowledge on the pathogenesis of obesity and can add clinical value to the development of pharmaceutical compounds against obesity in the future.

REFERENCES

1. Stunkard, A.J., T.T. Foch, and Z. Hrubec, *A Twin Study of Human Obesity.* Jama-Journal of the American Medical Association, 1986. **256**(1): p. 51-54.
2. Bouchard, C., et al., *The Response to Long-Term Overfeeding in Identical-Twins.* New England Journal of Medicine, 1990. **322**(21): p. 1477-1482.
3. Sha, B.Y., et al., *Genome-wide association study suggested copy number variation may be associated with body mass index in the Chinese population.* J Hum Genet, 2009. **54**(4): p. 199-202.
4. Wang, K., et al., *Large copy-number variations are enriched in cases with moderate to extreme obesity.* Diabetes, 2010. **59**(10): p. 2690-4.
5. Jarick, I., et al., *Novel common copy number variation for early onset extreme obesity on chromosome 11q11 identified by a genome-wide analysis.* Human Molecular Genetics, 2011. **20**(4): p. 840-852.
6. Olson, H., et al., *Copy number variation plays an important role in clinical epilepsy.* Ann Neurol, 2014. **75**(6): p. 943-58.
7. Zhang, F., et al., *Copy number variation in human health, disease, and evolution.* Annu Rev Genomics Hum Genet, 2009. **10**: p. 451-81.
8. Pinto, D., et al., *Functional impact of global rare copy number variation in autism spectrum disorders.* Nature, 2010. **466**(7304): p. 368-72.
9. Locke, A.E., et al., *Genetic studies of body mass index yield new insights for obesity biology.* Nature, 2015. **518**(7538): p. 197-206.
10. Yang, J., et al., *Genome partitioning of genetic variation for complex traits using common SNPs.* Nat Genet, 2011. **43**(6): p. 519-25.
11. Frayling, T.M., et al., *A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity.* Science, 2007. **316**(5826): p. 889-94.
12. Manolio, T.A., et al., *Finding the missing heritability of complex diseases.* Nature, 2009. **461**(7265): p. 747-53.
13. Yulyaningsih, E., et al., *NPY receptors as potential targets for anti-obesity drug development.* Br J Pharmacol, 2011. **163**(6): p. 1170-202.

VI. ABBREVIATIONS

LIST OF ABBREVIATIONS

AgRP	Agouti-related peptide
AMY1	Amylase 1
ATP	Adenosine triphosphate
BAT	Brown adipose tissue
BMI	Body Mass Index
bp	Base pairs
cAMP	Cyclic adenosine monophosphate
CART	Cocaine and amphetamine regulated transcript
cDNA	Complementary DNA
CNV	Copy number variation
CRE	cAMP response element
DQ	Dosage Quotient
DMEM	Dulbecco's modified Eagle medium
DNA	Deoxyribonucleic acid
FGF21	Fibroblast growth factor 21
FTO	Fat mass and obesity associated
GWAS	Genome-wide association studies
GPRC5B	G protein coupled receptor 5B
GPRIN2	G protein regulated inducer of neurite outgrowth 2
HEK	Human embryonic kidney
HRM	High-resolution melting curve analysis
HWE	Hardy-Weinberg equilibrium
IL15	Interleukin-15
JAK2	Janus Kinase-2
LD	Linkage disequilibrium
LDL	Low density lipoprotein
LEP	Leptin
MAF	Minor allele frequency
MAQ	Multiplex Amplicon Quantification
mCNV	Multi-allelic copy number variation
MLPA	Multiplex Ligation-dependent Probe Amplification

MC4R	Melanocortin-4 receptor
NAHR	Nonallelic homologous recombination
NEGR1	Neuronal growth regulator 1
NEHJ	Nonhomologous end joining
NFκβ	Nuclear factor kappa B
NGS	Nex generation sequencing
NPY	Neuropeptide Y
NPY2R	Neuropeptide Y2 receptor
NPY4R	Neuropeptide Y4 receptor
OR	Olfactory receptor
OR4C6	Olfactory receptor family 4 subfamily C member 6
OR4P4	Olfactory receptor family 4 subfamily P member 4
OR4S2	Olfactory receptor family 4 subfamily S member 2
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
PP	Pancreatic Polypeptide
PPY	Peptide YY
PPYR1	Pancreatic Polypeptide Y1 receptor
SD	Standard deviation
SH2B1	Src homology 2 binding adaptor protein 1
SNP	Single nucleotide polymorphism
UCP1	Uncoupling Protein 1
WAT	White adipose tissue
WHR	Waist-to-hip ratio
WT	Wild type

VII. CURRICULUM VITAE

CURRICULUM VITAE – EVI AERTS

PERSONAL DETAILS

Name/First name	Aerts Evi
Address	Stationsplein 2, 2590 Berlaar (Antwerp), Belgium
Date of birth	16.02.1988
Email	Aerts.Evi@gmail.com

EDUCATION AND ACADEMIC CAREER

Education

2010 – 2012 Master of Biochemistry and Biotechnology: Molecular and Cellular Gene Biotechnology, University of Antwerp, Belgium

Master dissertation entitled: "*Functional and genetic analysis of 2 candidate genes for monogenic obesity*"
› Graduated with Great Distinction

Promotor: Prof. Dr. Wim Van Hul

2006 – 2010 Bachelor in Biochemistry and Biotechnology

2000 – 2006 Latin-Mathematics-Science, Heilig Hart van Maria Instituut, Berlaar

Academic career

2012 – 2017 PhD in Sciences: Biochemistry and Biotechnology
Center of Medical Genetics, University of Antwerp, Belgium

PhD thesis entitled: *Study of copy number and DNA sequence variations in candidate genes for obesity*

Promotors: Prof. Dr. Wim Van Hul - Prof. Dr. Luc Van Gaal
Co-promotor: Dr. Sigri Beckers

ADDITIONAL SCIENTIFIC TRAINING

October 2012 **Laboratory Animal Science, FELASA, type C**
University of Antwerp, Belgium

GRANTS AND AWARDS

PhD fellowship

by the Agency for Innovation by Science and Technology (IWT)

December 2012, Antwerp, Belgium

PUBLICATION LIST

Evi Aerts, Sigri Beckers, Doreen Zegers, Jasmijn K Van Camp, Kim Van Hoorenbeeck, Guy Massa, An Verrijken, Ilse Mertens, Stijn L. Verhulst, Raoul P. Rooman, Luc F. Van Gaal and Wim Van Hul. Genetic and structural variation in the SH2B1 gene in the Belgian population. *Mol Genet Metab (2015) 115 (4): 193-198.*

Evi Aerts, Ellen Geets, Laure Sorber, Sigri Beckers, An Verrijken, Guy Massa, Kim Van Hoorenbeeck, Stijn L Verhulst, Luc F Van Gaal and Wim Van Hul. 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. *Annals of Human Genetics (2017) Accepted*

Evi Aerts, Sigri Beckers, Doreen Zegers, Kim Van Hoorenbeeck, Guy Massa, An Verrijken, Stijn L Verhulst, Luc F Van Gaal and Wim Van Hul. Combination of CNV analysis and mutation screening indicates an important role for the NPY4R gene in human obesity. *Obesity (Silver Spring) (2016) 24 (4): 970-976.*

Evi Aerts, Ellen Geets, Sigri Beckers, An Verrijken, Guy Massa, Kim Van Hoorenbeeck, Stijn L Verhulst, Luc F Van Gaal and Wim Van Hul. Brown adipose tissue induction and obesity; genetic aspects of UCP1. *Work in progress*

Evi Aerts, Ellen Geets, Sigri Beckers, An Verrijken, Guy Massa, Kim Van Hoorenbeeck, Stijn L Verhulst, Luc F Van Gaal and Wim Van Hul. Identification and functional characterization of FGF21 mutations in obese individuals. *Submitted*

Evi Aerts and Wim Van Hul. The role of copy number variation in the pathogenesis of obesity. *Submitted*

Doreen Zegers, An Verrijken, Sven Francque, Fenna de Freitas, Sigri Beckers, **Evi Aerts**, Martin Ruppert, Guy Hubens, Peter Michielsen, Wim Van Hul and Luc Van Gaal. Screening for rare variants in the PNPLA3 gene in obese liver biopsy patients. *Clinics and Research in Hepatology and Gastroenterology 2016; 40(6): 715-721*.

Doreen Zegers, An Verrijken, Sigri Beckers, Sven Francque, Jasmijn Van Camp, **Evi Aerts**, Martin Ruppert, Guy Hubens, Peter Michielsen, Wim Van Hul and Luc Van Gaal.. *Clinics and Research in Hepatology and Gastroenterology 2016; 40(3): 333-339*.

ORAL PRESENTATIONS

Belgian Association for the study of Obesity (BASO) – Annual Meeting 2016, Brussels, Belgium.

Presentation entitled: "*Combination of CNV analysis and mutation screening indicates an important role for the NPY4R gene in human obesity*"

Belgian Association for the study of Obesity (BASO) – Annual Meeting 2017, Brussels, Belgium.

Presentation entitled: "*Evaluation of a role for UCP1 in the pathogenesis of obesity by mutation and copy number variation analysis in obese children and adolescents*"

Research day of the faculty of Pharmaceutical, Biomedical and Veterinary Sciences, 30 October, 2015, University of Antwerp, Antwerp, Belgium.

Presentation entitled: "The role of NPY and its receptors in the pathogenesis of obesity"

POSTER PRESENTATIONS

American Society of Human Genetics (ASHG) – Annual Meeting 2017, Orlando, Florida, United States of America.

Poster presentation: "*New insights into the role of genetic variation within FGF21 in the pathogenesis of obesity*"

Belgian Society of Human Genetics (BeSHG) 17th Annual Meeting 2017, Louvain-la-Neuve, Belgium.

Poster presentation: "*Screening for genetic and structural variation in the UCP1 gene in obese children and adolescents.*"

Belgian Society of Human Genetics 16th annual meeting 2016, Leuven, Belgium.

Poster presentation: "*Combination of CNV analysis and mutation screening indicates an important role for the NPY4R gene in human obesity.*"

Belgian Society of Human Genetics 15th Annual meeting 2015, Charleroi, Belgium.

Poster presentation: "*Screening for genetic and structural variation in the NPY2R gene in obese children and adolescents.*"

Belgian Medical Genomics Innitiative (BeMGI) Annual Meeting 2015, Gent, Belgium.

Poster presentation: "*Screening for genetic and structural variation in the NPY2R gene in obese children and adolescents.*"

European Congress of Endocrinology (ECE) Annual Meeting 2015 Dublin, Ireland

Poster presentation: "*Screening for genetic and structural variation in the NPY2R gene in obese children and adolescents*"

European Society of Human Genetics (ESHG) Annual Meeting 2014 ", Milan, Italy

Poster presentation: "*Identification of new missense variations in the SH2B1 gene in obese and lean individuals*."

Belgian Medical Genomics Innitiative (BeMGI) Annual Meeting 2014 ", Gent, Belgium

Poster presentation: "*Identification of new missense variations in the SH2B1 gene in obese and lean individuals*."

Programming Obesity 2013 "Central and peripheral contributors", Cambridge, England

Poster presentation: "*Identification of new missense variations in the SH2B1 gene in obese and lean individuals*."

EDUCATIONAL ACTIVITIES

Assistance Project Proposal: "*Genetic Research into the role of the NPY2R gene in the development of obesity*" by Laure Sorber, 2nd Master in Biomedical Sciences, University of Antwerp, Belgium. Academic Year 2013-2014.

Assistance Master Thesis: "*Genetic Research into the role of the NPY2R gene and the NPY4R gene in the development of obesity*" by Laure Sorber, 2nd Master in Biomedical Sciences, University of Antwerp, Belgium. Academic Year 2013-2014.

Assistance Honours College: "*Neuropeptide Y en zijn receptoren: hun rol in obesitas andere pathologieën*" by Marlon van Loo, 3rd Bachelor in Biomedical Sciences, University of Antwerp, Belgium. Academic Year 2014-2015.

Assistance Honours College: *"Variaties in het uncoupling protein 1 (UCP1) gene die mogelijk een pathogeen effect uitoefenen in het ontwikkelen van obesitas"* by Hanne Leysen, 2nd Bachelor Biochemistry and Biotechnology, University of Antwerp, Belgium. Academic Year 2014-2015.

Assistance Bachelor Thesis: "*Genetische studie naar de rol van twee kandidaatgenen in de pathogenese van obesitas*" by Evelien van Dijck, 3rd Bachelor in Biochemistry and Biotechnology, University of Antwerp, Belgium. Academic Year 2015-2016.

Assistance Bachelor Thesis: "*De rol van bruine en beige adipocyten in de pathogenese van obesitas*" by Sander Eens, 3rd Bachelor in Biomedical Sciences, University of Antwerp, Belgium. Academic Year 2016-2017.

Assistance Bachelor Thesis: "*De rol van bruine en beige adipocyten in de pathogenese van obesitas*" by Sander Eens, 3rd Bachelor in Biomedical Sciences, University of Antwerp, Belgium. Academic Year 2016-2017.

Assistance Bachelor Thesis: "*Genetische analyse van kandidaatgenen voor monogene obesitas*" by Britt Poppe, 3rd Bachelor in Biochemistry and Biotechnology, University of Antwerp, Belgium. Academic Year 2016-2017.

VIII. DANKWOORD

DROMEN

Iedereen heeft ze en toch zijn ze voor iedereen heel verschillend. Vijf jaar geleden kwam er één van mijn dromen uit: "De start van mijn doctoraat aan het Centrum Medische Genetica". Heel erg gelukkig en vol ambitie was ik dan ook klaar om me **vollen bak** in te zetten voor een IWT beurs!

Iedereen weet hoe verrassend dat afliep…

Ik werd er dubbel en na 9 maanden ook redelijk "dik" voor beloond, een tweede droom werd plots waargemaakt met de geboorte van ons dochtertje. Vele jaren, een paar kleine rampen en heel wat dromen later, is het nu tijd om mijn doctoraat af te ronden. Hoog tijd dus om de nodige mensen te bedanken. Want zonder de steun, hulp en vooral ook **vriendschap** van velen, zou het mij nooit gelukt zijn.

Allereerst wil ik heel erg graag mijn promotor **Wim** bedanken voor de mooie kansen die hij mij gegeven heeft. Wim, bedankt voor je vertrouwen, je waardevolle steun en voor de vele "ok's" op heel wat vragen en ideeën. Daarnaast wil ik je ook nog even speciaal bedanken voor jouw oprechte medeleven, begrip en voor de flexibiliteit die je me gaf op momenten dat het nodig was. Ik heb de voorbije jaren heel erg graag in jouw onderzoeksgroep gewerkt. **Bedankt**!
Ik wil ook mijn tweede promotor, **Prof. Van Gaal**, bedanken om de brug te vormen met de dienst Endocrinologie, Diabetologie en Metabolisme. Bedankt voor de waardevolle wetenschappelijke discussies en ideeën van de afgelopen jaren.

Een grote dank gaat ook uit naar mijn **co-promotor**, **Sigri**. Bedankt om me tijdens het eerste jaar onder je vleugels te nemen en me in te wijden in de wereld van obesitas. Door jouw persoonlijke en intensieve begeleiding tijdens de "IWT maanden" werd een stevige basis voor de rest van mijn doctoraat gelegd. Mede dankzij jouw jarenlange ervaring en gedrevenheid heb ik een mooi project kunnen afwerken.

De leden van de jury, **Prof. Vanden Berghe, Prof. Van der Zee, Prof. Gies, Dr. Dirinck en Dr. Van Haelst**, wil ik graag bedanken voor de interesse die ze toonden in mijn onderzoeksproject en voor de kritische evaluatie van dit proefschrift.

Graag zou ik ook de artsen willen bedanken die betrokken waren bij dit project. Alle bloedstalen van kinderen met obesitas die door hen werden opgestuurd waren essentieel in mijn doctoraatsonderzoek. **Prof. Verhulst, Prof. Massa en Dr. Van Hoorenbeeck** wil ik dan ook heel erg graag bedanken voor de aangename en vlotte samenwerking.
Dr. Van Hoorenbeeck heb ik verder leren kennen als een uitstekende kinderlongarts. Kim, ik wil je heel graag in het bijzonder bedanken om Lora zo nauwgezet op te volgen met je uitgebreide kennis en expertise. Jij maakte een einde aan een bange periode vol onzekerheden. Bedankt om meer dan je job te doen!

Een speciaal woordje wil ik verder ook richten aan **Dr. An Verrijken**. Bedankt voor het kritisch nalezen van alle manuscripten, voor je waardevolle opmerkingen en suggesties en voor de toffe mailtjes.

Liesbeth, wij zijn altijd een beetje een gek duo geweest. Zo verschillend, maar op veel vlakken toch hetzelfde. Na onze master, was ik dan ook super content dat jij ook in het CMG je doctoraat ging doen. Voor het beantwoorden van de "stomme" vraagjes die ik enkel maar aan jou durfde te stellen, voor de toffe gesprekken, voor de steun en het vertrouwen, voor de vele koekjes die je me toestopte als ik de mijne weer eens vergeten was, voor de lach en soms ook de traan die we samen konden delen, wil ik je bedanken. Bedankt om zo'n uitzonderlijk goeie collega en vriendin te zijn.

Ik wil ook heel erg graag een bijzonder **team** in mijn doctoraat bedanken: de **BoObies**!
Ellen, mijn obesitas-maatje, ik zal nooit je gezichtsuitdrukking vergeten toen ik je vertelde dat er toch echt wel "iets mis was" met ons hotel in Orlando. Vol goeie moed en doorzettingsvermogen, besloot je na 20 uur onderweg te zijn geweest, toch nog op zoek te gaan naar een bewoonbaar stukje van het hotel. Hard gezocht, maar niet gevonden. Een moment dat me voor altijd zal bijblijven. Bedankt Ellen, voor de toffe babbels, leuke baby-gesprekken, de steun en vriendschap van de voorbije jaren.
Samen met **Sara, Silke en Raphaël** was het altijd extra gezellig aan onze kant van de bureau. Wekenlang bikini's en brillen kiezen, redenen verzinnen om toch nog maar eens te gaan lunchen, online shoppen, namen facebooken, ongekende familieleden opzoeken of ander detectivewerk, freaky voorspellingen maken (vooral Silke dan 😊), het gekste eerst! Maar daarnaast werd er ook heel hard gewerkt en efficiënt SAMEN-gewerkt. Zo werden heel wat samenvattingen, overzichten en eten (vooral Raphaël dan 😊) gedeeld. Bij jullie was het eigenlijk echt "thuiskomen". Voor de vele zotte en soms ook heel serieuze babbels, voor de grenzeloze steun, vriendschap en het hippe kleding- en media-advies wil ik jullie heel erg graag bedanken! Jullie zijn toppertjes!
Net als de rest van ons team eigenlijk. **Eveline, Ellen** en **Esther** jullie waren de voorbije jaren ontzettend lieve collega's. Onze dagelijkse coffee-time was telkens supergezellig en werd al snel onmisbaar. Bedankt voor de toffe babbels en de vriendschap van de voorbije jaren. Esther, ik wil jou ook nog even speciaal bedanken voor de leerrijke renovatie-tips. Jij begreep de renovatie-frustraties als geen ander. Bedankt!

Bedankt ook aan alle **CMG'ers** en **ex-CMG'ers** voor de bijzonder aangename werksfeer, de toffe babbels, de vele baby-bezoekjes. Het was fantastisch om met jullie te mogen samenwerken!

De laatste paragraaf gaat naar de meest **bijzondere** personen in mijn leven.
Jeff, mijn bolleke/hubby, bedankt voor je grenzeloze liefde, je vertrouwen en steun die je me de voorbije jaren gegeven hebt. Deze waren van onschatbare waarde tijdens mijn doctoraat! Je hebt mijn leven compleet gemaakt en gaf me het meest kostbare ter wereld: onze twee kapoentjes.
Lora, jij bent mama's kleine babbelkous, een lief prinsesje dat met de nodige kwajongensstreken ons leven heel wat spannender en vooral ook veel mooier maakt.
Stan/Stanneke/Stanneman/Frangipan/Protteman, als mama in je mooie en altijd vrolijke oogjes kijkt, dan gloeit mijn hart van trots. Jullie brengen samen leven in onze brouwerij. Jullie

omringen me dagelijks met vrolijke gezelligheid, vriendschap en liefde die geen grenzen kent. Ik hou zo ontzettend veel van jullie!
"Last but not least", **Mama** en **papa**! Wat heb ik zó enorm veel aan jullie te danken! Ik wil jullie vandaag dan ook heel erg graag in de bloemetjes zetten. Bedankt voor jullie nooit aflatende steun, voor de vele potjes zelfvertrouwen, de geluks-reuk, de talloze keren van hulp bij computerstress (lees: bij "ik ga bijna mijn laptop buiten smijten"-frustraties) en voor de grenzeloze liefde. Of het nu 's middags of in het midden van de nacht is, ik weet dat ik altijd met alles bij jullie terecht kan. Jullie zijn het mooiste voorbeeld dat ik maar volgen kan om een goeie ouder voor Stan en Lora te zijn. Bedankt voor alles!