In depth analysis of obesity candidate genes in the chromosomal region 16p11.2 identified in genome-wide association studies

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

Dr. rer. nat.

der Fakultät für

Biologie

an der

Universität Duisburg-Essen

vorgelegt von

Anna-Lena Volckmar

aus Gelsenkirchen

Juni 2013

Die der vorliegenden Arbeit zugrunde liegenden Experimente wurden in der Klinik für Psychiatrie, Psychosomatik und Psychotherapie des Kindes- und Jugendalters an der LVR Klinik/Institut der Universität Duisburg-Essen durchgeführt.

1. Gutachter: Prof. Dr. Johannes Hebebrand
2. Gutachter: Prof. Dr. Bernhard Horsthemke
3. Gutachter: -

Vorsitzender des Prüfungsausschusses: Prof. Dr. Michael Ehrmann
Tag der mündlichen Prüfung: 30.10.2013

Publications

Volckmar AL, Pütter C, Song JY, Graninger J, Knoll N, Wolters B, Hebebrand J, Scherag A, Reinehr T, Hinney A. Analyses of Non-Synonymous Obesity Risk Alleles in SH2B1 (rs7498665) and APOB48R (rs180743) in Obese Children and Adolescents Undergoing a 1-year Lifestyle Intervention. Exp Clin Endocrinol Diabetes. 2013; 121:1-4.

Volckmar AL, Bolze F, Jarick I, Knoll N, Scherag A, Reinehr T, Illig T, Grallert H, Wichmann HE, Wiegand S, Biebermann H, Krude H, Fischer-Posovszky P, Rief W, Wabitsch M, Klingenspor M, Hebebrand J, Hinney A. Mutation screen in the GWAS derived obesity gene SH2B1 including functional analyses of detected variants. BMC Med Genomics. 2012;5:65.

Albayrak O, Pütter C, **Volckmar AL**, Cichon S, Hoffmann P, Nöthen MM, Jöckel KH, Schreiber S, Wichmann HE, Faraone SV, Neale BM, Herpertz-Dahlmann B, Lehmkuhl G, Sinzig J, Renner TJ, Romanos M, Warnke A, Lesch KP, Reif A, Schimmelmann BG, Scherag A, Hebebrand J, Hinney A; Psychiatric GWAS Consortium: ADHD Subgroup. Common obesity risk alleles in childhood attention-deficit/hyperactivity disorder. Am J Med Genet B Neuropsychiatr Genet. 2013 Mar 26. doi: 10.1002/ajmg.b.32144. [Epub ahead of print]

Knoll N, Jarick I, **Volckmar AL,** Klingenspor M, Illig T, Grallert H, Gieger C, Wichmann HE, Peters A, Hebebrand J, Scherag A, Hinney A. Gene Set of Nuclear-Encoded Mitochondrial Regulators Is Enriched for Common Inherited Variation in Obesity. PLoS ONE. 2013;8: e55884.

Jarick I, **Volckmar AL**, Pütter C, Pechlivanis S, Nguyen TT, Dauvermann MR, Beck S, Albayrak O, Scherag S, Gilsbach S, Cichon S, Hoffmann P, Degenhardt F, Nöthen MM, Schreiber S, Wichmann HE, Jöckel KH, Heinrich J, Tiesler CM, Faraone SV, Walitza S, Sinzig J, Freitag C, Meyer J, Herpertz-Dahlmann B, Lehmkuhl G, Renner TJ, Warnke A, Romanos M, Lesch KP, Reif A, Schimmelmann BG, Hebebrand J, Scherag A, Hinney A. Genome-wide analysis of rare copy number variations reveals PARK2 as a candidate gene for attention-deficit/hyperactivity disorder. Mol Psychiatry. 2012; doi: 10.1038/mp.2012.161. [Epub ahead of print]

Knoll N, **Volckmar AL**, Pütter C, Scherag A, Kleber M, Hebebrand J, Hinney A, Reinehr T. The fatty acid amide hydrolase (FAAH) gene variant rs324420 AA/AC is not associated with weight loss in a 1-year lifestyle intervention for obese children and adolescents. Horm Metab Res. 2012;44:75-7.

Hinney A, Scherag A, Jarick I, Albayrak Ö, Pütter C, Pechlivanis S, Dauvermann MR, Beck S, Weber H, Scherag S, Nguyen TT, **Volckmar AL**, Knoll N, Faraone SV, Neale BM, Franke B, Cichon S, Hoffmann P, Nöthen MM, Schreiber S, Jöckel KH, Wichmann HE, Freitag C, Lempp T, Meyer J, Gilsbach S, Herpertz-Dahlmann B, Sinzig J, Lehmkuhl G, Renner TJ, Warnke A, Romanos M, Lesch KP, Reif A, Schimmelmann BG, Hebebrand J; Psychiatric GWAS Consortium: ADHD subgroup. Genome-wide association study in German patients with attention deficit/hyperactivity disorder. Am J Med Genet B Neuropsychiatr Genet. 2011;156B:888-97.

Scherag A, Jarick I, Grothe J, Biebermann H, Scherag S, **Volckmar AL**, Vogel CI, Greene B, Hebebrand J, Hinney A. Investigation of a genome wide association signal for obesity: synthetic association and haplotype analyses at the melanocortin 4 receptor gene locus. PLoS One. 2010;5:e13967.

Reviews

Volckmar AL, Knoll N, Hinney A. Genetische Prädikatoren der Adipositas. In: Ardelt-Gattinger, E, Ring-Dimitriou, S, Weghuber, D (eds). Der gesunde Adipöse. Hogrefe Verlag. 2013 (in press).

Hebebrand J, Hinney A, Knoll N, **Volckmar AL**, Scherag A. Molekulargenetische Aspekte der Körpergewichtsregulation. Dtsch Arztebl Int 2013; 110: 338-44.

Hinney A, **Volckmar AL**, Knoll N. Melanocortin-4 receptor in energy homeostasis and obesity pathogenesis. Prog Mol Biol Transl Sci. 2013;114:147-91.

Hebebrand J, **Volckmar AL**, Knoll N, Hinney A. Chipping away the 'missing heritability': GIANT steps forward in the molecular elucidation of obesity - but still lots to go. Obes Facts. 2010;3:294-303.

Talks

„*SH2B1*: Von der Adipositasmutation zur therapeutischen Option“; 28. Jahrestagung der Deutschen Adipositasgesellschaft 2012, October 4th-6th in Stuttgart, Germany

„Effect of variants in *SH2B1* and *APOB48R* in obesity and weight loss“; 2nd Annual Meeting of the BIOME Core “Genetics and Cell Biology” 2012, November 29th in Essen, Germany

Posters

„*SH2B1*- A new, rare mutation leading to monogenic obesity?“ Volckmar AL, Bolze F, Jarick I, Knoll N, Scherag A, Illig T, Grallert H, HE, Biebermann H, Krude H, Fischer-Posovszky P, Rief W, Klingenspor M, Wabitsch M, Hebebrand J, Hinney A

- 1st Annual Meeting of the BIOME Core “Genetics and Cell Biology” and GRK 1431 “Gene Transcription” 2011, May 3-4 in Dormagen/Zons, Germany
- 4th Annual Meeting of NGFN-Plus and NGFN-Transfer in the Program of Medical Genome Research 2011, September 27-29 in Berlin, Germany
- 27. Jahrestagung der Deutschen Adipositasgesellschaft 2011, October 6-8 in Bochum, Germany
- Forschungstag des Universitätsklinikums Essen 2011, November 25 in Essen, Germany
- 19th European congress on Obesity 2012, May 9-12 in Lyon, France
- EMBO/EMBL Symposium Diabetes and Obesity 2012, September 13-16 in Heidelberg, Germany

„Effect of variants in *SH2B1* and *APOB48R* in obesity and weight loss“ Volckmar AL, Song JY, Graninger J, Bolze F, Jarick I, Pütter C, Knoll N, Scherag A, Illig T, Grallert H, HE, Biebermann H, Krude H, Fischer-Posovszky P, Rief W, Klingenspor M, Wabitsch M, Wolters B, Wang H, Hebebrand J, Reinehr T, Hinney A

- Forschungstag des Universitätsklinikums Essen 2012, November 23rd in Essen, Germany

Table of Contents

List of Abbreviations

(-)-HEP	1-(α-hydroxyethyl)pyrene
ADCY3	adenylate cyclase 3
ADCY9	adenylate cyclase 9
ADPGK	ADP-dependent glucokinase
AGA	Arbeitsgemeinschaft Adipositas im Kindes- und Jugendalter
AKT	AKT1 kinase
APOB48R	Apolipoprotein B48 receptor
ATP	Adenosine-5'-triphosphate
ATP2A1	ATPase, Ca++ transporting, fast twitch 1 isoform
ATXN2L	ataxin 2 related protein isoform C
BDNF	brain derived neurotrophic factor
BDNFR	brain-derived neurotrophic factor receptor
BEPOC	Berlin Paediatric Obese Cohort
BMI	body mass index
BMI SDS	standard deviation from the medium BMI percentile
BSA	Bovine serum albumin
CCDC101	coiled-coil domain containing 101
CD19	CD19 antigen precursor
CDKAL1	CDK5 regulatory subunit associated protein 1-like 1
CI	confidence interval
CLN3	ceroid-lipofuscinosis, neuronal 3
CNV	Copy number variation
DAPOC	Datteln Paediatric Obese Cohort
dNTP	desoxynucleotide
ddNTP	didesoxynucleotide
DGAV	Deutsche Gesellschaft für Allgemein- und Viszeralchirurgie e.V.
dHPLC	Denaturing High Performance Liquid Chromatography
DNA	desoxy-ribonucleic acid
DNBA	2,4-dinitrobenzylalcohol
dsDNA	double stranded DNA
EAF	effect allele frequency
EC50	half-maximal activation capacity of the enzyme
EGFR	growth hormone receptors
EIF3CL/EIF3C	eukaryotic translation initiation factor 3
Emax	maximal activation capacity of the enzyme
En	total energy
ESE	exonic splicing enhancer
ESS	exonic splicing silencers
FANCL	Fanconi anemia, complementation group L
FOXO1	forkhead box O1
FTO	fat mass and obesity associated gene
GHR	growth hormone receptors
GIANT	Genetic Investigation of Antropomorphic Traits
GIPR	gastric inhibitory polypeptide receptor
GP2	glycoprotein 2
GPRC5B	G protein-coupled receptor, family C, group 5, member B
GWAS	Genome wide association study

HEK293	human embryonic kidney cells 293
HMP	1-hydroxymethylpyrene
HSD17B12	hydroxysteroid (17-beta) dehydrogenase 12
HOMA IR	Homeostasis Model Assessment of insulin resistance
HOXB5	homeobox B5
IGF1R	insulin-like growth factor I receptor
IL27	interleukin 27
INSR	insulin receptor
INSULA	Ulm Paediatric Obese Cohort A and B; Ulm children's study 2 and 3
IRS	Insulin receptor substrate
JAK	Janus kinase
KiGGS	bundesweiter Kinder- und Jugendgesundheitssurvey
KORA	Cooperative Health Research in the Region of Augsburg
LAT	linker for activation of T cells isoform b
LBXCOR1	transcriptional co repressor Corl1
LD	linkage disequilibrium
LEPRb	leptin receptor
MAF	minor allele frequency
MALDI TOF	matrix-assisted laser desorption/ionization – time of flight analysis
MAP2K5	mitogen-activated protein kinase kinase 5
MC4R	melanocortin 4 receptor
MONICA	monitoring trends and determinants in cardiovascular disease
MSRA	methionine sulfoxide reductase A
MTCH2	mitochondrial carrier 2
MTIF3	mitochondrial translational initiation factor 3
NFATC2IP	nuclear factor of activated T-cells, cytoplasmic 2-interacting protein
NPC1	Niemann-Pick disease, type C1 gene
NUPR1	p8 protein isoform a
OH-AAF	N-hydroxy-2-acetylaminofluorene
OH-AAP	2-hydrolyamino-5-phenylpyridine
OLFM4	olfactomedin 4
OR	odds ratio
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PCSK1	proprotein convertase subtilisin/kexin type 1
PDGF	platelet-derived growth factor alpha
PDK4	pyruvate dehydrogenase kinase, isozyme 4
PH	pleckstrin homology
PI3K	phosphoinositol-3 kinase
POC5	POC5 centriolar protein homolog
PPARγ	Peroxisome proliferator-activated receptor gamma
RABEP2	rabaptin, RABGTPase binding effector protein 2
RAF	risk allele frequency
RFLP	restriction fragment length polymorphism
SDCCAG8	serologically defined colon cancer antigen 8 ()
SEC16B	SEC16 homolog B
SH2B1	SH2B adaptor protein 2 isoform 1
SH2B2	SH2B adaptor protein 2 form 2, also APS

SH2B3	SH2B adaptor protein 2 form 3, also LNK
SHC	Src homology 2 domain containing protein
SLC39A8	solute carrier family 39
SNP	single-nucleotide polymorphism
SPNS1	spinster homolog 1 isoform 1
SSCP	Single Stranded Conformation Polymorphism
ssDNA	single stranded DNA
STAT3	Signal transducer and activator of transcription 3
SULT	sulfotranferase
SULT1A1	sulfotransferase family, cytosolic, 1A, member 1
SULT1A2	sulfotransferase family, cytosolic, 1A, member 2
T2D	Type 2 *Diabetes mellitus*
TCF7L2	transcription factor 7-like 2
TAE buffer	Tris-Acetat-EDTA buffer
TDT	Transmission Disequilibrium Test
TEAA	N,N-Diethylethanamine acetate
TMEM18	transmembrane protein 18
TNKS	tankyrase
TRKA	nerve growth factor receptor
TSF	Transcription factor binding sites
TUFM	Tu translation elongation factor, mitochondrial
URMEL	Ulm Research on Metabolism, Exercise and Lifestyle in Children; Ulm children's study 1
USA	United States of America
UTR	untranslated region
VFA	visceral fat area
VAT	visceral adipose tissue
WHO	world health organization
WHR	waist-hip-ratio

List of Tables

-Table will be continued-

-Table continued-

List of Figures

1. Introduction

1.1 Obesity

Obesity is characterized by an excess of body fat in relation to body composition. The World Health Organization (WHO) defines obesity by a body mass index (BMI) above 30 kg/m², while a BMI above 40 kg/m² defines severe obesity (WHO 1995, 2000, 2004). Obesity imposes a major risk for global health as it can lead to a range of co-morbidities like type 2 diabetes mellitus (T2D, Kloting et al. 2007), increased blood pressure (Nishina et al. 2003), atherosclerosis (Van Gaal et al. 1989), and increased risk for cardiovascular disease (Despres 2007). Additionally, obesity is associated with other diseases like Parkinson's disease (Hu et al. 2006), Alzheimer's disease (Whitmer et al. 2007) and several cancers (for example of the digestive system or kidney carcinoma). A BMI above 40 kg/m² increases the risk to die of cancer by 1.5 compared to normal weight individuals. Not only do obese individuals die from cancer more often, they also have higher lifetime rates of cancer (Calle et al. 2007).

According to the German Mikrozensus, 36.3% of all Germans are overweight, while 15.8% are obese. For men the numbers are higher than for women, as 44% of all German men are overweight and 16.1% obese, but only 28.8% of the German women are overweight and 15.6% are obese (German Mikrozensus 2011). These numbers are lower compared to the Nationale Verzehrsstudie II (Nationale Verzehrsstudie II 2008) which measured the participants directly instead of surveying. In 2010, the obesity rate among adult Americans was estimated at about 36% for both genders (Flegal et al. 2012). Obesity prevalence in the USA for adults aged 20 to 74 years increased by 7.9% for men and by 8.9 percentage points for women between 1976-1980 and 1988-1994. During the same time, overweight rates increased from 45% to 67%. Only recently, this increase has reached a plateau, as the mean adult BMI in the US has shown no significant increase since 2003 (Flegal 2012).

Obesity and its co-morbidities have large economic consequences not only on the individual level. Obesity is a chronic condition and one of the leading preventable causes of death worldwide (Allison et al. 1999). The cost of obesity on public health are high with an estimated $190.2 billion or 20.6% of all medical expenditures in the United States of America (Cawley and Meyerhoefer 2012) and about €13 billion or 8.4% of all medical costs of the statutory health insurance in Germany (Knoll and Hauner 2008).

1.1.1 Body Mass Index (BMI)

Body weight is a quantitative trait, whose extremes (extreme over- or underweight) can lead to several secondary co-morbidities. To classify body weight categories, the heuristic

measure of BMI or Quetelet-index is used internationally. BMI is calculated by the body weight in kilogram divided by body height in meter, squared:

$$\text{BMI (kg/m}^2\text{)} = \text{body weight [kg]} / \text{body height}^2\ [\text{m}^2]$$

To define thresholds for under- and overweight, the WHO regards a BMI below 18.5 kg/m² as underweight, between 18.5 kg/m² and 25 kg/m² as normal weight, and above 25 kg/m² as overweight (Table 1; WHO 1995, 2000, 2004).

The BMI provides several advantages: (i) BMI correlates strongly with body fat mass (Keys et al. 1972), but less strong with body height (Rolland-Cachera et al. 1982). (ii) While both body weight and height fluctuate diurnally and during menstrual cycle in women, BMI is highly reproducible and can be obtained easily. (iii) Since correlation with body fat is high, BMI is a good predicator for body composition (Jeukendrup and Gleeson 2005).

Table 1: Classification of weight categories in BMI thresholds as defined by the World Health Organization (modified from WHO 2004)

BMI range [kg/m²]	**Weight category**
<16.00	Severe thinness
16.00 - 16.99	Moderate thinness
17.00 - 18.49	Mild thinness
<18.50	Underweight
18.50 - 24.99	Normal range (healthy weight)
≥25.00	Overweight
≥30.00	Obesity
30 - 35	Obesity Class I
35 - 40	Obesity Class II
>40	Obesity Class III

BMI is used to define weight categories at the population level. The index has low predictive value for the amount of body fat at the individual level. The highest predictive value for the amount body fat applies to individuals with average body composition and low to medium activity levels (Jeukendrup and Gleeson 2005). Body composition of individuals deviating from the average composition like athletes, children or elderly is not represented accurately by BMI. In athletes, muscle weight contributes to body weight and increases BMI often into the overweight range, although athletes have low body fat. Lower bone density and decreased body height in elderly people decreases predictive value of BMI for body fat percentage in proportion to total body mass.

1.1.2 BMI in children and adolescents

In children, the use of BMI as a predictor for co-morbid diseases is not useful since body composition changes depending on age and sex. During development, BMI continually increases during the first 6 months because of an increase in fat mass, decreases then due to increased linear growth during age of 6 months to 8 years, to increase again at approximately 6-8 years of age. This period is referred to as the obesity rebound (Rolland-Cachera et al. 1982). An increased proportion of muscle mass in males of any age leads to gender differences in BMI scale.

Table 2: Classification of children and adolescents into weight categories by BMI percentiles according to the Arbeitsgemeinschaft Adipositas im Kindes- und Jugendalter (AGA; http://www.aga.adipositas-gesellschaft.de).

Percentile	Weight category
< 3rd	Extreme underweight
< 10th	Underweight
10th to 90th	Normal (healthy weight)
> 90th	Overweight
> 97th	Obesity
> 99.5th	Extreme obesity

In children and adolescents, age-, race- and sex adjusted BMI percentile curves are used worldwide. For these, a reference population is necessary equally sampled from all socioeconomic strata, all age groups and both genders (Hebebrand et al. 1994; Kromeyer-Hauschild et al. 2001). This data can be used to classify German children and adolescents into weight categories with a specific health risk prediction.

In addition to the BMI percentiles, the BMI standard deviation score (BMI SDS) is commonly used in children. It indicates a deviation of the BMI of an individual from the BMI median of a population with the same sex and age. A BMI SDS value above 2 is considered as obese, a BMI SDS below -2 as underweight (de Onis et al. 2010).

1.1.3 Obesity prevalence in children and adolescents

The WHO estimated that worldwide more than 20 million children below age 5 are overweight. In Europe an overall BMI increase is reported for Denmark (Petersen et al. 2002), England (Stamatakis et al. 2005), The Netherlands (Fredriks et al. 2000), Belgium (Hulens et al. 2001), and Spain (Moreno et al. 2001). In Germany, the most recent study on BMI in children and adolescents is the KiGGS (bundesweiter Kinder- und Jugendgesundheitssurvey). A total of 14,747 individuals aged three to 17 years were analyzed. Comparison of the KiGGS data which was collected 2003-2006 with the former

data from Kromeyer-Hauschild et al. (2001), which was not recruited epidemiologically, classifies 15% of the sampled children and adolescents as overweight (BMI ≥ 90th age and sex specific BMI percentile) and 6.3% as obese (BMI ≥ 97th age and sex specific BMI percentile; Kurth und Schaffrath-Rosario 2007).

This trend to increased obesity rates in children and adolescents can also be observed in the United States of America. In the USA, prevalence rates have been monitored regularly since 1960. While there were no significant changes of obesity prevalence rates of American children and adolescent (age 6-19 years) until 1980, rates have increased since then (Ogden et al. 2006). Currently, 33.6% of all children have a BMI above their 85th age- and sex specific BMI percentile and are therefore regarded as overweight, while 17.1% are obese (above the 95th age- and sex specific BMI percentile, Ogden et al. 2006;reference population by Himes and Dietz 1994).

Only recently, the increase in obesity rates has come to a halt. Data of 467,294 children aged 2-19 years coming from nine countries (Australia, China, England, France, Netherlands, New Zealand, Sweden, Switzerland and USA) showed that overweight and obesity rates stabilized between 1995 to 2008 (Olds et al. 2011).

1.1.4 Causes of obesity

A combination of genetic and environmental factors leads to obesity. A positive energy balance (more calories consumed than spent for basic metabolic rate and physical activity) over a long period of time leads to obesity. Increased obesity rates are partially caused by recent lifestyle changes, namely increased intake of energy dense food and lower physical activity levels. But an analysis of data from the USA showed that only a part of the population responded to these changes with increased weight (Troiano and Flegal 1998). While the lower to medium BMI percentiles (3rd to 50th) remained stable despite the lifestyle changes, BMI values comprising the 90th and 97th BMI percentile increased in children and adolescents over a period of 30 years. Also, the individual genetic background plays a role in the development of overweight (reviewed in Hinney et al. 2010).

The human genetic makeup evolved during times when high caloric food was sparse and individuals with energy efficient metabolisms had higher chances of procreation. The “thrifty genotype” hypothesis predicts that a genetic predisposition to obesity was advantageous in periods of low food availability (Neel 1962). This resulted in a human genome enriched for genetic variants that favor energy storage and uptake as well as lowered energy expenditure. In the modern high caloric environment, these genetic variations represent a maladaptation which leads to obesity (Hebebrand et al. 2010, Volckmar et al. in press).

1.1.5 Therapy

Today, obesity and its co-morbidities are a major cause of morbidity and mortality worldwide (Allison et al. 1999). Despite the large scale of the problem, successful prevention and treatment options for obesity remain limited (Cochrane et al. 2012). Therapy options for obesity include behavioral therapy, dietary changes, increased physical activity, pharmacotherapy and bariatric surgery. Generally, a therapy is indicated if the adult patient has a BMI ≥ 30 kg/m² or if the BMI is between 25-29.9 kg/m² and one of the following symptoms applies: (i) health problems related to overweight, (ii) high visceral fat mass distribution and/or (iii) substantial psychological suffering related to the overweight (Leitlinie 088 – 001, Deutsche Gesellschaft für Allgemein- und Viszeralchirurgie e.V. (DGAV)).

Dietary modifications like altered macronutrient composition and restricted caloric intake are the most common obesity treatments. However, diet programs result only in moderate weight-loss and poor maintenance of weight-loss. For example, a randomized control trial compared weight loss over six months for reduced carbohydrate intake (“Dr Atkins' new diet revolution”), formula diet (“Slim-Fast plan”), formula diet with physical activity (“Rosemary Conley's eat yourself slim diet and fitness plan”), and caloric restriction (“Weight Watchers pure points programme”). The 300 participants had an average weight loss of 4.4 kg; all diets were comparably successful (Truby et al. 2006). The recent follow up on the same study groups included general practioner’s care. Here, the weight loss ranged from 1.37 kg (general practice) to 4.43 kg (“Weight Watchers”; Jolly et al. 2011). Weight loss maintenance over a longer period of time was not reported for both studies; it is the main issue for all weight loss approaches.

While in adults, main focus of the therapy is to reduce weight to decrease BMI and therefore health risks, in children and adolescents the aim is to keep their weight stable. As they still grow, a decrease in BMI SDS score can be achieved even if the BMI itself increases over time. Treatment effects in terms of weight loss are generally small, so that ongoing efforts are needed to achieve long term results. Focusing on other interventions in children and adolescents, there are few randomized controlled trials that can be compared to ‘Obeldicks’. The ‘Obeldicks’ program is a 1-year lifestyle intervention for children aged 6 to 16 years with a BMI above the 90th percentile. In short, ’Obeldicks‘ is based on nutritional education, behavioral therapy including individual psychological care of the child and his or her family, and physical exercise. The nutritional education was based on an “optimized mixed diet”, containing 55% carbohydrates (5% sugar) of total dietary energy, 30 En% fat and 15 En% proteins (Reinehr et al. 2007, Reinehr et al. 2009, Dobe et al. 2011). In a study of Holzapfel et al. (2011) on the ‘Long-term effects of a lifestyle intervention in obesity and genetic influence in children (LOGIC)’, the BMI SDS loss over a period of 4-6 weeks was 0.36 ± 0.1,

which is comparable to the BMI SDS reduction of 0.3 over 1 year in the 'Obeldicks' sample. In a 5 months life-style intervention of Centis et al. (2012), BMI SDS decreased only by 0.08 in the intervention group.

1.2 Genetic mechanisms in obesity

Heritability estimates for BMI variance mostly range from 40-80% (Stunkard et al. 1986, Stunkard et al. 1999; Hinney et al., 2010), with estimates higher in adolescents than in adults (Pietiläinen et al. 1999). The estimates were derived from twin, family and adoption studies where twin studies usually result in the highest and family studies the lowest estimates of heritability (Maes et al. 1997, Hartz et al. 1977).

In twins raised apart, the BMI of the adopted twins showed no correlation to the adoptive parents (Stunkard et al. 1986). Up to 80% of the variability of the BMI was attributed to genetic factors (Pietiläinen et al. 1999) in a study that compared 4,884 twins and 2,509 singletons below 17 years of age. While all twins share the intrauterine environment, about two thirds of monozygotic twins share the same amnion. This factor cannot be excluded from the analysis, thus raising the estimates for genetic influences on trait heritability. Also, there may be unknown non-additive or epistatic effects (Stunkard et al. 1990).

Heritability estimates derived from family studies are much more diverse. In family studies, where the correlation of body weight between children and biological parents was analyzed, heritability is estimated mainly to be only 20-30%. Other studies report estimates of 70% which is close to the data derived from twin studies (Moll et al. 1991, Maes et al 1997). A single study detected a heritability of BMI variance of only 5% (Bouchard et al. 1988).

Adoption studies showed a higher BMI correlation of adoptees with their biological parents, while the BMI correlation with the adoptive parents was small (Sørensen et al. 1992). For family studies, the age of the analyzed individuals and the geographical setting of the study play a large role (Elks et al. 2012a).

1.2.1 Syndromal forms of obesity

Obesity is a symptom of several infrequent genetic syndromes; the most common example is the Prader-Willi syndrome with a frequency of 1:15,000 to 1:25,000 (Buiting 2010). A deletion of the paternal chromosome 15q11-q13 is the most common reason for Prader-Willi syndrome, but a maternal disomy of chromosome 15 or an imprinting defect can lead to the same phenotype (Horsthemke and Buiting 2006). A causal gene for the hyperphagia that leads to obesity in these patients has not yet been discovered, although a knockout mouse model of the small nucleolar RNA SNORD116 located on chr15q11-q13 was shown to be lean (Powell et al. 2013). The exact mechanisms of SNORD116's influence on energy expenditure is not yet known.

Obesity resulting from genetic deletions or loss of function mutations in major obesity genes can also be regarded as syndromal obesity as they have diverse other effects on the organism. For example, obesity elicited by a chromosomal deletion on chr16p11.2 is often accompanied by other symptoms like autism or developmental delay (e.g. Bochukova et al. 2011, Bachman-Gagescu et al. 2011) and therefore regarded as a syndromal form of obesity (Shinawi et al. 2009, Perrone et al. 2010). Monogenic obesity caused by mutations in the gene encoding leptin or the leptin receptor are accompanied by delayed puberty and fertility defects as leptin is involved in the gonadotropin-gonadal axis and the regulation of energy homeostasis (Farooqi et al. 1999, Ramachandrappa and Farooqi 2011).

1.2.2 Monogenic forms of obesity and major gene effects

Genes with a strong influence on weight regulation and obesity development are termed major obesity genes. A monogenic form of obesity is present if a mutation in a single gene is sufficient to cause obesity. Besides the monogenes Leptin and Melanocortin 4 receptor described in detail below, variants in NPY, CART and AGRP can lead to obesity (Hebebrand et al. 2013). Monogenic obesity only plays a minor role in the general population as these forms are rare.

1.2.2.1 Leptin

One of the obesity monogenes originally detected in the mouse (ob/ob gene; Zhang et al. 1994) is the leptin gene encoding the satiety hormone leptin. It is mainly produced by adipocytes and signals the size of body fat mass via the blood stream to the hypothalamus. Amongst other pathways, signalling is passed on from the leptin receptor by inducing pro-opiomelanocortin (POMC) transcription. POMC is cleaved e.g. into adrenocorticotrophic hormone (ACTH) and alpha and beta melanocyte stimulating hormone (α-MSH and β-MSH) which both activate the melanocortin 4-receptor (MC4R). If leptin levels are high, the pathway is anorexigenic and increases satiety, leading to decreased food intake (reviewed in Hinney et al. 2010).

Only few recessively inherited mutations in the leptin or leptin receptor gene are known in humans. Affected individuals develop an early-onset extreme obesity and derive with one exception (Fischer-Posovszky et al. 2010) from consanguineous families (Montague et a. 1997, Mazen et al. 2009, Strobel et al. 1998). Patients with these mutations lack satiety and consume large amounts of food. For example, a nine year old mutation carrier consumed 1600 kcal in a test meal which is comparable to her total daily caloric requirement (D-A-CH Referenzwerte für die Nährstoffzufuhr, 2008). Treatment with subcutaneous leptin injections reduced this caloric uptake to 1000 kcal in a test meal and led to weight loss of 16.4 kg in one year (Farooqi et al. 1999). Besides hyperphagia and obesity, individuals with loss of function mutations in the leptin gene also have various other symptoms like delayed puberty

or failure of normal pubertal development, and lowered inflammatory response which indicates a deficient immune system (Ramachandrappa and Farooqi 2011).

Mutations in the leptin receptor gene cannot be treated with leptin supplementation as the receptor does not react to its substrate. Mutation carriers in the leptin receptor are often hyperleptinemic. Here, the adipocytes produce more leptin to signal increased fat mass, but do not receive negative feedback as the leptin signal is not received in the hypothalamus (Farooqi et al. 2007). Up to 3% of patients with extreme obesity harbor loss of function mutations in the leptin receptor which potentially contribute to the phenotype (Farooqi et al. 2007).

1.2.2.2 Melanocortin-4 receptor (MC4R)

The melanocortin-4 receptor (MC4R) is downstream of leptin in the leptinergic pathway (Hinney et al. 2013). For the MC4R, major- and polygenic effects on obesity are known. About 2.5% of extremely obese children and adolescents carry mutations in the *MC4R* gene (Hinney et al. 1999, Hinney et al. 2003), while the mutation frequency is lower in obese adults (0.6%; Hinney et al. 2006). MC4R mutations follow a co-dominant mode of inheritance. A mutation leading to a reduced function in just a single copy of the gene already leads to an increased body weight, if both alleles harbor a mutation the effect on body weight is elevated further. Today, 166 mutations and variants in the *MC4R* are known. Most of the variants that show functional reduction or loss of function *in vitro* affect MC4R expression or trafficking of the receptor to the cell surface (Hinney et al. 2013). The quantitative effect of these variants is gender specific; female mutation carriers are approximately 30 kg heavier while male mutation carriers only weigh approximately 15 kg more than unaffected family members (Dempfle et al. 2004).

In contrast to these major gene effects, MC4R also harbors weight lowering alleles. The two variant alleles of Val103Ile (Geller et al. 2004, Wang et al. 2005, Heid et al. 2005) and Ile251Leu (Stutzmann et al. 2007) lead to improved receptor signalling; carriers of the polygenic allele show a by 30% decreased risk to develop obesity. The weight of a heterozygous, adult male is reduced by approximately 1.5kg compared to wild type carriers (Geller et al. 2004; Young et al. 2007; Stutzmann et al. 2007).

1.2.3 Genome wide association studies and polygenic obesity

Genome wide association studies (GWAS) identified 61 obesity associated loci to date (published in peer-reviewed journals; Table 3). GWAS compare the allele frequencies of single nucleotide polymorphisms (SNPs) between (e.g. normal weight) controls and (e.g. obese) cases. A higher allele frequency in normal weight individuals shows that the allele is protective against obesity, while a higher allele frequency in obese cases defines it as the

obesity risk allele of the SNP. SNPs with genome wide significant association (significance level $p=5*10^{-8}$) are called lead SNPs and mark genetic loci associated to an analyzed trait like obesity (Kwon und Goate 2000, Attia et al. 2009a-c). These lead SNPs are usually non-coding and can cover large genetic regions with several genes. All in all, between 1-2 mio SNPs are used to analyze the whole human genome.

1.2.3.1 BMI/obesity loci in adults

GWAS were used to identify BMI associated loci already in 2006, where the *INSIG2* (insulin induced protein 2) locus was identified (Herbert et al. 2006). While many genetic loci were robustly replicated (see Table 3), other loci like *INSIG2* could not be replicated in independent study groups. In 2010, the „Genetic Investigation of Antropomorphic Traits“ (GIANT) consortium published a GWAS meta analysis of BMI associated variants. They initially analyzed GWAS data on 123,865 individuals and confirmed their findings in 125,931 independent individuals. They confirmed 14 known BMI and waist-hip-ratio (WHR) loci and identified 18 new variants (Speliotes et al. 2010). While *FTO* and *MC4R* were already known obesity genes/loci (Loos et al. 2007, Hinney et al. 2007, Thorleifsson et al. 2009, Willer et al. 2009), many other loci were detected for which the causal gene had not yet been identified. Altogether, the risk alleles at the 32 BMI associated loci have small effects on body weight; each risk allele leads to an approximate gain of 0.17 kg/m² which equals about 453-551 g increase in an individual with a height of 1.6-1.8 m, respectively (Speliotes et al. 2010).

Recently, the same group published preliminary data on a larger set of 236,231 individuals for the initial meta-analysis and 103,046 individuals for the confirmation (Berndt et al. 2012). They reportedly confirmed 31 of the 32 BMI associated loci from their previous work (*ZNF608* was not replicated) and identified 67 new independent loci (separated by at least 500 kb). They also found independent secondary signals for the five loci *MC4R*, brain derived neurotrophic factor (*BDNF*), G protein-coupled receptor, family C, group 5, member B (*GPRC5B*), Fanconi anemia, complementation group L (*FANCL*), and adenylate cyclase 9 (*ADCY9*). Many of the detected loci contain genes related to glucose and insulin homeostasis (e.g., transcription factor 7-like 2 (*TCF7L2*), gastric inhibitory polypeptide receptor *(GIPR*), ADP-dependent glucokinase *(ADPGK*)) or lipid metabolism (e.g., the apolipoprotein gene cluster, Niemann-Pick disease, type C1 gene (*NPC1*)). They also found genetic loci involved in mitochondrial processes (e.g., mitochondrial translational initiation factor 3 *(MTIF3*), pyruvate dehydrogenase kinase, isozyme 4 *(PDK4*)), and hormone regulation (hydroxysteroid (17-beta) dehydrogenase 12 (*HSD17B12);* Berndt et al. 2012).

Table 3: GWAS derived BMI associated loci (in order of identification, modified from Sandholt et al. 2011).

Chr	SNP*	Regional gene(s)	Gene name	Effect Allele	Effect Allele Frequency	Effect size BMI (kg/m²)	Ethnicities	Age	Identification (bold) and Replication of Locus
2	rs7566605	*INSIG2*	insulin induced protein 2	C	0.37	1.00	Caucasian, African	adult, children	**Herbert et al. 2006**, Lui et al. 2008
16	rs9939609	*FTO*	fat mass and obesity associated gene	A	0.45	0.36	Caucasian, African, Asian	adult, children	**Frayling et al. 2007**, Hinney et al. 2007, Scuteri et al. 2007, Willer et al. 2009, Thorleifsson et al. 2009, Meyre et al. 2009, Speliotes et al. 2010, Scherag et al. 2010, Bradfield et al. 2012, Wen et al. 2012, Okada et al. 2012, Monda et al. 2013
10	rs6602024	*PFKP*	phosphofructokinase, platelet	A	0.10	0.84	Caucasian	adult	**Scuteri et al. 2007**, Lui et al. 2008
20	rs6013029	*CTNNBL1*	catenin β-like 1	T	0.05	0.12	Caucasian	adult	**Lui et al. 2008**
18	rs17782313	*MC4R*	melanocortin 4 receptor		0.24	0.22	Caucasian, Asian	adult, children	**Loos et al. 2008,** Willer et al. 2009, Thorleifsson et al. 2009, Meyre et al. 2009, Speliotes et al. 2010, Scherag et al. 2010, Bradfield et al. 2012, Wen et al. 2012, Okada et al. 2012, Monda et al. 2013
2	rs6548238	*TMEM18*	transmembrane protein 18	G	0.84	0.26	Caucasian, African	adult, children	**Willer et al. 2009,** Thorleifsson et al. 2009, Speliotes et al. 2010, Scherag et al. 2010, Bradfield et al. 2012

-Table 3 continued-

-Table 3 continued-

Chr	SNP*	Regional gene(s)	Gene name	Effect Allele	Effect Allele Frequency	Effect size BMI (kg/m²)	Ethnicities	Age	Identification (bold) and Replication of Locus
16	rs7498665	*SH2B1*	SH2B adaptor protein 2 isoform 1	G	0.41	0.15	Caucasian, African	adult	**Willer et al. 2009,** Thorleifsson et al. 2009, Speliotes et al. 2010
19	rs11084753	*KCTD15*	potassium channel tetramerisation domain	C	0.67	0.06	Caucasian, African	adult	**Willer et al. 2009,** Thorleifsson et al. 2009, Speliotes et al. 2010
1	rs2815752	*NEGR1*	neuronal growth regulator 1	G	0.62	0.10	Caucasian, African	adult	**Willer et al. 2009,** Thorleifsson et al. 2009, Speliotes et al. 2010
4	rs10938397	*GNPDA2*	glucosamine-6-phosphate deaminase 2	G	0.45	0.19	Caucasian, Asian	adult	**Willer et al. 2009**, Speliotes et al. 2010, Wen et al. 2012, Monda et al. 2013
11	rs10838738	*MTCH2*	mitochondrial carrier 2	G	0.34	0.07	Caucasian	adult	**Willer et al. 2009**, Speliotes et al. 2010
11	rs925946	*BDNF*	brain-derived neurotrophic factor	G	0.30	0.19	Caucasian, African, Asian	adult	**Thorleifson et al. 2009,** Speliotes et al. 2010, Jiao et al. 2011, Wen et al. 2012, Okada et al. 2012
1	rs10913469	*SEC16B*	leucine zipper transcription regulator 2	C	0.20	0.50	Caucasian, African, Asian	adult, children	**Thorleifson et al. 2009**, Speliotes et al. 2010, Scherag et al. 2010, Bradfield et al. 2012, Wen et al. 2012, Okada et al. 2012, Monda et al. 2013

-Table 3 continued-

-Table 3 continued-

Chr	SNP*	Regional gene(s)	Gene name	Effect Allele	Effect Allele Frequency	Effect size BMI (kg/m²)	Ethnicities	Age	Identification (bold) and Replication of Locus
12	rs7138803	*FAIM2*	Fas apoptotic inhibitory molecule 2	A	0.37	0.54	Caucasian, African	adult, children	**Thorleifson et al. 2009,** Speliotes et al. 2010, Bradfield et al. 2012
3	rs7647305	*ETV5*	ets variant gene 5	C	0.77	0.54	Caucasian, African	adult	**Thorleifson et al. 2009,** Speliotes et al. 2010
18	rs1805081	*NPC1*	Niemann-Pick disease, type C1 precursor	A	0.44	−0.06	Caucasian	adult	**Meyre et al. 2009**
16	rs1424233	*MAF*	v-maf musculoaponeurotic fibrosarcoma oncogene	A	0.43	0.03	Caucasian	adult	**Meyre et al. 2009**
10	rs10508503	*PTER*	phosphotriesterase related	C	0.09	0.02	Caucasian	adult	**Meyre et al. 2009**
6	rs4712652	*PRL*	prolactin	A	0.41	−0.08	Caucasian	adult	**Meyre et al. 2009**
2	rs713586	*POMC*	proopiomelanocortin	C	0.47	0.14	Caucasian, Asian	adult, children	**Speliotes et al. 2010**, Bradfield et al. 2012, Wen et al. 2012
16	rs12444979	*GPRC5B*	G protein-coupled receptor, family C, group 5,	C	0.87	0.17	Caucasian	adult	**Speliotes et al. 2010**
15	rs2241423	*MAP2K5*	mitogen-activated protein kinase kinase 5	G	0.78	0.13	Caucasian, Asian	adult	**Speliotes et al. 2010**, Wen et al. 2012
19	rs2287019	*GIPR*	gastric inhibitory polypeptide receptor	C	0.80	0.15	Caucasian, Asian	adult	**Speliotes et al. 2010**, Wen et al. 2012, Okada et al. 2012
1	rs1514175	*TNNI3K*	TNNI3 interacting kinase	A	0.43	0.07	Caucasian	adult	**Speliotes et al. 2010**
4	rs13107325	*SLC39A8*	solute carrier family 39 (zinc transpor ter) member 8	T	0.07	0.19	Caucasian	adult	**Speliotes et al. 2010**
5	rs2112347	*HMGCR*	3-hydroxy-3-methylglutaryl-Coenzyme A reductase	T	0.63	0.10	Caucasian	adult	**Speliotes et al. 2010**

-Table 3 continued-

-Table 3 continued-

Chr	SNP*	Regional gene(s)	Gene name	Effect Allele	Effect Allele Frequency	Effect size BMI (kg/m²)	Ethnicities	Age	Identification (bold) and Replication of Locus
9	rs10968576	*LRRN6C*	leucine rich repeat and Ig domain co ntaining 2	G	0.31	0.11	Caucasian	adult	**Speliotes et al. 2010**
19	rs3810291	*TMEM160*	transmembrane protein 160	A	0.67	0.09	Caucasian	adult	**Speliotes et al. 2010**
2	rs887912	*FANCL*	Fanconi anemia, complementation gr oup L isoform	T	0.29	0.10	Caucasian	adult	**Speliotes et al. 2010**
3	rs13078807	*CADM2*	immunoglobulin superfamily, member 4D	G	0.20	0.10	Caucasian	adult	**Speliotes et al. 2010**
14	rs11847697	*PRKD1*	protein kinase D1	T	0.04	0.17	Caucasian	adult	**Speliotes et al. 2010**
2	rs2890652	*LRP1B*	low density lipoprotein-related protein 1B	C	0.18	0.09	Caucasian	adult	**Speliotes et al. 2010**
1	rs1555543	*PTBP2*	polypyrimidine tract binding protein 2	C	0.59	0.06	Caucasian	adult	**Speliotes et al. 2010**
13	rs4771122	*MTIF3*	mitochondrial translational initiation f actor 3	G	0.24	0.09	Caucasian	adult	**Speliotes et al. 2010**
5	rs4836133	*ZNF608*	zinc finger protein 608	A	0.48	0.07	Caucasian	adult	**Speliotes et al. 2010**
11	rs4929949	*TUB*	tubby gene	C	0.52	0.06	Caucasian	adult	**Speliotes et al. 2010**
6	rs206936	*NUDT3*	nudix-type motif 3	G	0.21	0.06	Caucasian	adult	**Speliotes et al. 2010**
14	rs10150332	*NRXN3*	neurexin 3	C	0.21	0.13	Caucasian	adult	**Speliotes et al. 2010**
6	rs987237	*TFAP2B*	transcription factor AP-2 beta	G	0.18	0.13	Caucasian, Asian	adult	**Speliotes et al. 2010**, Wen et al. 2012
8	rs17150703	*TNKS, MSRA*	tankyrase, methionine sulfoxide reductase A	A	0.10	−0.10	Caucasian	children and adolescents	**Scherag et al. 2010**
1	rs12145833	*SDCCAG8*	serologically defined colon cancer antigen 8	T	0.87	0.05	Caucasian	children and adolescents	**Scherag et al. 2010**
10	rs2116830	*KCNMA1*	large conductance calcium-activated potassium channel	G	0.80	1.00	Caucasian	adult	**Jiao et al. 2011**

-Table 3 continued-

-Table 3 continued-

Chr	SNP*	Regional gene(s)	Gene name	Effect Allele	Effect Allele Frequency	Effect size BMI (kg/m²)	Ethnicities	Age	Identification (bold) and Replication of Locus
13	rs9568856	*OLFM4*	olfactomedin 4	A	0.16	1.22# (1.14–1.29)	Caucasian	children and adolescents	**Bradfield et al. 2012**
17	rs9299	*HOX5B*	homeobox B5	T	0.65	1.14# (1.09–1.20)	Caucasian	children and adolescents	**Bradfield et al. 2012**
6	rs9356744	*CDKAL1*	CDK5 regulatory subunit associated protein 1-like 1	T	0.58	0.34	Asian	adult	**Wen et al. 2012**, Okada et al. 2012
5	rs261967	*PCSK1*	proprotein convertase subtilisin/kexin type 1	C	0.41	0.38	Asian	adult	**Wen et al. 2012**
16	rs12597579	*GP2*	glycoprotein 2	C	0.80	0.41	Asian	adult	**Wen et al. 2012**
9	rs11142387	*KLF9*	Kruppel-like factor 9	C	0.46	0.48	Asian	adult	**Okada et al. 2012**
3	rs6794092	*TMEM212*	transmembrane protein 212	G	0.90	0.27	African	adult	**Ng et al. 2012**
5	rs268972	*CDH12*	cadherin 12, type 2	C	0.66	0.15	African	adult	**Ng et al. 2012**
20	rs2033195	*FER1L4*	fer-1-like 4	C	0.57	0.15	African	adult	**Ng et al. 2012**
5	rs6088887	*MFAP3, GALNT10*	microfibrillar-associated protein 3 - UDP-N-acetyl-alpha-D-galactosamine:polypeptide, N-acetylgalactosaminyltransferase 10	G	0.79	0.18	African	adult	**Ng et al. 2012,** Monda et al. 2013
13	rs7989336	*HS6ST3*	heparan sulfate 6-O-sulfotransferase 3	A	0.47	1.10#	Caucasian	adult	**Berndt et al. 2013**
1	rs17381664	*ZZZ3*	zinc finger, ZZ-type containing 3	C	0.39	1.09#	Caucasian	adult	**Berndt et al. 2013**
1	rs17024258	*GNAT2*	guanine nucleotide binding protein alpha	T	0.04	1.25#	Caucasian	adult	**Berndt et al. 2013**
8	rs4735692	*HNF4G*	hepatocyte nuclear factor 4 gamma	A	0.58	1.06#	Caucasian	adult	**Berndt et al. 2013**
20	rs13041126	*MRPS33P4*	Mitochondrial Ribosomal Protein S33 Pseudogene 4	T	0.72	1.06#	Caucasian	adult	**Berndt et al. 2013**

-Table 3 continued-

-Table 3 continued

Chr	SNP*	Regional gene(s)	Gene name	Effect Allele	Effect Allele Frequency	Effect size BMI (kg/m²)	Ethnicities	Age	Identification (bold) and Replication of Locus
16	rs2531995	*ADCY9*	adenylate cyclase 9	T	0.61	1.07#	Caucasian	adult	**Berndt et al. 2013,** Monda et al. 2013
17	rs7503807	*RPTOR*	raptor	A	0.57	1.04#	Caucasian	adult	**Berndt et al. 2013**
6	rs974417	*KLHL32*	kelch-like 32	C	0.66	0.31	African	adult	**Monda et al. 2013**
7	rs10261878	*MIR148A-NFE2L3*	microRNA 148a, nuclear factor (erythroid-derived 2)-like 3	A	0.44	0.32	African	adult	**Monda et al. 2013**

*Only the lead SNP of the discovery study is listed here, for replicated loci the lead SNPs often vary. Independent loci within the same chromosomal region are not listed separately.
Odd's ratio

To reduce the amount of missing heritability of the phenotype obesity, studies also focus on the extremes of the weight distribution. When analyzing individuals with a weight in the highest and lowest 5% of the population, 7 loci for extreme obesity (hepatocyte nuclear factor 4 gamma (*HNF4G*), raptor (*RPTOR*), guanine nucleotide binding protein alpha (*GNAT2*), Mitochondrial Ribosomal Protein S33 Pseudogene 4 (*MRPS33P4*), *ADCY9*, heparan sulfate 6-O-sulfotransferase 3 (*HS6ST3*), and zinc finger, ZZ-type containing 3 (*ZZZ3*)) were identified (Berndt et al. 2013). Again, some of these genes are involved in lipid metabolism (*ADCY9*) and mitochondrial processes (*MRPS33P4*).

Many BMI associated loci replicate well in other ethnicities. Wen et al. (2012) replicated seven previously identified loci (*FTO*, SEC16 homolog B (*SEC16B*)*, MC4R*, *GIPR,* adenylate cyclase 3 (*ADCY3*), *BDNF*, and mitogen-activated protein kinase kinase 5 (*MAP2K5)*) and identified three novel loci in or near the CDK5 regulatory subunit associated protein 1-like 1 (*CDKAL1*), proprotein convertase subtilisin/kexin type 1 (*PCSK1)*, and glycoprotein 2 (*GP2*) genes when analyzing 27,715 East Asians and following the SNPs up in 37,691 and 17,642 additional East Asians.

Another BMI GWAS meta-analysis by Okada et al. (2012) replicated the European BMI loci at *SEC16B*, *BDNF*, *FTO*, *MC4R* and *GIPR* and the Asian *CDKAL1* locus from Wen et al. (2012). Additionally, they detected a new association signal at the Kruppel-like factor 9 (*KLF9*) locus which could not be replicated in the European cohorts (Speliotes et al. 2010).

GWAS (meta-) analyses in African or African-American populations also replicate known BMI loci from European populations (*FTO*, *SEC16B, MC4R*, *GNPDA2* and the recently identified extremes locus *ADCY9*; Monda et al. 2013). On the other hand, the locus harboring the genes encoding microfibrillar-associated protein 3 (*MFAP3)* and UDP-N-acetyl-alpha-D-galactosamine polypeptide, N-acetylgalactosaminyltransferase 10 (*GALNT10),* which is robustly associated with obesity in African populations (Ng et al. 2012, Monda et al. 2013), showed no association in Caucasian BMI GWAS (Speliotes et al. 2010). Other BMI associated loci in African populations vary between studies; like the loci harboring transmembrane protein 212 (*TMEM212*), cadherin 12 (*CDH12)* and fer-1-like 4 (*FER1L4*) which were exclusively detected by Ng et al. (2012) and the locus harboring kelch-like 32 (*KLHL32)* and the combined locus near microRNA 148a (*MIR148A*) and nuclear factor (erythroid-derived 2)-like 3 (*NFE2L3*; Monda et al. 2013). All in all, while many regulators of BMI and weight gain may be comparable across ethnicities, each ethnicity also comprises independent association signals.

1.2.3.2 BMI/obesity loci in extremely obese children and adolescents

In children and adolescents, additional loci play a role in weight regulation. Scherag et al. (2010) identified serologically defined colon cancer antigen 8 (*SDCCAG8*) and a locus

between tankyrase (*TNKS*) and methionine sulfoxide reductase A (*MSRA*) as associated with early-onset obesity. SDCCAG8 is expressed in the gut and displays anti cancer capacity (Kenedy et al. 2003). It also plays a role in development of the Bardet-Biedl syndrome which is a syndromic form of obesity (Billingsley 2012, Schaefer et al. 2011). The locus seems to be specific for early-onset obesity as variants in *SDCCAG8* were associated with reduced weight loss in 401 children participating in the lifestyle intervention 'Obeldicks' but not in 626 adults in a comparable lifestyle intervention (Scherag et al. 2012).

A recent GWAS meta-analysis on 5,530 obese cases (BMI ≥ 95th percentile) and 8,318 controls (BMI <50th percentile) detected eight previously unknown loci of which 2 were confirmed in further 2,818 cases and 4,083 controls (Bradfield et al. 2012). The two new confirmed loci associated with early-onset obesity are near the genes olfactomedin 4 (*OLFM4*) and homeobox B5 (*HOXB5*), which both also show (not genome-wide significant) BMI association in adults (Speliotes et al. 2010). Again, the *FTO* and *MC4R* loci were confirmed in this study (Bradfield et al. 2012). Hence, common genetic mechanisms are relevant for body weight regulation throughout lifetime.

1.2.4 Copy number variation (CNV)

Besides SNPs, common genetic polymorphisms that affect only one nucleotide position, other genetic variations can also show association with diseases and traits. Recently, copy number variations (CNVs) came into the focus of genetic obesity research. CNVs are structural genetic variations of at least 1kb length, whose number is deviant from the original two copies either by deletion or duplication/multiplication (Stankiewicz and Lupski 2010). These genetic regions of up to several megabases of DNA can harbor genes or regulatory regions and therefore change the amount of protein transcribed.

Jarick et al. (2011) detected 20 common copy number variable regions (CNVRs) that are associated to early-onset obesity in two German study groups (424 case-parents obesity trios and an independent sample of 453 extremely obese children and adolescents and 435 normal-weight and lean adult controls). Of these, for only one region BMI association was previously unknown; this region covers several olfactory genes.

A rare deletion of 200-500 kbps of chr16p11.2 was associated with severe early-onset obesity in children with developmental delay. Depending on the exact position of the deletion, the CNVs were associated with (i) cognitive deficits and early onset overweight or obesity (Walters et al. 2010), (ii) developmental delay of varying severity and obesity (Bachman-Gagescu et al. 2010), (iii) mild developmental delay, hyperphagia, insulin resistance and obesity (Bochukova et al. 2010) or (iv) severe early-onset obesity as the only symptom (Jarick et al. 2011). In the light of these observations, the chromosomal region 16p11.2 is regarded to be associated with a rare syndromal form of obesity with additional symptoms

like developmental delay or autism (Shinawi et al. 2009, Perrone et al. 2010). The corresponding reciprocal duplication of the same region is associated with underweight (Jacquemont et al. 2011). Individuals with the duplication also show an unusually high frequency of selective and restrictive eating behaviors, as well as developmental or intellectual disabilities or psychiatric disorders (Jacquemont et al. 2011).

1.2.5 Missing heritability

Although the use of larger study groups for GWAS analyses and the incorporation of CNVs and other markers of inheritance increased the knowledge about genetic influences on the variance of BMI, most of the 50% BMI heritability expected from twin-, adoption and family studies remains unexplained (Hebebrand et al. 2010). Even including BMI association signals not reaching genome-wide significance in current GWAS studies, a maximum of 17% of the genetic BMI inheritance or 8.5% of the total BMI variance can be explained (Yang et al. 2011).

Several potential explanations for this missing or hidden heritability are currently discussed. (i) BMI is a non-specific quantitative phenotype and both measures, height and weight, one inherited differentially from each other. As explained above, BMI is an accurate measure for adiposity only in the upper weight range and within a specific height range (Revicki and Israel 1986). Also, both measures vary over the course of time which leads to systematic errors. Height decreases during the day and lifespan, while weight fluctuates with seasons and increases with age (Watson et al. 1979, Visscher and Seidell 2004). Comparing BMI of persons at age 10 with the BMI of the same individuals at age 35 results in a correlation of only 0.5, while the BMI at age 18 gives a better estimate for the BMI later in life (0.75; Guo et al. 1994).

(ii) The BMI effect of each variant by itself is small. However, the combined effect is substantial. Increasing the sample size from GWAS data on 123,865 individuals (Speliotes et al. 2010) to 339,277 individuals for the GIANT consortium studies led to the confirmation and discovery of 98 BMI associated loci which still only explain 4.5% of the inherited variance of BMI (Berndt et al. 2012). Further increases of the sample size will probably lead to the discovery of more BMI associated variants, but with further decreases of effect sizes for every single variant.

(iii) Another problem of GWAS studies is the use of common SNPs to tag large areas of the genome. Infrequent to rare variants with potential large effect sizes will not be detected with the used setup of BMI association studies, as they are only partially inherited with the common lead SNPs (Speliotes et al. 2010). In addition, potentially antagonistic effects will mask each other when tagged by the same common SNP. This effect could also diminish the

effect of GWAS lead SNPs if these are not corrected for independent signals at the same locus.

(iv) Heritability might be overestimated from twin, family and adoption studies. Here, the estimates ranged from mostly 40-70% with twin studies giving the highest estimates and family studies consistently giving lower heritability estimates. For twin studies, alternative hypotheses like shared intrauterine environment, a higher dependency and a more equal treatment of monozygotic than dizygotic twins could contribute to an overestimated heritability (Farber 1981, Eysenck 1981). Some family studies calculated inheritance estimates as low as 20% (Maes et al. 1997) or even 5% (Bouchard et al. 1993).

(v) There are genetic mechanisms that cannot be analyzed by the current GWAS methods. Gene environment interactions potentially play a large role in weight regulation, as well as epigenetic changes in the DNA that lead to changes in gene expression. For example, CNVs are estimated to explain up to 18% of the missing heritability (Bochukova et al. 2010).

1.3 Chromosomal Region 16p11.2

The currently published largest GWAS meta-analysis and replication study for body weight variation identified 32 obesity associated loci. One of the SNPs (rs7359397) is located in the chromosomal region 16p11.2 near the SH2B adapter protein gene (*SH2B1*; Speliotes et al. 2010). An effect size of 0.15 BMI (kg/m²) units for the tenth best hit rs7359397 was described, translating into a weight increase of approx. 577 g per risk allele for an individual of 1.8 m of body height. Association with obesity had also been shown for a non-synonymous SNP in *SH2B1* (rs7498665: g.8164A/G [Thr484Ala]; GWAS: Thorleifsson et al. 2009, Willer et al. 2009) with high linkage disequilibrium (LD) with rs7359397 (r^2=0.965, D'=1; HapMap, http://hapmap.ncbi.nlm.nih.gov/). Speliotes et al. (2010) also list other coding SNPs in high LD with the lead SNP rs7359397 in chr.16p11.2 (supplementary materials), including variants in the genes apolipoprotein B48 receptor (*APOB48R*) and sulfotransferase family, cytosolic, 1A, member 2 (*SULT1A2*; see Figure 1).

The association of the risk allele G at rs7498665 which tags the entire chromosomal region from chr16:28,177,800 to chr16:28,944,400 has been robustly replicated in (i) 4,923 Swedish adults (Renström et al. 2009), (ii) in 4,157 Chinese women (Shi et al. 2010), (iii) in 18,264 Japanese adults (Takeuchi et al. 2011), (iv) in 12,462 individuals from the German MONICA/KORA study (Holzapfel et al. 2010), (v) in 1,045 obese adults and 317 healthy lean individuals from Belgium (Beckers et al. 2011), (vi) in 4,992 subjects from six African-American cohorts (Hester et al. 2011), (vii) 8,050 nondiabetic French individuals (Robiou-du-Pont et al. 2012), and (viii) 2,537 individuals from a longitudinal British birth cohort (Elks et al. 2012b).

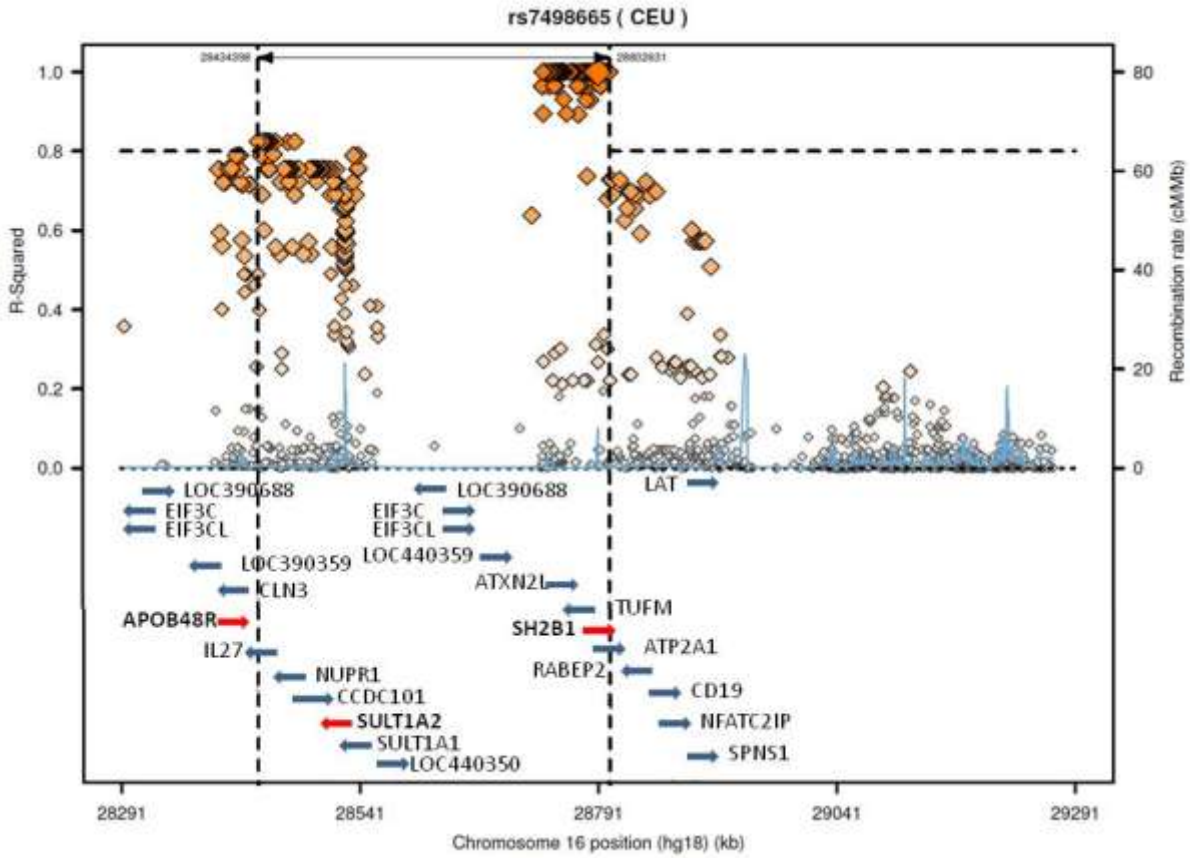

Figure 1: Regional association and linkage disequilibrium plot of data from the 1000 genome project data centered to SNP rs7498665 (http://www.broadinstitute.org/mpg/snap/; adapted from Volckmar et al. 2012). Displayed are recombination rate (blue), r² to rs7498665 (range of orange, increased intensity shows higher linkage disequilibrium) and genes in this region. Dashed lines mark a region in high LD ($r^2 > 0.8$) with rs7498665. The GWAS lead SNPs rs7498665 and rs7359397 are in close proximity (distance 2.4 kb) and represent the same association signal ($r^2=0.965$, D'=1). The genes that were analyzed in this thesis are marked with red arrows. Gene abbreviations: *EIF3CL/EIF3C* (eukaryotic translation initiation factor 3), *CLN3* (ceroid-lipofuscinosis, neuronal 3), *APOB48R* (apolipoprotein B48 receptor), *IL27* (interleukin 27), *NUPR1* (p8 protein isoform a), *CCDC101* (coiled-coil domain containing 101), *SULT1A1* (sulfotransferase family, cytosolic, 1A, member 1), *SULT1A2* (sulfotransferase family, cytosolic, 1A, member 2), *ATXN2L* (ataxin 2 related protein isoform C), *TUFM* (Tu translation elongation factor, mitochondrial), *SH2B1* (SH2B adaptor protein 1 isoform 1), *ATP2A1* (ATPase, Ca++ transporting, fast twitch 1 isoform), *RABEP2* (rabaptin, RABGTPase binding effector protein 2), *CD19* (CD19 antigen precursor), *NFATC2IP* (Nuclear factor of activated T-cells, cytoplasmic 2-interacting protein), *SPNS1* (spinster homolog 1 isoform 1), *LAT* (linker for activation of T cells isoform b).

As mentioned above, additional obesity association signals for the genetic region chr16p11.2 come from CNV analysis, where carriers of a deletion in this region are obese (Bochukova et al. 2010, Walters et al. 2010, Perrone et al. 2010, Bachmann-Gagescu et al. 2010, Jarick et al. 2011) and duplication carriers are underweight (Jacquemont et al. 2011). These large genetic variants are relatively common for chr16p11.2, because the complete region was duplicated in the human genome (Loftus et al. 1999). The high amount of analogous

sequence from the ancient duplication makes the chromosomal region prone to replication errors (Martin et al. 2004, Mefford and Eichler 2009).

The duplication leads to large protein families with high sequence similarities like the sulfotranferases (SULTs) or the Src-homology 2 domain containing binding protein (SH2B) gene family located in a cluster. These duplicated genes present a good target for evolution as several genes originally have the same function and therefore underlie lower genetic conservation. In case of the SULT family, SULT1A1 and SULT1A2 have the same function, but are differentially expressed (Freimuth et al. 2004, Glatt et al. 2001). Proteins of the SH2B family can form heterodimers which show reduced function (SH2B1 and SH2B2beta: Li et al. 2007; SH2B3 and SH2B1beta: Wang et al. 2011).

1.3.1 Obesity candidate gene *SH2B1*

SH2B1 is a likely obesity gene as suggested by studies in humans and animal models. In humans, variants (namely the risk allele G at rs7498665) in the SH2B1 protein lead to an increase in whole body fat mass and serum leptin levels in females (Jamshidi et al. 2007). The influence of SH2B1 variants on the distribution of body fat and the amount of visceral adipose tissue is still under discussion. Hotta et al. (2011) found association for the G allele of rs7498665 to an increased amount of visceral fat area (VFA) determined by computer tomography. Haupt et al. (2010) reported association of the same allele with increased visceral adipose tissue (VAT) mass, as assessed with magnetic resonance tomography. In contrast, Renström et al. (2009) detected no association of rs7498665 alleles with total or abdominal adipose mass determined by DEXA scans. In addition to the above mentioned BMI association, SH2B1 variants have also been associated with type 2 diabetes independently of BMI (Sandtholt et al. 2011).

The tyrosine kinase receptor adapter protein SH2B1 is part of pathways mediated by Janus kinase (JAK; Rui et al. 1997, O'Brien et al., 2002) and several other receptor tyrosine kinases like the (i) insulin-like growth factor I receptor (IGF1R; Carter-Su et al. 2000), (ii) insulin receptor (INSR see Figure 2; Riedel et al. 1997, Kotani et al. 1998), (iii) leptin receptor (LEPRb see Figure 2; Ren et al. 2005) and (iv) brain-derived neurotrophic factor receptor (BDNFR), which take part in energy homeostasis (Hinney et al. 2010). SH2B1 forms a complex with insulin receptor substrates (IRS1 and 2) which increase signalling activity in downstream pathways, like insulin (Morris et al. 2009), via the insulin/IGF1 receptor-Akt-Foxo1-PPARγ pathway (Yoshiga et al. 2007). While IGF1 participates in the regulation of adipose tissue growth and differentiation of pre-adipocytes to their matured form, insulin regulates energy functions like the lipid or glucose metabolism in the cell (Tseng et al. 2005). Sh2b1 knockout mice (complete, regional or functionally limited to mutations in the SH2 and PH domains) are obese, hyperphagic, hyperlipidemic, hyperleptinaemic, hyperglycaemic,

glucose intolerant and insulin resistant (Ren et al. 2005, 2007, Morris et al. 2010). Selective neuronal rescue eliminated both obesity and the insulin resistance phenotype (Morris et al. 2010). Both SH2 and PH domain seem to be required for maintenance of normal glucose metabolism and body weight. Although there is a striking likelihood for the neuronal repression of the leptin receptor (de Luca et al. 2005, Kowalski et al. 2001), the Sh2b1 knockdown mice showed no complete leptin resistance. Hence it is unlikely that SH2B1 is the only moderator in this pathway. An interaction with the cytokine signalling 3 suppression of leptin receptor might explain the obesity protection from over expression of neuronal Sh2b1 (Chua 2009).

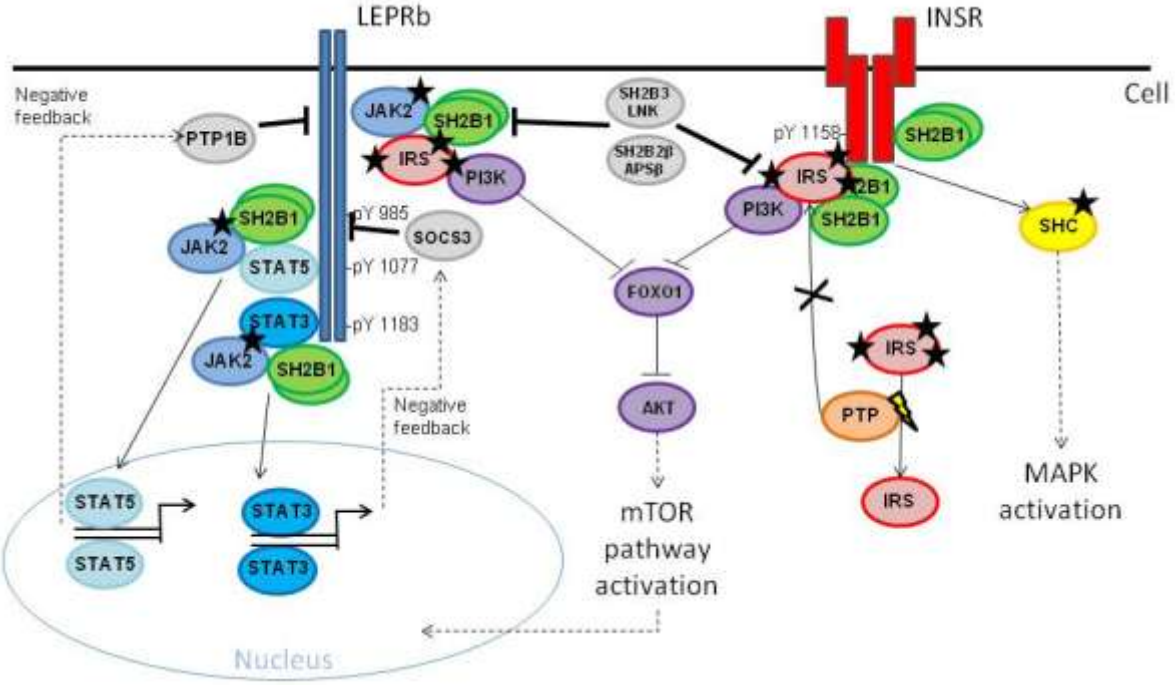

Figure 2: Overview of the functional interactions of SH2B1 (green) in leptin- (LEPRb; left) and insulin receptor (INSR; right) signalling. In leptin receptor signalling, SH2B1 dimers mainly increase the autophosphorylation of JAK2 which increases signalling of the leptin receptor via STAT3 and STAT5. Negative feedback is given by PTP1B and SOCS3. In insulin receptor signalling, SH2B1 binds directly to the receptor and increases INSR phosphorylation which in turn increases downstream signalling of Src homology 2 domain containing protein (SHC). Additionally, SH2B1 dimers increase insulin receptor substrate 1 and 2 (IRS) phosphorylation (marked by stars) which cannot be attenuated by protein tyrosine phosphatase (PTP) when SH2B1 is bound to IRS. IRS can also be phosphorylated by a complex of an SH2B1 dimer with JAK2. In both cases, IRS phosphorylation leads to activation of the mTOR pathway by phosphoinositol-3 kinase (PI3K), forkhead box O1 (FOXO1) and AKT1 kinase (AKT). SH2B2β (SH2B adaptor protein 2 form 2, also APSβ) and SH2B3 (SH2B adaptor protein 2 form 3, also LNK) can both dimerize with SH2B1 and negatively regulate its function (adapted from Morris et al. 2009, Morris and Rui 2010).

The evolutionary conservation of SH2B1 function was analyzed in the knockout model of dSh2b in *Drosophila melanogaster*. While the regulation of energy homeostasis via lipid and glucose has been consistent for the ubiquitously expressed dSh2b, neuronal dSh2b played

only a minor role for modulation of haemolymph carbohydrate levels and whole body lipid levels. Therefore, the role of neuronal SH2B1 must have gained impact on metabolic traits during evolution from insect to mammals, while the basic function of energy homeostasis regulation is highly conserved (Song et al. 2010).

Humans express three different splice forms of SH2B1- alpha, beta and gamma (Nishi et al. 2005, see Figure 2) – while mice have an additional forth splice variant delta (Riedel et al. 1997, Rui et al. 1997, Nelms et al. 1999, Yousaf et al. 2001). Recently, Doche et al. (2012) presented data indicating a fourth splice variant delta in humans. They analyzed differential expression of all four SH2B1 variants in 24 human tissues (whole brain, hypothalamus, hippocampus, cerebellum, cerebral cortex, salivary gland, trachea, thyroid gland, thymus, lung, small intestine, colon, pancreas, spleen, liver, kidney, heart, sceletal muscle, white adipose tissue, adrenal gland, prostate, testis, uterus, and placenta) and fetal brain and liver. Splice variants beta and gamma were ubiquitously expressed, while the expression of alpha and delta limited to the analyzed (fetal) brain tissues (Doche et al. 2012). This is in contrast to the findings in animals, where differential expression is reported. All splice variants contain three domains with different functions. The dimerization domain is located close to the amino terminal end and is needed to dimerize SH2B1 localized to plasma membranes (Nishi et al. 2005, Rui et al. 1999a). The central pleckstrin homology (PH) domain may serve to bind SH2B1 and its closely related adapter protein APS to cellular phosphatidyl inositides, while they can also mediate binding to inactive JAK2 (Li et al. 2007). The highly conserved amino-terminal SH2 domain is needed for autophosphorylation and substrate phosphorylation of JAK2 by SH2B1 dimers (Rui et al. 1999b, Nishi et al. 2005).

The alpha splice variant is mainly expressed in murine brain, heart and liver but low expression was shown in all tissues except for spleen (Riedel et al. 1997). The phosphorylation of three tyrosine motifs in the N-terminal part of the alpha splice variant regulates interaction with the insulin receptor (Zhang et al. 2008). This function has also been shown for the gamma splice variant in mice (Nelms et al. 1999). The alpha splice variant constitutively interacts with GrB2 (Yousaf et al. 2001).

Main expression of the beta splice variant is found in murine hypothalamus, an important regulatory region for energy balance. The beta splice variant of SH2B1 was found to be a substrate and cytoplasmic activator of JAK2 after responding to IGF1 (Li et al. 2007). At least in mice this variant seems to be relevant for SH2B1 regulation of energy homeostasis, as knockout and rescue for this variant elicited and reversed the overall knockout phenotype (Ren et al. 2007, Morris et al. 2010). It also plays a role in cell division and migration (Lanning et al. 2011).

The gamma splice variant of SH2B1 is also expressed ubiquitously. In contrast to beta, murine gamma interacts with the insulin signalling cascade. The insulin receptor is activated by SH2B1 binding with its third tyrosine residue (at amino acid position 1158) (Morris et al. 2009). The gamma splice variant elicits the strongest increase of IGF- and platelet-derived growth factor alpha (PDGF)-mediated mitogenesis of all four SH2B1 variants (Yousaf et al. 2001, Zhang et al. 2007).

The recently detected human delta splice variant (Doche et al. 2012) presumably contains a unique nuclear localization signal in the C-terminus. In the mouse, it shares high sequence similarity with the alpha splice variant (Yousaf et al. 2001). No data on the human sequence is available yet.

1.3.2 Obesity candidate gene *APOB48R*

Another gene inherited in the same linkage block of more than 2Mb tagged by rs7359397 on chr16p11.2 is the apolipoprotein B48 receptor (*APOB48R* or *APOBR*). Speliotes et al. (2010) attributed the BMI association signal to several genes as they report coding variants in high LD (r^2>0.75) with the initial lead SNP. The SNP rs180743 (Pro419Ala) in *APOB48R* is one of the coding variants that could explain the BMI association.

The receptor for APOB48 is involved in lipid homeostasis (function reviewed in Dominiczak and Caslake 2011). However, association to obesity has not been described yet besides the data of Speliotes et al. as the BMI association of the coding common rs7498665 was attributed to *SH2B1* (2010). The macrophage receptor APOB48R binds dietary triglyceride-rich lipoproteins after activation by its substrate APOB48 (Fujita et al. 2005). APOB48R is anchored in the membrane and highly expressed in lung and placenta. The receptor is predicted to provide lipid-soluble vitamins and dietary lipids to reticuloendothelial cells like monocytes and accessible macrophages of the immune system. Increased blood lipid levels are discussed to impair receptor function. APOB48R could be involved in foam cell formation, endothelial dysfunction, and atherothrombogenesis, as the receptor ensures rapid uptake of lipoproteins with high specificity which leads to visible cellular triglyceride and cholesterol accumulation *in vivo* (Brown et al. 2000).

Expression of *APOB48R* is increased in adipose tissue macrophages in obese C57BL/6 mice fed a high fat diet. Macrophage infiltration into fat tissue is an inflammatory marker. Hence, higher lipid intake of these cells could link APOB48R and inflammation in obesity (Lumeng et al. 2007). The minor allele of the coding SNP rs180743 (Pro419Ala) shows association to hypercholesterolemia. A recent study attributes the association to hypercholesterolemia of variants in the linkage block on chr16p11.2 to other nearby genes (*SH2B1, SPNS1*) in this region (Västermark et al. 2012).

Animal knockout or over expression models for APOB48R are not yet available; the gene has no known homologue in other species besides *Mus musculus*.

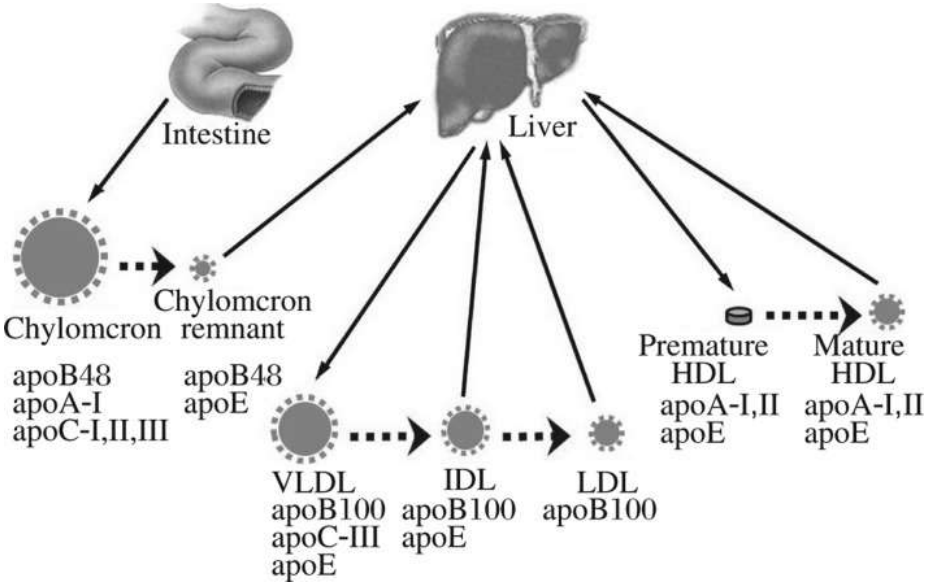

Figure 3: APOB mediated uptake of blood circulating lipids in chylomicrons. Dietary lipids taken up by gut epithelial cells are embedded into chylomicrons coated with apolipoprotein apoB48, apoA1 and apoC1-3. These chylomicrons can bind to the chylomicron remnant receptor of the hepatocytes in the liver. Here, the lipids are broken down further. The very low density lipoproteins (VLDL) are released into vesicles with APOB100 which interacts with the low density lipoprotein (LDL) receptor. These vesicles are released into the blood stream but can also be re-merged with the hepatocytes for energy storage. In the end, high density lipoproteins mature and are excreted (Jiang et al. 2013),

1.3.3 Obesity candidate gene *SULT1A2*

The sulfotransferase *SULT1A2* is another gene with a coding SNP (rs1059491, Asn235Thr) in high LD (r^2>0.75) with the lead SNP rs7359397 from the chromosomal region 16p11.2 which showed association with BMI/obesity (Speliotes et al. 2010). Sulfotransferases can be distinguished into two classes; a membrane bound family that metabolises macromolecular endogenous structures at the Golgi-apparatus and a cytosolic family that metabolizes xenobiotics and small endogenous compounds. Both groups transfer a sulpho moeity from the co-factor 5'-phosphoadenosine-3'-phosphosulfate to nucleophilic groups of xenobiotics and small endogenous compounds like hormones and neurotransmitters. The addition of a sulpho group modifies the molecule to include a sulphate, thiosulphate or sulphamate group which is ionized in the physiological pH range and thus increases water solubility and decreases their permeation through cell membranes (Glatt et al. 1994).

The gene *SULT1A2* encodes one of two phenol sulfotransferases with thermostable enzyme activity, the other being SULT1A1, a very similar member of the same gene family. Both sulfotranferases are highly similar in genetic and amino acid sequence and share all

substrates and a high promutagenic activity. Both genes are expressed in the same tissues (Freimuth et al. 2004, Glatt et al. 2001). Usually homodimers of the SULTs are formed for activation, but heterodimers between SULT1A1 and SULT1A2 are possible (Glatt et al. 2001). A pseudogene (*SULT1D2P*) of SULT1A2 was found on chr3q22.2 (Freimuth et al. 2004).

An impact of sulfotransferase action on body weight regulation could be mediated by the regulation of sex hormones by SULT1A2 substrates that modulate steroids like estrogens and androgens. After sulfonation of "estrogenic" alkylphenols and 17beta-estradiol, these hormones lose their function and are excreted from the body (Harris et al. 2000). Obesity leads to altered metabolism of hormones; for example increased serum estrogen levels have been associated with obesity (Ghose et al. 2011). Obesity increases the serum concentration of the steroid hormones estradiol (17-β-estradiol), estron and also estron sulfate, which are all substrates of SULT1A2 (Mahabir 2006, Emaus et al. 2008). Prolonged high levels of estrogen increases risk for breast carcinogenesis, as the estrogens are oxidized to form CE-semiquinones and CE-quinones. These can attach to DNA and cause mutations (Hui 2008). SULT1A2 mutations are associated with high mammographic density which is the main risk factor for breast cancer development (Ellingjord-Dale et al. 2012). Mammographic density is also associated with increased breast cancer risk (Spink et al. 2000).

Additionally, childhood obesity is associated with premature puberty onset (Wolff 1955) which in turn is associated with an increased risk of adult obesity and T2D (Rodríguez-Moran et al. 2010). Puberty is driven by hormone excretion which leads to an increased body fat percentage of up to 17% in girls which is necessary for menarche (Frisch 1987). In early puberty, weight gain in females results from changed behavior, less physical activity and altered eating habits elicited by the hormonal changes (Sallis 2000, Pate et al. 2009). These can potentially also lead to obesity in adult life (Jasik et al. 2008). A coding SNP (rs1059491 –Asn235Ser-) in SULT1A2 shows a trend for increased weight, although this result was not significant in a study group of 692 obese and normal weight individuals (Glatt et al. 2002).

Also, SULT1A2 activates the compounds 2,4-dinitrobenzylalcohol (DNBA), 1-α-hydroxyethyl)pyrene ((-)-HEP), 1-hydroxymethylpyrene (HMP), N-hydroxy-2-acetylaminofluorene (OH-AAF) and 2-hydrolyamino-5-phenylpyridine (OH-AAP), which lead to breast, prostate and hepatic cancer in rat animal models and in *in vitro* experiments (Glatt et al. 1994). The potential mode of action of SULT1A2 in breast or other cancers is therefore dual by changed excretion of hormones and by activating carcinogenic compounds.

As for APOB48R, animal knockout or over expression models for SULT1A2 are not available. The gene has orthologues in the kingdom *animalia* and *plantae*.

2. Aims of the project

The currently largest published GWAS meta-analysis for body weight identified 32 obesity-susceptibility loci (18 of them novel) in a total of 249,796 individuals of European descent. One of the BMI associated SNPs (rs7359397) is located in the chromosomal region 16p11.2 (Speliotes et al. 2010). Despite the robust BMI association of this locus, which was confirmed in many GWAS studies and by CNV findings in the region, the underlying causal variation(s) is (are) not known. The overall aim of this study was a detailed analysis of obesity candidate genes in the region chr16p11.2 including functional analyses of single variants. Specific aims were:

- **Mutation screens and association studies:**
  - A mutation screen of the coding region of three obesity candidate genes (*SH2B1* 7,726 bp, *APOB48R* 3,824 bp and *SULT1A2* 3,903 bp) in 95 extremely obese children and adolescents.
  - Obesity association analysis of the detected variants in up to 11,406 obese or overweight individuals and 4,568 controls independent of the initial screening sample.
  - Analysis of these data together with data on two more obesity candidate genes (*SULT1A1* and *TUFM*) which were screened for mutations by our group independently from this thesis (Horn 2011, Göbel in prep., Struwe in prep., co-supervised by the author of this thesis).

- **Functional assessment of detected variants:**
  - *In silico* analyses for all detected variants in *SH2B1*, *APOB48R* and *SULT1A2.*
  - *In vitro* studies of the impact of *SH2B1* variants with strong *in silico* prediction of functional changes or obesity association on leptin signaling in a human cell culture assay (Volckmar et al. 2012).

- **Analysis of impact on weight loss intervention:**
  - To analyze the impact of obesity associated coding SNPs in *SH2B1* and *APOB48R* pertaining to their effect on weight loss and related anthropometric parameters in a 1-year lifestyle intervention in children and adolescents (Volckmar et al. 2013).

3. Materials and Methods

3.1 Materials

3.1.1 Used chemicals

Table 4: List of used chemicals (alphabetical order)

Chemical	Method	Seller
100bp DNA Ladder	Gelelectrophoresis	Karl Roth GmbH, Karlsruhe, Germany
Acetic acid	SSCP	Karl Roth GmbH, Karlsruhe, Germany
Acetonitrile	dHPLC	Karl Roth GmbH, Karlsruhe, Germany
Acrylamid-Bisacrylamid (37,5:1)	SSCP	Karl Roth GmbH, Karlsruhe, Germany
Agarose	Gelelectrophoresis	Invitrogen, Karlsruhe, Germany
Ammoniumperoxidsulfate 10% (APS)	SSCP	Karl Roth GmbH, Karlsruhe, Germany
AmpliTaq® Gold (5U/µl) with according buffers	PCR	Applied Biosystems by Life Technologies, California, USA
Azotic acid	SSCP	Karl Roth GmbH, Karlsruhe, Germany
Boric acid	SSCP	Karl Roth GmbH, Karlsruhe, Germany
Bovine Serum Albumin 100x (BSA, 100µg/ml)	SSCP	New England Biolabs, Frankfurt, Germany
Bromphenol blue	Gelelectrophoresis	Karl Roth GmbH, Karlsruhe, Germany
Calcium chloride	STAT3 mediated leptin signalling	Karl Roth GmbH, Karlsruhe, Germany
Dimethylsulfoxid (DMSO)	SSCP	BIORAD, Munich, Germany
Deoxy Nucleoside triphosphates (dNTPs)	PCR	Sigma-Aldrich Chemie GmbH, Munich, Germany
Beta human SH2B1 clone (inserted into pCMV-XL5 expression vectors)	STAT3 mediated leptin signalling	Origene via AMS Bio, Abingdon, UK
Beta human SH2B1 clone containing Thr484Ala (inserted into pCMV-XL5 expression vectors)	STAT3 mediated leptin signalling	Z I E L Research Center for Nutrition and Food Sciences (Department of Molecular Nutritional Medicine, Else Kroener-Fresenius Center, Technical University of Munich, Germany)
Beta human SH2B1 clone containing βThr656Ile (inserted into pCMV-XL5 expression vectors)	STAT3 mediated leptin signalling	Z I E L Research Center for Nutrition and Food Sciences (Department of Molecular Nutritional Medicine, Else Kroener-Fresenius Center, Technical University of Munich, Germany)
Gamma human SH2B1 clone (inserted into pCMV-XL5 expression vectors)	STAT3 mediated leptin signalling	Origene via AMS Bio, Abingdon, UK
Gamma SH2B1 clone containing Thr484Ala (inserted into pCMV-XL5 expression vectors)	STAT3 mediated leptin signalling	Z I E L Research Center for Nutrition and Food Sciences (Department of Molecular Nutritional Medicine, Else Kroener-Fresenius Center, Technical University of Munich, Germany)

-Table 4 will be continued-

-Table 4 continued-

Chemical	Method	Seller
Gamma SH2B1 clone containing γPro674Ser (inserted into pCMV-XL5 expression vectors)	STAT3 mediated leptin signalling	Z I E L Research Center for Nutrition and Food Sciences (Department of Molecular Nutritional Medicine, Else Kroener-Fresenius Center, Technical University of Munich, Germany)
LEPRb	STAT3 mediated leptin signalling	Rosenblum et al. 1998
pcDNA3 (empty vector)	STAT3 mediated leptin signalling	Z I E L Research Center for Nutrition and Food Sciences (Department of Molecular Nutritional Medicine, Else Kroener-Fresenius Center, Technical University of Munich, Germany)
STAT3 reporter construct (pAH32)	STAT3 mediated leptin signalling	Rosenblum et al. 1998
phrG-B	STAT3 mediated leptin signalling	Z I E L Research Center for Nutrition and Food Sciences (Department of Molecular Nutritional Medicine, Else Kroener-Fresenius Center, Technical University of Munich, Germany)
Dual-Luciferase Reporter Assay System	STAT3 mediated leptin signalling	Promega, Mannheim, Germany
Dulbecco's modified Eagle's medium (DMEM)	STAT3 mediated leptin signalling	BIORAD, Munich, Germany
Ethanol	STAT3 mediated leptin signalling	Karl Roth GmbH, Karlsruhe, Germany
Ethylendiamintetraacetat (EDTA)	DNA extraction	Karl Roth GmbH, Karlsruhe, Germany
Ethidium bromide	Gelelectrophoresis	Karl Roth GmbH, Karlsruhe, Germany
Fetal calf serum (FCS)	STAT3 mediated leptin signalling	Biochrom AG, Berlin, Germany
Ficoll Type 400	Gelelectrophoresis	Karl Roth GmbH, Karlsruhe, Germany
Formaldehyde	SSCP	Karl Roth GmbH, Karlsruhe, Germany
Formamide deion.	SSCP	Karl Roth GmbH, Karlsruhe, Germany
Glacial acetic acid	SSCP	Karl Roth GmbH, Karlsruhe, Germany
Glycerol	SSCP	Karl Roth GmbH, Karlsruhe, Germany
HEPES-buffered saline (HBS buffer)	STAT3 mediated leptin signalling	BIORAD, Munich, Germany
HEK293 Cells	STAT3 mediated leptin signalling	Z I E L Research Center for Nutrition and Food Sciences (Department of Molecular Nutritional Medicine, Else Kroener-Fresenius Center, Technical University of Munich, Germany)
Leptin	STAT3 mediated leptin signalling	Z I E L Research Center for Nutrition and Food Sciences (Department of Molecular Nutritional Medicine, Else Kroener-Fresenius Center, Technical University of Munich, Germany)
$MgCl_2$ (25mM)	PCR	Sigma-Aldrich Chemie GmbH, Munich, Germany Sigma-Aldrich Chemie GmbH, Munich, Germany

-Table 4 will be continued-

-Table 4 continued-

Chemical	Method	Seller
Optimase® (5U/µl) with according buffers	PCR	Transgenomics Limited, Glasgow, UK
PCR buffer 10x	PCR	Sigma-Aldrich Chemie GmbH, Munich, Germany
Phosphate buffered saline (PBS buffer)	STAT3 mediated leptin signalling	BIORAD, Munich, Germany
Penicilin	STAT3 mediated leptin signalling	BIORAD, Munich, Germany
Poly-D-Lysin	STAT3 mediated leptin signalling	BIORAD, Munich, Germany
Primer for PCR (Oligonucleotides; 25pmol/µl)	PCR	Sigma-Aldrich Chemie GmbH, Munich, Germany
Restriction enzyme buffer 10x (NEB Puffer 1, 2, 3 or 4)	RFLP	New England Biolabs, Frankfurt, Germany
Restriction enzymes	RFLP	New England Biolabs, Frankfurt, Germany
Silver Nitrate ($AgNo_3$)	SSCP	Karl Roth GmbH, Karlsruhe, Germany
SYBR® Gold	Gelelectrophoresis	Invitrogen, Karlsruhe, Deutschland
Taq DNA Polymerase (5U/µl)	PCR	Sigma-Aldrich Chemie GmbH, Munich, Germany
TaqMan® SNP Genotyping Assay	TaqMan	Applied Biosystems by life technologies, Carlsbad, California, USA
TaqMan® Genotyping Master Mix	TaqMan	Applied Biosystems by life technologies, Carlsbad, California, USA
Tetramethylethylendiamide (TEMED	SSCP	Karl Roth GmbH, Karlsruhe, Germany
Triethylammoniumacetate	dHPLC	Transgenomics Limited, Glasgow, UK
Tris-HCL	SSCP	Karl Roth GmbH, Karlsruhe, Germany
Trypsin	STAT3 mediated leptin signalling	BIORAD, Munich, Germany
Xylencyanol	SSCP, Gelelectrophoresis	Karl Roth GmbH, Karlsruhe, Germany

3.1.2 Study groups

Written informed consent for molecular genetic analyses was given by all participants and in case of minors by at least one of their parents. All studies were approved by the Ethics Committees of the respective Universities (University of Marburg, University of Duisburg-Essen, Rheinische Friedrich-Wilhelms-Universität Bonn) and Institutions (Vestic Clinic of Children and Adolescent Medicine, Datteln) and were performed in accordance with the *Declaration of Helsinki*. For all of the study groups used in this study, case control (by Fisher's exact test) or family-based association analyses (by transmission disequilibrium test) were performed. Descriptive statistics of all participants are listed in Table 5.

Table 5: Phenotypic description of analyzed study groups (modified from Volckmar et al. 2012).

Sample[a]	status	n	% male [%]	age [mean ± SD]	BMI [mean ± SD]	BMI SDS[b] [mean ± SD]
mutation screen	Cases	95	48.42	13.43 ± 3.37	31.87 ± 5.04	4.10 ± 1.71
family-based GWAS	Cases	705	45.11	13.44 ± 3.02	32.02 ± 5.82	4.23 ± 1.96
	Parents	1,410	50.00	42.54 ± 6.02	30.28 ± 6.33	1.65 ± 1.84
subset of family-based GWAS (for association testing)	Cases	355	54.63	13.72 ± 3.12	31.75 ± 6.06	4.13 ± 2.06
	Parents	710	50.00	42.85 ± 6.14	29.95 ± 6.15	1.57 ± 1.81
case-control GWAS	Cases	453	42.60	14.37 ± 3.75	33.15 ± 6.68	4.51 ± 2.15
	Controls	435	39.08	26.08 ± 5.75	18.09 ± 1.14	-1.45 ± 0.34
subset of case-control GWAS (for association testing)	Cases	179	49.16	14.27 ± 2.39	35.56 ± 6.13	5.3 ± 2.09
	Controls	185	55.68	25.56 ± 3.94	18.39 ± 1.09	-1.47 ± 0.33
obese adults	Cases	988	37.25	47.17 ± 14.23	35.70 ± 5.43	3.22 ± 1.66
BEPOC	Cases	1,046	47.99	10.89 ± 3.57	29.46 ± 5.91	3.54 ± 1.82
DAPOC	Cases	1,185	44.22	10.72 ± 2.78	27.69 ± 5.11	3.07 ± 1.56
Obeldicks subgroup	Cases	454	44.83	10.8 ± 2.58	27.28 ± 4.67	2.38 ± 0.5
INSULA (Ulm children's study 3)	Cases	129	43.41	14.90 ± 1.84	40.45 ± 8.00	7.05 ± 2.92
Ulm children's study 2	Cases	271	51.29	11.07 ± 3.69	29.75 ± 6.04	3.62 ± 1.88
URMEL ICE (Ulm children's study 1)	Cases	97	57.73	7.57 ± 0.42	20.68 ± 1.71	1.06 ± 0.46
	Controls	685	54.31	7.56 ± 0.42	15.63 ± 1.38	-0.25 ± 0.37
KORA	Cases	6.633	56.43	56.62 ± 13.13	29.42 ± 3.70	1.19 ± 1.09
	Controls	3.444	36.76	47.18 ± 13.37	22.72 ± 1.68	-0.53 ± 0.51

[a]**Mutation screen sample:** part of the family-based and the case-control GWAS samples' cases; **Family-based GWAS sample:** 705 index patients with early-onset extreme obesity and their biological parent (Scherag et al. 2010); **case-control GWAS sample:** GWAS of early-onset extremely obese children and adolescents in comparison to lean, adult controls (Hinney et al. 2003); **case-control sample for association testing:** early-onset extremely obese children and adolescents in comparison to lean, adult controls; independent of initial screening sample; subset of cases-control GWAS sample; **Obese adults** (Hinney et al. 2006), **DAPOC:** Datteln Paediatric Obese Cohort (Reinehr et al. 2006); **Obeldicks subgroup**: Children who completed the 1 year lifestyle intervention Obeldicks (Volckmar et al. 2013); **KORA:** Cooperative Health Research in the Region of Augsburg (Wichmann et al. 2005); **Ulm children's study 1:** Ulm Research on Metabolism, Exercise and Lifestyle in Children (Nagel et al. 2009); **BEPOC:** Berlin Paediatric Obese Cohort (Bau et al. 2009); **Ulm children's study 2 and 3:** Ulm Paediatric Obese Cohort A and B (INSULA) (Wabitsch et al. 2004).

[b]Calculation of BMI SDS values is based on population reference values following the National Nutrition Survey I (Kromeyer-Hauschild et al. 2004).

3.1.2.1 Family-based GWAS

The family-based GWAS sample (705 trios) consisted of one extremely obese child or adolescent (index patient) with both biological parents. These core families will be referred to as “trios” from here on. Detailed interviews were obtained from index patient (>12 years) and both parents for sociodemographic and anthropometric data as well as known organic diseases.

A total of 705 trios were recruited by staff of the following clinics:

- Department of Child and Adolescent Psychiatry of the Phillips University in Marburg, Germany
- Department of Child and Adolescent Psychiatry, University of Duisburg-Essen, Germany
- Vestic Clinic of Child and Adolescent Medicine, Datteln, Germany
- Center of Pediatrics, Department of General Pediatrics and Policlinic, Rheinische Friedrich-Wilhelms-Universität Bonn, Germany

Inclusion criteria for the index patients were early-onset, BMI ≥ 90^{th} age and sex specific percentile with regard to the German population at the time of sample recruitment (Hebebrand et al. 2004), and exclusion of somatic diseases that affect body weight. No restrictions were given for recruitment of parents. Trios with Mendelian inconsistencies according to genotyping information were excluded from further analyses. This study group was described in detail by Hinney et al. (2007).

3.1.2.2 Case Control GWAS

For the case control GWAS (CC GWAS), recruitment of the cases was done with the same inclusion criteria as for the index patients of the family-based trios GWAS: BMI ≥ 90^{th} age and sex specific percentile with regard to the German population at the time of sample recruitment (Hebebrand et al. 2004), and exclusion of somatic diseases that affect body weight.

Additionally, normal weight or lean adult controls were recruited with the following inclusion criteria: BMI ≥ 40^{th} and ≤ 60^{th} age and sex specific percentile with regard to the German population at the time of sample recruitment (Hebebrand et al. 2004) for normal weight or BMI ≤ 20^{th} age specific percentile with regard to the German population at the time of sample recruitment (Hebebrand et al. 2004) for lean controls, no somatic diseases, and no medication that affects body weight.

Recruitment of 453 extremely obese cases and 435 normal weight or lean controls was performed by staff of the following clinics:

- Department of Child and Adolescent Psychiatry of the Phillips University in Marburg, Germany

- Department of Child and Adolescent Psychiatry, University of Duisburg-Essen, Germany
- Vestic Clinic of Child and Adolescent Medicine, Datteln, Germany
- Center of Pediatrics, Department of General Pediatrics and Policlinic, Rheinische Friedrich-Wilhelms-Universität Bonn, Germany

This study group was described in detail by Scherag et al. (2010).

3.1.2.3 Mutation screen

From the two GWAS samples, a selection of individuals for the mutation screen was based on genotypes at SNP rs2008514 (proxy of rs7359397) near *SH2B1*. We screened for mutations in 95 individuals, of whom 90 were likely enriched for the presence of rare mutations in *SH2B1.* These extremely obese children and adolescents from the family-based GWAS sample were homozygous for the T-allele of rs2008514 and had at least one heterozygous parent, and therefore substantially contributed to the observed over-transmission of the rs2008514 risk-allele T. The additional five extremely obese individuals (one index patient from the Trio GWAS, four cases from the case-control GWAS) carried a deletion on chr16p11.2, which did not harbor any of the screened genes (compare Table 6 for position of deletion; SH2B1 position chr16:28,782,725-28,793,027; APOB48R position chr16:28,413,493-28,417,783; SULT1A2 position chr16:28,510,766-28,515,892; according to hg18, NCBI 36).

Table 6: Positions of the deletions on Chr. 16p11.2 in the five extremely obese individuals from the mutation screen (adapted from Jarick et al. 2011)

Individual	Study group	Length [kb]	Start*	End*	Distance to [bp]		
					SH2B1	***APOB48R***	***SULT1A2***
Case 1	CC GWAS	894	29,333,901	30,227,809	540,874	916,118	818,009
Case 2	CC GWAS	709	29,425,200	30,134,432	632,173	1,007,417	909,308
Case 3	CC GWAS	864	29,345,828	30,210,117	552,801	928,045	829,936
Case 4	CC GWAS	709	29,425,200	30,134,432	632,173	1,007,417	909,308
Case 5	705 Trios	709	29,425,200	30,134,432	632,173	1,007,417	909,308

*Start and endpoint of the CNV and genes are according to hg18, NCBI 36

3.1.2.4 Subset of case-control GWAS (for association testing)

Detected genetic variants from the mutation screen study group were genotyped in a sample of 179 extremely obese (age- and sex-adjusted BMI percentile ≥ 99th; Hebebrand et al. 2004) children or adolescents and 185 lean adult (age- and sex-adjusted BMI percentile ≤ 20th; Hebebrand et al. 2004) controls (Table 5). These individuals were independent of the mutation screening sample and were part of the above mentioned case-control GWAS sample (Genome-Wide Human SNP Array 6.0, see: Hinney et al. 2007, Scherag et al. 2010).

3.1.3 Additional study groups

The following study groups were not recruited in our laboratory, but belong to cooperating partners in the NGFNplus network “Molecular Mechanisms in Obesity”. To confirm initial molecular genetic findings, replication in independent study groups is necessary. Written informed consent was given by all participants and in case of minors by their parents. These studies were approved by the Ethics Committees of the respective universities and institutions and were performed in accordance with the *Declaration of Helsinki*.

3.1.3.1 Obese adults

The first independent obese cohort comprised 988 obese adults (Hinney et al. 2006, Table 6). The recruitment took place in Marburg, where several local physicians provided age, gender, body weight, and height of the individuals.

3.1.3.2 BEPOC

The BEPOC (Berlin Pediatric Obese Cohort) study group comprised 1,046 obese children and adolescents recruited by the Institute of Experimental Pediatric Endocrinology at the Charité Berlin in Germany (Bau et al. 2009, Table 6). The children were recruited during 2006 and 2007 in 68 Berlin primary and secondary schools. Body weight and height were measured; information about nutrition, sports, and leisure habits; the frequency of sporting activities, the quality of life, ethnic, and migration background as well as mother’s language were collected based on a standardized questionnaire covering 80 items.

3.1.3.3 DAPOC

The DAPOC study group consisted of 1,185 independent obese children or adolescents of the ‘Datteln Paediatric Obese Cohort’ (Reinehr et al. 2007, Table 6) that participated in the “Obeldicks” program. The ascertainment and intervention strategy has previously been described in detail (Reinehr et al. 2006, 2007). In short, “Obeldicks” is based on physical exercise, nutrition education, and behavioral therapy including individual psychological care of the child and his or her family. The nutritional course was based on an “optimized mixed diet”, containing 55% carbohydrates (5% sugar) of total dietary energy (En %), 30 En % fat and 15 En% proteins. Participants needed to prove their motivation before they were included in the intervention. Recruitment took place at the Vestic Clinic of Child and Adolescent Medicine in Datteln, Germany.

3.1.3.3.1 Obeldicks subgroup

A subset of the DAPOC study group was analyzed for the impact of BMI risk alleles at rs7498665 (coding SNP in *SH2B1*, Thr484Ala) and rs180743 (coding SNP in *APOB48R*, Pro419Ala) on weight loss and related parameters. We analyzed 454 overweight and obese children and adolescents, who completed the 1–year lifestyle intervention (‘Obeldicks’

program, Table 6; Volckmar et al. 2013). Data on anthropometrics (BMI, BMI SDS), blood pressure (systolic and diastolic) and plasma parameters (total cholesterol, LDL-cholesterol, HDL-cholesterol, triacylglycerides, glucose, insulin, and HOMA) were available for baseline and after the intervention.

3.1.3.4 INSULA

The INSULA study group consisted of the Ulm Children's Studies 2 and 3, including 271 and 129 obese children and adolescents, respectively (Wabitsch et al. 2004, Table 6). These children showed reduced insulin sensitivity in an oral glucose tolerance test or were referred to the weight loss study with T2D. The T2D status for all children was additionally analyzed with the detection of antibodies relevant for the disease. Recruitment took place in the Division of Pediatric Endocrinology and Diabetes of the Department of Children and Adolescent Medicine at the University Medical Center in Ulm, Germany.

3.1.3.5 URMEL

The population-based study group ‘Ulm Children's Study 1’ (formerly known as URMEL ICE - Ulm Research on Metabolism, Exercise and Lifestyle Intervention in Children) consisted of 782 of children and adolescents (Nagel et al. 2009, Table 6). They were recruited by the Division of Pediatric Endocrinology and Diabetes at the Department of Child and Adolescent Medicine of the University Medical Center in Ulm, Germany. To test the effectiveness of the life style intervention, 32 elementary schools were recruited in Ulm, Neu-Ulm and within the Alb-Donau-Kreis region and split equally into intervention- or control group. Recorded parameters were anthropometric measures, cardiovascular parameters, physical ability testing, exercise- and dietary pattern of children and parents and a series of psychological factors. The cooperating teachers trained the children via provision of short (5-7 min) physical exercise sessions twice a day. Consecutive follow-up examinations were performed a year later. Here, the samples were divided into cases (children and adolescents: ≥ 90^{th} BMI percentile (www.mybmi.de)) and controls (children and adolescents: BMI ≤ 90^{th} percentile (www.mybmi.de)) for analyses.

3.1.3.6 KORA

The population-based adult cohort, the ‘Cooperative Health Research in the Region of Augsburg’ (KORA; Wichmann et al. 2005, Table 6) consisted of 10,077 individuals recruited at the Institute of Epidemiology of the Helmholtz-Centrum Munich in Germany. These individuals are regularly recontacted after initially assessment of anthropometrical data with questionnaires and blood sample collection for DNA extraction. A subsample (follow-up studies F3 and F4) additionally gave further DNA and serum samples. For analyses, the cohort was split into cases (adults: BMI ≥ 25) and controls (adults: BMI ≤ 25).

3.2 Methods

3.2.1 DNA extraction and storage

Genomic DNA was isolated from EDTA-anticoagulated blood using standard procedures (Miller et al. 1988). The DNA used here was already extracted from whole blood or saliva and stored at -80°C in TE buffer (recipe see Table 7 below). All samples were available for usage before this study started. The stored original concentration of DNA was quantified with a photometer via UV light absorption at 260nm (NanoDrop® ND-1000; NanoDrop Technologies, Inc., Wilmington, USA) and diluted with aqua bidest. to a concentration of 20ng/µl for use. The concentration (C) of double stranded (ds) DNA was calculated from the dilution factor (D) and the multiplication factor specific for dsDNA (F) with the following formula:

$$C = OD_{260nm} * D * F$$

C = concentration of dsDNA (µg/ml)
OD_{260nm} = optical density (Absorption at 260nm)
D = dilution factor
F = multiplication factor specific for dsDNA (50µg/ml)

Table 7: Materials needed for TE buffer

TE-buffer (pH 8)	g/1000ml
Tris-HCL (10mM)	1,21
EDTA (1mM)	0,37

3.2.2 Polymerase chain reaction (PCR)

To amplify the amount of the region of interest, the polymerase chain reaction (PCR) was used. This method to amplify small fragments of DNA was developed by Mullis et al. (1986) and is based on cyclic repetition of a denturation-, an annealing- and an elongation-step:

1. **Denaturation at 94°C**
   The hydrogen bonds of dsDNA are destroyed (denaturated) by heating DNA to 94°C, so that the complementary DNA strands are split.

2. **Annealing of the Primers**
   Two synthetic oligonucleotides define region of interest by binding complementarily to both ends of the DNA for amplification. The forward primer binds to the 5' end and binds to the coding strand of the DNA while the reverse primer binds the 3' end of the non-coding strand (note that the reverse primer has to be designed reverse and complementary to the coding DNA sequence). Annealing temperature is dependent on the primer sequence and can be approximately calculated with the following formula:

$$T_{Ann} = 2 * n_A + 2 * n_T + 4 * n_G + 4 * n_C$$

T_{Ann} = annealing temperature of the primer (°C)
n_A = number of adenine nucleotides in the primer
n_T = number of thymine nucleotides in the primer
n_G = number of guanine nucleotides in the primer
n_C = number of cytosine nucleotides in the primer

Despite this calculation, the determination of the optimal annealing temperature also varies with use of different PCR buffers and cyclers, so an additional temperature gradient PCR is recommended.

3. **Elongation at 72°C**

   The optimal working temperature for the thermostable *Taq* polymerase isolated from the bacterium *Thermus aquaticus* is 72°C. This enzyme binds double stranded DNA and elongates from the 3' end complementary bases to single stranded (ss) DNA if desoxynucleosidtriphosphates (dNTP) and the cofactor Mg^{2+} are available. A DNA strand complementary to the presented is synthesized from 5' to 3' end. This synthesis occurs on both complementary strands at the same time (since either the forward or the reverse primer bind to each ssDNA strand) and leads to exponential amplification of the targeted DNA fragment. Proofreading *Taq* polymerases such as AmpliTaq® Gold or Optimase® can increase the yield of correctly amplified PCR product.

Procedure

For each amplicon, a PCR was run with a final volume of 12.5µl per sample, including 1.25µl DNA with a concentration of 20ng/µl (25 ng in total). A master mix containing water, buffer, magnesium chloride, primers, and nucleotides was prepared for each 96 well plate (Abgene, Hamburg, Germany; n = 100 reactions to supplement for inevitable loss due to pipetting). Each plate contained one negative control for contamination of the master mix or single ingredients of the master mix which contains 1.25µl water instead of DNA. 10.75µl master mix was filled into each well with a multichannel pipette, and then the DNA was added.

Before amplification, the plate containing master mix and DNA was sealed with self-adhesive foil (Abgene, Hamburg, Germany) and centrifuged for 1 minute at 2000 Upm. The target regions were then amplified using thermocyclers (PD-200; BioRad GmbH, Munich, Germany; Veriti; Applied Biosystem by Life Technologies, California, USA) programmed with an initial heat activation step at 94°C, a denaturation step at 94°C, an annealing step with primer specific temperature and an elongation step at 72°C which were repeated cyclically for 35

times (Table 9). A final elongation step at 72°C concluded the amplification protocol. After PCR, samples were stored at 4°C.

Table 8: Chemicals and volumes used for PCR master mix

Chemicals	Volume per reaction (µl)	Final concentration per sample
aqua bidest.	8.675	
10x buffer	1.250	1x
MgCl2 (25mM)	0.750	1.5mM/µl
Primer F (25pmol/µl)	0.125	0.25pmol/µl
Primer R (25pmol/µl)	0.125	0.25pmol/µl
dNTP (20mM)	0.125	0.2mM/µl
Taq Polymerase (5U/µl)	0.200	1U/reaction
DNA (20ng/µl)	1.250	2ng/µl

Table 9: Conditions and cycles for PCR.

Number of cycles	Step	Duration	Temperature
1	Denaturation	10min	94°C
	Denaturation	30sec	94°C
	Annealing	30sec	T_{anneal}
35	Elongation	60sec	72°C
1	Elongation	10min	72°C
1	Cooling	∞	4°C

T_{anneal}= primer specific annealing temperature

3.2.2.1 Nested PCR

If a PCR amplicon showed unwanted byproducts in gel electrophoresis, which could not be eliminated by further adaptation of PCR conditions like annealing temperature or Mg2+ concentration, a nested PCR was used to improve results. For this method, a larger DNA region was chosen with primers further up- and downstream of the target amplicons (outer primer) and amplified with PCR. The resulting PCR product was then diluted with aqua bidest. 1:100 and used as sample material for a second PCR with the original primers for the targeted amplicons.

3.2.3 Primer design

For primer design, the genomic sequence of the region of interest was extracted from the database ENSEMBL (http://www.ensembl.org). Primers were designed mainly with help of the online software `Primer3´ (http://frodo.wi.mit.edu/; You et al. 2008), using the following specifications:

1. Length of about 20 nucleotides
2. No known polymorphisms in the primer region
3. Guanine and cytosine (GC) content should be between 40-60%, leading to an annealing temperature between 55-65°C

4. The primer should not contain (long) palindromic sequences to avoid hairpin formation
5. The 3' end of the forward and the 5'end of the reverse primer should not contain more than 3 guanines or cytosines

Specificity of each primer pair was tested with the BLAT and *in silico* PCR functions of the `UCSC Genome Browser´ (http://genome.ucsc.edu/). These tools search the whole human genome for complementary sequences (BLAT) or regions amplified by the given primers (*in silico* PCR). Designed primers were only used if *in silico* PCR revealed no other amplified regions than the targeted region.

Primers used for this study were synthesized by Sigma-Aldrich Chemie GmbH (Munich, Deutschland). Primers were diluted with *aqua bidest.* to a concentration of 1000pmol/µl for stock and 25pmol/µl for PCR and stored at -20°C.Optimal annealing temperature for each primer pair was tested with a temperature gradient PCR ranging from ± 5 °C of the calculated annealing temperature (see above). The highest temperature was chosen if several temperatures revealed the same amount of PCR amplicon.

Table 10: Sizes of the coding region and protein of the analyzed genes *SH2B1*, *APOB48R*, and *SULT1A2.*

Gene	Coding region [bp]	Amino acids [aa]	Fragments
SH2B1	7,726	756	12
APOB48R	3,824	1,088	14
SULT1A2	3,903	295	5

3.2.3.1 SH2B1

For the mutation screen analyzing the coding region of *SH2B1*, the gene was covered with 12 fragments (Table 10). For two of these fragments (2h and 10a), nested PCR was done to eliminate unspecific amplification. The screening method dHPLC prohibited analyzing only the coding region, most fragments contain at least a small part of intronic sequence. Exons 2 and 10 were covered with several fragments, overlapping about 50bp. All used primers with experimentally derived optimal annealing temperature, magnesium chloride concentration and GC content are shown in Table S1 in the appendix.

3.2.3.2 APOB48R

For the mutation screen of *APOB48R*, the coding region was covered with 14 fragments (Table 10). The screening methods dHPLC and SSCP prohibited analyzing only the coding region; most fragments contain at least a small part of intronic sequence. Exons 2 and 3 were covered with fragments overlapping about 50bp. All used primers with experimentally

derived optimal annealing temperature, magnesium chloride concentration and GC content are shown in Table S2 in the appendix.

Mutation screen and analysis of the *APOB48R* gene were supervised by the author of the thesis and conducted in cooperation with Jie-Yun Song (PhD student from the Institute of Child and Adolescent Health at the Peking University Health Science Center, Bejing, China) and Jessica Graninger (Medical Biologogy, Master student from the University of Duisburg-Essen, Germany). The data presented in this thesis refer to half of the screened gene, encompassing the fragments 2a, 2c, 3b, 3d, 3e, 5, and 6.

3.2.3.4 SULT1A2

For the mutation screen analyzing the coding region of *SULT1A2*, the gene was covered by five fragments (Table 10). The screening methods dHPLC and SSCP prohibited analyzing only the coding region; each fragment contains at least a small part of intronic sequence. All used primers with experimentally derived optimal annealing temperature, magnesium chloride concentration and GC content are shown in Table S3 in the appendix.

Mutation screen and analysis of the *SULT1A2* were supervised by Anna-Lena Volckmar and conducted partially by Katharina Haas (Medical Biologogy, Master student from the University of Duisburg-Essen, Germany).

3.2.4 Gel electrophoresis

Gel electrophoresis of agarose or acrylamid gels was used to separate amplified PCR products by size. In this study, agarose gels with a concentration of 5% were prepared and included into a gel chamber with TAE buffer (Tris-Acetat-EDTA buffer; composition see Table 11) with an applied current. The negatively charged phosphate groups of the phosphodiester spine of the DNA strands travels from cathode (minus) to anode (plus). The agarose polymers of the gel hinder free passage of the DNA through the gel, imposing stronger adhesive forces for longer DNA molecules and therefore separating DNA by size. A coloring agent like ethidium bromide (Karl Roth GmbH) or SYBR® gold (Invitrogen) which intercalate in the DNA double strand is then used to visualize the resulting DNA bands under UV light.

Table 11: Materials needed for TAE buffer

TAE-Puffer (pH 8)	g/1000ml
Tris-HCL (10mM)	1,21
EDTA (1mM)	0,37

Procedure:
A 2.5% agarose gel is made by heating a mixture 12.5g agarose with 500ml TAE buffer (recipe see Table 11) until agarose has dissolved completely. After cooling down to 65°C, ethidium bromide or SYBR® gold is added to a final concentration of mmol/ml. The gel is

then cast without bubbles into a gel chamber with combs and cooled down to room temperature.

After comb removal, the gel is placed into the electrophoresis chamber and overlaid with 1x TAE buffer. A mix of 5µl PCR product and 2µl bromphenol blue is added to the gel slots, accompanied by one slot with 100bp ladder to estimate PCR product sizes.

A potential of 240V and 400mA was applied to the gel for 30-45min. After this time, a UV chamber (PEQLAB Biotechnologie GmbH, Erlangen, Germany) was used to visualize the DNA with intercalated ethidium bromide or SYBR® gold. A picture for documentation was taken with a digital camera and printed.

3.2.5 Screening methods

The coding sequence of the genes *SH2B1, APOB48R*, and *SULT1A2* was screened for variants (rare mutations, small InDels and polymorphisms) that may be associated with early-onset obesity. Depending on the size of the screened fragment, one of the following two methods was used: Single stranded conformation polymorphism analyses for PCR amplicons up to 300bp (Hayashi and Yandell 1993) or denaturing high-performance liquid chromatography for PCR amplicons up to 600bp (Liu and Steward 1997; O'Donovan et al. 1998). These fragment sizes were recommended by manufacturers, so both methods achieve a high sensitivity (5% error rate, comparable to Sanger sequencing) given a correct optimization (Jones et al. 1999, Kurelac et al. 2012).

3.2.5.1 Denaturing high-performance liquid chromatography (dHPLC)

Mutation screening with dHPLC is based on the temperature-modulated analysis of heteroduplex formation of PCR amplicons. For this, PCR amplicons are heat denatured at 94°C and slowly cooled to room temperature, leading to formation of DNA double strands. If one strand contains a heterozygous variant, the newly formed complementary DNA double strands differ in this position and form heteroduplexes. Therefore, the amplicons build two different homoduplexes (one wild type and one carrying the variant) and the heteroduplexes in 1:1 ratio (compare Figure 4). Heteroduplexes vary from the original amplicons by melting temperature since the double strand formation is not complete at the position of the variant. A reverse-phase chromatography can separate them from homoduplexes. In contrast to heterozygous variants, dHPLC is biased in detection of homozygous variants (Taliani et al. 2001). This bias can be avoided by mixing wild type DNA into the sample in a 0.5:1 ratio before PCR amplification.

For this study, a system with a chromatographic column (alkylized, hydrophobic polysteren-divinylbenzol particles) stationary phase and a mobile phase containing triethylammoniumacetate (TEAA) and acetonitrile was used (WAVE® DNA Fragment

Analysis System, Transgenomics Limited, Glasgow, UK). The amphiphile TEAA binds to the column particles by its hydrophobic alkyl side chain and presents its positively charged ammonium as bait for the negatively charged phosphate spine of the DNA. Increasing the ratio of acetonitrile leads to a wash off of the TEAA-bound DNA where the thermo labile heteroduplexes are washed of earlier than the homoduplexes at a given temperature.

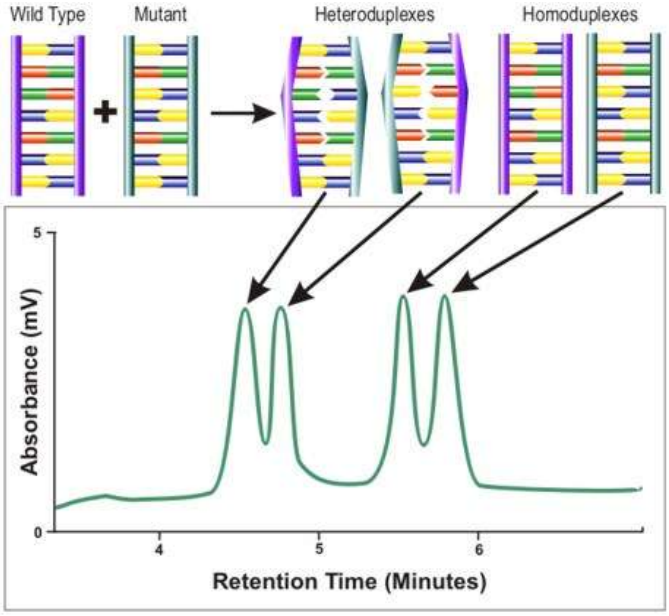

Figure 4: Separation of DNA into homo- and heteroduplexes after melting of the DNA double strand during dHPLC. The curve below depicts the absorption curve during retention time with separate peaks for both hetero- and homoduplexes (source: http://pharmatech.co.kr/pimg/wave2/image007.jpg)

A UV detector constantly measures the absorption of the solution at a wavelength of 254nm, giving out a wavelength per time chromatogram (compare Figure 4). The first two absorption peaks are elicited from the heteroduplexes; the two additional peaks are elicited from the two different homoduplexes.

Procedure:

A mutation screen of *SH2B1*, *APOB48R*, *and SULT1A2* fragments was done with the WAVE® DNA Fragment Analysis System (Transgenomic Limited, Glasgow, UK). The sample specific temperature for optimal analysis of the sample was determined with WAVEmaker® Software Version 4.0 (Transgenomic Limited, Glasgow, UK), depending on the size and GC content of the fragment.

After PCR amplification, the samples were denatured for 5 min at 94°C and cooled down to room temperature for 30 minutes, allowing for the formation of homoduplexes and heteroduplexes. Then, 5µl of each sample were applied to a 96 well plate and inserted into the WAVE® DNA Fragment Analysis System. The buffers listed in Table 12 were used to

bind the DNA to the chromatography column first (buffer A and B, contain N,N-Diethylethanamine acetate (TEAA)) and after elevating the temperature to the level determined by the WAVEmaker® Software Version 4.0 to dissolve the DNA again by applying a linear acetonitrile concentration gradient. The flow rate of the solution was kept constant at 0,9ml/min, while the mixing ratios for the buffers A, B, and C and the flow times for each buffer are given in Table 13. After each sample, the column was cleaned for 30 sec with buffer B and equilibrated to starting conditions. Buffer D is used to clean the injection needle before and after each sample run.

A full list of fragments from *SH2B1*, *APOB48R*, *and SULT1A2* that were screened for genetic variants with dHPLC can be found in Table S4 in the appendix.

Table 12: Buffers used for mutation screen via dHPLC

	Buffer [ml per liter]			
	A	**B**	**C**	**D**
TEAA (2M)	50	50	-	-
Acetonitrile	0.25	250	750	80
aqua bidest.	949.75	700	250	920

Table 13: Conditions at the chromatographic column during the mutation screen with dHPLC

Step	Time [min]	Puffer A [%]	Puffer B [%]
Loading	0.0	53	47
Gradient start	0.5	48	52
Gradient finish	4.5	39	61
Cleaning start	0.1	0	100
Cleaning finish	0.5	0	100
Equilibration start	0.1	53	47
Equilibration finish	0.9	53	47

3.2.5.2 SSCP

The single stranded conformation polymorphism analysis (Orita et al. 1989) is a method to detect sequence variants in amplified DNA fragments. The method is based on the specific tertiary structure of single stranded DNA when cooled quickly after denaturation. Since the primary structure of ssDNA affects the build of the three-dimensional tertiary structure, variants like SNPs or InDels change the tertiary structure of ssDNA. These changes can be detected by measuring the resistance and velocity of migration of the ssDNA in a polyacrylamide gel electrophoresis (PAGE).

After the run the polyacrylamide gel is stained with silver nitrate. The positively charged silver ions bind the negatively charged phosphate spine of the DNA and are reduced by alkaline formaldehyde to silver atoms. This leads to a brownish color of DNA silver complexes.

Comparable to dHPLC, SSCP does not allow direct ascertainment of genetic variants; additional Sanger re-sequencing is needed to determine to position and sequence of the variant. The formation and stability of tertiary DNA structures is influenced by the temperature at which the PAGE run is performed. The use of two different temperatures in this study (Orita et al. 1989) and a fragment size of 250bp max. (Xian et al. 1997, Sheffield et al. 1993) ensured a high sensitivity of this method for sequence variants.

Procedure:

Table 14: Buffers and solutions used for mutation screen via SSCP

Fixing solution	**Volume (ml)**
Ethanol (99,9%)	100
glacial acetic acid	5
aqua bidest.	895

1% azotic acid	**Volume (ml)**
azotic acid (65%)	15
aqua bidest.	985

10x Silver nitrate solution (0,12M)	
Silver nitrate	2.04g
aqua bidest.	100ml

Developing solution (280mM)	**Volume (ml)**
5x Sodium carbonide solution (1,4M)	100
Formaldehyde (37%)	0,257

Loading buffer	
EDTA (0,5M)	8ml
Formamide, deinonised	190ml
Glyzerine	2ml
Bromphenolblue	50mg
Xylenxyanole	50mg

1x Silver nitrate solution (12mM)	**Volume (ml)**
10x Silver nitrate solution (0,12M)	50
aqua bidest.	450

5x Sodium carbonide solution (1,4M)	
sodium carbonide	148.4g
aqua bidest.	1000ml

Stop solution	**Volume (ml)**
glacial acetic acid (99,9%)	100
aqua bidest.	900

PAGE setup

Glass plates (160mm x 178mm; Amersham Biosciences, Buckinghamshire, UK) were cleaned with isopropanol and fixed with 0.75mm spacers and a comb for gel pockets. The prepared slides were inserted into the SSCP gel chamber (Amersham Biosciences, Buckinghamshire, UK). PAGE gels consist of a network of acrylamide and N,N-methylenbisacrylamide; the ratio of these influences the pore size of the gel. To get the 15% PAGE gel, 28.75ml aqua bidest. were mixed with 18.75ml acrylamide-bisacrylamide (37.5:1) and 2.5ml 5x TBE buffer (see Table 14) in a 50ml Falkon tube. Polymerization of the acrylamide bisacrylamide network is initiated by the oxidation with ammoniumpersulfate

(APS) in presence of the catalysator tetramethylendiamin (TEMED). Therefore, 500µl APS (10%) and 34µl TEMED are added to the solution. After mixing, the solution is poured into the prepared gel chamber without bubbles. After polymerization, the gel slots are filled with aqua bidest. and the gels are placed into the SSCP gel chamber where they are overlaid with 10x TBE buffer.

PAGE

After PCR amplification of the target region and –if necessary- digestion of the resulting amplicons with restriction enzymes, 5-10µl PCR amplicons mixed with 5-0µl aqua bidest. and 8µl loading buffer (see Table 14) were denatured for 5 min at 94°C. Next, the samples were immediately placed on ice to prevent renaturation and filled into the gel slots. The SSCP gel chamber was then closed and an electrical current was applied to the gel. This separates tertiary structures of each sample by travelling of the negatively charged phosphate groups of the phosphodiester spine of the DNA strands from cathode (minus) to anode (plus). Since the sensitivity of the method is temperature dependent, two conditions were applied per run: One SSCP was run at room temperature with a current of 500V and 300mA applied to the gel for 4.5h, another SSCP was run at 4°C with a current of 300V and 240mA applied to the gel for 14-18h.

Silver nitrate stain

A modified protocol by Budowle and Baechtel (1991) was used to visualize the DNA bands by silver staining. In short, the gel was carefully removed from the glass slides and fixated for 5min in 10% ethanol and 0.5% glacial acetic acid, stopping DNA movement or band dissolving. Shaking the gel in 1% azotic acid for 10min increases binding of the silver ions Ag^+ to the negatively charged phosphodiester spine of the DNA. In the next step, 12mM silver solution was applied to the PAGE gel for 20-30min. After three wash steps with aqua bidest., the PAGE gel was pivoted in an alkaline formamide-sodium carbonide solution, which reduces the silver ion to silver atoms and makes to DNA bands visible. When the staining was visible, the reaction was stopped in a bath of 99.9% glacial acetic acid and aqua bidest. Shaking the gel in glycerol for 10min decreases cracking of the gel during drying in a vacuum dryer for 3h.

A full list of fragments from *SH2B1*, *APOB48R*, *and SULT1A2* that were screened for genetic variants with SSCP can be found in Table S5 in the appendix.

3.2.6 Sanger re-sequencing

The chain-termination method of Sanger re-sequencing (Sanger et al. 1977) has an error rate of approximately 1-5% (Richter et al. 2008). This method is based on intermitted DNA polymerization if didesoxynucleotides (ddNTPs) are inserted into the replicated DNA strand. The missing 3'OH group which is necessary for phosphodiester binding of the next

nucleotide of the DNA strand leads to chain termination. Differential fluorescent labeling of the four ddNTPs can be displayed in a chromatogram after the capillary electrophoresis is read by a fluorescence detector. Each chain terminating ddNTP (A, T, G, and C) is emitting at a different wavelength, so at each DNA chain termination a different fluorescent signal can be detected.

3.2.6.1 PCR purification

PCR products contain primers and small length unfinished PCR amplicons which can decrease sequencing quality. Therefore, the amplified PCR products were purified with the QIAquick® PCR Purification Kit (Qiagen GmbH, Hilden, Germany) according to manufacturers manual. In short, this kit uses anion exchange columns whose positive charge binds the negatively charged phosphate groups of the phosphodiester spine of the DNA strands to a dextrane matrix while other byproducts of the PCR are washed off. A change in pH decreases the charge of the DNA and releases DNA from the column.

3.2.6.2 Sanger re-sequencing

The purified PCR products were bidirectional re-sequenced by SeqLab (Göttingen, Germany) or LGCgenomics (Berlin, Germany). The resulting fragments were analyzed with the software DNAStar Version 10.1.0 (Lasergene, Madison, Wisconsin, USA) independently by two individuals.

3.2.7 Genotyping and Mutation verification

The association of a genetic variant to a specific trait or phenotype can only be analyzed in large study groups. Since both dHPLC and SSCP only detect the presence of a mutation but cannot be used to accurately determine genotypes, fast and cheap methods are used to genotype variants. In this study, PCR-based methods were used to determine genotypes in larger study groups.

3.2.7.1 Gel electrophoresis to determine deletion and insertion genotypes

The gene *APOB48R* contains several small deletions of about 20-30bp. To determine the genotype of the tested individuals, a gel electrophoreses was conducted. Changes to above procedure was the use of a 3.5% agarose gel (15mg agarose mixed with 500ml TAE buffer) and a longer time in the gel electrophoreses chamber (1h-1.5h).

3.2.7.2 Restriction fragment length polymorphism (RFLP)

To determine SNPs or point mutations, restriction enzymes are used that cut specific palindromic DNA sequences that harbor either the wild type or mutation allele. After binding to the sequence, these enzymes cut the phosphodiester bonds of the DNA strand. The target amplicons is amplified with PCR before the sequence is cut by the enzyme. A gel

electrophoresis separation of the resulting fragments is used to determine the genotype of each sample.

Restriction enzymes with cleavage sites affected by the genetic variant were determined with the online software NEBCutter2 (New England Biolabs Inc., Ipswich, MA, USA; http://tools.neb.com/NEBcutter2/index.php). The enzymes and experimental conditions used in this study are listed in Table 15.

Table 15: Composition of the used 10x buffer for RFLP digests from New England Biolabs (NEB)

NEB 10x buffer	1	2	3	4
NaCl (mM)	n/a	500	1000	n/a
Tris-HCL (mM)	n/a	100	500	n/a
$MgCl_2$ (mM)	100	100	100	n/a
Dithiothreitol (mM)	10	10	10	10
Bis-Tris-Propan-HCL (mM)	100	n/a	n/a	n/a
Tris-Acetate (mM)	n/a	n/a	n/a	200
Potassium Acetate (mM)	n/a	n/a	n/a	500
Magnesiumacetate (mM)	n/a	n/a	n/a	100

Procedure:

Per reaction 5-10µl PCR product was digested, which correlates to an amount of 200-250ng DNA. Total reaction volume of the enzymatic digest was 25µl, containing 2 Units of enzyme, 2.3µl buffer and –if necessary for reaction- 0.25µl Bovine serum albumin (BSA). Each reaction mix was brought to the temperature noted in Table 15 and digested overnight. The fragment sizes and therefore genotypes were determined by gel electrophoresis.

A list of the used enzymes with cleavage site for the corresponding SNPs and mutations can be found in Table 16. The table also contains buffer name, temperature and duration of the digest, and product sizes of the digested fragment.

Table 16: Overview of the restriction enzymes used to genotype SNPs and mutations in *SH2B1* and *SULT1A2*

Gene	Variant	Restriction enzyme	specific palindromic DNA sequence	Temp. [°C]	Product sizes [bp]
SH2B1	rs28433345	Nla3	5′...CATG▼...3′ 3′...▲GTAC...5′	37	C: 422 T: 381+41
SH2B1	rs62037369	Fok1	5′...GGATG(N)$_9$▼...3′ 3′...CCTAC(N)$_{13}$▲...5′	37	C: 188+112+53+34 T: 188+146+53
SH2B1	rs62037368	Hae3	5′...GG▼CC...3′ 3′...CC▲GG...5′	37	G:365+22 A:387
SULT1A2	rs4115668	PspG1	5′...▼CCWGG...3′ 3′...GGWCC▲...5′	75	C: 329+125 T: 454
SULT1A2	rs3743963	Ban1	5′...G▼GYRCC...3′ 3′...CCRYG▲G...5′	37	T:311+143 C: 272+143+39

All restriction enzyme digests were done in NEB 4 10x buffer.

3.2.7.3 TaqMan® SNP genotyping

If SNPs or mutations are genotyped on a larger level, the TaqMan® SNP allelic discrimination assay (TaqMan® SNP Genotyping, Applied Biosystems by life technologies, Carlsbad, California, USA) provides a fast, sensitive and reproducible method. Basis principle of the assay is a quantitative PCR as depicted in Figure 5 below. Each assay contains two allele specific probes labeled with either VIC or FAM dye which bind to the targeted region. After binding of the probe to the complementary DNA, an exonuclease cleaves of the non-fluorescent quencher that prevents the fluorescent signal elicited from the VIC or FAM dye at the end of the probe. This allows fluorescent signal when the probe binds correctly to the DNA. Since each allele elicits only one signal, either VIC or FAM, homo- or heterozygous state of the sample can be analyzed by quantifying the fluorescence elicited at a wavelength of 488nm and 520nm, respectively. The minor groove binder increases specificity by binding to the minor groove of the DNA helix.

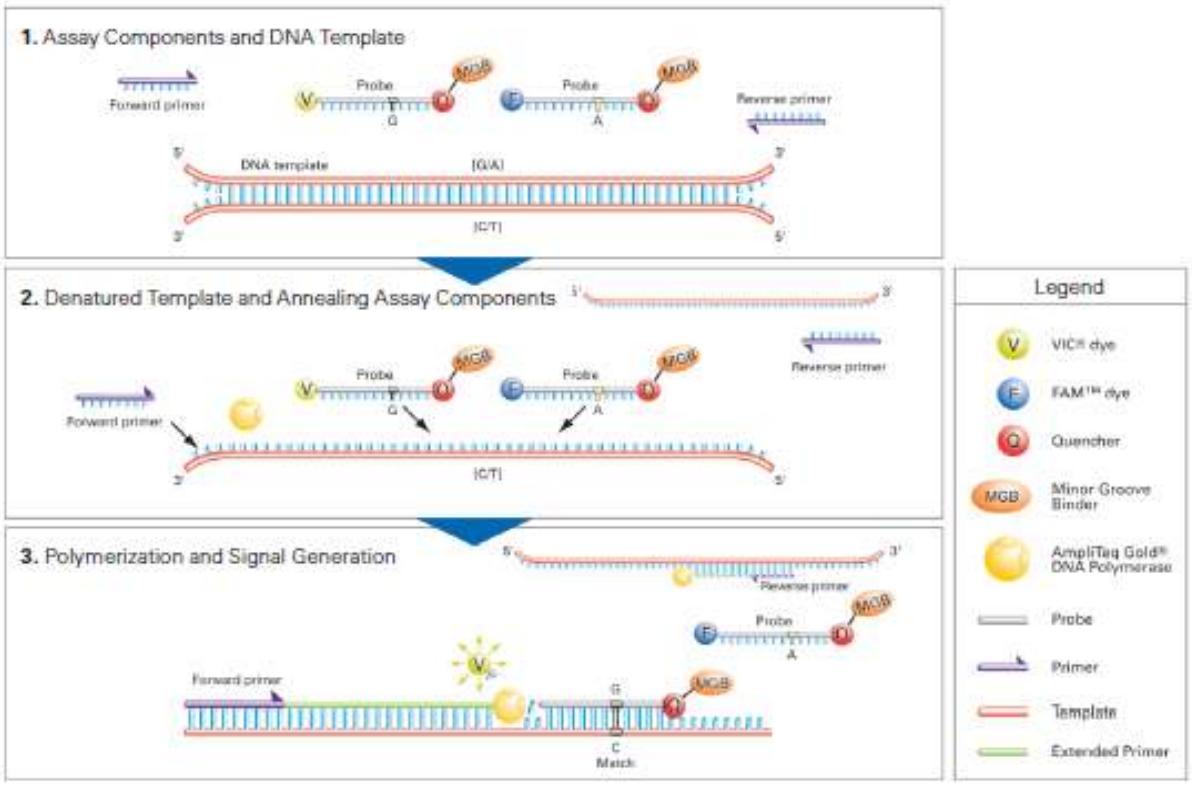

Figure 5: Reaction scheme of TaqMan® SNP allelic discrimination assay. Two allele specific probes containing either FAM or VIC dye bind to DNA with help of the MGB. Cleaving of the quencher by an exonuclease elicits the fluorescent signal (© Copyright 2007, 2010. Applied Biosystems).

A list of the TaqMan® SNP allelic discrimination assays for the corresponding SNPs and mutations can be found in Table 17 below.

Table 17: List of SNPs and mutations from *SH2B1*, *APOB48R*, *and SULT1A2* that were genotyped by TaqMan® SNP Genotyping Assay

Gene	SNP/mutation	Amino acid exchange	Assay
SH2B1	rs147094247	Thr175Asp	AHCS0BY
SH2B1	rs7498665	Thr484Ala	C__25999166_10
SH2B1	g.9483C>T	*β*Thr656Ile/*γ*Pro674Ser	AHMSHDX
APOB48R	rs180743	Pro419Ala	C____789270_10
SULT1A2	rs4987024	Tyr62Phe	C__31470561_10

Procedure:

TaqMan® SNP Genotyping Assays were performed on a scale of 6µl total volume. The mix for the real time PCR contained 2.4µl aqua bidest., 2.5µl TaqMan® Genotyping Master Mix, 0.1µl TaqMan® SNP Genotyping Assay and 1µ of sample material (DNA, c=20ng/ml). PCR conditions are listed in Table 18 below.

Table 18: PCR conditions for TaqMan® SNP Genotyping Assays

Step	Temperature	Time
Melting Temp. accessment	60°C	30 sec
Taq activation	95°C	10 min
Denaturation	92°C	15 sec
Annealing and elongation	60°C	90 sec
Melting Temp. accessment	60°C	30 sec

3.2.7.4 Matrix-assisted laser desorption/ionization time of flight analysis (MALDI TOF)

A non-PCR based approach to analyze genotypes on a large scale is the Matrix Assisted Laser Desorption Ionization – Time of Flight Mass spectrometry (MALDI TOF). This method utilizes the molecular mass of biological particles by measuring the time of flight that ionized bio-molecules take for a defined flight distance in the spectrometer. Since the mass of the nucleotides building the DNA is different (adenine 491.2g/mol, thymine 482.2g/mol, cytosine 467.2g/mol, guanine 507.2g/mol), MALDI TOF can be used to analyze genetic variants (Storm et al. 2003, Pusch et al. 2002). For the analysis, the DNA molecules are embedded into 3-hydroxypicolinic acid and submitted to laser rays in vacuum. This ionizes the DNA particles which are accelerated by an electrostatic field to constant kinetic energy (E_{kin}) and reach the detector after traveling for a defined distance (l; see Figure 6).

The time (t) for this flight is measured and the velocity (v) of the ions can be calculated from the following formula:

$$v = l / t$$

Since the E_{kin} is constant, the mass of the ionized particles can be calculated from the velocity as follows:

$$m= 2(E_{Kin}*v^2)$$

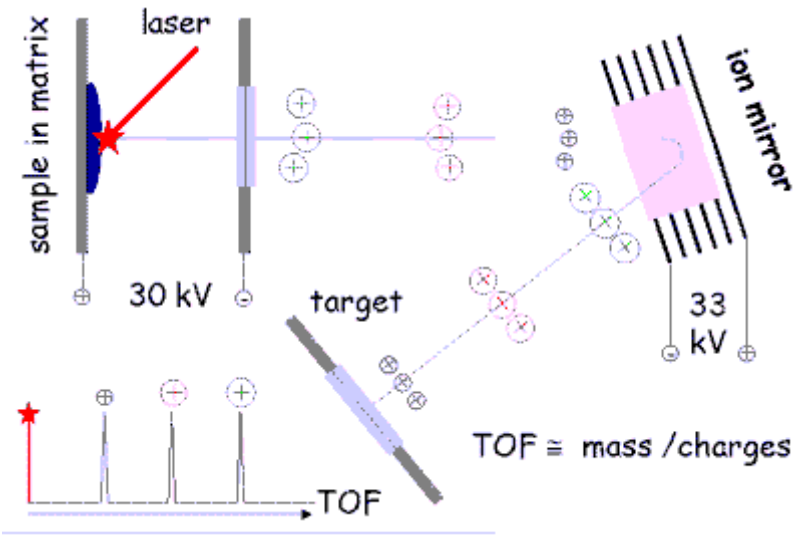

Figure 6: MALDI TOF method in short: A laser shoots ionized particles from the sample which are drawn to the ion mirror. There, they are reflected and fly to the target of the electric field. The time of flight for each particle is then calculated. (source:http://edoc.hu-berlin.de/dissertationen/xie-jing-2003-12-15/HTML/xie_html_284c12c8.png)

Table 19: SNPs in *APOB48R* and *SULT1A2* genotyped with MALDI TOF

Gene	SNP	Amino acid exchange
APOB48R	rs180743	Pro419Ala
APOB48R	rs180744	Gln553His
SULT1A2	rs4149404	Ile7Thr
SULT1A2	rs10797300	Pro19Leu
SULT1A2	rs145008170	Ser44Asn
SULT1A2	rs4987024	Tyr62Phe
SULT1A2	rs142241142	Ala164Val
SULT1A2	rs1059491	Asn235Ile
SULT1A2	rs75191166	Glu282Lys

In this study, the SNPs and mutations listed in Table 19 above were genotyped in larger collectives in the laboratory of Prof. Dr. T. Illig and Prof. Dr. H.E. Wichmann at the Helmholtz-Zentrum Munich (Germany) under the supervision of Dr. H. Grallert. Primers were developed to bind 5bp 5' lateral to the targeted SNP. Elongation for two or three nucleotides with either Desoxyribonukleotides (dNTP) or Didesoxyribonukleotides (ddNTPs) was then done to analyze genotypes with MALDI TOF.

3.2.8 *In silico* analyses of mutations/SNPs

In silico prediction of the impact of genetic variants on the expression or functionality of proteins is a useful tool to predict potential causality for the trait observed. While for coding variants, the software usually calculates differences in the side chains of the changed amino acids, calculating differences for non-coding variants is more difficult. Usually, the programs analyze whether new splice sites open (or existing splice sites are removed) or if new splicing enhancers or silencers are introduced in the sequence. These would change the

formation of a protein and can lead to loss of protein function if important domains are no longer expressed. A change in the regulatory domain of a gene might also affect protein expression without being coding. Another potential non-coding way to modify protein expression is by changing (destroying existing or building new) transcription factor binding sites. If these are affected, protein expression would stop at transcription which decreases total protein amount in the cell.

To analyze the potential impact of detected variants in *SH2B1*, *APOB48R*, *and SULT1A2* on protein function and expression, these were analyzed *in silico* with the following online available software:

3.2.8.1 Regulation of Splicing

Exonic splicing enhancer (ESE) and exonic splicing silencers (ESS) are 6 bp long DNA sequence motifs within exonic regions of a gene. They regulate splicing of pre-mRNA into mRNA. These conserved sites are necessary for correct splicing to ensure correct build of the resulting protein. Abrogation of ESEs or ESSs can result from mutations of the genetic code which potentially leads to false splicing of the mRNA and hence unstable or non functional protein.

ESE finder (http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi, Cartegni et al. 2003), ESRsearch (http://esrsearch.tau.ac.il/ Goren et al. 2006), PESX (http://cubweb.biology.columbia.edu/pesx/, Zhang et al. 2005) and RESCUE-ESE (http://genes.mit.edu/burgelab/rescue-ese/, Fairbrother et al. 2003) are tools provided online which predict ESEs and ESSs by comparing known splicing enhancers to the genetic sequence. Variants disrupting these predicted ESEs and ESSs can potentially affect protein splicing.

3.2.8.2 Regulation of transcription

The regulation of DNA transcription is an important mediator of the amount of expressed protein in the cell. Transcription factor binding sites (TSFs) are 5-7bp long conserved sequence motifs in the genetic region which can be bound by transcription factors such as TFIIA, TFIIB, TFIID (TATA binding protein), TFIIE, TFIIF, and TFIIH (common) or hormones like estrogen or testosterone, heat shock factors, or cell differentiation signals like Myc (specific; Funnel and Crossley 2012). These factors bind with their DNA-binding domain to conserved sequence motifs and can stabilize or block RNA binding to DNA, catalyze histone modification or recruit other proteins necessary for DNA transcription. Variants disrupting TSFs can interfere with correct regulation of transcription leading to increased or decreased amount of total protein.

The free online tools TFSEARCH (http://www.cbrc.jp/research/db/TFSEARCH.html, Heinemeyer et al. 1998) and CONSITE (http://asp.ii.uib.no:8090/cgi-bin/CONSITE/consite/, Sandelin et al. 2006) were used to predict disruption of transcription factor binding sites by the variants detected in the genes *SH2B1*, *APOB48R*, and *SULT1A2.*

3.2.8.3 Changes in protein folding and amino acid sequence

If a mutation or SNP affects the amino acid sequence of the protein, important domains with functional aspects like binding or dimerization, catalytic centers of enzymes or other functions might be disrupted. A polar amino acid built in instead of an unpolar amino acid might disrupt a transmembrane domain which in turn could prohibit a receptor taking correct position in the membrane or derange the transmembrane part of an ion channel. On the other hand, unpolar amino acids substituting polar amino acids can lead to malformation of elicitor binding or false/ no signalling of a receptor after binding of the signal (Dobson 2003). The impact of coding variants on protein function was predicted with the following online tools:

Mutation Taster

Mutation Taster (http://www.mutationtaster.org/, Schwarz et al. 2010) calculates probabilities whether an alteration in a given sequence may be a disease causing mutation or a simple polymorphism. The closer the probability of the calculation comes to 1, the higher is the reliability of the prediction. The conservation of the position of the exchange in five different species is taken into account (human, *Pan troglodytes, Mus musculus, Danio rerio, and Caenorhabditis elegans*). Additionally, the program analyzes the disruption of known functional domains in the protein. Prediction of splice site alteration is included in the analysis.

Panther

Panther (Protein ANalysis THrough Evolutionary Relationships; http://www.pantherdb.org/; Thomas et al. 2003) calculates whether a particular non-synonymous amino-acid substitution causes a functional alteration of the protein with a background of conservation at this position in several species. It computes the subPSEC (substitution position-specific evolutionary conservation) score based on an alignment of related proteins in other species (gene families classified by their function in the Gene Ontology Reference Genome Project). The subPSEC score is the negative decadic logarithm of the probability ratio of the wild-type and mutant amino acids at the position of the exchange. These scores value from 0 to about -10 with increasing predicted effect on correct protein function. A functional effect of the variant is predicted if the subPSEC score reaches -3 corresponding to a 50% probability that a score is deleterious (Pdeleterious of 0.5).

PolyPhen-2

PolyPhen2 (http://genetics.bwh.harvard.edu/pph2/; Adzhubei et al. 2013) calculates the absolute difference of profile scores (PSIC score) for both allelic variants within the polymorphic site. These PSIC scores are dependent on polarity, charge and size of the analyzed amino acids and have been calculated from known effects of amino acid substitutions. The outcome of the substitution is calculated as a score between zero and one which is then evaluated as benign (0>0.15), possibly damaging (0.15>0.85), or probably damaging (>0.85). In the study of Bonnefond et al. (2012), the sensitivity and accuracy of two PolyPhen2 software versions (V2.022 and V2.023) were compared. In MTRN1B, the correctness of PolyPhen2 prediction was 62.5% or 60% when validated with *in vitro* data, respectively.

SIFT

SIFT (Sorting Intolerant From Tolerant; http://sift.bii.a-star.edu.sg/; Ng et al. 2003) predicts whether an amino acid substitution in a protein will be neutral (tolerated) or functional (intolerated) by sequence homology. Conserved functional regions, so called domains, are rarely affected by evolution. Hence, amino acid substitutions in such conserved domains, detected by multi-alignment of protein families have a stronger prediction of functional relevance (Lipman et al. 2002). Normalized probabilities for all possible substitutions from the alignment are calculated; these range from one (tolerated) to zero, were a cutoff of 0.05 is defined for deleterious variants.

SNAP

SNAP evaluates effects of single amino acid substitutions on protein function with a neural-network based method that uses *in silico* derived protein information (http://rostlab.org/services/snap/; Bromberg and Rost 2007). For example, protein secondary structure, conservation of amino acid position, and solvent accessibility of the mutated position are used to determine predictions regarding functionality of mutated proteins. SNAP calculates the accuracy of its own prediction based on given protein annotation from several data bases. Predictions with lower accuracy than 50% are regarded as unsafe. In a cross-validation test on over 80,000 mutants, SNAP identified 80% of the non-neutral substitutions at 77% accuracy and 76% of the neutral substitutions with 80% accuracy.

3.2.8.4 Conservation analysis

Conservation was analyzed by aligning sequences of a maximum of 21 species in total (21 α, eight β and six γ sequences *SH2B1*) with the software BioEdit Version 7.1.11. Species were human (*Homo sapiens*; α NP_001139267.1, β NP_001139268.1, γ NP_001139269.1), chimp (*Pan troglodytes*; α ENSPTRP00000053639), gorilla (*Gorilla gorilla*; α ENSGGOP00000022688), mouse (*Mus musculus*; α NP_001074928.1, β NP_035493.2, γ

AF421139, δ AF380422), rat (*Rattus norvegicus*; α NP_604451.2, β NP_001041645.1), horse (*Equus caballus*; α XP_003362765.1, β XP_001502284.1, γ XP_003362766.1), cattle (*Bos taurus*; α NP_001192458.1, β XP_872072.3), panda bear (*Ailuropoda melanoleuca*; α XP_002927372.1, β XP_002927373.1), Northern white-cheeked gibbon (*Nomascus leucogenys*; α XP_003261644.1, β XP_003261646.1, γ XP_003261648.1), Hoffmann's Two-toed Sloth (*Choloepus hoffmanni*; β ENSCHOT00000007709), African bush elephant (*Loxodonta africana*; α ENSLAFP00000014023), Northern Treeshrew (*Tupaia belangeri*; α ENSTBEP00000013184), cat (*Felis catus*; α ENSFCAP00000002859), Large Flying Fox (*Pteropus vampyrus*; α ENSPVAP00000003750), Common Bottlenose Dolphin (*Tursiops truncatus*; α ENSTTRP00000007486), Common Dog (*Canis lupus familiaris*; α, β, γ), Rhesus macaque (*Macaca mulatta*; α ENSMMUP00000030963, γ ENSMMUP00000030964), little brown bat (*Myotis Lucifugus*; α ENSMLUP00000009040), Nine-Banded Armadillo (*Dasypus novemcinctus*; α ENSDNOP00000000328), Rock Hyrax (*Procavia capensis*; α ENSPCAP00000014368) , common marmoset (*Callithrix jacchus*; α ENSCJAP00000011426, β ENSCJAP00000005535, γ ENSCJAP00000011454), guinea pig (*Cavia porcellus*; α ENSCPOP00000007490), pig (*Sus scrofa*; α ENSSSCP00000008333).

3.2.9 *In vitro* leptin receptor sensitivity assay

To quantify the impact of SH2B1 variants on leptin signalling, an *in vitro* assay was conducted. SH2B1 interacts with downstream signalling of the leptin receptor mainly by increasing autophosphorylation of the Janus kinase 2 (JAK2) which in turn binds STAT3 and STAT5. Both STAT complexes are transported into the nucleus where they activate transcription of several genes including PTP1B and SOCS3 which are responsible for negative feedback. The assay we used to analyze changes in leptin signalling by mutating SH2B1 with the detected variants rs7498665 (Thr484Ala) and βThr656Ile/γPro674Ser focused on STAT3 mediated signalling.

The protocol for this method was adapted from Rosenblum et al. (1998). HEK293 cells were co-transfected via calcium phosphate transfection with LEPRb, a STAT3 reporter construct and SH2B1 splice variants with or without the mutations. After stimulation with leptin which activates STAT3 mediated leptin signalling, the STAT3 reporter construct leads to luciferase transcription. Luciferase emits fluorescence when provided with ATP and can be measured photometrically with the Promega Dual – Luciferase Reporter Assay system according to manufactures' instruction.

Experiments and analysis of the STAT3 mediated leptin signalling assay were done under the supervision of Prof. Dr. M. Klingenspor and Dr. F. Bolze at the Z I E L Research Center for Nutrition and Food Sciences (Department of Molecular Nutritional Medicine, Else Kroener-Fresenius Center, Technical University of Munich Germany).

Procedure:

HEK293 cells were kept at 37°C in discs (6 or 10cm) or plates (24 or 48 well plates with flat bottom) coated with poly-D-lysin in Dulbecco's modified Eagle's medium with 10% FCS and 1% Penicillin. Transfection with the SH2B1 clones by calcium chloride was done with a mix of 10ng DNA with 31µl $CaCl_2$ and 500µl HBS buffer ad 1000µl total volume (ad aqua bidest. to achieve). The DNA mix contained 2.5ng of the STAT3 reporter construct kindly provided from Rosenblum et al. (1998), 2.5ng of LEPRb (also kindly provided from Rosenblum et al. 1998) since the HEK293 cells to not produce the leptin receptor internally, 2.5ng of phrG-B as the internal control for transfection, and 2.5ng of one of the various SH2B1 clones to be tested. In the negative control, the SH2B1 clone was replaced with pcDNA3; an empty vector as the transfection control. The transfection medium was added drop wise to the cells and incubated over night.

The next day, cells were distributed equally to plates (24 or 48 well plates with flat bottom) coated with poly-D-lysin after trypsination and grown for 24 hours. For each SH2B1 clone and the negative control condition, leptin was added to the medium (Dulbecco's modified Eagle's medium with 10% FCS and 1% Penicillin) in eight increasing concentrations (0ng/ml, 0.5ng/ml 1ng/ml, 5ng/ml, 10ng/ml, 50ng/ml, 100ng/ml, 500ng/ml), one concentration per well.

After 16h of incubation, the cells were washed with PBS buffer and killed by freezing them at -80°C. The photometric measurements were done with a photometer (Flouroscan Ascent FL, Thermo Scientific) in three steps according to the manufacturer's manual of the Dual-Luciferase Reporter Assay System (Promega). At first, the cells were lysed with 100µ lysis buffer for 20 minutes on a shaker and separated into duplicates for each leptin concentration and SH2B1 clone. Then, 50µl LARII (PLuc) was added to the lysed cells and the fluorescence was measured with the photometer. At last, 50µl Stop&Glow (RLuc) were added to the mix and the fluorescence was measured with the photometer. Dose-response curves, EC50 and Emax values were calculated with Prism version 5 (GraphPad Software, Inc., La Jolla, California, USA).

3.2.10 Statistics

3.2.10.1 Hardy-Weinberg-Equilibrium

In early 1900, Hardy and Weinberg simultaneously discovered that allele and genotype frequencies of a bi-allelic variant are constant in (ideal) populations. The formula to describe the probability of this phenomenon is

$$p^2 + 2pq + q^2 = 1$$

where p is the frequency of the wild type allele and q the frequency of the mutated allele. The combined frequency of both alleles at the locus is 1. Deviation from the Hardy-Weinberg-

Equilibrium (given by a deviation of the allele frequencies from the ideal population with a significance of $p<0.05$) is either a sign for evolutionary pressure on the analyzed allele or genotyping errors.

3.2.10.2 Transmission Disequilibrium Test

The Transmission Disequilibrium Test (TDT) tests for association when genetic linkage is present (Spielman et al. 1993). The TDT is robust against possible bias by population stratification in case-control association testing. The TDT measures the over-transmission of one allele from heterozygous parents to offspring that carry the target trait. Each parent has two alleles which can be inherited by the child according to the following scheme:

	Non-transmitted Allele		
Transmitted Allele	*WT*	*MUT*	*total*
WT	a	b	a+b
MUT	c	d	c+d
total	a+c	b+d	*2n*

The relevant transmission number is the amount of passed on mutated alleles (MUT) from heterozygous parents (b + c in the scheme). Any deviation from the expected mendelian transmission rate of 50% per allele is a sign for association of the observed trait with the mutated allele. If the transmission rate is above 0.5, the mutated allele increases chance for the child to carry the trait, if the inheritance is below 0.5, the mutated allele is protective for the observed trait. Association of the allele of interest with the trait is based on the following test statistics (adapted chi-square test):

$$\chi^2 = (b - c)^2/(b + c)$$

which is the chi square distribution under the null hypothesis of a transmission rate of 0.5. For this study, a two tailed, asymptotic TDT was calculated with the software PLINK (Purcell et al. 2007). If not stated otherwise, all p-values are asymptotic, two-sided and not corrected for multiple testing.

3.2.10.3 Fisher's exact test

In addition, Fisher's exact test was used to analyze allelic association in case control setups. The test is calculating the deviation from the null hypothesis ("both study groups contain the same allele frequencies") and advantages above the chi square which could also be used for the same task by providing an exact deviation rather than an approximation. The two main assumptions of the test are a fixed number of events (in this case alleles that can be inherited), and a low number of degrees of freedom (1-10). In contrast to TDT, sparseness in the cell counts of the cross table is unproblematic. For this study, a two tailed Fisher's exact

test was calculated with the software PLINK (Purcell et al. 2007). If not stated otherwise, all p-values are asymptotic, two-sided and not corrected for multiple testing.

3.2.10.4 Linear regression

To calculate the impact of risk alleles at SNPs on the outcome of the 1-year lifestyle intervention “Obeldicks”, linear regression with sex, age with and without baseline measurement as covariates was used. In regression it is analyzed whether a dependent variable (i.e. the weight loss over time) can be explained by other independent variables (i.e. the genotype at a specific SNP). The analyses were performed with the software PLINK (Purcell et al. 2007).

3.2.10.5 Power analysis

When designing an experiment, the number of cases and controls to be analyzed is an important factor for the validity of test results and conclusions and depends among other things on the expected effect size of the outcome. We tested power with QUANTO Version 1.2.4 (http://hydra.usc.edu/gxe).

4. Results

4.1 Chromosomal region Chr16p11.2

We analyzed our GWAS data for the 32 loci associated to obesity (Speliotes et al. 2010). Our study groups comprised 453 extremely obese cases and 435 normal weight or lean controls and 705 family trios (extremely obese child or adolescent with both biological parents). As not all lead SNPs were available on the SNP chip (Genome-Wide Human SNP Array 6.0, see: Hinney et al. 2007, Scherag et al. 2010) used in our study groups, imputed data were also used.

The locus at SNP rs7359397 near *SH2B1* was the fourth best hit in the combined analysis of our two GWAS (Scherag et al. 2010). The best hits were SNPs near *FTO*, *MC4R* and *TMEM18* and were already/are currently analysed (Scherag et al. 2010, Hinney et al. 2006). SNP rs2008514 (proxy of rs7359397), 50 kb upstream of *SH2B1*, showed a directionally consistent effect in both GWAS for early onset extreme obesity (p_{trios} = 0.009; p_{cc}= 0.025).

Individuals who contributed to the positive TDT for rs2008514 are most likely enriched for putative (infrequent) variants in the chromosomal region of chr16p11.2 harboring *SH2B1*, *APOB48R* and *SULT1A2*. Of these individuals, 90 were selected for the initial mutation screen of these genes. Five additional extremely obese individuals were chosen as they carried a deletion on chr16p11.2 which, however, did not harbour any of the screened genes (Table 6). They were not enriched for (coding) variants in high linkage disequilibrium with rs2008514 (proxy of rs7359397) as they were not selected because of their genotype at rs2008514.

Speliotes et al. (2010) also reported coding SNPs in high linkage disequilibrium (r^2>0.75) with the lead SNPs from their study, but give no association data for the SNPs. Hence, we analyzed the association with obesity for the lead SNPs and coding SNPs in a sample of 705 family based trios (Table 20). Three of the lead SNPs showed nominal association with obesity; rs713586 (p = 0.046) in the *POMC* locus, rs2241423 (p = 0.042) in the *MAP2K5* locus and rs7359397 (p = 0.015) in the *SH2B1* locus. But only for the SH2B1 locus, the coding variants rs180743 (APOB48R [Pro419Ala]) and rs7498665 (SH2B1 [Thr484Ala]) also showed nominal association with obesity in our sample (p = 0.003 and 0.009, respectively).

Interestingly, the coding variant (rs11676272 [Ser107Pro]) in high LD with the *POMC* locus is located in the adenylate cyclase 3 (*ADCY3*). This protein is a member of the adenylate cyclase family which has recently been shown to be associated with obesity (Berndt et al. 2012) and extreme classes of obesity (Berndt et al. 2013).

Table 20: BMI association signals of non-synonymous variants in high LD (r^2>0.75) with lead SNPs (Speliotes et al. 2010) in 705 family trios (Volckmar et al. in prep.).

Lead SNP / nearest gene(s)	Risk allele / frequency	Odds Ratio	p-value	Coding SNP	Gene	Coding change	Risk allele / frequency	Odds Ratio	p-value
rs713586 *POMC*	A / 0.52	1.16	**0.046**	rs11676272	*ADCY3*	S107P	G / 0.48	1.12	0.114
rs13107325 *SLC39A8*	T / 0.08	1.12	0.438	rs13107325	*SLC39A8*	A391T	T / 0.08	1.12	0.438
rs2112347 *FLJ35779/ HMGCR*	T / 0.65	0.96	0.583	rs2307111	*POC5*	H11R	C / 0.60	0.98	0.760
rs10767664 *BDNF*	T / 0.21	1.04	0.710	rs6265	*BDNF*	V66M	T / 0.19	0.98	0.845
rs3817334 *MTCH2*	G / 0.42	1.09	0.249	rs1317149	*MTCH2*	-	A / 0.36	1.10	0.221
rs2241423 *MAP2K5*	A / 0.78	0.83	**0.042**	rs7170185	*LBXCOR1*	W200R	T / 0.78	0.85	0.073
rs7359397 *SH2B1*	A / 0.40	1.21	**0.015**	rs180743	*APOB48R*	P419A	C / 0.43	1.27	**0.003**
rs7359397 *SH2B1*	A / 0.40	1.21	**0.015**	rs7498665	*SH2B1*	T484A	T / 0.41	1.21	**0.009**
rs2287019 *GIPR*	G / 0.79	0.88	0.192	rs1800437	*GIPR*	E354Q	G / 0.80	0.88	0.188

For rs1064608 [Pro290Ala] in MTCH2, genotyping of the coding variant was not possible with the chosen method, this SNP was substituted by a non-coding variant (rs1317149) in high LD (r^2=1) with both lead and coding SNP. Results with a nominal p-value < 0.05 are shown bold. Abbreviations: *ADCY3* (adenylate cyclase 3), *SLC39A8* (solute carrier family 39), *POC5* (POC5 centriolar protein homolog), *BDNF* (brain-derived neurotrophic factor), *MTCH2* (mitochondrial carrier 2), *LBXCOR1* (transcriptional corepressor Corl1), *APOB48R* (apolipoprotein B48 receptor), *SH2B1* (SH2B adaptor protein 1 isoform 1), *SULT1A2* (sulfotransferase family, cytosolic, 1A, member 2), *GIPR* (gastric inhibitory polypeptide receptor).

4.2.1 SH2B1

The *SH2B1* gene is the most likely obesity candidate gene within the chromosomal region 16p11.2. It is involved, as signal amplifier, in several pathways that affect energy homeostasis like leptin (Rui et al. 2007) and insulin signalling (Morris et al. 2010). The murine knockout model is obese and hyperphagic and shows hyperleptinemia and hyperinsulinemia (Rui et al. 2005). When screening the coding region of the *SH2B1* gene encoding for the tyrosine kinase receptor adapter protein in 95 extremely obese children and adolescents, we identified two previously unknown mutations (g.9483C/T: *β*Thr656Ile/*γ*Pro674Ser; g.10182C/A) and detected five known SNPs (rs7498665: Thr484Ala, rs147094247: Thr175Asp, rs28433345 (formerly rs60604881), rs62037368, and rs62037369) in SH2B1 (see Table 21, Figure 7).

One of these previously unknown variants is located at position g.9483 (C/T) of SH2B1 and results in non-synonymous, non-conservative exchanges in two of the four human splice variants (β and γ) of SH2B1. Since the two variants result in a shifted reading frame, the

mutation entails two different non-conservative exchanges and non-synonymous (βThr656Ile or γPro674Ser) for both splice variants. Since the *β*Thr656Ile/*γ*Pro674Ser mutations were predicted to change the protein structure (Table 23), we genotyped the mutation in a total of 11,406 (extremely) obese or overweight children, adolescents and adults and 4,568 controls (for children and adolescents BMI < 90th percentile, for adults BMI < 25kg/m²).

While we detected two additional obese cases carrying the mutation *β*Thr656Ile/*γ*Pro674Ser, no mutation carrier was found among the controls. Due to the low frequency of the mutation, the association analysis is grossly underpowered (p=1; Table 21). According to power calculation (for a p-value below 0.05 with statistical power above 80%), more than 545,757 individuals would need to be genotyped for a significant p-value if the observed trend (only mutation carriers among the overweight or obese individuals) would remain stable. Especially, the number of controls would need to be increased largely (by 394,695 individuals) to validly show that the mutation *β*Thr656Ile/*γ*Pro674Ser is only present in obese individuals and hence associated with obesity.

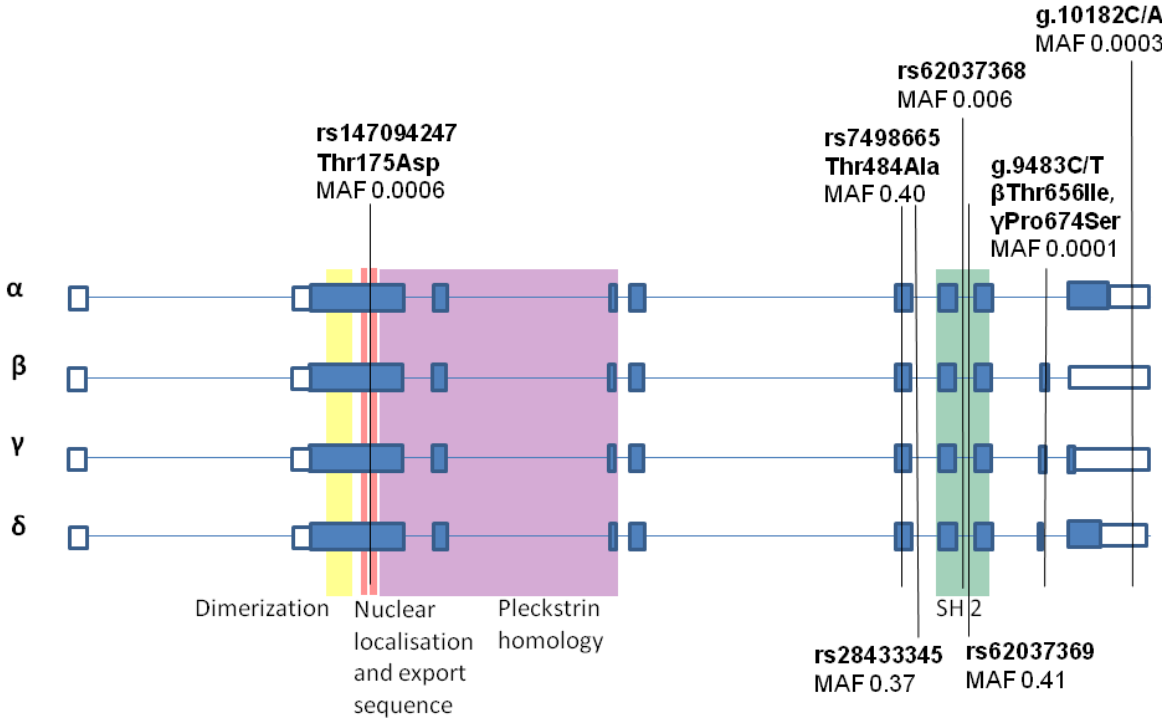

Figure 7: Identified variants in the four splice variants (α, β, γ and δ) of human *SH2B1.* *SH2B1* mRNA – coding parts as filled blocks – (Ensembl sequences α: ENST00000322610, β: ENST00000359285, γ: ENST00000337120 and δ: AF380422, extrapolated from the murine sequence as the human sequence is not available yet). The domain structure with dimerization (yellow), nuclear localization and export sequence (red), Pleckstrin homology (purple) and SH2 (green) domain is shown as underlying colored boxes (Quian and Ginty 2001, Doche et al. 2012). Positions of detected variants are marked with lines. Available rs-numbers, if applicable amino acid exchanges and minor allele frequencies in obese cases (MAF according to Table 21) are given for each variant (adapted from Volckmar et al. 2012).

Table 21: Frequencies of the detected SNPs and mutations in *SH2B1* excluding the screening group (Volckmar et al. 2012)

			Genotypes cases				Genotypes controls						
Position	**rs-Number**	**Amino acid exchange**	**11***	**12***	**22***	**Minor allele frequency cases [%]**	**11***	**12***	**22***	**Minor allele frequency controls [%]**	**Odds Ratio[d]**	**95% Confidence Interval**	**Nominal p-value**
g.2749C/A	rs147094247	Thr175Asp	11, 257	11	0	0.05	4,511	1	0	0.01	4.4	0.57 - 34.13	0.199[c]
g.8164A/G	rs7498665	Thr484Ala	512	1,526	1,101	40.62	58	195	181	35.83	1.2	1.06-1.42	**0.007**[a]
g.8250C/T	rs28433345		70	87	22	36.59	73	75	37	40.27	0.86	0.63 - 1.15	0.323[b]
g.8738A/G	rs62037368		176	2	0	0.56	182	3	0	0.81	0.69	0.11 - 4.16	1[b]
g.8764C/T	rs62037369		65	81	32	40.73	79	82	23	34.78	1.29	0.95 - 1.74	0.107[b]
g.9483C/T		βThr656Ile, γPro674Ser	11,206	2	0	0.01	4,506	0	0	0.00	NA	NA	1[c]
g.10182C/A			178	1	0	0.28	184	0	0	0.00	NA	NA	0.493[b]

[a] genotyped in 3,230 (extremely) obese cases and 439 lean controls

[b] genotyped in a total of 179 extremely obese cases and 185 lean controls

[c] genotyped in a total of 11,406 (extremely) obese and overweight cases and 4,568 (mainly population based) controls

[d] Odds ratio is given with respect to the minor allele.

*1 = the major allele, 2 = the minor allele (11 homozygous for wild type alleles, 12 heterozygous, 22 homozygous for mutant alleles)

All three identified *β*Thr656Ile/*γ*Pro674Ser mutation carriers were female. The initially detected mutation carrier 1 in the screening sample (age 12.7 years) and the mutation carrier 2 from the DAPOC collective (age 9.9 years) had a BMI > 99th percentile (Table 22). As depicted in Table 22, the overweight or obese mother (BMI 25.76 kg/m² and 32.61 kg/m², respectively) transmitted the mutation to the extremely obese child in both cases.

The third mutation carrier only had a BMI slightly above the 90th age and sex adjusted percentile (age 7.2 years, Table 22). For this mutation carrier, no genotypic information about the parents was available (mother BMI 19.81 kg/m², father BMI 26.7 kg/m²; Table 22). Unfortunately, extensive family analysis was not possible as additional family members were not recruited. Haplotype analysis of the risk alleles of the mutation *β*Thr656Ile/*γ*Pro674Ser (T allele) and rs7498665 (G allele) was not possible as the data on inheritance is limited (Table 22). A founder effect of the mutation is likely.

Table 22: Phenotypes and genotypes of mutation carriers of *β*Thr656Ile/*γ*Pro674Ser and their family members.

Individual	BMI [kg/m²]	BMI percentile (SDS)	*β*Thr656Ile/ *γ*Pro674Ser	rs7498665
Mutation carrier 1	32.44	>99th (2.75)	12	22
Mother 1	25.76		12	12
Father 1	27.17		11	12
Mutation carrier 2	26.38	>99th (2.50)	12	12
Mother 2	32.61		12	12
Father 2	28.54		NA	NA
Mutation carrier 3	18.64	91th (1.32)	12	NA
Mother 3	19.81		NA	NA
Father 3	26.70		NA	NA

The genotype for the lead SNP for BMI association rs7498665 was included for haplotype analysis. 1 = minor allele, 2 = major allele of each variant.

Due to SH2B1 interaction with the insulin receptor and the role of *Sh2b1* in a hyperinsulinemic mouse model (Morris et al 2008), phenotypic information was analyzed. Only for mutation carrier 2, insulin levels (9.4 mU/l) were available which were in the upper normal age range (<10 mU/l in children younger than 10 years, Levy-Marchal et al. 2010). The HOMA IR value (2.14) of the same mutation carrier was increased, but still does not depict an insulin resistance. In regard of the recent findings that variants in SH2B1 potentially lead to maladaptive behaviour including social isolation and aggression (Doche et al. 2012), psychological data on mutation carrier 3 were analyzed, who showed no abnormalities in the KINDL R test which is a method of measuring quality of life in children (Hullmann 2011), or in

additional tests for anxiety (SCARY in URMEL ICE; Nagel et al. 2009). These data were not available for the other two mutation carriers.

4.2.2 *In silico* functional analyses of SH2B1

On a molecular level, the amino acid exchanges in both affected splice variants βThr656Ile/γPro674Ser are located outside the domain structure relevant for SH2B1 function (Figure 8). While the outcome of the *γ*Pro674Ser exchange in the *γ* splice variant is predicted to be neutral by *in silico* programs, the exchange in the *β* splice variant *β*Thr656Ile was predicted to be functionally relevant in three of four programs (Table 23). The exchange in the *β* splice variant would destroy a predicted O-glycosylation site (53.5 % accuracy) of the SH2B1 protein. In contrast, amino acid conservation was stronger for *γ*Pro674Ser (100%) than for *β*Thr656Ile (86%) as shown in Table 20. Additionally, several potential changes for splicing enhancers and transcription factor binding sites were predicted for this exchange on the DNA sequence level (Table 24).

The second previously unknown mutation is located in the 3' UTR at base pair position g.10182 (C/A). The variant is not coding in any of the three splice variants. The mutation was detected twice within the screening sample, and once in an obese case in the association testing step, but it showed no association to obesity in a small case control comparison (p=0.49; Table 21). *In silico* analyzes predicted a possible change in splice sites and transcription factor binding sites for this variant (Table 24).

The obesity association of the previously described risk allele G at SNP rs7498665 (e.g. Thorleifsson et al. 2009, Willer et al. 2009) was confirmed in our 705 obesity trios (p=0.009) and in 3,139 independent obese or overweight cases and 434 normal- or underweight controls (p = 0.007, odds ratio (OR) = 1.22, 95% confidence interval (CI) 1.06-1.42; Table 21). Further analysis of the obesity association of this SNP was omitted as the association was very robustly replicated in several ethnicities (e.g. Thorleifsson et al. 2009, Willer et al. 2009). The coding common SNP rs7498665 results in the non-conservative amino acid exchange of a polar but uncharged threonine to an hydrophobic alanine at position 484 which shows low conservation (5%; Table 23). *In silico* analyses predicted no functional consequences for rs7498665 (Table 24). This reflects previous analyses (Jamshidi et al. 2007, Hotta et al. 2011).

The coding SNP rs147094247 leads to a non-synonymous exchange (Thr175Asp) of two amino acids with polar, uncharged side chains. The SNP showed no obesity association in a sample of 11,268 obese and overweight cases and 4,512 lean or normal weight controls (p = 0.199, odds ratio (OR) = 4.4, 95% confidence interval (CI) 0.57 - 34.13; Table 21).The

position of the exchange is conserved (71%; Table 23), but *in silico* methods predicted a neutral outcome for this SNP (Table 24).

The three remaining detected SNPs are non-coding and show no evidence for obesity association in a small case control sample of 179 obese cases and 185 lean controls (Table 21). For the SNP rs28433345 (c.1513+23T/C, MAF: 0.316), *in silico* analyses predicted a potential splicing enhancer site change. The SNP rs62037368 (c.1726-40G/A, MAF: 0.004) is predicted to change potential splicing enhancer binding sites and transcription factor binding sites. The last SNP was rs62037369 (c.1726-14C/T, MAF: 0.218) for which *in silico* analyzes predicted splice site or transcription factor binding site changes (namely GATA1 and 2, MZF1; Table 24).

Table 23: *In silico* functional prediction of detected non-synonymous mutations in *SH2B1* (adapted from Volckmar et al. 2012)

		PolyPhen-2			SNAP			PMUT		Mutation Taster		Panther	
Amino acid changes	**DNA position**	*Delta_Score*	*Prediction*	*RI*	*Expected Accuracy (%)*	*Prediction*	*Score*	*Reliability*	*Prediction*	*Prediction*	*Probability*	*Pred.*	**Conservation**
Thr175Asp	g.2749C/A	0.000	Benign	0	53	neutral	0.2582	4	neutral	neutral	0.6181	0.35568	71%
Thr484Ala	g.8164A/G	0.219	Benign	4	85	neutral	0.3210	3	neutral	neutral	0.9999	0.12427	5%
βThr656Ile	g.9483C/T	0.107	Benign	1	63	not neutral	0.9386	8	pathological	disease causing	0.9992	NA*	86%
γPro674Ser	g.9483C/T	0.038	Benign	0	53	neutral	0.2451	5	neutral	disease causing	0.9992	NA*	100%

*NA: Variants that are not present in the alpha splice variant of the gene cannot be analyzed with Panther.

Conservation was analyzed as percentage of species (given in the materials and methods section) carrying the same amino acid on the position of the exchange. As not all species express all splice variants, 21 species were analyzed for Thr175Asp and Thr484Ala, 8 species for βThr656Ile and 6 species for γPro674Ser (for details, see Methods).

Table 24: *In silico* prediction of splice sites, transcription factor binding sites and o-glycosylation sites of detected variants in *SH2B1* (adapted from Volckmar et al. 2012)

PROGRAM			ESEfinder	ESRSearch	RESCUE_ESE	TFSearch	Consite	OGPET	Mutation Taster	
SNPs	Amino acid changes	DNA position	*Splice sites*	*Splice sites*	*Splice sites*	*transcription factor binding sites*	*transcription factor binding sites*	*O-glycosylation site prediciton [%]*	*Prediction*	*Prob.*
rs147094247	Thr175Asp	g.2749C/A	changed	changed	not changed	not changed	not changed	33.1829		
rs7498665	Thr484Ala	g.8164A/G	changed	changed	not changed	not changed	not changed	---		
rs28433345	---	g.8250C/T	changed	SRp40, FOX1-FOX2[a]	not changed	not changed	not changed	---		
rs62037368	---	g.8738A/G	changed	not changed	not changed	not changed	COUP-TF, c-REL[b]	---		
rs62037369	---	g.8764C/T	changed	changed	not changed	GATA1, GATA2, MZF1[b]	not changed	---		
g.9483C/T	βThr656Ile, γPro674Ser	g.9483C/T	changed	not changed	not changed	GATA1, GATA2[b]	not changed	53.5329		
g.10182C/A	---	g.10182C/A	changed	PESS[a]	not changed	not changed	not changed	---		

[a] particular changed predicted splice sites, [b] particular changed predicted transcription factor binding sites

4.2.2 SH2B1 *in vitro* studies

In vitro data on the variants detected in our mutation screen of the coding region of *SH2B1* in 95 extremely obese children and adolescents was not available when we initiated our functional analyses. For rs7498665 [Thr484Ala], a low potential to have a functional effect on SH2B1 was predicted (Jamshidi et al. 2007, Volckmar et al. 2012). Leptin signalling is the most prominent mediator of energy homeostasis (Morris and Rui 2010). Hence, we analyzed the impact on leptin signalling *in vitro* using a luciferase reporter assay of rs7498665 (484Ala allele) due to the association with obesity and the mutation alleles β656Ile/γ674Ser with *in silico* predicted functional impact.

The heterologous cell system of the Human Embryonic Kidney 293 cells (HEK293) was chosen, as these cells are known for straightforward transfection. HEK293 intrinsically express SH2B1, but lack the long form of the leptin receptor (LEPRb) that functions in energy homeostasis. We therefore transfected the cells with LEPRb and the SH2B1 wild type clones β and γ. The two splice variants were chosen as the variant rs7498665 [Thr484Ala] affects all splice variants, but the mutation βThr656Ile/γPro674Ser with higher *in silico* potential for functional changes of SH2B1 is only present in SH2B1 β and γ.

Transfection of HEK293 cells with the wild type human SH2B1 clones increased leptin response for both splice variants (Figure 8, Table S6). This indicates that the combination of the heterologous cell system and the STAT3 assay allows characterization of SH2B1 function. This is the first analysis on the interaction of leptin receptor and human SH2B1 splice variants. In the murine model, only the α splice variant was tested for the interaction with leptin signalling (Ren et al. 2005, 2007).

Although only the self-dimerization and SH2 binding domain are known to be involved in leptin signalling (Rui et al. 2007), potential folding changes of the protein because of βThr656Ile/γPro674Ser could impact leptin signalling. We therefore tested the influence of the obesity risk alleles in both splice variants on STAT3 mediated leptin signalling *in vitro*. For both splice variants β656Ile and γ674Ser, there was no significant effect on leptin signalling detectable. Both Emax and EC50 were non-significantly reduced for the risk allele in both β and γ splice variants of SH2B1 (Figure 8, Table S6) which could indicate both a loss of function by the decreased maximal activation capacities as well as a reduced function since more leptin is necessary to elicit the same signal at half maximum capacity.

When tested for its impact on STAT3 mediated leptin signalling *in vitro*, the obesity risk variant of rs7498665 484Ala showed unaltered leptin signalling for both splice variants β and γ. In both β and γ splice variants of SH2B1, Emax and EC50 were non-significantly reduced for the risk allele G (Figure 8, Table S6). Again, this could hint to both a gain of function and

reduced function. These findings reflect the *in silico* prediction from our own study (Volckmar et al. 2012) and others (Jamshidi et al. 2007).

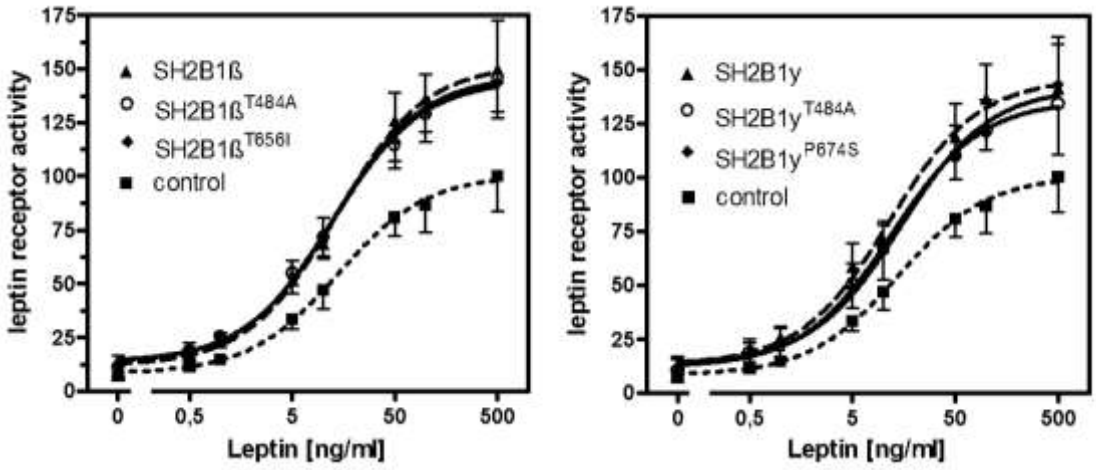

Figure 8: Dose-response curve of STAT3 mediated leptin signalling with (mutated) *SH2B1*. HEK293 cells were co-transfected with the long form of the leptin receptor (LEPRb), a STAT3 responsive element and SH2B1 splice variants beta (left) and gamma (right) with and without the risk/effect alleles at rs7498665 (Thr484Ala) and βThr656Ile/γPro674Ser (n=8 separate experiments). The dose response curves show LEPRb activity after stimulation with leptin (exact values for each data point see Table S6; Volckmar et al. 2012).

4.3 APOB48R

The second obesity candidate gene in the chromosomal region 16p11.2 is *APOB48R* which encodes for the apolipoprotein 48 receptor. Unlike *SH2B1*, for which obese animal knockout models exist (Rui et al. 2005) and direct interaction with pathways involved in energy homeostasis is known (Morris and Rui 2010), data on *APOB48R* is scarce. The macrophage receptor APOB48R regulates fat uptake into cells like adipocytes (Brown et al. 2000). Higher macrophage lipid uptake is associated with inflammation of fat tissue which is a symptom of the metabolic syndrome (Lumeng et al. 2007). Indeed, variants in APOB48R are also associated with hypercholesterolemia (Fujita et al. 2007) which contributes to higher risk for cardiac diseases. Hence, we analyzed the gene for coding variants with potential obesity association.

In the mutation screen of *APOB48R* in the 95 extremely obese individuals of the initial screening sample, we detected a total of 13 variants (Table 26; Graninger 2011, Figure 9). Three of them were non-synonymous coding variants; the SNP rs180743 (Pro419Ala) and the InDels p.Ser323delSerGlyGlyGluGluAlaGlyThrAla and rs3833080 (p.Ala345_Gly346insdelAlaGlyThrAlaSerGlyGlyGluGluAlaGly).

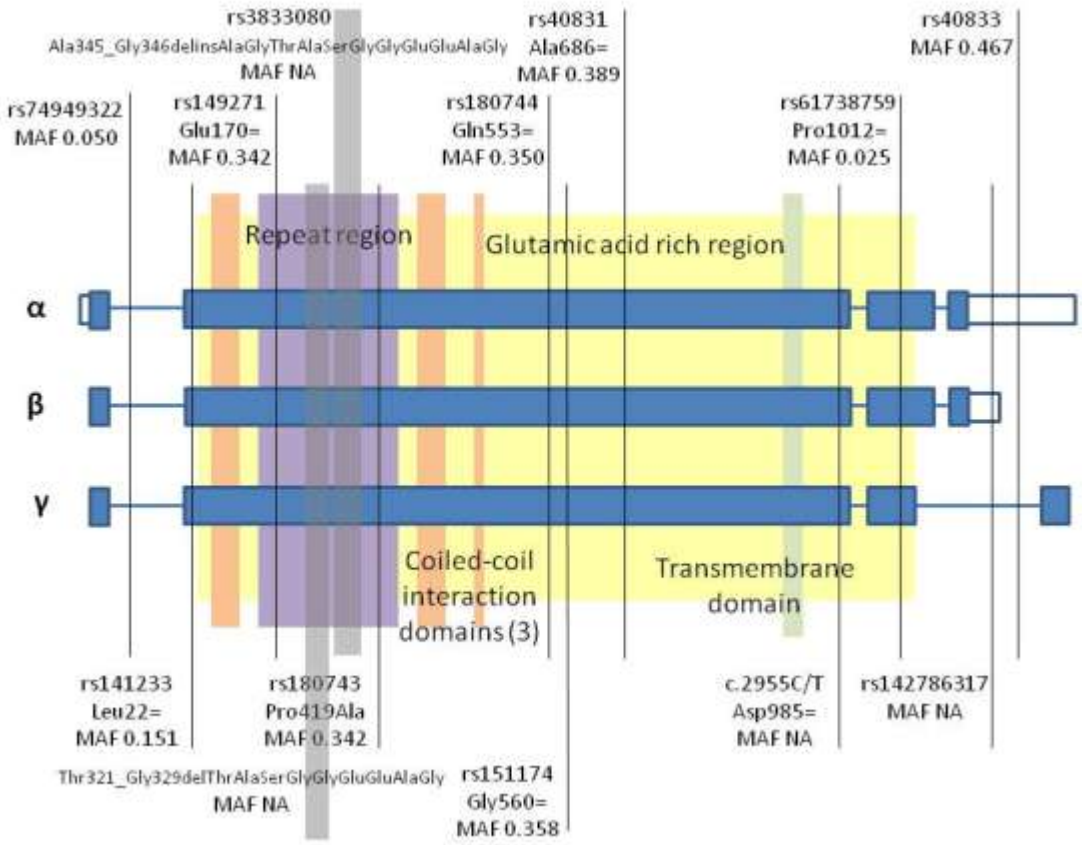

Figure 9: **Identified variants in the three splice variants (α, β, γ) of human *APOB48R*.** *APOB48R* mRNA – coding parts as filled blocks – (Ensembl sequences α: ENST00000564831, β: ENST00000431282 and γ: ENST00000328423). The functional domains – glutamic acid rich region (yellow), three coiled-coil interaction domains (orange), the glycin-rich repeat region (purple) and the transmembrane domain (green) - are shown as underlying boxes (Brown et al. 2000). Positions of detected variants (partially from Graninger 2011) are marked with lines or in case of deletions or insertions as wide bars. Available rs-numbers, if applicable amino acid exchanges and minor allele frequencies (MAF according to CEU) are given for each variant. NA = not available.

The deletion Del1 (Ser323delSerGlyGlyGluGluAlaGlyThrAla) was detected only once among 2,179 obese cases and 435 normal weight or lean controls. The mutation was transmitted from the obese mother (BMI 38.23 kg/m²) to the extremely obese female mutation carrier (height 169 cm, weight 104.7 kg, BMI 36.66 kg/m², age 13.1 years, BMI > 99th age and sex specific percentile). The normal weight father (BMI 24.77 kg/m²)) did not harbor the deletion. The carrier of Del1 was also homozygous for risk alleles of rs180743 [Pro419Ala] and rs3833080 [Ala345_Gly346insdelAlaGlyThrAlaSerGlyGlyGluGluAlaGly]. The deletion Del1 is located within a glutamic acid-rich region (AA 158-970) which covers most of the protein and the glycin-rich repeat region (AA 204-369). Here, it deletes several amino acids with polar uncharged side chains and amino acids with hydrophobic side chains. As the variant is not located in any of the functional domains of APOB48R (Brown et al. 2000), it is hard to predict a functional effect of this variant (Table 26). The *in silico* prediction programs used do not evaluate deletions other than Mutation Taster which predicted this variant to be “disease

causing" which is equal to functionally relevant (Table 25). Rarity of this variant prevented an obesity association analysis.

The previously unknown variant c.2955C/T is a synonymous coding SNP (Asp985=). Here, the mutation was again transmitted from the obese mother (BMI 38.23 kg/m²) to the extremely obese daughter (height 169 cm, weight 104.7 kg, BMI 36.66 kg/m², BMI SDS 3.11, age 13.1 years). This variant is predicted (Mutation Taster) to lead to a potential new splice site. Changes in splicing enhancer or silencer as well as transcription factor binding sites are also predicted *in silico* (Table 26).

The second deletion Del2 (p.Ala345_Gly346insdelAlaGlyThrAlaSerGlyGlyGluGluAlaGly; rs3833080) is also located in the repeat region of *APOB48R*. The repeated motif is [GlyGlyGluGluAlaGluThrAlaSer] which is repeated 3 times exactly and 8 times with small changes of the sequence. Therefore, it is impossible to determine the exact position of the deletion (Figure 10). One of the repeats in the sequence is intronic, hence the shortened DNA sequence does not have to be coding if the intron is not spliced and replaces the coding sequence. On the other hand, dbSNP lists a coding variant (rs3833080) of the same nucleotide sequence and MAF in the repetitive region which is most likely our variant Del2.

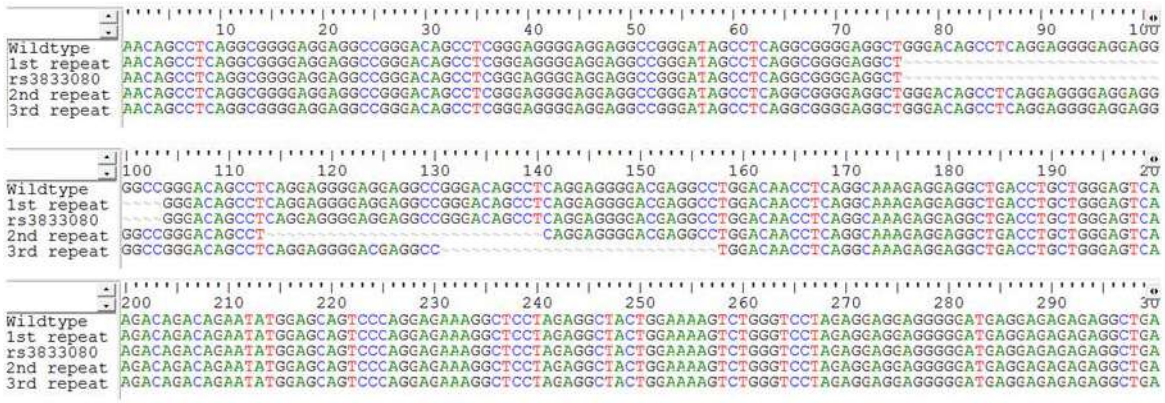

Figure 10: **DNA sequence of APOB48R at the position of Del2 (rs3833080; Volckmar et al. in prep.)**. Several potential positions of the deletion are displayed below the wild type sequence of the repeat region. The sequence marked with rs3833080 is the taken from the SNP database of the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/snp/).

Like Del1, Del2 lies within the glutamic acid-rich region and the glycin-rich region of APOB48R, but not in any functional domain. Mutation Taster, the only *in silico* prediction tool for deletions used here, predicted this deletion as a non functional polymorphism (Table 25). While genotyping Del2 in our samples, we also detected a known insertion at the same position (p.Ala345_Gly346insAlaGlyThrAlaSerGlyGlyGluGluAlaGly; rs3833080) with a much lower frequency and exclusively in (extremely) obese cases (MAF 0.001 in 2,540 cases) but not in 481 lean or normal weight controls.

The common (MAF 0.38) polymorphism Pro419Ala (rs180743) locates in the first exon of *APOPB48R* in high LD (r^2 = 0.98 in cc and trios study) with the deletion 2. The variant exchanges a hydrophobic alanine to an aromatic proline in a position that is conserved (conservation 66% over 29 species, ENSEMBL). Functional *in silico* predictions for this SNP vary; they predict the SNP to be deleterious (PolyPhen2), or neutral (SIFT; Table 25). One *in silico* program predicted the introduction of a new splice site by this variant and a disruption of the glutamate-rich region and a potential loss of a phosphoserine if the introduced splice site is used (Mutation Taster). The SNP showed association to hypercholesterolemia in the study of Fujita et al. (2005).

Besides the InDel variants, seven synonymous variants (rs151233 [Leu22=], rs149271 [Glu170=], rs180744 [Gln553=], rs151174 [Gly560=], rs40831 [Ala686=], and rs61738759 [Pro1012=]; Table 26) and three intronic variants (rs74949322, rs40833 and rs142786317; Table 26) were detected. *In silico* analysis predicted a potential change of splice site enhancers or silencers and/or transcription factor binding sites for all of these variants (Table 26).

For two of the non-synonymous variants (rs180743 [Pro419Ala] and rs3833080 [Ala345_Gly346delAlaGlyThrAlaSerGlyGlyGluGluAlaGly]), obesity association analyses in a sample of 453 extremely obese children and adolescents and 435 normal weight or lean adult controls were performed (Graninger 2011, Volckmar et al. in prep.). Both variants did not deviate from Hardy-Weinberg equilibrium in the controls (p = 0.411 for rs180743 [Pro419Ala]; p = 0.221 for rs3833080 [Ala345_Gly346delAlaGlyThrAla-SerGlyGlyGluGluAlaGly]), and showed nominal obesity association in the case control sample (p = 0.002 for rs180743 [Pro419Ala]; p = 0.003 for rs3833080 [Ala345_Gly346delAlaGlyThrAlaSerGlyGlyGluGluAlaGly]). These findings were replicated in a family based association study in 615 independent family trios comprising an extremely obese child or adolescent and both biological parents, where the risk alleles (G and deletion allele) of both variants were strongly (nominally) associated with obesity (p = 0.003 for rs3833080 [Ala345_Gly346delAlaGlyThrAlaSerGlyGlyGluGluAlaGly]; p = 0.002 for rs180743 [Pro419Ala]; Table 27 adapted from Graninger 2011, Volckmar et al. in prep.). During genotyping of rs3833080 [p.Ala345_Gly346delAlaGlyThrAlaSerGlyGlyGluGluAlaGly], we also detected the known insertion at the same position (Ala345_Gly346insAlaGlyThrAlaSerGlyGlyGluGluAlaGly; rs3833080) exclusively in (extremely) obese cases (MAF 0.001 in 2,540 cases). The same *in silico* functional predictions for the deletion also apply for the insertion, as it possibly elongates the intron or increases the repeat region by one repeat.

Table 25: *In silico* functional prediction of detected non-synonymous mutations in *APOB48R* (adapted from Graninger 2011, Volckmar et al. in prep.).

		PolyPhen-2		SNAP			PMUT			Mutation Taster		PANTHER
Amino acid changes	**DNA position**	*Delta_Score*	*Prediction*	*RI*	*Expected Accuracy (%)*	*Prediction*	*Score*	*Reliability*	*Prediction*	*Prediction*	*Probability*	*Prediction*
Thr321_Gly329delThrAlaSerGlyGlyGluGluAlaGly	c.961_996delACAGCCTCAGGCGGGGAGGAGGCCGGGACAGCCTC?*	NA			NA			NA		Disease causing	0.72	NA
Ala345_Gly346insdelAlaGlyThrAlaSerGlyGlyGluGluAlaGly	g.28507397_28507398insdelGGGACAGCCTCAGGAGGGGAGGAGGCC	NA			NA			NA		Polymorph.	1	NA
Pro419Ala	c.1255C/G	0.600	possibly damaging	3	0.78	Non-neutral	0.600	1	PATHOLOGICAL	Polymorph.	1	0.59057

Variants colored in grey were detected by Graninger 2011 and are listed here only for a complete overview.
* exact position of the deletion within the repetitive region cannot be determined, hence the questionnaire in the nomenclature as recommended by Antonarakis (www.hgmd.cf.ac.uk/docs/mut_nom.html)

Table 26: Detected variants and *in silico* analysis of potential functional changes in *APOB48R* (adapted from Graninger 2011, Volckmar et al. in prep.).

PROGRAM			ESEfinder	ESRSearch	RESCUE_ESE	TFSearch	Consite	OGPET	Mutation Taster	
SNPs	**Amino acid changes**	**DNA position**	*Splice sites*	*Splice sites*	*Splice sites*	*transcription factor binding sites*	*transcription factor binding sites*	*O-glycosylation site prediciton [%]*	*Prediction*	*Prob.*
rs74949322	-	c.57+50C/T	changed	changed	not changed	not changed	changed	---	Polymorp.	0.73
rs151233	Leu22=	c.75C/T	not changed	not changed	not changed	not changed	not changed	---	Polymorp.	0.71
rs149271	Glu170=	c.520A/G	changed	changed	changed	not changed	not changed	---	Polymorp.	0.71
-	Thr321_Gly329delThrAlaSerGlyGlyGluGluAlaGly	c.961_996delACAGCCTCAGGCGGGGAGGAGGCCGGGACAGCCTC?*	NA	NA		changed	changed	---	Polymorp.	0.73
rs3833080	Ala345_Gly346insdelAlaGlyThrAlaSerGlyGlyGluGluAlaGly	g.28507397_28507398insdelGGGACAGCCTCAGGAGGGGAGGAGGCC	NA	NA		changed	changed	---	Polymorp.	1
rs180743	Pro419Ala	c.1255C/G	changed	not changed	not changed	changed	changed	---	Polymorp.	1
rs180744	Gln553=	c.1659A/G	changed	changed	changed	not changed	not changed	---	Polymorp.	0.71
rs151174	Gly560=	c.1680C/T	changed	changed	changed	not changed	possible	---	Polymorp.	0.72
rs40831	Ala686=	c.2058A/G	changed	not changed	not changed	changed	changed	---	Polymorp.	0.71
-	Asp985=	c.2955C/T	changed	changed	changed	not changed	possible	---	Polymorp.	0.71
rs61738759	Pro1012=	c.3036G/A	changed	changed	changed	not changed	not changed	---	Polymorp.	0.72
rs40833	-	c.*218C/G[a]	changed	not changed	changed	changed	changed	---	Polymorp.	0.73
rs142786317	-	c.*118_*119delCA	changed	changed	changed	changed	changed	---	Polymorp.	0.71

Variants colored in grey were detected by Graninger 2011 and are listed here only for a complete overview.
* exact position of the deletion within the repetitive region cannot be determined, hence the questionnaire in the nomenclature as recommended by Antonarakis (www.hgmd.cf.ac.uk/docs/mut_nom.html)

73

Table 27: Association analyses of polymorphism in *APOB48R* in extremely obese children and adolescents and lean controls (Graninger 2011, Volckmar et al. in prep.).

Position / rs-Number	Amino acid change		Genotypes			Allelic association		Genotypic association	
			11[a]	12[a]	22[a]	MAF [%]	OR (95%CI)	Nominal p-value in Case-Control sample[#]	Nominal p-value in 705 Trios[*]
rs180743	Pro419Ala	Control	169	193	66	37.97	1.00		
		Case	939	1,370	574	43.67	1.27(1.09-1.47)	0.002	0.002
rs3833080	Ala345_Gly346delinsAla GlyThrAlaSerGlyGlyGlu GluAlaGly	Control	170	191	69	38.26	1.00		
		Case	855	1,213	527	43.68	1.25(1.08-1.45)	0.003	0.003

Variants colored in grey were detected by Graninger 2011 and are listed here only for a complete overview.
[#]sample of 453 extremely obese children and adolescents and 435 normal weight or lean adult controls
[*]615 family trios comprising an extremely obese child or adolescent and both biological parents, independent of the initial screening sample

MAF, minor allele frequency; OR, odds ratio; 95%CI, 95% confidence interval.
[a] 1 = the major allele, 2 = the minor allele at each position.

74

4.4 Impact of variants in SH2B1 and APOB48R on weight loss parameters

The mutation screen of *SH2B1* and *APOB48R* revealed two common SNPs which showed nominal association to obesity. While we detected no functional impact on STAT3 mediated leptin signalling for rs7498665 [Thr484Ala] (Volckmar et al. 2012), the SNP shows robust association with BMI and obesity (e.g. Willer et al. 2009, Thorleifsson et al. 2009, Speliotes et al. 2010, Volckmar et al. 2012). SH2B1 is involved in several pathways that regulate weight (Morris and Rui 2010), the knockout mouse model shows hyperphagia, obesity and hyperinsulinemia. We are the first group to describe a BMI association signal for variants in *APOB48R*. The obesity associated ($p = 0.002$ in a sample of 610 trios and 453 independent extremely obese cases and 434 normal- or underweight controls; Graninger 2011, Volckmar et al. in prep.) SNP rs180743 [Pro419Ala] was previously shown to be associated with hypercholesterolemia (Fujita et al. 2005).

Here, we analyzed the impact of the two coding variants in the obesity candidate genes *SH2B1* and *APOB48R* in the chromosomal region 16p11.2 on the change of anthropometric parameters during a 1-year lifestyle intervention. Special focus was a potential differentiation of risk allele carriers by insulinemic or blood fat parameters. We therefore measured total cholesterol, LDL-cholesterol, HDL-cholesterol, triacylglycerides, glucose, insulin, and HOMA. According to the hypothesis, risk allele carriers (G) at rs7498665 [Thr484Ala] in *SH2B1* should have higher insulin levels and decreased glucose tolerance. Carriers of the risk allele (G) at rs180743 [Pro419Ala] in *APOB48R* should be prone to increased total cholesterol and triacylglycerides levels in serum.

After completion of the lifestyle intervention “Obeldicks”, both BMI and BMI-SDS were on average lower (data not shown; explorative $p < 0.05$, $n = 454$). The mean systolic blood pressure of the subjects was also decreased, and the mean level of HDL-cholesterol was increased; both predict a lower risk of infarction and dyslipidemia (explorative $p < 0.05$, $n = 454$).

When analyzing the impact of the coding SNPs rs7498665 (*SH2B1*, Thr484Ala) and rs180743 (*APOB48R*, Pro419Ala) on baseline data and changes in anthropometrics (body-mass index (BMI), and BMI standard deviation score (BMI-SDS)), blood pressure (systolic and diastolic) and plasma parameters (total cholesterol, LDL-cholesterol, HDL-cholesterol, triacylglycerides, glucose, insulin, and HOMA IR), we did not find an effect of the genotypes on any of the analyzed phenotypes (additive mode of inheritance as suggested by Speliotes et al. (2010); Tables 28 and 29; Volckmar et al. 2013). The study group of 454 overweight and obese children and adolescents has a power of 0.8 to detect a change of 0.2 BMI-SDS changes per risk allele (SNP with MAF 0.38, p-value= 0.05, two-sided) under the additive mode of inheritance suggested by Speliotes et al. (2010).

Table 28: Changes in anthropometric and plasma variables for rs7498665 (Thr484Ala) at *SH2B1* in 454 obese children or adolescents undergoing a 1-year intervention (Volckmar et al. 2013).

	AA n=138		AG n=217		GG n=99		
	T0	**T1**	**T0**	**T1**	**T0**	**T1**	**p-value#**
BMI (kg/m²)*	27.3 ± 4.4	26.7 ± 4.8	27.7 ± 4.7	27.1 ± 4.7	27 ± 3.7	26.6 ± 4	0.772
BMI-SDS*	2.4 ± 0.5	2.1 ± 0.6	2.4 ± 0.5	2.1 ± 0.6	2.3 ± 0.5	2 ± 0.6	0.947
RR dias (mm Hg)*	65.2 ± 12.2	64.7 ± 12.7	66.1 ± 12.8	64.6 ± 12.6	65 ± 12.4	66 ± 13	0.515
RR sys (mm Hg)*	115.6 ± 17.4	112 ± 16.7	116.5 ± 17.1	112.7 ± 15.6	113.3 ± 18.1	111.3 ± 16.1	0.996
TAG (mg/dl)	105.3 ± 48.5	102.8 ± 52.2	108.2 ± 64.8	104 ± 53.7	102.7 ± 52.1	98 ± 50	0.681
TC (mg/dl)	171.1 ± 33.6	170.4 ± 34.6	171.8 ± 31.4	170.1 ± 33.2	172.1 ± 38.5	171 ± 34.5	0.840
LDL-cholesterol (mg/dl)	109.5 ± 98.9	98.5 ± 36.1	101 ± 35.9	97.3 ± 34.2	105 ± 38.7	99.6 ± 32.9	0.390
HDL-cholesterol (mg/dl)*	48.8 ± 13.3	51.5 ± 13.8	48.6 ± 14.8	50.6 ± 15.1	50.1 ± 13.4	51.1 ± 14.2	0.161
Glucose (mg/dl)*	83.6 ± 17.6	84.7 ± 16.3	84.9 ± 13.3	89.3 ± 37.5	82 ± 20.7	83.2 ± 18.7	0.914
Insulin (mU/l)	14.5 ± 8.9	15 ± 9.9	15.3 ± 14.5	16.2 ± 11.9	14.4 ± 10.1	15.3 ± 8.6	0.877
HOMA*	2.7 ± 3.1	4 ± 3.5	2.7 ± 4.6	4.6 ± 4.1	2.6 ± 4	4.2 ± 3.6	0.569

Values are mean ± SD; BMI-SDS: BMI standard deviation score; HOMA: Homeostasis Model Assessment; TAG: triacylglycerol; TC: total cholesterol, * nominal $p \leq 0.05$ different in whole group compared to baseline # For the Wald test statistics of the linear regression with untransformed outcomes and sex, age, and baseline measurement as covariates

Table 29: Changes in anthropometric and plasma variables for rs180743 (Pro419Ala) at *APOB48R* in 454 obese children or adolescents undergoing a 1-year intervention (Volckmar et al. 2013).

	CC n=146		CG n=206		GG n=102		
	T0	**T1**	**T0**	**T1**	**T0**	**T1**	**p-value#**
BMI (kg/m²)*	27.6 ± 4.7	27 ± 4.9	27.4 ± 4.5	26.9 ± 4.7	27.1 ± 3.8	26.7 ± 4	0.983
BMI-SDS*	2.4 ± 0.5	2.2 ± 0.6	2.3 ± 0.5	2.1 ± 0.6	2.3 ± 0.5	2.1 ± 0.6	0.881
RR dias (mm Hg)*	65.6 ± 12.4	65 ± 12.6	65.6 ± 12.6	64.6 ± 12.5	65.4 ± 12.6	65.4 ± 13.4	0.833
RR sys (mm Hg)*	116.9 ± 17.6	113.3 ± 18	115.3 ± 16.7	112.4 ± 14.5	114 ± 18.5	110.3 ± 15.8	0.206
TAG (mg/dl)	105.1 ± 48.1	101.5 ± 48.4	106.4 ± 61.1	102.9 ± 50	107.1 ± 62.8	102.3 ± 62.3	0.985
TC (mg/dl)	173 ± 34	172.3 ± 35.3	170.9 ± 30.8	169.3 ± 31.8	171.2 ± 38.5	169.8 ± 35.9	0.814
LDL-cholesterol (mg/dl)	109.3 ± 96.8	99.8 ± 37.2	101.4 ± 36.1	96.9 ± 33.2	103.6 ± 37.4	98.4 ± 33.2	0.977
HDL-cholesterol (mg/dl)*	49 ± 14.3	51.6 ± 14.5	48.6 ± 14	50.9 ± 14.6	49.8 ± 13.8	50.5 ± 14.4	0.109
Glucose (mg/dl)*	84.1 ± 17.3	85.1 ± 16	84.9 ± 13.8	89.3 ± 38.4	81.6 ± 19.8	83.1 ± 18.6	0.895
Insulin (mU/l)	14.5 ± 9.4	15.3 ± 10.4	15.1 ± 14.4	16 ± 11.8	15 ± 10.6	15.5 ± 8.3	0.643
HOMA*	2.7 ± 3.3	4.2 ± 3.8	2.7 ± 4.7	4.4 ± 4	2.7 ± 3.6	4.3 ± 3.7	0.377

Values are mean ± SD; BMI-SDS: BMI standard deviation score; HOMA: Homeostasis Model Assessment; TAG: triacylglycerol; TC: total cholesterol, *nominal $p \leq 0.05$ different in whole group compared to baseline # For the Wald test statistics of the linear regression with untransformed outcomes and sex, age, and baseline measurement as covariates

4.5 SULT1A2

In addition to *SH2B1* and *APOB48R,* the gene family of the sulfotransferases harbors coding SNPs in high LD with the BMI associated lead SNP rs7359397 (Speliotes et al. 2010). The sulfotransferase SULT1A2 sulfonates hormones like estrogen and several androgens. This leads to excretion of the hormones and lowers for example levels of estrogenic alkylphenols and 17-β-estradiol (Harris et al. 2000). Obesity is associated with higher levels of 17-β-estradiol, estron, and estron sulfate which are substrates of SULT1A2 (Ghose et al. 2011). A non-significant association of a coding SNP (rs1059491 [Asn235Ser]) in a small study group of 300 individuals was described previously (Glatt et al. 2002). Hence, we deem SULT1A2 an interesting obesity candidate gene.

In the mutation screen of the coding region of the sulfotransferase *SULT1A2* in 95 extremely obese children and adolescents, we detected seven known non-synonymous SNPs (rs4149404 [Ile7Thr], rs10797300 [Pro19Leu], rs145008170 [Ser44Asn], rs4987024 [Tyr62Phe], rs142241142 [Ala164Val], rs1059491 [Asn235Thr], rs75191166 [Lys282Gly]), two known synonymous SNPs (rs1690407 [Ser8=], rs139896537 [Ala164=]) and five non-coding variants (rs4149406, rs3743963, rs710410, rs762634, rs145790611; Table 31, Figure 11).

The coding SNP rs1136703 [Ile7Thr] has long been suggested as a functionally relevant variant (Zhu et al. 1996, Raftogianis et al. 1999, Glatt et al. 2004). All *in silico* analyses predicted no functional effect of this variant (Table 31). For rs10797300 [Pro19Leu], in silico programs predict a functional outcome for this variant *in silico* (PolyPhen2 and PANTHER, Table 30). The polymorphism rs145008170 [Ser44Asn] locates close to the amino acid Lys48, which is relevant for binding of the xenobiotic p-nitrophenol to the binding pocket of SULT1A2 (Lu et al. 2010). *In silico* analyses predict it to be “non-neutral” (SNAP) or “deleterious” (PANTHER; Table 30).

The rare coding variants rs4987024 [Tyr62Phe] and rs142241142 [Ala164Val] are both conservative amino acid exchanges that are not located close to the binding pocket. *In silico* prediction shows a higher probability of functional changes for Ala164Val, although the analyses revealed mixed results (Table 31). *In silico* prediction mostly interpreted the SNP rs1059491 [Asn235Thr] as functionally relevant (“probably damaging” PolyPhen2, “non-neutral” SNAP and “pathological” PMUT; Table 30).

For the two known synonymous SNPs (rs1690407 – Ser8=, rs139896537 – Ala164=) and the five non-coding variants (rs4149406, rs3743963, rs710410, rs762634, rs145790611), *in silico* prediction vary. For each variant, at least the change in either transcription factor binding

sites or splicing enhancer or silencer sites was predicted (Table 31). No variant disrupted a glycosylation site.

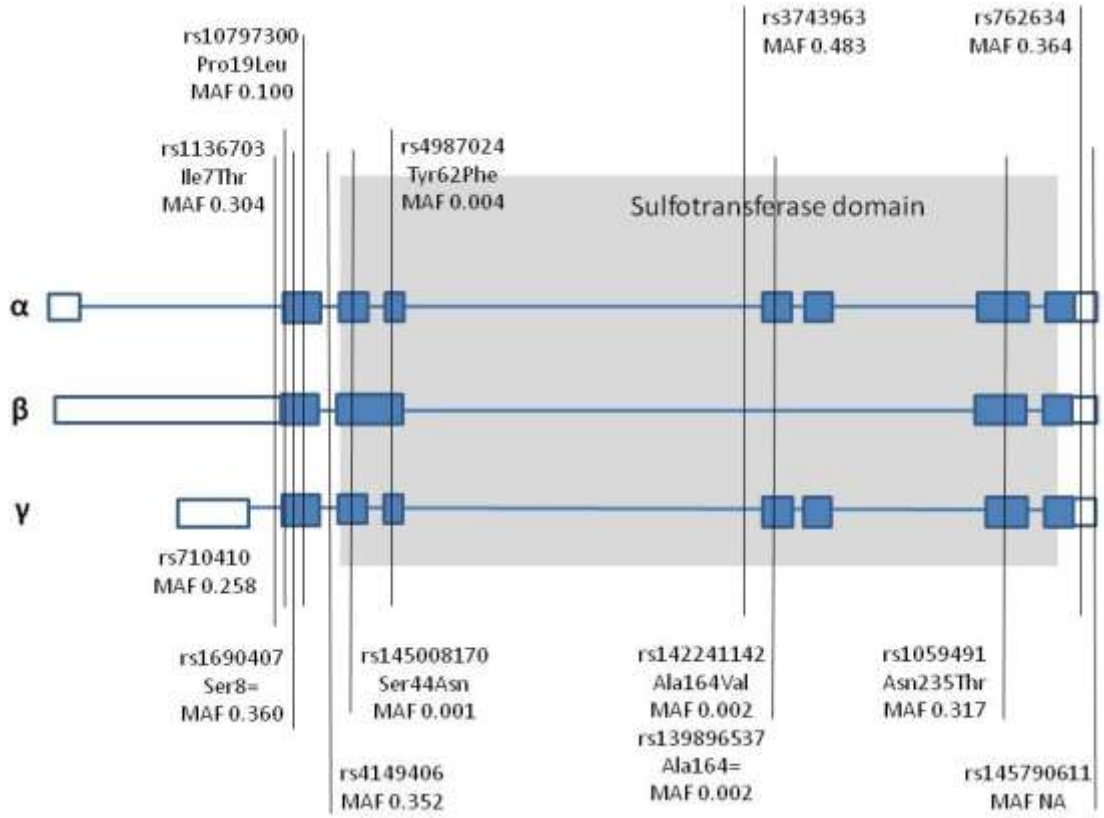

Figure 11: Identified variants in the three splice variants (α, β, γ) of human *SULT1A2.* *SULT1A2* mRNA – coding parts as filled blocks – (Ensembl sequences α: ENST00000335715, β: ENST00000533150 and γ: ENST00000395630). The sulfotransferase domain is shown as underlying grey box (Lu et al. 2010). Positions of detected variants are marked with lines. Available rs-numbers, if applicable amino acid exchanges and minor allele frequencies (MAF according to CEU), are given for each variant. NA = not available.

The coding SNPs detected in *SULT1A2* were analyzed for obesity association in a sample of 355 obesity trios (extremely obese child or adolescent with both biological parents; Table 32). The SNP rs1059491 [Asn235Thr] displayed a high mendelian error rate of 6.84% which is in contrast to previous GWAS analyses of the same study group (Hinney et al. 2006, Scherag et al. 2010). For the variant rs145008170 [Ser44Asn], the risk allele could not be determined as the minor allele frequency of the variant is very low in the sample (MAF = 0.19%). Only two children and adolescents and two respective parents carried one copy of the variant.

Table 30: *In silico* functional prediction of detected non-synonymous mutations in *SULT1A2*

		PolyPhen-2		SNAP			PMUT		Mutation Taster		PANTHER
Amino acid changes	**DNA position**	*Delta_Score*	*Prediction*	*RI*	*Expected Accuracy (%)*	*Prediction*	*Score*	*Reliability*	*Prediction*	*Prediction*	*Probability*	*Prediction*
Ile7Thr	c.20T/C	0	benign	1	0.6	Neutral	0.0641	8	neutral	Polymorphism	1	0.09179
Pro19Leu	c.56C/T	1	probably damaging	0	0.53	Neutral	0.4100	1	neutral	Polymorphism	0.99	0.80791
Ser44Asn	c.131G/A	0.036	benign	1	0.63	Non-neutral	0.1386	7	neutral	Polymorphism	0.97	0.3034
Tyr62Phe	c.185A/T	0.999	probably damaging	2	0.69	Neutral	0.0626	8	neutral	Polymorphism	0.99	NA
Ala164Val	c.491C/T	0.006	benign	4	0.85	Neutral	0.6043	2	pathological	Polymorphism	1	0.19468
Asn235Ile	c.704A/C	1	probably damaging	4	0.82	Non-neutral	0.8050	6	pathological	Polymorphism	0.04	NA

Table 31: *In silico* prediction of splice sites, transcription factor binding sites and o-glycosylation sites of detected variants in *SULT1A2*

PROGRAM			ESEfinder	ESRSearch	RESCUE_ESE	TFSearch	Consite	OGPET	Mutation Taster	
SNPs	Amino acid changes	DNA position	*Splice sites*	*Splice sites*	*Splice sites*	*transcription factor binding sites*	*transcription factor binding sites*	*O-glycosylation site prediciton [%]*	*Prediction*	*Prob.*
rs710410	---	c.*7T/C	changed	changed	not changed	not changed	not changed	---	Polymorphism	0.71
rs1136703	Ile7Thr	c.20T/C	changed	not changed	not changed	not changed	not changed	---	Polymorphism	1
rs1690407	Ser8=	c.24T/C	changed	not changed	not changed	not changed	p50	---	Polymorphism	0.72
rs10797300	Pro19Leu	c.56C/T	not changed	changed	changed	GATA1	Snail	---	Polymorphism	0.99
rs4149406	---	c.148+34T/C	changed	not changed	not changed	not changed	not changed	---	Polymorphism	0.72
rs145008170	Ser44Asn	c.131G/A	not changed	changed	not changed	deltaE	cFOS	---	Polymorphism	0.97
---	---	c.241+39G/A	changed	not changed	not changed	not changed	not changed	---	Polymorphism	0.71
rs4987024	Tyr62Phe	c.185A/T	changed	changed	not changed	Nkx-2	Snail	---	Polymorphism	0.99
rs3743963	---	c.500-19T/C	not changed	not changed	not changed	not changed	RXR-VDR	---	Polymorphism	0.73
rs142241142	Ala164Val	c.491C/T	changed	changed	changed	not changed	not changed	---	Polymorphism	1
rs139896537	Ala164=	c.492T/C	changed	not changed	not changed	not changed	E74A	---	Polymorphism	0.73
rs1059491	Asn235Ile	c.704A/C	not changed	changed	not changed	not changed	CFI-USP	---	Polymorphism	0.04
rs762634	---	c.1091A/G	changed	not changed	not changed	not changed	not changed	---	Polymorphism	0.73

[a] particular abolished predicted splice sites, [b] particular abolished predicted transcription factor binding sites

Table 32: Association analyses of polymorphism in *SULT1A2* in 355 obesity trios (comprising an extremely obese child or adolescent and both biological parents; Volckmar et al. in prep.).

Position / rs-Number	Amino acid change	Alleles		Effect allele frequencies [%]		Mendelian error rate [%]	Nominal p-value
		Effect	other	Index	Parents		
rs4149404	Ile7Thr	C	T	57.14	56.65	0.00	0.6337
rs10797300	Pro19Leu	T	C	11.90	11.71	0.29	0.7518
rs145008170	Ser44Asn	NA	NA	-	-	0.00	0.1573
rs4987024	Tyr62Phe	A	T	98.45	98.37	0.00	0.6698
rs142241142	Ala164Val	A	G	0.85	0.57	0.00	0.1573
rs1059491	Asn235Thr	A	C	96.49	93.69	6.78	NA

Effect allele refers to the allele that is transmitted from parents to the obese index patient with higher frequency than expected (50%). The effect allele frequencies refer to the frequencies of the overtransmitted alleles in either the indexes or the parents.

4.5 Combined analysis of region chr16p11.2

After detecting obesity associated variants in the obesity candidate genes *SH2B1*, *APOB48R* and *SULT1A2* within the chromosomal region 16p11.2, we analyzed the combined impact of coding variants in this region on BMI. Additionally, data on two other genes were used for this analysis that were screened for mutations independently from this study but in the same 95 extremely obese children and adolescents (SULT1A1 (Horn 2011) and TUFM (Göbel in prep., Struwe in prep.). Low BMI variance in the sample (as the index patients were recruited only if they had a BMI ≥ 97th age and sex specific percentile) prevented the analysis of quantitative effects.

The initial screening sample was enriched for variants in linkage disequilibrium with the BMI associated lead SNP rs7359397 (Speliotes et al. 2010) by using extremely obese children and adolescents that contributed substantially to the positive TDT of the proxy rs2008514 (linkage to rs7359397 $r^2 = 0.965$, D' = 1). Hence, variants in high linkage disequilibrium with both rs7359397 and rs2008514 showed a higher frequency in the initial screening sample than expected from population samples (Tables 21 and 27). Variants in high LD were rs7498665 (*SH2B1* Thr484Ala), rs3833080 (*APOB48R* deletion allele of Ala345_Gly346insdelAlaGlyThrAlaSerGlyGlyGluGluAlaGly, Del2), and rs180743 (*APOB48R* Pro419Ala). The five carriers of a heterozygous deletion of chr16p11.2 which does not harbor any of the analyzed genes deviate from this pattern.

All of these variants in high LD with the original association signal at rs2008514 were also associated with early-onset extreme obesity (p (*SH2B1* Thr484Ala) = 0.007, p (*APOB48R* Ala345_Gly346insdelAlaGlyThrAlaSerGlyGlyGluGluAlaGly) = 0.003, and p (*APOB48R*

Pro419Ala) = 0.002) in the trio sample upon exclusion of the individuals form the intial mutation screen.

Besides these variants which all represent the same obesity associated haplotype, several other infrequent variants were observed in the mutation screens. For the variant βThr656Ile, γPro674Ser in SH2B1, the minor allele frequency is 0.01 and it shows no association to obesity (p = 1, odds ratio (OR) = NA; Table 21). The *SH2B1* variant Thr175Ala (rs181294111) has a MAF of 0.05 in our obese cases (no minor allele frequency for the CEU group is given). When screening the variant in a larger sample of 11,268 obese and overweight cases and 4,512 lean or normal weight controls, neither allele showed association to obesity (p = 0.199, odds ratio (OR) = 4.4, 95% confidence interval (CI) 0.57 - 34.13; Table 21). None of the variants in SULT1A2 showed association with obesity (Table 30).

In addition to the genes *SH2B1*, *APOB48R* and *SULT1A2*, two more genes from the chromosomal region 16p11.2 were screened for mutations. The mutation screen of the sulfotransferase *SULT1A1* was conducted by Lucie Horn, while Maria Göbel and Christoph Struwe screened the gene encoding for the Tu translation elongation factor mitochondrial (*TUFM*). Both projects were co-supervised by the author of this thesis. For a comprehensive analysis of combined effects of variants in the region chr16p11.2, variants from both mutation screens are taken into account here.

Of the coding SULT1A1 variants detected in the mutation screen of Horn (2011), only two could be verified with independent methods in this study. Of these, the SNP rs1042014 [Glu151Asp] showed a trend for association with obesity in a sample of 355 obesity family trios (p = 0.083). As the effect allele is largely overrepresented in our sample (Effect allele frequency (EAF) in CEU 0.867, EAF in 355 family trios 1), this trend could potentially be verified in larger study groups. The previously unknown variant Phe222Leu (g.17420T/C) was not associated with obesity in our trio sample (p = 0.44)

The 95 extremely obese children and adolescents of the initial screening group did not harbor coding variants in *TUFM*; hence no mutations could be integrated into the analysis here.

5. Discussion

5.1 SH2B1

SH2B1 is a mediator of energy homeostasis by increasing leptin and insulin potency in downstream signalling pathways (Morris and Rui 2010). The knockout mouse model is obese and hyperphagic, but also shows hyperlipidemia, leptin resistance, hyperglycemia, insulin resistance and glucose intolerance (Ren et al. 2007). Hence, the gene coding for SH2B1 is a robust obesity candidate gene which is also implicated by GWAS data (Willer et al. 2009, Thorleifsson et al. 2009, Speliotes et al. 2010).

In a mutation screen of the coding region of SH2B1 in 95 extremely obese children and adolescents, we identified a new infrequent mutation g.9483C/T [*β*Thr656Ile/*γ*Pro674Ser] in the *SH2B1* gene for which all three identified mutation carriers (among 11,303 (extremely) obese or overweight children, adolescents and adults and 4,568 normal weight controls) were overweight or obese. The mutation was transmitted from the overweight or obese mothers in the two cases for which parental genotypes were available, at least once together with the obesity risk allele (G) of rs7498665, a coding variant which shows association with obesity (Table 21). Since the minor allele frequency of the variant is low, our study sample was underpowered to identify association with obesity. Comparable to a recent study on Type 2 diabetes mellitus (Bonnefond et al. 2012), our study implies that genes indicated by lead SNPs in large GWAS harbour rare variants that possibly affect gene function.

Two (β and γ variant) of the four splice variants of SH2B1 are affected by the non-synonymous, non-conservative mutation g.9483C/T [*β*Thr656Ile/*γ*Pro674Ser] in their C-terminal part. The C-terminus is the only diverse part of the three splice variants, they all share known functional domains (self-dimerization, nuclear localisation and export sequence, Pleckstrin-homology and SH2 domain; Quian and Ginty 2001, Doche et al. 2012). The functional impact of variants outside the functional domains is hard to determine. *In silico* analyses suggested an abolished or reduced function for the *β* splice variant but not for the *γ* splice variant although the conservation is higher for the *γ* splice variant (Table 23).Our *in vitro* studies for the mutation in both SH2B1 splice variants showed no effect on STAT3 mediated leptin signalling (Figure 9).

Previous *in vitro* analyses suggested different functions for the SH2B1 splice variants. For example in mice a rescue of the *β* splice variant was sufficient to prevent the Sh2b1 knockout phenotype in mice (Rui et al. 2008). While in mice, the *β* splice variant of SH2B1 is most prominently expressed in the brain and hypothalamus (Nishi et al. 2005), in humans splice variants *α* and *δ* show higher expression in these tissues (Doche et al. 2012). For SH2B1*β*, recruiting of the insulin receptor substrates 1 and 2 (IRS1 and 2) to the LEPRb/JAK2 complex was shown in mice (Morris and Rui 2009). Murine Sh2b1*β* enhances JAK2 activity

and promotes the activation of additional downstream networks like STAT3, STAT5 and phosphatidylinositol (PI) 3-kinase pathways (Morris et al. 2009, Morris and Rui 2009). For the murine Sh2b1 *γ*, a direct interaction with Tyr1158 in the activation loop of the insulin receptor was shown which enhances receptor auto-phosphorylation (Morris et al. 2010). Additionally, mSH2B1 *γ* prohibits dephosphorylation of IRS1 and IRS2 which enhances insulin signalling, leading in turn to activation of downstream pathways (Duan et al. 2004).

In addition to *β*Thr656Ile/*γ*Pro674Ser, we detected one rare non-coding mutation and observed five known SNPs in *SH2B1*. Of these, only the non-synonymous coding SNP rs7498665 [Thr484Ala] showed association (p = 0.007 in 3,139 obese cases and 434 lean controls) with increased BMI and obesity. This is in agreement with several previous studies (e.g. Thorleifsson et al. 2009, Willer et al. 2009, Speliotes et al. 2010). Like the mutation alleles *β*656Ile/*γ*674Ser, the obesity risk allele 484Ala (rs7498665) showed non-significantly reduced Emax and EC50 in leptin receptor signalling in an *in vitro* assay (Figure 9). The variant lies outside of the known functional domains of SH2B1 on a position with low conservation which suggests a low impact on protein function (Jamshidi et al. 2007). As the obesity risk allele of rs7498665 increases BMI by only approximately 0.15 BMI units kg/m², as calculated in a population of 125,931 European individuals (Speliotes et al 2010), only subtle functional alterations were expected. We observed no association to obesity for the remaining four known SNPs (rs147094247 [Thr175Ala], rs60604881, rs62037368 and rs62037369) in our study groups (Table 21).

5.2 SH2B1 functional studies

Both *in vitro* analysed variants (rs7498665 and *β*Thr656Ile/*γ*Pro674Ser) showed no significant reduction of STAT3 mediated leptin signalling by the risk alleles. They both displayed a non-significant decrease in EC50 and Emax in splice variants *β* and *γ*, which indicates a potential gain of function and reduced function, respectively. Reduced EC50 indicates a lower amount of leptin needed to achieve half-maximal activation of the leptin receptor, while a reduced Emax represents a lower overall activity of the leptin receptor after activation. Hence, we were not able to determine whether these variants lead to a gain of function or reduced function. Since the changes are not significant in both cases, the functional biological impact of the risk alleles at rs7498665 [Thr484Ala] and *β*Thr656Ile/*γ*Pro674Ser remains to be solved. A potential minor functional effect could only be confirmed by a much larger number of replicates (e.g. ~ 2x270 replicates determined by using the Emax point estimates of SH2B1γ vs. SH2B1γP674S and their variances, applying the Satterthwaite/Welch t-Test aiming at 80% power for two-sided α error of 5%). On the other hand, the results could indicate that the two variants are not functionally relevant.

Yet for a polygenic variant like rs7498665 [Thr484Ala], a large functional effect is unlikely. Comparably, the melanocortin 4 receptor gene (*MC4R*), a well known obesity gene (e.g. Speliotes et al. 2010, Hinney et al. 2003 and 2006), harbors two polymorphisms (Val103Ile and Ile251Leu) whose minor alleles are negatively associated with obesity. Minor allele carriers have a reduced BMI (ca. 0.5 kg/m²) compared to carriers of the wild type alleles (Geller et al. 2004, Stutzmann et al. 2007). Initially, no evidence for functional impact of the minor alleles of these variants was detected in *in vitro* tests, but an increase in test numbers revealed a potential small gain of function for both SNPs (Xiang et al. 2006). Therefore, a potential functional impact of the analyzed *SH2B1* variants could be revealed by a larger number of functional tests or *in vitro* analyses of other homeostatic pathways influenced by SH2B1. The STAT3 mediated leptin signalling assay suggests a potential small effect of the variants, so that testing of the impact on interaction with the insulin receptor or other substrates like IRS or STAT5 would be of high interest. Several small functional changes with different SH2B1 interaction partners could add up to a substantial overall functional effect.

In regard of the low frequency of *β*Thr656Ile/*γ*Pro674Ser, the results of our mutation screen cannot explain our positive TDT for rs2008514 with obesity. A strong functional impact of rs7498665 is regarded as unlikely (Jamshidi et al. 2007, Hotta et al. 2011), which we confirm as we find no significant functional effect *in vitro*. The region on chr16p11.2 shows low recombination rates for a large linkage block of 1Mb (chr16:28,177,800 displays a recombination peak of 37cM/Mb and chr16:28,944,400 a recombination peak of 36 cM/Mb; HapMap, http://hapmap.ncbi.nlm.nih.gov/). The area tagged by both BMI GWAS derived SNPs rs7498665 [Thr484Ala] and rs7359397 (Thorleifsson et al. 2009, Willer et al. 2009, Speliotes et al. 2010) harbours 17 genes (Figure 1). Hence, relevant mutations in one or more of the remaining 16 genes besides *SH2B1* could account for all or a proportion of the GWAS results. We therefore analyzed additional candidate genes from the region of chr16p11.2, *APOB48R* and *SULT1A2*.

5.3 APOB48R

Since the mutation screen of the coding region of *SH2B1* could not explain the initial GWAS result of a strong association of rs2008514 with obesity, the *APOB48R* gene was screened for variants associated with obesity. The gene encodes for a macrophage receptor that regulates fat uptake into the cells (Brown et al. 2000). A defective receptor could increase blood lipids and hence lead to increased risk for cardiovascular diseases. Again, the coding region of the gene was screened in the same sample of 95 extremely obese children and adolescents. Thirteen variants were detected in the *APOB48R* coding region; of these three were non-synonymous (p.Ser323delSerGlyGlyGluGluAlaGlyThrAla, rs180743 [Pro419Ala]

and rs3833080 [p.Ala345_Gly346insdelAlaGlyThrAlaSerGlyGlyGluGluAlaGly]). The two non-synonymous variants rs180743 and rs3833080 were genotyped in the family-based (705 core family trios) and case-control study groups (453 extremely obese cases and 435 normal weight or lean controls; Volckmar et al. in prep., Graninger 2011).

The first identified deletion is a previously unknown variant with a length of 27bp (c.961_996delACAGCCTCAGGCGGGGAGGAGGCCGGGACAGCCTC), which results in an in frame shortened amino acid (p.Ser323delSerGlyGlyGluGluAlaGlyThrAla) sequence of the repeat region of APOB48R. Despite no described function for the repeat region, *in silico* prediction of the variant is functional (Mutation Taster "disease causing"; Table 25). The deletion was transmitted from the obese mother (BMI 39.06 kg/m²) to the obese female mutation carrier (height 147 cm, weight 49 kg, BMI 22.68 kg/m², BMI SDS 1.44, age 10.5 years). The deletion was inherited together with the risk alleles for rs180743 and rs3833080 (G and deletion allele, respectively). Rarity of the variant Del 1 prevented obesity association analysis.

Like Del1, the second identified deletion Del2 (p.Ala345_Gly346insdelAlaGlyThrAlaSerGlyGlyGluGluAlaGly; rs3833080) is also located in the repeat region of *APOB48R*. The deletion of Del2 leads to a shortened repeat [GlyGlyGluGluAlaGluThrAlaSer] region with possible coding outcome (Figure 10). Del2 is most likely identical to the dbSNP listed rs3833080, which shares the deleted nucleotide sequence and MAF with the variant detected in our sample. The deletion allele of Del2 is predicted to be a non functional polymorphism (Mutation Taster; Table 25).

The second common coding polymorphism Pro419Ala (rs180743) is in high LD with the Del2 in the first exon of *APOPB48R*. The position of the variant is conserved and the amino exchange leads to a change from hydrophobic to aromatic which was predicted to be deleterious by two *in silico* prediction programs (Mutation Taster and PolyPhen2; Table 25).
Besides the chr16p11.2 deletion carriers, all initially screened individuals are carriers of the risk alleles of the rs180743 [Pro419Ala] and rs3833080 [Del2] polymorphisms. Del2 is in strong LD with rs180743 [Pro419Ala] ($r^2 = 0.98$) which explains the shared previously known association to hypercholesterolemia (Fujita et al. 2005) and the association to obesity detected here ($p = 0.003$ in 705 family based trios, $p = 0.003$ in the case control sample). Since both variants are in close proximity and in high LD, a genetic differentiation of both signals to identify the underlying causal variant of the BMI association is not possible. Both variants are located in a series of 9 amino acid repeats in the extracellular domain of APOB48R for which Brown et al (2000) suggested a contribution to binding of the specific ligand apoB48. Therefore, these alterations possibly lead to reduced uptake of chilomicron remnants and increased blood lipid levels. *In vitro* testing of this hypothesis by lipoprotein-

uptake assays or ligand blotting would be possible, as suggested by Brown et al. (2000) and Daniel et al. (1983).

This study is the first to show BMI/obesity association of variants in *APOB48R* which might completely explain or contribute to the BMI association signal of lead SNPs in or nearby *SH2B1*. An *in vitro* functional validation of both deletion rs3833080 [Del2] and SNP rs180743 [Pro419Ala] are warranted. Further (infrequent to rare) variants in other genes tagged by the same original BMI associated SNPs could potentially additionally account for the observed effect.

5.4 Impact of variants in SH2B1 and APOB48R on weight loss parameters

As we detected coding SNPs in *SH2B1* (rs7498665 [Thr484Ala]) and *APOB48R* (rs180743 [Pro419Ala]) that showed association to obesity, we were interested in their effect on weight loss and associated parameters. Since the SNP in *SH2B1* is associated with T2D after BMI adjustment in an adult study group (Sandtholt et al. 2011), we were interested in the serum insulin levels and HOMA of rs7498665 [Thr484Ala] risk allele carriers. The SNP rs180743 [Pro419Ala] in *APOB48R* is associated with hypercholesterolemia in adults (Fujita et al. 2005). Hence, we expected the risk allele carriers to have increased total cholesterol and triacylglycerides levels in serum.

While a BMI association of SNPs in the chromosomal region 16p11.2 has often been replicated (e.g. Thorleifsson et al. 2009, Willer et al. 2009, Speliotes et al. 2010) for adults, data on the association in children is sparse. Speliotes et al. (2010) report association for rs7359397 ($p = 0.023$ in obese children and adolescents versus controls; $n = 9{,}081$), we detected association for the coding SNPs rs7498665 ([SH2B1 Thr484Ala] $p = 0.007$ in 3,139 obese cases and 434 lean controls; Volckmar et al. 2012) and rs180743 ([APOB48R Pro429Ala] $p = 0.002$ both in 705 obesity trios and 453 extremely obese cases and 435 normal weight or lean controls; Graninger et al. 2011, Volckmar et al. in prep.). Although our finding demonstrates strong association of rs7498665 to obesity in children and adolescents, the effect of the BMI risk allele seems to increase during maturation as obesity association is stronger for adults than for children (Graff et al. 2013). Up to now, longitudinal studies pertaining to the impact of the risk alleles of variants in SH2B1 and APOB48R on weight change and related parameters in obese or overweight individuals participating in a lifestyle intervention have been lacking.

The non-synonymous SNPs rs7498665 ([Thr484Ala] risk allele A) in *SH2B1* and rs180743 ([Pro429Ala] risk allele G) in *APOB48R* (linkage disequilibrium of both SNPs: $D'=0.921$, $r^2=0.789$) showed no association with any of the observed change parameters in 454 overweight and obese children or adolescents who underwent a 1-year lifestyle intervention program. Although both SNPs are associated with obesity in our samples, the analysis

indicates that both variants have at most a minor impact on weight loss in the subsample of children and adolescents that completed the 'Obeldicks' intervention. A previous longitudinal study on weight changes related to SNPs in *SH2B1* also showed no association to weight gain during treatment with antipsychotics (Perez-Iglesias et al. 2010) although this finding gives no information about the effect on weight loss within a conventional weight loss program. Of interest might also be the finding that the obesity risk allele of rs7498665 is nominally associated with increased snacking behavior (Robiou-du-Pont et al. 2013). For rs7498665, *in silico* analyses predicted no altered SH2B1 function (Jamshidi et al. 2007) for the less frequent allele of this coding SNP, although the region flanking the position is highly conserved (Hotta et al. 2011). The only i*n vitro* study on the variant showed no functional effect of the obesity risk allele on STAT3 meditated leptin signalling (Volckmar et al. 2012). This could indicate that both variants are not causally involved in the obesity or influence the weight loss.

'Obeldicks' is a highly effective lifestyle intervention which includes diet, physical activity, and behavioral approaches over a long period of time. The weight loss during 'Obeldicks' (BMI SDS reduction of 0.3 in this study group, BMI SDS reduction of 0.41 in another sample; Reinehr et al. 2007) is comparable to other studies (e.g. BMI SDS reduction of 0.36 over a period of 4-6 weeks; Holzapfel et al. 2012). While the other studies did not follow up on weight loss, for 'Obeldicks' the reduction of BMI SDS sustained 3 years after the end of the intervention (Reinehr et al. 2007). Hence, the intervention is effective for decreasing the BMI SDS in obese and overweight children and adolescents over a long time period. This effectiveness could mask genotype effects on weight loss or other change parameters, as subtle effects of genotype on the outcome are concealed. A recent study in the same study group detected an effect of variants in *SDCCAG8* on weight loss, but not for risk alleles in the known obesity associated genes *FTO*, *MC4R*, *TMEM18*, and *TNKS/MSRA* (Scherag et al. 2012) which show stronger effect on BMI than *SH2B1* (Speliotes et al. 2010).

A limitation of our study is the small size of the study group. Still, this is the first analysis of the impact of genetic variants in chr16p11.2 on weight loss and change parameters during a 1-year lifestyle intervention in overweight to obese children and adolescents. The power was sufficient to detect moderate and strong effects (a change of 0.2 BMI SDS changes per risk allele ($p = 0.05$, two-sided)).

5.5 SULT1A2

The last obesity candidate gene analyzed in this study was *SULT1A2*. This sulfotransferase influences hormone levels like 17-β-estradiol, estron and estron sulfate (Harris et al. 2000), which in turn are associated with obesity (Ghose et al. 2011). The same 95 individuals enriched for variants in the region of high LD to rs2008514 were screened for variants in the

coding region of *SULT1A2*. Of the 14 detected variants, seven were non-synonymous (rs4149404 [Ile7Thr], rs10797300 [Pro19Leu], rs145008170 [Ser44Asn], rs4987024 [Tyr62Phe], rs142241142 [Ala164Val], rs1059491 [Asn235Thr], rs75191166 [Lys282Gly]), while seven had no effect on the coding sequence of SULT1A2 (rs1690407 [Ser8=], rs139896537 [Ala164=], rs4149406, rs3743963, rs710410, rs762634, rs145790611).

The coding SNP rs1136703 [Ile7Thr] has long been suggested as a functionally relevant variant due to its association with increased cancer risk (Zhu et al. 1996, Raftogianis et al. 1999, Glatt et al. 2004). First functional studies showed a lower expression of this variant compared to wild type *in vitro* and lower sulfonation of xenobiotic toxic compounds 1HMP (1-hydroxymethylpyrene) and OH-AAF (N-hydroxylamino-2-acetylanimofluorene) but not OH-APP (2-hydroxy-5-phenylpyridine) in *in vitro* experiments (Meinl et al. 2002), although it is not located near the binding pocket of SULT1A2 (Lu et al. 2010). In contrast to the *in vitro* findings, all *in silico* analyses predicted no functional effect of this variant (Table 31).

For rs10797300 [Pro19Leu], *in vitro* experiments showed lower expression, lower sulfonation of xenobiotic substrates 1HMP and OH-AAF (Meinl et al. 2002). Studies regarding binding of SULT1A2 to the xenobiotic p-nitrophenol suggested the neighboring amino acid Tyr23 to be located in the binding pocket (Lu et al. 2010). A functional relevance for a non-conservative amino acid exchange in close proximity is likely. In silico programs predict a functional outcome for this variant *in silico* (Table 31).

The rare coding variants rs4987024 [Tyr62Phe] and rs142241142 [Ala164Val] are both conservative amino acid exchanges that are not located close to the binding pocket. Their low frequency (MAF in CEU 0.001 compared to MAF in CEU ~0.3 for Ile7Thr, Pro19Leu, and Asn235Ile) prevented them from being included in the early functional studies (Meinl et al. 2002). *In silico* prediction shows a higher probability of functional changes for Ala164Val, although the analyses revealed mixed results (Table 31).

The non conservative amino acid exchange rs1059491 [Asn235Thr] showed lower expression and lower sulfonation of the xenobiotic substrates with an overall stronger effect of the variant than the 7Thr allele (Meinl et al. 2002). Structural studies of the binding pocket of SULT1A2 for p-nitrophenol suggested the neighboring amino acid Tyr240 to be involved directly in the binding. Since Asn235Thr is a non-conservative amino acid exchange close to the binding site, a functional relevance is likely (Lu et al. 2010). Glatt et al. (2002) found non-significantly increased MAFs for the 235Ser allele of this triallelic variant in obese individuals. *In silico* prediction mostly interpreted this SNP correctly as functionally relevant (Table 30).

Of the detected non synonymous variants, several have known functional effects on xenobiotic sulfonation, e.g. rs4149404 [Ile7Thr], rs10797300 [Pro19Ser], and rs1059491

[Asn235Thr] (Meinl et al. 2002). None of the *in vitro* functional variants showed obesity association in our sample of 355 family trios. The polymorphism rs1059491 [Asn235Thr] showed high mendelian error rates which stand in contrast to previous mendelian error rates in the GWAS analyzed 355 trios. The use of families instead of a case control design for this analysis revealed an error that would have otherwise been missed. A possible explanation for this error might be the high similarity between sulfotransferases; both SULT1A1 and SULT1A2 harbor the same variant Asn235Thr according to dbSNP (SULT1A1: rs35728980, SULT1A2: rs1059491). It is possible that one or both variants were mischaracterized and either both genes only possess one allele (either Asn or Thr) or that the variant is falsely genotyped in both genes at the same time which results in mixed genotypes at both positions.

As the variants in SULT1A2 detected in our mutation screen of 95 extremely obese children and adolescent show no association to obesity, they do not contribute to our initial TDT finding.

5.6 Chromosomal region 16p11.2

The rate of associated loci detection in large GWAS (meta-) analyses is high, the identification of 18 new BMI associated loci doubled the amount of known obesity associated loci (Speliotes et al. 2010). Lead SNPs, tagging large regions in the human genome, are favoured for these analyses as a relatively low number of SNPs can span the whole genome for a given expected minor allele frequency. One major problem of this method is the general expectancy that each lead SNP corresponds with one mutation in one nearby gene which in turn explains the observed traits. If the lead SNP is indeed a coding variant of a gene or at least in the intronic or UTR region of a gene (as for example in *FTO*), this hypothesis is likely. But most of the lead SNPs tag a large number of genes whereas most GWAS studies report only single or few genes for each locus. Even if a candidate gene for the observed trait is within the locus tagged by the lead SNP, other nearby genes can contribute to the signal.

In the case of the chromosomal region chr16p11.2, the lead SNP rs7359397 tags a large linkage block of 1Mb (chr16:28,177,800 displays a recombination peak of 37cM/Mb and chr16:28,944,400 a recombination peak of 36 cM/Mb; HapMap, http://hapmap.ncbi.nlm.nih.gov/). The linkage disequilibrium block habors 16 genes besides the obesity candidate gene *SH2B1*. The SH2B1 protein is a tyrosine kinase involved in several pathways relevant for energy homeostasis (Rui et al. 2007, Morris et al. 2010). The murine knockout model for Sh2b1 is obese and shows hyperphagia, hyperglycemia and insulin resistance (Rui et al. 2007, Morris et al. 2010). Therefore, Speliotes et al. (2010) attributed the association signal mainly to *SH2B1*. In this study we present evidence that two coding variants in SH2B1 had no functional effect on leptin signalling which is one of the

main regulators of energy homeostasis. One of these variants is a GWAS derived coding SNP (rs7498665 [Thr484Ala]) which has no predicted functional effect (Jamshidi et al. 2007, Hotta et al. 2011). The variants detected in the mutation screen of SH2B1 did not explain the initial association signal observed in our own study groups (p_{trios} = 0.0087; p_{cc}= 0.025 for rs2008514). These findings are in concordance with data from Doche et al. (2012) who detected only one functionally relevant frameshift mutation, but saw no functional changes in leptin or insulin signalling for three non-synonymous coding variants in *SH2B1* detected in an extremely obese sample.

In addition to the common SNP rs7498665 [SH2B1 Thr484Ala], two further coding variants in the gene encoding for the APOB48R showed association with obesity (rs180743 [Pro419Ala] and rs3833080 [Ala345_Gly346insdelAlaGlyThrAlaSerGlyGlyGluGluAlaGly]). These two variants are in high linkage disequilibrium with each other and both rs7498665 and the initial association signal of the lead SNP rs7359397 in our GWAS samples. Together they build an obesity risk haplotype. Differentiation of this combined association signal is not possible, so that attribution of the causal association with increased body weight to only one of the variants remains unclear. Further functional *in vitro* studies on these three variants are warranted. Besides *SH2B1*, *APOB48R*, and *SULT1A2*, our group also analyzed further genes in the region of chr16p11.2: *SULT1A1* and *TUFM*.

5.6.1 SULT1A1

The sulfotransferase 1A1 (SULT1A1) has a high similarity to SULT1A2; both genes are located closely to each other on chr16p11.2. Due to the high sequence identity of 96% (Figure 13) and the higher expression of *SULT1A1* than *SULT1A2* in several tissues, *SULT1A2* is discussed to be a pseudogene of *SULT1A1* (Freimuth et al. 2004, Glatt et al. 2001 of the same work group; leader H. Glatt), although the detection of expression in some tissues is contradictory to the definition of a pseudogene. *SULT1A1* shows ubiquitous expression (Glatt et al. 2001), while the expression of *SULT1A2* is limited to liver, blood platelets, heart, brain and skin. Both sulfotransferases share their substrates; they sulfonate molecules that modulate sex hormones like estrogen or xenobiotics that are involved in cancer risk (Glatt et al. 1994). Glatt et al. (2001) report an obesity association for the minor allele of *SULT1A1* SNP CM973382 in unpublished data from another group.

During a mutation screen of the coding region of *SULT1A1* in the same 95 individuals analyzed in this thesis, Horn (2012) detected a total of 15 mutations and SNPs, of these eight coding ((g.390A/C) Met1Leu, rs1042011 [Glu151Gln], rs1042014 [Glu151Asp], (g.17420T/C) Phe222Leu, rs1801030 [Val223Met], (g.17429C/G) Gln225Glu, rs35728980 [Asn235Thr], (g.17487C/G) Pro244Arg, (g.17490 A/G) Gln245Arg). All coding variants were genotyped for association analysis in 355 family trios.

Several of the variants detected in the mutation screen and characterized by Sanger sequencing turned out to be artifacts ((g.390A/C) Met1Leu, rs1042011 [Glu151Gln], rs1801030 [Val223Met], (g.17429C/G) Gln225Glu, rs35728980 [Asn235Thr], (g.17487C/G) Pro244Arg, (g.17490 A/G) Gln245Arg), potentially because of the high sequence similarity within the *SULT* family. Although the primers for PCR amplification were chosen to match only one position in the human genome, the amplification of other unwanted fragments can never be excluded. For the variant Met1Leu (g.390A/C), independent validation was attempted using PCR-RFLP and TaqMan assay which both could not verify the variant. Also for variants (g.17429C/G) Gln225Glu and (g.17490 A/G) Gln245Arg, an enzymatic digest showed that both variants are non-existent. Comparing the sequences of SULT1A1 and SULT1A2 reveals that the amino acid exchanges Gln225Glu and Gln245Arg predicted from Sanger re-sequencing are the deviant amino acids between both proteins (Figure 12). The same is true for the variants rs1042011 [Glu151Gln], rs35728980 [Asn235Thr], and (g.17487C/G) Pro244Arg, which could not be verified with MALDI TOF in 355 family based trios.

```
SULT1A1          1 MELIQDTSRP PLEYVKGVPL IKYFAEALGP LQSFQARPDD LLISTYPKSG
SULT1A2            MELIQDISRP PLEYVKGVPL IKYFAEALGP LQSFQARPDD LLISTYPKSG
                   ****** *** ********** ********** ********** **********

SULT1A1         50 TTWVSQILDM IYQGGDLEKC HRAPIFMRVP FLEFKAPGIP SGMETLKDTP
SULT1A2            TTWVSQILDM IYQGGDLEKC HRAPIFMRVP FLEFKVPGIP SGMETLKNTP
                   ********** ********** ********** *****.**** *******:**

SULT1A1        100 APRLLKTHLP LALLPQTLLD QKVKVVYVAR NAKDVAVSYY HFYHMAKVHP
SULT1A2            APRLLKTHLP LALLPQTLLD QKVKVVYVAR NAKDVAVSYY HFYHMAKVYP
                   ********** ********** ********** ********** ********:*

SULT1A1        150 EPGTWDSFLE KFMVGEVSYG SWYQHVQEWW ELSRTHPVLY LFYEDMKENP
SULT1A2            HPGTWESFLE KFMAGEVSYG SWYQHVQEWW ELSRTHPVLY LFYEDMKENP
                   .****:**** ***.****** ********** ********** **********

SULT1A1        200 KREIQKILEF VGRSLPEETV DFVVQHTSFK EMKKNPMTNY TTVPQEFMDH
SULT1A2            KREIQKILEF VGRSLPEETV DLMVEHTSFK EMKKNPMTNY TTVRREFMDH
                   ********** ********** *::*:***** ********** *** :*****

SULT1A1        250 SISPFMRKGM AGDWKTTFTV AQNERFDADY AEKMAGCSLS FRSEL
SULT1A2            SISPFMRKGM AGDWKTTFTV AQNERFDADY AKKMAGCSLS FRSEL
                   ********** ********** ********** *:******** *****
```

Figure 12: Multiple sequence alignment of SULT1A1 and SULT1A2. The amino acid sequence of the alpha splice variants of SULT1A1 (ENST00000314752) and SULT1A2 (ENSG00000197165) were aligned using the program T-Coffee (http://www.ebi.ac.uk/Tools/msa/tcoffee/). Here, an asterisk below the amino acid alignment marks complete sequence identity, while dots below the sequence mark amino acids that share similar side chains or charges. The positions of non-synonymous variants detected in the mutation screen are marked in yellow, green marks the variants that could not be verified with independent methods (adapted from Horn 2012).

Only two of the variants detected in *SULT1A1* could be verified with an independent method. The amino acid exchange at position Glu151 (rs1042014 [Glu151Asp]) is located in close proximity of Tyr149, which is directly interacting with the substrate in the binding pocket. Surprisingly, the position is not strongly conserved (Figure 12). The variant is not significantly associated with obesity ($p = 0.08$ in 355 family based trios), but larger sample sizes could increase the power for association analysis. The variant Phe222Leu was not associated with obesity ($p = 0.44$).

None of the two variants (rs1042014 [Glu151Asp] and Phe222Leu) that could be verified by independent methods besides Sanger sequencing, was associated with obesity. Even if they showed biological functional changes on protein level, other sulfotransferases can most likely compensate for the function of each other *in vivo* in case only one is affected (Glatt et al. 2004).

5.6.2 TUFM

Another possible obesity candidate gene in the chromosomal region of chr16p11.2 is the Tu translation elongation factor mitochondrial gene (*TUFM*). *TUFM* encodes a transcription factor that is necessary for mitochondrial gene expression (Ling et al. 1997) by regulating type 1 interferon activity (Lei et al. 2012). As the BMI of a child is more strongly correlated with its mother's BMI (Hebebrand et al. 2001) and mitochondria are inherited exclusively maternally, mitochondrial DNA could partially explain the heritability of the BMI variance. Also, the mitochondria produce ATP which is equivalent to energy used for most cellular processes. Changes in mitochondrial activity pertaining to oxidative phosphorylation could lead to relatively increased or decreased ATP production for a given caloric intake. Hence, individuals with hyperfunctional oxidative phosphorylation would gain a higher amount of ATP and gain weight from the same amount of caloric consumption. A mutation (rs121434452 [Arg339Gln]) in *TUFM* is associated with a combined oxidative phosphorylation deficiency type 4 (COXPD4) which results in lactic acidosis in young children (Valente et al. 2007). The expression of TUFM is up regulated in DIO rats on a high fat diet (Gutierrez-Aguilar et al. 2012). In man, TUFM expression is potentially regulated by an intronic SNP (rs4788099) in B cells and monocytes (Guo et al. 2012). Santos and Duarte (2008) described differential expression of TUFM in hippocampal neurons elicited by BDNF, a known BMI associated protein (e.g. Speliotes et al. 2010).

Therefore, the coding region of TUFM was screened for mutations in the same mutation screen sample as the genes in this thesis (Struwe in prep., Göbel in prep.). All detected variants were non-coding and with the exception of c.3536C>G in the 3' untranslated region previously known (rs7187776, rs4788099, rs8061877 and rs61737565).

For rs8061877, the disruption of a transcription factor binding site was predicted (TFSearch and Consite), although for two different transcription factors. For the new variant g.28854194C>G, *in silico* prediction of all three programs was change in splice enhancer binding sites and splicing silencer sites. Also, a change in splice sites for this variant was predicted (Mutation Taster) despite the variant being in the non-coding 3' UTR of *TUFM*. If any of these predicted changes has an impact on the expression of TUFM under biological conditions, needs to be tested *in vivo*.

Since we detected no coding variant in the 95 extremely obese individuals of the initial screening sample in *TUFM*, we did not analyze any variant for obesity association. As the initial screening sample was enriched for variants contributing to the obesity association signal in chr16p11.2 and we did not find any coding variants within the gene, *TUFM* is the gene with the lowest potential as an obesity candidate gene for future studies.

5.6.3 Combined data on chromosomal region 16p11.2

Besides the common coding variants in *SH2B1* and *APOB48R*, two sulfotransferase genes (*SULT1A1* and *SULT1A2*) harbored coding variants. Of the detected mutations in our enriched screening sample, none was associated with obesity although complete validation of the polymorphism rs1042014 [SULT1A1 Glu151Asp] in a larger study population is warranted. For some of the detected infrequent variants in these genes, *in vitro* functional data was available that indicates reduced function for the potentially affected SULT1A2 and SULT1A1 proteins. The low frequency of the variant rs1042014 [*SULT1A1* Glu151Asp] could prevent obesity association analysis in the small sample size in this study; a larger study group would provide more power to verify an association. The mitochondrial gene expression factor TUFM contained no coding variants in the enriched screening sample.

Recent studies on cis-regulatory effects of SNPs outside the *SH2B1* coding region indicate a potential expression regulating effect of genetic variation. An intronic SNP (rs4788099) in *SH2B1* regulated the expression of several nearby genes in B cells and monocytes (*TUFM*, coiled-coil domain containing 101 (*CCDC101*), Homo sapiens spinster homolog 1 (*SPNS1*), *SULT1A1* and sulfotransferase family, cytosolic, 1A, phenolpreferring, member 4 (*SULT1A4*); Guo et al. 2012). Some of these genes showed differential regulation in diet induced obese (DIO) rats on a high fat diet, which increases evidence for regulatory elements in chr16p11.2 that influence energy homeostasis (Gutierrez-Aguilar et al. 2012). Another study showed that SH2B1 expression was increased in total brain but not hypothalamus of mice fed a high fat high sucrose diet (Yonagathan et al. 2012), while the expression of SH2B1 in the aforementioned rats fed a high fat diet is decreased (Gutierrez-Aguilar et al. 2012). Differential regulation of SH2B1 expression depending on diet would be of high interest as SH2B1 increases signalling for the leptin receptor which signals the amount body fat mass to

the brain and the insulin receptor which influences blood glucose levels. The association signal of SNP rs4788099 which influences regulation is congruent with the SNP rs2008514 that was used to enrich our initial screening collective (LD $r^2 = 1$, D' = 1).

Several genes could contribute to the initial association signal of rs7359397 by changed expression or protein function. Small effects in many pathways can sum up to large overall effects but may be hard to detect with current methods. From the results of this study, we would recommend analyzing not only the most likely candidate genes in a region tagged by a GWAS lead SNP but also examine other nearby genes with variants in linkage disequilibrium with the original lead signal for a potential contribution to the observed trait.

5.7 Methods

5.7.1 Mutation screening

For this thesis, the coding sequence of the genes *SH2B1, APOB48R*, and *SULT1A2* were screened for variants (rare mutations, small InDels and polymorphisms) that may be associated with early-onset extreme obesity. The most suitable screening method was chosen by the size of the PCR amplicons: SSCP for PCR amplicons up to 300bp (Hayashi and Yandell 1993, Xie et al. 1997, Sheffield et al. 1993) and dHPLC for PCR amplicons up to 600bp (Liu and Steward 1997; O'Donovan et al. 1998). Both methods do only detect presence or absence of variants but not the exact location or sequence alterations which then have to be determined by Sanger re-sequencing. For SSCP, the build and stability of the tertiary DNA structure influences whether a variant can be detected or not. Here, two different temperatures (room temperature ~20°C and 4°C) were used to increase specificity and sensitivity of the method (Orita et al. 1989). The dHPLC on the other hand is less sensitive for detection of homozygous variants (Taliani et al. 2001). This bias was decreased by adding wild type DNA to each sample. With these protocol modifications, both methods achieve a high sensitivity (5% error rate, comparable to Sanger sequencing) given a correct optimization (Jones et al. 1999, Kurelac et al. 2012).

5.7.2 Determination of Deletion positions

Sanger sequencing is the gold standard of sequencing techniques although Next Generation Sequencing methods like Pyrosequencing or whole genome/exome sequencing gain popularity due to cost and time efficiency (Janitz 2008). For all of these methods, the exact determination of the position of deletions in repetitive regions of the genome remains problematic. In Sanger re-sequencing, an InDel shifts the sequence of one strand of DNA in heterozygous samples, leading to double signals for each nucleotide after the InDel which promote early amplification determination. In homozygous deletions, only the comparison of the sequenced DNA with the wild type sequence allows detection of the InDel. Larger InDels

can be separated by gel electrophoresis before sequencing, so that for heterozygous carriers both alleles can be sequenced independently. In this thesis, the InDel p.Ala345_Gly346insdelAlaGlyThrAlaSerGlyGlyGluGluAlaGly (rs3833080) was detected during the mutation screen of *APOB48R*. As this InDel is located in a repetitive DNA motive [GlyGlyGluGluAlaGluThrAlaSer], which is repeated 3 times exactly and 8 times with small changes of the sequence, the deletion could be of any of the three exact repeats. For this thesis, the position of the InDel is taken from the SNP database (dbSNP; http://www.ncbi.nlm.nih.gov/snp/) where the above position of the variant is listed.

5.7.3 *In silico* functional analyses

The *in silico* studies used in this thesis do vary in their prediction depending on the respective algorithms underlying the tools. To make the overall prediction more reliable, several in silico prediction methods were used. SNAP, Mutation Taster and Panther take conservation of the protein over several animal species into account. Regions with high conservation are more likely to be involved in the function of the analyzed protein, as deviate genes and proteins seldom thrive in the gene pool. The other *in silico* prediction programs SIFT and PolyPhen2 rely on the chemical properties of the exchanged amino acids. Here, conservation is not taken into account, but changes from uncharged to charged amino acid side chains have a stronger predictive effect on protein function. Using several in silico prediction tools, we avoid being biased by either species conservation dependent or amino acid property dependent tools. No *in silico* method can replace actual *in vitro* studies.

5.7.4 *In vitro* functional analyses

The assay to measure STAT3 mediated leptin response showed increased leptin response after co-transfection with wild type SH2B1 splice variants *β* and *γ*. This indicates that a combination of HEK293 cells, that intrinsically express SH2B1, and the STAT3 assay allow functional characterization of SH2B1. While in mice only the alpha splice variant was tested for leptin signalling (Li et al. 2007) and a higher occurrence of the splice variants α and δ was observed in human brain tissues (Doche et al. 2012), an effect on leptin signalling for both other splice variants (*β* and *γ*) was observed in the human cell system.

6. Conclusion and Outlook

In conclusion, in this study only frequent coding variants in *SH2B1* (rs7498665 [Thr484Ala]) and *APOB48R* (rs180743 [Pro419Ala], rs3833080 [Ala345_Gly346insdelAlaGlyThrAlaSerGlyGlyGluGluAlaGly]) showed association to obesity. For other analyzed genes in the region of chr16p11.2 (*SULT1A1*, *SULT1A2*, and *TUFM*), association to early-onset extreme obesity was not detected in a sample of 355 family trios comprising one extremely obese child or adolescent with both biological parents.

In vitro functional analyzes did not reveal impairments in STAT3 mediated leptin signaling for the *SH2B1* variant rs7498665 and the rare mutation βThr656Ile/γPro674Ser in *SH2B1*. Further studies on the impact of both variants in functional assays on the leptinergic and the insulinergic systems are of high interest. We applied for funding to assess the impact of further SH2B1 variants on leptin and insulin receptor signalling in homologous SH2B1 knockout cell lines (murine adult hypothalamic cells, insulin sensitive muscle cells and white adipocytes). Of interest are also further interaction partners of SH2B1 which may play a role in weight regulation and energy homeostasis like insulin-like growth factor receptor (IGFR), growth hormone receptors (EGFR and GHR), brain-derived neurotrophic factor receptor (BDNFR), and nerve growth factor receptor (TRKA) (Morris and Rui 2009). As SH2B1 is involved in many neurodevelopmental, energy homeostasis and hormonal pathways, it is not a valid target for pharmaceutic interventions against obesity.

Additional functional studies for example on the capabilities of APOB48R to bind to apoB48 and the associated chilomicrons by lipoprotein-uptake assays or ligand blotting as suggested by Brown et al. (2000) and Daniel et al. (1983) are warranted. A recent study showed increased APOB48R expression after the intake of a high fat meal which could be detected *in vivo* in the postprandial phase (Varela et al. 2011) which is in line with the finding that the polymorphism rs180743 [APOB48R Pro419Ala] shows association to hypercholesterolemia (Fujita et al. 2007) as well as obesity.

Several of the *SULT1A1* and *SULT1A2* variants detected in the initial mutation screen sample showed reduced expression and changed anti-cancerogenic capabilities in *in vitro* studies (Meinl et al. 2002). Although the variants detected here did not display obesity association in the small sample of 355 trio families in this study, association tests with larger sample sizes should follow (e.g. for rs1059491 [SULT1A2 Asn235Thr]). If any variants show association with obesity, further *in vitro* analyses of the impact of the obesity associated variants on hormone metabolism especially of estrogens are warranted. Analyses pertaining potential risk haplotypes that could cover one or several of the sulfotransferases within chr16p11.2 would provide additional information about the involvement of *SULT* genes in obesity.

Pertaining to the findings that differential expression of several proteins encoded in the chromosomal region 16p11.2 is connected to the composition of the diet of the analysed rodents (Guo et al. 2012, Yonagathan et al. 2012, Gutierrez-Aguilar et al. 2012), the effect of (non-coding) variants in interaction with various feeding conditions would be of high interest. The influence of the detected obesity associated variants on the protein expression of *SH2B1* and *APOB48R* has not yet been analyzed. Here, *in vitro* experiments on homologous cell lines under culture conditions with high glucose or high fat content would help to analyze whether the presence or absence of both conditions and mutations would affect gene expression and protein levels. If these tests show an interaction effect of gene variants in *SH2B1*, *APOB48R*, *SULT1A1*, *SULT1A2* and/or *TUFM* described in this thesis and medium condition, further experiments *in vivo* should analyze if animals with this specific genetic background also react on feeding conditions.

Although we find at most a minor effect of the obesity risk alleles of non-synonymous SNPs at *SH2B1* and *APOB48R* on weight loss-related phenotypes in overweight children after a 1-year lifestyle intervention, a potential small effect cannot be excluded. As the association of the GWAS lead SNPs rs7359397 proved to be larger in adults than in children and adolescents (Speliotes et al. 2010), an effect on weight loss might be more pronounced in an adult lifestyle intervention study. According to our findings, the use of variants in *SH2B1* and *APOB48R* as predictors of weight loss success is not recommended in children and adolescents.

7. Summary

Heritability of the variance of body weight is high. Large-scale meta-analyses of genome-wide association studies identified association to obesity of two SNPs (rs7498665, rs7359397) in the chromosomal region 16p11.2. The locus harbors the *SH2B1* gene, the *APOB48R* gene and the *SULT1A2* gene. *SH2B1* is a solid candidate gene for obesity as it is an important mediator of leptin and insulin receptor activity and influences several downstream pathways that regulate energy homeostasis. *Sh2b1* knockout mice show a phenotype of obesity, hyperlipidemia, leptin resistance, hyperphagia, hyperglycemia, insulin resistance and glucose intolerance. The macrophage receptor APOB48R binds to dietary triglyceride-rich lipoproteins after activation by APOB48 and is predicted to provide dietary lipids and lipid-soluble vitamins to reticuloendothelial cells. The effect of SULT1A2 on body weight regulation could be mediated by the regulation of sex hormones like estrogens and androgens which are metabolized as SULT1A2 substrates.

Here we screened the coding region of *SH2B1*, *APOB48R* and *SULT1A2* via dHPLC and SSCP for (infrequent) mutations in 95 children and adolescents with early onset extreme obesity. Detected coding variants were tested for association (RFLP, ARMS PCR, TaqMan SNP genotyping, MALDI TOF) in independent large study groups of up to 11,406 obese and overweight cases versus 4,568 normal weight controls. For SH2B1, two coding variants (rs7498665 [Thr484Ala], βThr656Ile/γPro674Ser) were analyzed *in vitro* for their impact on STAT3 mediated leptin signalling. Additionally, the impact of common coding variants (rs7498665 and rs180743 [APOB48R Pro419Ala]) in both genes on the outcome of weight loss and several related parameters after a 1-year life style intervention study (“Obeldicks”) was investigated.

All genes contained known variants and new mutations. Association with obesity was detected only for frequent coding variants in *SH2B1* (rs7498665 [Thr484Ala]; p=0.019, OR= 1.19) and *APOB48R* (rs180743 [Pro419Ala]; p=0.002, OR=1.17 and rs3833080 [Ala345_Gly346insdelAlaGlyThrAlaSerGlyGlyGluGluAlaGly]; p=0.003, OR= 1.25). The *SH2B1* variant rs7498665 and the rare mutation βThr656Ile/γPro674Ser showed no functional impairment in *in vitro* functional analyzes of leptin signalling. Regarding weight loss and related parameters like blood pressure, blood lipids and HOMA IR, the obesity risk alleles of non-synonymous SNPs at *SH2B1* and *APOB48R* had no effect in overweight and obese children after a 1-year lifestyle intervention, despite an overall BMI SDS reduction and improvement of other related parameters of the individuals. Other genes in the region (*SULT1A1* and *SULT1A2*, *TUFM*) harbored no obesity associated variants. Additional functional studies on the obesity associated variants detected here are warranted.

Zusammenfassung

Die Heritabilität der BMI-Varianz ist hoch. Durch große Genom-weite Assoziationsstudien wurden in der chromosomalen Region 16p11.2 zwei SNPs (rs7498665, rs7359397) identifiziert, die mit Adipositas assoziiert sind. In dieser Region befinden sich die Gene *SH2B1, APOB48R* und *SULT1A2*. *SH2B1* ist ein solides Adipositas-Kandidatengen, da es ein wichtiger Vermittler der Leptin- und Insulinrezeptor-Aktivität ist und dabei mehrere nachgeschaltete Signalwege beeinflusst. Die *Sh2b1* knockout Maus ist hyperphag und adipös, außerdem weist sie Hyperlipidemie, Leptinresistenz, Hyperglycämie, Insulinresistenz und Glukoseintoleranz auf. Der Makrophagenrezeptor APOB48R bindet Chymolmikrone, die Tryglycerid-reiche Lipoproteine enthalten, nach Aktivierung durch apoB48. Zusätzlich soll er Nahrungsfette und fettlösliche Vitamine für reticuloendotheliale Zellen bereitstellen. SULT1A2 könnte das Körpergewicht über die Regulation von Geschlechtshormonen wie Östrogenen und Androgenen beeinflussen. Viele Hormone werden von SULT1A2 sulfoniert.

In dieser Arbeit wurden die kodierenden Regionen der Gene *SH2B1*, *APOB48R* und *SULT1A2* bei 95 extrem adipösen Kindern und Jugendlichen mit dHPLC und SSCP auf (seltene) Mutationen untersucht. Gefundene kodierende Varianten wurden in unabhängigen Studiengruppen von bis zu 11,406 adipösen und übergewichtigen Fällen und 4,568 normalgenwichtigen Kontrollen auf Assoziation zu Adipositas analysiert (RFLP, ARMS PCR, TaqMan SNP Genotypisierung, MALDI TOF). Der Einfluss zweier kodierender Varianten in SH2B1 (rs7498665 [Thr484Ala], βThr656Ile/γPro674Ser) auf STAT3-vermitteltes Leptinsignalling wurde *in vitro* untersucht. Zusätzlich wurde der Einfluss von häufigen kodierenden Varianten (rs7498665 und rs180743 [APOB48R Pro419Ala]) auf die Ergebnis einer einjährigen Lebensstil-Interventionsstudie (“Obeldicks”) bezüglich Gewichtsreduktion und anthropometrische Parameter analysiert.

Alle untersuchten Gene enthielten bekannte Varianten und neue Mutationen. Nur für häufige kodierende Varianten ließ sich Assoziation zu Adipositas nachweisen: *SH2B1* (rs7498665 [Thr484Ala]; p=0.019, OR= 1.19) und *APOB48R* (rs180743 [Pro419Ala]; p=0.002, OR=1.17 und rs3833080 [Ala345_Gly346insdelAlaGlyThrAlaSerGlyGlyGluGluAlaGly]; p=0.003, OR= 1.25). Die *SH2B1*-Variante rs7498665 und die seltene Mutation βThr656Ile/γPro674Ser zeigten *in vitro* beide keinen Einfluss auf Leptinsignalling. Die Adipositas-Risikoallele der kodierenden SNPs in *SH2B1* und *APOB48R* zeigten keinen Effekt auf Gewichtsverlust und verwandte Parameter wie Blutdruck, Blutfette und HOMA IR nach der einjährigen Lebensstil-Interventionsstudie, obwohl die übergewichtigen und adipösen Kinder insgesamt den BMI SDS reduzierten und andere Parameter verbesserten. Andere Gene in der Region (*SULT1A1* und *SULT1A2*, *TUFM*) enthielten keine Varianten, die mit Adipositas assoziiert waren. Weitere funktionelle Studien zu den hier beschriebenen Varianten mit Adipositas-Assoziation wären von großem Interesse.

8. References

Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet. 2013 Jan;Chapter7:Unit7.20.

Allison DB, Fontaine KR, Manson JE, Stevens J, VanItallie TB. Annual deaths attributable to obesity in the United States. JAMA. 1999 Oct;282:1530-8.

Attia J, Ioannidis JP, Thakkinstian A, McEvoy M, Scott RJ, Minelli C, Thompson J, Infante-Rivard C, Guyatt G. How to use an article about genetic association: A: Background concepts. JAMA. 2009;301:74-81.

Attia J, Ioannidis JP, Thakkinstian A, McEvoy M, Scott RJ, Minelli C, Thompson J, Infante-Rivard C, Guyatt G. How to use an article about genetic association: B: Are the results of the study valid? JAMA. 2009;301:191-7.

Attia J, Ioannidis JP, Thakkinstian A, McEvoy M, Scott RJ, Minelli C, ThompsonJ, Infante-Rivard C, Guyatt G. How to use an article about genetic association: C: What are the results and will they help me in caring for my patients? JAMA. 2009;301:304-8.

Bachmann-Gagescu R, Mefford HC, Cowan C, Glew GM, Hing AV, Wallace S, Bader PI, Hamati A, Reitnauer PJ, Smith R, Stockton DW, Muhle H, Helbig I, Eichler EE, Ballif BC, Rosenfeld J, Tsuchiya KD. Recurrent 200-kb deletions of 16p11.2 that include the SH2B1 gene are associated with developmental delay and obesity. Genet Med. 2010;12:641-7.

Bau AM, Ernert A, Schenk L, Wiegand S, Martus P, Grüters A, Krude H. Is there a further acceleration in the age at onset of menarche? A cross-sectional study in 1840 school children focusing on age and bodyweight at the onset of menarche. Eur J Endocrinol 2009, 160:107-13.

Beckers S, Zegers D, Van Gaal LF, Van Hul W. Replication of the SH2B1 rs7498665 association with obesity in a Belgian study population. Obes Facts 2010, 4:473-477.

Berndt SI, Vedantam S, Day F, Gustafsson S, Locke AE, Powell C, Kahali B, Croteau-Chonka DC, Winkler TW, Scherag A, Barroso I, Beckmann JS, Justice A, Lindgren CM, Pers T, Visscher P, Yang J, Boehnke M, Abecasis G, Willer CJ, Mohlke KI, North KE, Hirschhorn JN, Ingelsson E, Loos RJF, Speliotes EK for the GIANT Consortium. Meta-analysis of genetic associations in up to 339,224 individuals identify 67 new loci for BMI, confirming a neuronal contribution to body weight regulation and implicating several novel pathways. (Poster at the Annual Meeting of Human Genetics, San Francisco 2012).

Billingsley G, Vincent A, Deveault C, Héon E. Mutational analysis of SDCCAG8 in Bardet-Biedl syndrome patients with renal involvement and absent polydactyly. Ophthalmic Genet. 2012;33:150-4.

Bochukova EG, Huang N, Keogh J, Henning E, Purmann C, Blaszczyk K, Saeed S, Hamilton-Shield J, Clayton-Smith J, O'Rahilly S, Hurles ME, Farooqi IS. Large, rare chromosomal deletions associated with severe early-onset obesity. Nature. 2010;463:666-70.

Bonnefond A, Clément N, Fawcett K, Yengo L, Vaillant E, Guillaume JL, Dechaume A, Payne F, Roussel R, Czernichow S, Hercberg S, Hadjadj S, Balkau B, Marre M, Lantieri O, Langenberg C, Bouatia-Naji N; Meta-Analysis of Glucose and Insulin-Related Traits Consortium (MAGIC), Charpentier G, Vaxillaire M, Rocheleau G, Wareham NJ, Sladek R, McCarthy MI, Dina C, Barroso I, Jockers R, Froguel P. Rare MTNR1B variants impairing melatonin receptor 1B function contribute to type 2 diabetes. Nat Genet. 2012;44:297-301.

Bouchard C, Pérusse L. Genetic aspects of obesity. Ann N Y Acad Sci. 1993;699:26-35.

Bradfield JP, Taal HR, Timpson NJ, Scherag A, Lecoeur C, Warrington NM, Hypponen E, Holst C, Valcarcel B, Thiering E, Salem RM, Schumacher FR, Cousminer DL, Sleiman PM, Zhao J, Berkowitz RI, Vimaleswaran KS, Jarick I, Pennell CE, Evans DM, St Pourcain B, Berry DJ, Mook-Kanamori DO, Hofman A, Rivadeneira F, Uitterlinden AG, van Duijn CM, van der Valk RJ, de Jongste JC, Postma DS, Boomsma DI, Gauderman WJ, Hassanein MT, Lindgren CM, Mägi R, Boreham CA, Neville CE, Moreno LA, Elliott P, Pouta A, Hartikainen AL, Li M, Raitakari O, Lehtimäki T, Eriksson JG, Palotie A, Dallongeville J, Das S, Deloukas P, McMahon G, Ring SM, Kemp JP, Buxton JL, Blakemore AI, Bustamante M, Guxens M, Hirschhorn JN, Gillman MW, Kreiner-Møller E, Bisgaard H, Gilliland FD, Heinrich J, Wheeler E, Barroso I, O'Rahilly S, Meirhaeghe A, Sørensen TI, Power C, Palmer LJ, Hinney A, Widen E, Farooqi IS, McCarthy MI,

Froguel P, Meyre D, Hebebrand J, Jarvelin MR, Jaddoe VW, Smith GD, Hakonarson H, Grant SF; Early Growth Genetics Consortium. A genome-wide association meta-analysis identifies new childhood obesity loci. Nat Genet. 2012;44:526-31.

Bromberg Y, Rost B. SNAP: predict effect of non-synonymous polymorphisms on function. Nucleic Acids Res. 2007;35:3823-35.

Brown ML, Ramprasad MP, Umeda PK, Tanaka A, Kobayashi Y, Watanabe T, Shimoyamada H, Kuo WL, Li R, Song R, Bradley WA, Gianturco SH. A macrophage receptor for apolipoprotein B48: cloning, expression, and atherosclerosis. Proc Natl Acad Sci U S A. 2000;97:7488-7493.

Budowle B, Baechtel FS. Modifications to improve the effectiveness of restriction fragment length polymorphism typing. Appl Theor Electrophor.1990;1:181-7.

Buiting K. Prader-Willi syndrome and Angelman syndrome. Am J Med Genet C SeminMed Genet. 2010;154C:365-76.

Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz (2008) Nationale Verzehrsstudie II - Ergebnisbericht Teil 1. http://www.bmelv.de/SharedDocs/Downloads/Ernaehrung/NVS_Ergebnisbericht.html

Calle EE. Obesity and cancer. BMJ. 2007;335:1107-8.

Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. ESEfinder: a web resource to identify exonic splicing enhancers. Nucleic Acid Res 2003, 31: 3568-71.

Carter-Su C, Rui L, Herrington J. Role of the tyrosine kinase JAK2 in signal transduction by growth hormone. Pediatr Nephrol. 2000;14:550-7.

Cawley J, Meyerhoefer C. The medical care costs of obesity: an instrumental variables approach. J Health Econ. 2012;31:219-30.

Centis E, Marzocchi R, Di Luzio R, Moscatiello S, Salardi S, Villanova N, Marchesini G; G. C. Croce School of Health Promotion. A controlled, class-based multicomponent intervention to promote healthy lifestyle and to reduce the burden of childhood obesity. Pediatr Obes. 2012;7:436-45.

Chua S Jr. Signal transduction pathways for leptin: an embarrassment of riches. Diabetes. 2009;58:513-4.

Cochrane T, Davey R, Iqbal Z, Gidlow C, Kumar J, Chambers R, Mawby Y. NHS health checks through general practice: randomised trial of population cardiovascular risk reduction. BMC Public Health. 2012;12:944.

D-A-CH (Deutsche Gesellschaft für Ernährung, Österreichische Gesellschaft für Ernährung, Schweizerische Gesellschaft für Ernährungsforschung, Schweizerische Vereinigung für Ernährung) (2008) Referenzwerte für die Nährstoffzufuhr. 1. Auflage, 3. vollständig durchgesehener und korrigierter Nachdruck, Frankfurt/Main: Umschau/Braus Verlagsgesellschaft.

Daniel TO, Schneider WJ, Goldstein JL, Brown MS. Visualization of lipoprotein receptors by ligand blotting. J Biol Chem. 1983;258:4606-11.

de Luca C, Kowalski TJ, Zhang Y, Elmquist JK, Lee C, Kilimann MW, Ludwig T, Liu SM, Chua SC Jr. Complete rescue of obesity, diabetes, and infertility in db/db mice by neuron-specific LEPR-B transgenes. J Clin Invest. 2005;115:3484-93.

de Onis M, Blössner M, Borghi E. Global prevalence and trends of overweight and obesity among preschool children. Am J Clin Nutr. 2010;92:1257-64.

Dempfle A, Hinney A, Heinzel-Gutenbrunner M, Raab M, Geller F, Gudermann T, Schäfer H, Hebebrand J. Large quantitative effect of melanocortin-4 receptor gene mutations on body mass index. J Med Genet. 2004;41:795-800.

Despres JP. Cardiovascular disease under the influence of excess visceral fat. Crit Pathw Cardiol. 2007;6:51-9.

Dobe M, Geisler A, Hoffmann D, Kleber M, von Köding P, Lass N, Müther S, Pohl B, Rose K, Schaefer A, Többens ML, Vierhaus R, Winkel K, Reinehr T. The Obeldicks concept: An example for a successful outpatient lifestyle intervention for overweight or obese children and adolescents. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2011; 54: 628-35.

Dobson CM. Protein folding and misfolding. Nature. 2003;426:884-90

Doche ME, Bochukova EG, Su HW, Pearce LR, Keogh JM, Henning E, Cline JM, Dale A, Cheetham T, Barroso I, Argetsinger LS, O'Rahilly S, Rui L, Carter-Su C, Farooqi IS. Human SH2B1 mutations are associated with maladaptive behaviors and obesity. J Clin Invest. 2012;122:4732-6.

Dominiczak MH, Caslake MJ. Apolipoproteins: metabolic role and clinical biochemistry applications. Ann Clin Biochem. 2011;48:498-515.

Duan C, Li M, Rui L. 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, 279:43684-91.

Duan C, Yang H, White MF, Rui L. Disruption of the SH2-B gene causes age-dependent insulin resistance and glucose intolerance. Mol Cell Biol. 2004;24:7435-43.

Elks CE, den Hoed M, Zhao JH, Sharp SJ, Wareham NJ, Loos RJ, Ong KK. Variability in the heritability of body mass index: a systematic review and meta-regression. Front Endocrinol (Lausanne). 2012;3:29.

Elks CE, Loos RJ, Hardy R, Wills AK, Wong A, Wareham NJ, Kuh D, Ong KK. Adult obesity susceptibility variants are associated with greater childhood weight gain and a faster tempo of growth: the 1946 British Birth Cohort Study. Am J Clin Nutr. 2012;95:1150-6.

Ellingjord-Dale M, Lee E, Couto E, Ozhand A, Qureshi SA, Hofvind S, Van Den Berg DJ, Akslen LA, Grotmol T, Ursin G. Polymorphisms in hormone metabolism and growth factor genes and mammographic density in Norwegian postmenopausal hormone therapy users and non-users. Breast Cancer Res. 2012;14:R135

Emaus A, Espetvedt S, Veierød MB, Ballard-Barbash R, Furberg AS, Ellison PT, Jasienska G, Hjartåker A, Thune I. 17-beta-estradiol in relation to age at menarche and adult obesity in premenopausal women. Hum Reprod. 2008;23:919-27.

Fairbrother WG, Yeh RF, Sharp PA, Burge CB. Predictive identification of exonic splicing enhancers in human genes. Science. 2002, 297:1007-13.

Farber SL (1981) Identic twins reared apart. New York, Basic Books

Farooqi IS, Jebb SA, Langmack G, Lawrence E, Cheetham CH, Prentice AM, Hughes IA, McCamish MA, O'Rahilly S. Effects of recombinant leptin therapy in a child with congenital leptin deficiency. N Engl J Med. 1999;341:879-84.

Farooqi IS, Wangensteen T, Collins S, Kimber W, Matarese G, Keogh JM, Lank E, Bottomley B, Lopez-Fernandez J, Ferraz-Amaro I, Dattani MT, Ercan O, Myhre AG, Retterstol L, Stanhope R, Edge JA, McKenzie S, Lessan N, Ghodsi M, De Rosa V, Perna F, Fontana S, Barroso I, Undlien DE, O'Rahilly S. Clinical and molecular genetic spectrum of congenital deficiency of the leptin receptor. N Engl J Med. 2007;356:237-47.

Farooqi IS. Monogenic human obesity syndromes. Prog Brain Res. 2006;153:119-25.

Fischer-Posovszky P, von Schnurbein J, Moepps B, Lahr G, Strauss G, Barth TF, Kassubek J, Mühleder H, Möller P, Debatin KM, Gierschik P, Wabitsch M. A new missense mutation in the leptin gene causes mild obesity and hypogonadism without affecting T cell responsiveness. J Clin Endocrinol Metab. 2010;95:2836-40.

Flegal KM, Carroll MD, Kit BK, Ogden CL. Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. JAMA. 2012;307:491-7.

Flegal KM, Carroll MD, Ogden CL, Curtin LR. Prevalence and trends in obesity among US adults, 1999-2008. JAMA. 2010;303:235-41.

Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, Perry JR, Elliott KS, Lango H, Rayner NW, Shields B, Harries LW, Barrett JC, Ellard S, Groves CJ, Knight B, Patch AM, Ness AR, Ebrahim S, Lawlor DA, Ring SM, Ben-Shlomo Y, Jarvelin MR, Sovio U, Bennett AJ, Melzer D, Ferrucci L, Loos RJ, Barroso I, Wareham NJ, Karpe F, Owen KR, Cardon LR, Walker M, Hitman GA, Palmer CN, Doney AS, Morris AD, Smith GD, Hattersley AT, McCarthy MI. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science. 2007;316:889-94.

Fredriks AM, van Buuren S, Wit JM, Verloove-Vanhorick SP. Body index measurements in 1996-7 compared with 1980. Arch Dis Child 2000;82:107-12.

Freimuth RR, Wiepert M, Chute CG, Wieben ED, Weinshilboum RM. Human cytosolic sulfotransferase database mining: identification of seven novel genes and pseudogenes. Pharmacogenomics J. 2004;4:54-65.

Frisch RE. Body fat, menarche, fitness and fertility. Hum Reprod. 1987;2:521-33.

Fujita Y, Ezura Y, Bujo H, Nakajima T, Takahashi K, Kamimura K, Iino Y, Katayama Y, Saito Y, Emi M. Association of nucleotide variations in the apolipoprotein B48 receptor gene (APOB48R) with hypercholesterolemia. J Hum Genet. 50: 2005;203-209.

Funnell AP, Crossley M. Homo- and heterodimerization in transcriptional regulation. Adv Exp Med Biol. 2012;747:105-21.

Geller F, Reichwald K, Dempfle A, Illig T, Vollmert C, Herpertz S, Siffert W, Platzer M, Hess C, Gudermann T, Biebermann H, Wichmann HE, Schäfer H, Hinney A, Hebebrand J. Melanocortin-4 receptor gene variant I103 is negatively associated with obesity. Am J Hum Genet. 2004;74:572-81.

German Mikrozensus, Statistisches Bundesamt Wiesbaden 2010, https://www.zensus2011.de/

Ghose R, Omoluabi O, Gandhi A, Shah P, Strohacker K, Carpenter KC, McFarlin B, Guo T. Role of high-fat diet in regulation of gene expression of drug metabolizing enzymes and transporters. Life Sci. 2011;89:57-64.

Glatt H, Boeing H, Engelke CE, Ma L, Kuhlow A, Pabel U, Pomplun D, Teubner W, Meinl W. Human cytosolic sulphotransferases: genetics, characteristics, toxicological aspects. Mutat Res. 2001;482:27-40.

Glatt H, Meinl W. Pharmacogenetics of soluble sulfotransferases (SULTs). Naunyn Schmiedebergs Arch Pharmacol. 2004;369:55-68.

Glatt H, Schneider H, Murkovic M, Monien BH, Meinl W. Hydroxymethyl-substituted furans: mutagenicity in Salmonella typhimurium strains engineered for expression of various human and rodent sulphotransferases. Mutagenesis. 2012;27:41-8.

Göbel (in preparation) "Molekulargenetische Analyse des TUFM-Gens (Tu translation elongation factor mitochondrial gene) bei extrem adipösen Kindern und Jugendlichen. Medical PhD thesis at the University of Duisburg-Essen.

Goldstone AP, Beales PL. Genetic obesity syndromes. Front Horm Res. 2008;36:37-60.

Goren A, Ram O, Amit M, Keren H, Lev-Maor G, Vig I, Pupko T, Ast G. Comparative analysis identifies exonic splicing regulatory sequences—The complex definition of enhancers and silencers. Mol Cell. 2006, 22:769-81.

Graff M, Ngwa JS, Workalemahu T, Homuth G, Schipf S, Teumer A, Völzke H, Wallaschofski H, Abecasis GR, Edward L, Francesco C, Sanna S, Scheet P, Schlessinger D, Sidore C, Xiao X, Wang Z, Chanock SJ, Jacobs KB, Hayes RB, Hu F, Van Dam RM; The GIANT Consortium, Crout RJ, Marazita ML, Shaffer JR, Atwood LD, Fox CS, Heard-Costa NL, White C, Choh AC, Czerwinski SA, Demerath EW, Dyer TD, Towne B, Amin N, Oostra BA, Van Duijn CM, Zillikens MC, Esko T, Nelis M, Nikopensius T, Metspalu A, Strachan DP, Monda K, Qi L, North KE, Cupples LA, Gordon-Larsen P, Berndt SI. Genome-wide analysis of BMI in adolescents and young adults reveals additional insight into the effects of genetic loci over the life course. Hum Mol Genet. 2013 Jun 5. [Epub ahead of print]

Graninger J. Mutationsanalyse im kodierenden Bereich des Apolipoprotein B48 Rezeptorgens (APOB48R) bei extrem adipösen Kindern und Jugendlichen. Master thesis at the University of Duisburg-Essen, 2011.

Guo SS, Roche AF, Chumlea WC, Gardner JD, Siervogel RM. The predictive value of childhood body mass index values for overweight at age 35 y. Am J Clin Nutr. 1994;59:810-9.

Guo Y, Lanktree MB, Taylor KC, Hakonarson H, Lange LA, Keating BJ; IBC 50K SNP array BMI Consortium. Gene-centric meta-analyses of 108 912 individuals confirm known body mass index loci and reveal three novel signals. Hum Mol Genet. 2013;22:184-201.

Gutierrez-Aguilar R, Kim DH, Woods SC, Seeley RJ. Expression of new loci associated with obesity in diet-induced obese rats: from genetics to physiology. Obesity (Silver Spring). 2012, 2:306-12.

Harris RM, Waring RH, Kirk CJ, Hughes PJ. Sulfation of "estrogenic" alkylphenols and 17beta-estradiol by human platelet phenol sulfotransferases. J Biol Chem. 2000;275:159-66.

Hartz A, Giefer E, Rimm AA. Relative importance of the effect of family environment and heredity on obesity. Ann Hum Genet. 1977;41:185-93.

Haupt A, Thamer C, Heni M, Machicao F, Machann J, Schick F, Stefan N, Fritsche A, Häring HU, Staiger H: Novel obesity risk loci do not determine distribution of body fat depots: a whole-body MRI/MRS study. Obesity (Silver Spring). 2010, 18:1212-17.

Hayashi K, Yandell DW. How sensitive is PCR-SSCP? Hum Mutat. 1993;2:338-46.

Hebebrand J, Bammann K, Hinney A. [Genetic determinants of obesity. Current issues]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2010;53:674-80.

Hebebrand J, Geller F, Dempfle A, Heinzel-Gutenbrunner M, Raab M, Gerber G, Wermter AK, Horro FF, Blundell J, Schäfer H, Remschmidt H, Herpertz S, Hinney A. Binge-eating episodes are not characteristic of carriers of melanocortin-4 receptor gene mutations. Mol Psychiatry. 2004;9:796-800.

Hebebrand J, Heseker H, Himmelmann GW, Schäfer H, Remschmidt H: Altersperzentilen für den Body Mass Index aus Daten der Nationalen Verzehrstudie einschließlich einer Übersicht zu relevanten Einflussfaktoren. Aktuel Ernährungsmed. 1994, 19:259–65.

Hebebrand J, Hinney A, Knoll N, Volckmar AL, Scherag A. Molekulargenetische Aspekte der Körpergewichtsregulation. Dtsch Arztebl Int 2013; 110: 338-44.

Hebebrand J, Sommerlad C, Geller F, Görg T, Hinney A. The genetics of obesity: practical implications. Int J Obes Relat Metab Disord. 2001;25 Suppl 1:S10-8.

Hebebrand J, Volckmar AL, Knoll N, Hinney A. Chipping away the 'missing heritability': GIANT steps forward in the molecular elucidation of obesity - but still lots to go. Obes Facts. 2010;3:294-303.

Heid IM, Vollmert C, Hinney A, Döring A, Geller F, Löwel H, Wichmann HE, Illig T, Hebebrand J, Kronenberg F; KORA Group. Association of the 103I MC4R allele with decreased body mass in 7937 participants of two population based surveys. J Med Genet. 2005;42:e21.

Heinemeyer T, Wingender E, Reuter I, Hermjakob H, Kel AE, Kel OV, Ignatieva EV, Ananko EA, Podkolodnaya OA, Kolpakov FA, Podkolodny NL, Kolchanov NA: Databases on Transcriptional Regulation: TRANSFAC, TRRD, and COMPEL. Nucleic Acids Res. 1998, 26: 364-70.

Herbert A, Gerry NP, McQueen MB, Heid IM, Pfeufer A, Illig T, Wichmann HE, Meitinger T, Hunter D, Hu FB, Colditz G, Hinney A, Hebebrand J, Koberwitz K, Zhu X, Cooper R, Ardlie K, Lyon H, Hirschhorn JN, Laird NM, Lenburg ME, Lange C, Christman MF. A common genetic variant is associated with adult and childhood obesity. Science. 2006;312:279-83.

Hester JM, Wing MR, Li J, Palmer ND, Xu J, Hicks PJ, Roh BH, Norris JM, Wagenknecht LE, Langefeld CD, Freedman BI, Bowden DW, Ng MC. Implication of European-derived adiposity loci in African Americans. Int J Obes (Lond). 2012;36:465-73.

Hinney A, Bettecken T, Tarnow P, Brumm H, Reichwald K, Lichtner P, Scherag A, Nguyen TT, Schlumberger P, Rief W, Vollmert C, Illig T, Wichmann HE, Schäfer H, Platzer M, Biebermann H, Meitinger T, Hebebrand J: Prevalence, spectrum, and functional characterization of melanocortin-4 receptor gene mutations in a representative population-based sample and obese adults from Germany. J Clin Endocrinol Metab. 2006, 91:1761-1769.

Hinney A, Hebebrand J. Three at one swoop! Obes Facts. 2009;2:3-8.

Hinney A, Hohmann S, Geller F, Vogel C, Hess C, Wermter AK, Brokamp B, Goldschmidt H, Siegfried W, Remschmidt H, Schäfer H, Gudermann T, Hebebrand J: Melanocortin-4 receptor gene: case-control study and transmission disequilibrium test confirm that functionally relevant mutations are compatible with a major gene effect for extreme obesity. J Clin Endocrinol Metab. 2003, 88:4258-67.

Hinney A, Nguyen TT, Scherag A, Friedel S, Brönner G, Müller TD, Grallert H, Illig T, Wichmann HE, Rief W, Schäfer H, Hebebrand J: Genome wide association (GWA) study for early onset extreme

obesity supports the role of fat mass and obesity associated gene (FTO) variants. PloS One. 2007, 2:e1361.

Hinney A, Remschmidt H, Hebebrand J. Candidate gene polymorphisms in eating disorders. Eur J Pharmacol. 2000;410:147-59.

Hinney A, Schmidt A, Nottebom K, Heibült O, Becker I, Ziegler A, Gerber G, Sina M, Görg T, Mayer H, Siegfried W, Fichter M, Remschmidt H, Hebebrand J: Several mutations in the melanocortin-4 receptor gene including a nonsense and a frameshift mutation associated with dominantly inherited obesity in humans. J Clin Endocrinol Metab. 1999, 84:1483-86.

Hinney A, Vogel CI, Hebebrand J. From monogenic to polygenic obesity: recent advances. Eur Child Adolesc Psychiatry. 2010;19:297-310.

Hinney A, Volckmar AL, Knoll N. Melanocortin-4 receptor in energy homeostasis and obesity pathogenesis. Prog Mol Biol Transl Sci. 2013;114:147-91.

Holzapfel C, Grallert H, Huth C, Wahl S, Fischer B, Döring A, Rückert IM, Hinney A, Hebebrand J, Wichmann HE, Hauner H, Illig T, Heid IM. Genes and lifestyle factors in obesity: results from 12,462 subjects from MONICA/KORA. Int J Obes (Lond). 2010;34:1538-45.

Holzapfel C, Siegrist M, Rank M, Langhof H, Grallert H, Baumert J, Irimie C, Klopp N, Wolfarth B, Illig T, Hauner H, Halle M. Association of a MTNR1B gene variant with fasting glucose and HOMA-B in children and adolescents with high BMI-SDS. Eur J Endocrinol. 2011;164:205-12.

Horn L. Mutation screen for Sulfotransferase 1A1 gene in extremely obese children and adolescents. Master thesis at the University of Duisburg-Essen, 2012

Horsthemke B, Buiting K. Imprinting defects on human chromosome 15. Cytogenet Genome Res. 2006;113(1-4):292-9.

Hotta K, Kitamoto T, Kitamoto A, Mizusawa S, Matsuo T, Nakata Y, Hyogo H, Ochi H, Kamohara S, Miyatake N, Kotani K, Komatsu R, Itoh N, Mineo I, Wada J, Yoneda M, Nakajima A, Funahashi T, Miyazaki S, Tokunaga K, Masuzaki H, Ueno T, Chayama K, Hamaguchi K, Yamada K, Hanafusa T, Oikawa S, Yoshimatsu H, Sakata T, Tanaka K, Matsuzawa Y, Nakao K, Sekine A: Computed tomography analysis of the association between the SH2B1 rs7498665 single-nucleotide polymorphism and visceral fat area. J Hum Genet. 2011;56:716-9.

Hu G, Jousilahti P, Nissinen A, Antikainen R, Kivipelto M, Tuomilehto J. Body mass index and the risk of Parkinson disease. Neurology. 2006;67:1955-9.

Hui Y, Yasuda S, Liu MY, Wu YY, Liu MC. On the sulfation and methylation of catecholestrogens in human mammary epithelial cells and breast cancer cells. Biol Pharm Bull. 2008;31:769-73.

Hulens M, Beunen G, Claessens AL, Lefevre J, Thomis M, Philippaerts R, Borms J, Vrijens J, Lysens R, Vansant G. Trends in BMI among Belgian children, adolescents and adults from 1969 to 1996. Int J Obes Relat Metab Disord 2001;25:395-9.

Hullmann SE, Ryan JL, Ramsey RR, Chaney JM, Mullins LL. Measures of general pediatric quality of life: Child Health Questionnaire (CHQ), DISABKIDS Chronic Generic Measure (DCGM), KINDL-R, Pediatric Quality of Life Inventory (PedsQL) 4.0 Generic Core Scales, and Quality of My Life Questionnaire (QoML). Arthritis Care Res (Hoboken). 2011;63:S420-30.

Jacquemont S, Reymond A, Zufferey F, Harewood L, Walters RG, Kutalik Z, Martinet D, Shen Y, Valsesia A, Beckmann ND, Thorleifsson G, Belfiore M, Bouquillon S, Campion D, de Leeuw N, de Vries BB, Esko T, Fernandez BA, Fernández-Aranda F, Fernández-Real JM, Gratacòs M, Guilmatre A, Hoyer J, Jarvelin MR, Kooy RF, Kurg A, Le Caignec C, Männik K, Platt OS, Sanlaville D, Van Haelst MM, Villatoro Gomez S, Walha F, Wu BL, Yu Y, Aboura A, Addor MC, Alembik Y, Antonarakis SE, Arveiler B, Barth M, Bednarek N, Béna F, Bergmann S, Beri M, Bernardini L, Blaumeiser B, Bonneau D, Bottani A, Boute O, Brunner HG, Cailley D, Callier P, Chiesa J, Chrast J, Coin L, Coutton C, Cuisset JM, Cuvellier JC, David A, de Freminville B, Delobel B, Delrue MA, Demeer B, Descamps D, Didelot G, Dieterich K, Disciglio V, Doco-Fenzy M, Drunat S, Duban-Bedu B, Dubourg C, El-Sayed Moustafa JS, Elliott P, Faas BH, Faivre L, Faudet A, Fellmann F, Ferrarini A, Fisher R, Flori E, Forer L, Gaillard D, Gerard M, Gieger C, Gimelli S, Gimelli G, Grabe HJ, Guichet A, Guillin O, Hartikainen AL, Heron D, Hippolyte L, Holder M, Homuth G, Isidor B, Jaillard S, Jaros Z, Jiménez-Murcia S, Helas GJ, Jonveaux P, Kaksonen S, Keren B, Kloss-Brandstätter A, Knoers NV, Koolen DA, Kroisel PM, Kronenberg F, Labalme A, Landais E, Lapi E, Layet V, Legallic S, Leheup B, Leube B, Lewis S, Lucas J,

MacDermot KD, Magnusson P, Marshall C, Mathieu-Dramard M, McCarthy MI, Meitinger T, Mencarelli MA, Merla G, Moerman A, Mooser V, Morice-Picard F, Mucciolo M, Nauck M, Ndiaye NC, Nordgren A, Pasquier L, Petit F, Pfundt R, Plessis G, Rajcan-Separovic E, Ramelli GP, Rauch A, Ravazzolo R, Reis A, Renieri A, Richart C, Ried JS, Rieubland C, Roberts W, Roetzer KM, Rooryck C, Rossi M, Saemundsen E, Satre V, Schurmann C, Sigurdsson E, Stavropoulos DJ, Stefansson H, Tengström C, Thorsteinsdóttir U, Tinahones FJ, Touraine R, Vallée L, van Binsbergen E, Van der Aa N, Vincent-Delorme C, Visvikis-Siest S, Vollenweider P, Völzke H, Vulto-van Silfhout AT, Waeber G, Wallgren-Pettersson C, Witwicki RM, Zwolinksi S, Andrieux J, Estivill X, Gusella JF, Gustafsson O, Metspalu A, Scherer SW, Stefansson K, Blakemore AI, Beckmann JS, Froguel P: Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. Nature 2011, 478:97-102.

Jamshidi Y, Snieder H, Ge D, Spector TD, O'Dell SD: The SH2B gene is associated with serum leptin and body fat in normal female twins. Obesity (Silver Spring) 2007, 15:5-9.

Janitz M. Next-Generation Genome Sequencing: Towards Personalized Medicine. 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, ISBN: 978-3-527-32090-5

Jarick I, Vogel CI, Scherag S, Schäfer H, Hebebrand J, Hinney A, Scherag A: Novel common copy number variation for early onset extreme obesity on chromosome 11q11 identified by a genome-wide analysis. Hum Mol Genet 2011, 20:840-52.

Jasik CB, Lustig RH. Adolescent obesity and puberty: the "perfect storm". Ann N Y Acad Sci. 2008;1135:265-79.

Jeukendrup A, Gleeson M, (2005). Sports Nutrition. Human Kinetics: An Introduction to Energy Production and Performance.

Jiang ZG, Robson SC, Yao Z. Lipoprotein metabolism in nonalcoholic fatty liver disease. J Biomed Res. 2013; 27: 1–13.

Jiao H, Arner P, Hoffstedt J, Brodin D, Dubern B, Czernichow S, van't Hooft F, Axelsson T, Pedersen O, Hansen T, Sørensen TI, Hebebrand J, Kere J, Dahlman-Wright K, Hamsten A, Clement K, Dahlman I. Genome wide association study identifies KCNMA1 contributing to human obesity. BMC Med Genomics. 2011;4:51.

Jolly K, Lewis A, Beach J, Denley J, Adab P, Deeks JJ, Daley A, Aveyard P. Comparison of range of commercial or primary care led weight reduction programmes with minimal intervention control for weight loss in obesity: lighten Up randomised controlled trial. BMJ. 2011;343:d6500.

Jones AC, Austin J, Hansen N, Hoogendoorn B, Oefner PJ, Cheadle JP, O'Donovan MC. Optimal temperature selection for mutation detection by denaturing HPLC and comparison to single-stranded conformation polymorphism and heteroduplex analysis. Clin Chem. 1999;45:1133-40.

Kenedy AA, Cohen KJ, Loveys DA, Kato GJ, Dang CV. Identification and characterization of the novel centrosome-associated protein CCCAP. Gene. 2003;303:35-46.

Keys A, Fidanza F, Karvonen MJ, Kimura N, Taylor HL. Indices of relative weight and obesity. J Chronic Dis. 1972;25:329-43.

Kloting N, Stumvoll M, Bluher M. [The biology of visceral fat]. Internist (Berl). 2007;48:126-33.

Knoll KP, Hauner N. Kosten der Adipositas in der Bundesrepublik Deutschland-eine aktuelle Krankheitskostenstudie. Adipositas – Ursachen, Folgeerkrankungen, Therapie. 2008;2:204-10.

Kotani K, Wilden P, Pillay TS. SH2-Balpha is an insulin-receptor adapter protein and substrate that interacts with the activation loop of the insulin-receptor kinase. Biochem J. 1998;335:103-9.

Kowalski TJ, Liu SM, Leibel RL, Chua SC Jr. Transgenic complementation of leptin-receptor deficiency. I. Rescue of the obesity/diabetes phenotype of LEPR-null mice expressing a LEPR-B transgene. Diabetes. 2001;50:425-35.

Kromeyer-Hauschild K, Wabitsch M, Kunze D, Geller F, Geiß HC, Hesse V, von Hippel A, Jaeger U, Johnsen D, Korte W, Menner K, Müller G, Müller JM, Niemann-Pilatus A, Remer T, Schaefer F, Wittchen HU, Zabransky S, Zellner K, Ziegler A, Hebebrand J. Perzentile für den Body-mass-Index für das Kindes-und Jugendalter unter Heranziehung verschiedener deutscher Stichproben. MonatsschrKinderheilkd 2001;149:807–18.

Kurelac I, Lang M, Zuntini R, Calabrese C, Simone D, Vicario S, Santamaria M, Attimonelli M, Romeo G, Gasparre G. Searching for a needle in the haystack: comparing six methods to evaluate heteroplasmy in difficult sequence context. Biotechnol Adv. 2012;30:363-71.

Kurth BM, Schaffrath Rosario A. [The prevalence of overweight and obese children and adolescents living in Germany. Results of the German health interview and examination survey for children and adolescents (KiGGS)]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2007;50:736-43.

Kwon JM, Goate AM. The candidate gene approach. Alcohol Res Health. 2000;24:164-8.

Lanning NJ, Su HW, Argetsinger LS, Carter-Su C. Identification of SH2B1β as a focal adhesion protein that regulates focal adhesion size and number. J Cell Sci. 2011;124:3095-105.

Lei Y, Wen H, Yu Y, Taxman DJ, Zhang L, Widman DG, Swanson KV, Wen KW, Damania B, Moore CB, Giguère PM, Siderovski DP, Hiscott J, Razani B, Semenkovich CF, Chen X, Ting JP. The mitochondrial proteins NLRX1 and TUFM form a complex that regulates type I interferon and autophagy. Immunity. 2012;36:933-46.

Li M, Li Z, Morris DL, Rui L. Identification of SH2B2beta as an inhibitor for SH2B1- and SH2B2alpha-promoted Janus kinase-2 activation and insulin signaling. Endocrinology. 2007;148:1615-21.

Li Z, Zhou Y, Carter-Su C, Myers MG Jr, Rui L: SH2B1 enhances leptin signaling by both Janus kinase 2 Tyr813 phosphorylation-dependent and –independent mechanisms. Mol Endocrinol 2007, 21:2270-81.

Ling M, Merante F, Chen HS, Duff C, Duncan AM, Robinson BH. The human mitochondrial elongation factor tu (EF-Tu) gene: cDNA sequence, genomic localization, genomic structure, and identification of a pseudogene. Gene. 1997;197:325-36.

Lipman DJ, Souvorov A, Koonin EV, Panchenko AR, Tatusova TA. The relationship of protein conservation and sequence length. BMC Evol Biol. 2002;2:20.

Liu J, Stewart JT. Quantitation of trimipramine enantiomers in human serum by enantioselective high-performance liquid chromatography and mixed-mode disc solid-phase extraction. J Chromatogr B Biomed Sci Appl. 1997;700:175-82.

Liu YJ, Liu XG, Wang L, Dina C, Yan H, Liu JF, Levy S, Papasian CJ, Drees BM, Hamilton JJ, Meyre D, Delplanque J, Pei YF, Zhang L, Recker RR, Froguel P, Deng HW. Genome-wide association scans identified CTNNBL1 as a novel gene for obesity. Hum Mol Genet. 2008;17:1803-13.

Loftus BJ, Kim UJ, Sneddon VP, Kalush F, Brandon R, Fuhrmann J, Mason T, Crosby ML, Barnstead M, Cronin L, Deslattes Mays A, Cao Y, Xu RX, Kang HL, Mitchell S, Eichler EE, Harris PC, Venter JC, Adams MD. Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q. Genomics. 1999;60:295-308.

Loos RJ, Lindgren CM, Li S, Wheeler E, Zhao JH, Prokopenko I, Inouye M, Freathy RM, Attwood AP, Beckmann JS, Berndt SI; Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, Jacobs KB, Chanock SJ, Hayes RB, Bergmann S, Bennett AJ, Bingham SA, Bochud M, Brown M, Cauchi S, Connell JM, Cooper C, Smith GD, Day I, Dina C, De S, Dermitzakis ET, Doney AS, Elliott KS, Elliott P, Evans DM, Sadaf Farooqi I, Froguel P, Ghori J, Groves CJ, Gwilliam R, Hadley D, Hall AS, Hattersley AT, Hebebrand J, Heid IM; KORA, Lamina C, Gieger C, Illig T, Meitinger T, Wichmann HE, Herrera B, Hinney A, Hunt SE, Jarvelin MR, Johnson T, Jolley JD, Karpe F, Keniry A, Khaw KT, Luben RN, Mangino M, Marchini J, McArdle WL, McGinnis R, Meyre D, Munroe PB, Morris AD, Ness AR, Neville MJ, Nica AC, Ong KK, O'Rahilly S, Owen KR, Palmer CN, Papadakis K, Potter S, Pouta A, Qi L; Nurses' Health Study, Randall JC, Rayner NW, Ring SM, Sandhu MS, Scherag A, Sims MA, Song K, Soranzo N, Speliotes EK; Diabetes Genetics Initiative, Syddall HE, Teichmann SA, Timpson NJ, Tobias JH, Uda M; SardiNIA Study, Vogel CI, Wallace C, Waterworth DM, Weedon MN; Wellcome Trust Case Control Consortium, Willer CJ; FUSION, Wraight, Yuan X, Zeggini E, Hirschhorn JN, Strachan DP, Ouwehand WH, Caulfield MJ, Samani NJ, Frayling TM, Vollenweider P, Waeber G, Mooser V, Deloukas P, McCarthy MI, Wareham NJ, Barroso I, Jacobs KB, Chanock SJ, Hayes RB, Lamina C, Gieger C, Illig T, Meitinger T, Wichmann HE, Kraft P, Hankinson SE, Hunter DJ, Hu FB, Lyon HN, Voight BF, Ridderstrale M, Groop L, Scheet P, Sanna S, Abecasis GR, Albai G, Nagaraja R, Schlessinger D, Jackson AU, Tuomilehto J, Collins FS, Boehnke M, Mohlke KL. Common variants near MC4R are associated with fat mass, weight and risk of obesity. Nat Genet. 2008;40:768-75.

Loos RJ, Rankinen T, Pérusse L, Tremblay A, Després JP, Bouchard C. Association of lipin 1 gene polymorphisms with measures of energy and glucose metabolism. Obesity (Silver Spring). 2007;15:2723-32.

Lu J, Li H, Zhang J, Li M, Liu MY, An X, Liu MC, Chang W. Crystal structures of SULT1A2 and SULT1A1 *3: insights into the substrate inhibition and the role of Tyr149 in SULT1A2. Biochem Biophys Res Commun. 2010;396:429-34.

Lumeng CN, Deyoung SM, Bodzin JL, Saltiel AR. Increased inflammatory properties of adipose tissue macrophages recruited during diet-induced obesity. Diabetes. 2007;56:16-23.

Maes HH, Neale MC, Eaves LJ. Genetic and environmental factors in relative body weight and human adiposity. Behav Genet 1997;27:325-351.

Mahabir S, Baer DJ, Giffen C, Clevidence BA, Campbell WS, Taylor PR, Hartman TJ. Comparison of energy expenditure estimates from 4 physical activity questionnaires with doubly labeled water estimates in postmenopausal women. Am J Clin Nutr. 2006;84:230-6.

Martin J, Han C, Gordon LA, Terry A, Prabhakar S, She X, Xie G, Hellsten U, Chan YM, Altherr M, Couronne O, Aerts A, Bajorek E, Black S, Blumer H, Branscomb E, Brown NC, Bruno WJ, Buckingham JM, Callen DF, Campbell CS, Campbell ML, Campbell EW, Caoile C, Challacombe JF, Chasteen LA, Chertkov O, Chi HC, Christensen M, Clark LM, Cohn JD, Denys M, Detter JC, Dickson M, Dimitrijevic-Bussod M, Escobar J, Fawcett JJ, Flowers D, Fotopulos D, Glavina T, Gomez M, Gonzales E, Goodstein D, Goodwin LA, Grady DL, Grigoriev I, Groza M, Hammon N, Hawkins T, Haydu L, Hildebrand CE, Huang W, Israni S, Jett J, Jewett PB, Kadner K, Kimball H, Kobayashi A, Krawczyk MC, Leyba T, Longmire JL, Lopez F, Lou Y, Lowry S, Ludeman T, Manohar CF, Mark GA, McMurray KL, Meincke LJ, Morgan J, Moyzis RK, Mundt MO, Munk AC, Nandkeshwar RD, Pitluck S, Pollard M, Predki P, Parson-Quintana B, Ramirez L, Rash S, Retterer J, Ricke DO, Robinson DL, Rodriguez A, Salamov A, Saunders EH, Scott D, Shough T, Stallings RL, Stalvey M, Sutherland RD, Tapia R, Tesmer JG, Thayer N, Thompson LS, Tice H, Torney DC, Tran-Gyamfi M, Tsai M, Ulanovsky LE, Ustaszewska A, Vo N, White PS, Williams AL, Wills PL, Wu JR, Wu K, Yang J, Dejong P, Bruce D, Doggett NA, Deaven L, Schmutz J, Grimwood J, Richardson P, Rokhsar DS, Eichler EE, Gilna P, Lucas SM, Myers RM, Rubin EM, Pennacchio LA. The sequence and analysis of duplication-rich human chromosome 16. Nature. 2004;432:988-94.

Mazen I, El-Gammal M, Abdel-Hamid M, Amr K. A novel homozygous missense mutation of the leptin gene (N103K) in an obese Egyptian patient. Mol Genet Metab. 2009;97:305-8.

Mefford HC, Eichler EE. Duplication hotspots, rare genomic disorders and common disease. Curr Opin Genet Dev. 2009; 19: 196–204.

Meinl W, Ebert B, Glatt H, Lampen A. Sulfotransferase forms expressed in human intestinal Caco-2 and TC7 cells at varying stages of differentiation and role in benzo[a]pyrene metabolism. Drug Metab Dispos. 2008;36:276-83.

Meinl W, Meerman JH, Glatt H. Differential activation of promutagens by alloenzymes of human sulfotransferase 1A2 expressed in Salmonella typhimurium. Pharmacogenetics. 2002;12:677-89.

Meyre D, Delplanque J, Chèvre JC, Lecoeur C, Lobbens S, Gallina S, Durand E, Vatin V, Degraeve F, Proença C, Gaget S, Körner A, Kovacs P, Kiess W, Tichet J, Marre M, Hartikainen AL, Horber F, Potoczna N, Hercberg S, Levy-Marchal C, Pattou F, Heude B, Tauber M, McCarthy MI, Blakemore AI, Montpetit A, Polychronakos C, Weill J, Coin LJ, Asher J, Elliott P, Järvelin MR, Visvikis-Siest S, Balkau B, Sladek R, Balding D, Walley A, Dina C, Froguel P. Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. Nat Genet. 2009;41:157-9.

Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988;16:1215.

Moll PP, Burns TL, Lauer RM. The genetic and environmental sources of body mass index variability: the Muscatine Ponderosity Family Study. Am J Hum Genet. 1991;49:1243-55.

Monda KL, Chen GK, Taylor KC, Palmer C, Edwards TL, Lange LA, Ng MC, Adeyemo AA, Allison MA, Bielak LF, Chen G, Graff M, Irvin MR, Rhie SK, Li G, Liu Y, Liu Y, Lu Y, Nalls MA, Sun YV, Wojczynski MK, Yanek LR, Aldrich MC, Ademola A, Amos CI, Bandera EV, Bock CH, Britton A, Broeckel U, Cai Q, Caporaso NE, Carlson CS, Carpten J, Casey G, Chen WM, Chen F, Chen

YD, Chiang CW, Coetzee GA, Demerath E, Deming-Halverson SL, Driver RW, Dubbert P, Feitosa MF, Feng Y, Freedman BI, Gillanders EM, Gottesman O, Guo X, Haritunians T, Harris T, Harris CC, Hennis AJ, Hernandez DG, McNeill LH, Howard TD, Howard BV, Howard VJ, Johnson KC, Kang SJ, Keating BJ, Kolb S, Kuller LH, Kutlar A, Langefeld CD, Lettre G, Lohman K, Lotay V, Lyon H, Manson JE, Maixner W, Meng YA, Monroe KR, Morhason-Bello I, Murphy AB, Mychaleckyj JC, Nadukuru R, Nathanson KL, Nayak U, N'diaye A, Nemesure B, Wu SY, Leske MC, Neslund-Dudas C, Neuhouser M, Nyante S, Ochs-Balcom H, Ogunniyi A, Ogundiran TO, Ojengbede O, Olopade OI, Palmer JR, Ruiz-Narvaez EA, Palmer ND, Press MF, Rampersaud E, Rasmussen-Torvik LJ, Rodriguez-Gil JL, Salako B, Schadt EE, Schwartz AG, Shriner DA, Siscovick D, Smith SB, Wassertheil-Smoller S, Speliotes EK, Spitz MR, Sucheston L, Taylor H, Tayo BO, Tucker MA, Van Den Berg DJ, Edwards DR, Wang Z, Wiencke JK, Winkler TW, Witte JS, Wrensch M, Wu X, Yang JJ, Levin AM, Young TR, Zakai NA, Cushman M, Zanetti KA, Zhao JH, Zhao W, Zheng Y, Zhou J, Ziegler RG, Zmuda JM, Fernandes JK, Gilkeson GS, Kamen DL, Hunt KJ, Spruill IJ, Ambrosone CB, Ambs S, Arnett DK, Atwood L, Becker DM, Berndt SI, Bernstein L, Blot WJ, Borecki IB, Bottinger EP, Bowden DW, Burke G, Chanock SJ, Cooper RS, Ding J, Duggan D, Evans MK, Fox C, Garvey WT, Bradfield JP, Hakonarson H, Grant SF, Hsing A, Chu L, Hu JJ, Huo D, Ingles SA, John EM, Jordan JM, Kabagambe EK, Kardia SL, Kittles RA, Goodman PJ, Klein EA, Kolonel LN, Le Marchand L, Liu S, McKnight B, Millikan RC, Mosley TH, Padhukasahasram B, Williams LK, Patel SR, Peters U, Pettaway CA, Peyser PA, Psaty BM, Redline S, Rotimi CN, Rybicki BA, Sale MM, Schreiner PJ, Signorello LB, Singleton AB, Stanford JL, Strom SS, Thun MJ, Vitolins M, Zheng W, Moore JH, Williams SM, Ketkar S, Zhu X, Zonderman AB; NABEC Consortium; UKBEC Consortium; BioBank Japan Project; AGEN Consortium, Kooperberg C, Papanicolaou GJ, Henderson BE, Reiner AP, Hirschhorn JN, Loos RJ, North KE, Haiman CA. A meta-analysis identifies new loci associated with body mass index in individuals of African ancestry. Nat Genet. 2013;45:690-6.

Montague CT, Farooqi IS, Whitehead JP, Soos MA, Rau H, Wareham NJ, Sewter CP, Digby JE, Mohammed SN, Hurst JA, Cheetham CH, Earley AR, Barnett AH, Prins JB, O'Rahilly S. Congenital leptin deficiency is associated with severe early-onset obesity in humans. Nature. 1997;387:903-8.

Moreno LA, Sarria A, Fleta J, Rodriguez G, Gonzalez JM, Bueno M. Sociodemographic factors and trends on overweight prevalence in children and adolescents in Aragon (Spain) from 1985 to 1995. J Clin Epidemiol 2001;54:921-27.

Morris DL, Cho KW, Rui L. Critical role of the Src homology 2 (SH2) domain of neuronal SH2B1 in the regulation of body weight and glucose homeostasis in mice. Endocrinology 2010, 151:3643-51.

Morris DL, Cho KW, Zhou Y, Rui L: SH2B1 enhances insulin sensitivity by both stimulating the insulin receptor and inhibiting tyrosine dephosphorylation of insulin receptor substrate proteins. Diabetes 2009, 58:2039-47.

Morris DL, Rui L. Recent advances in understanding leptin signalling and leptin resistance. Am J Physiol Endocrinol Metab 2009. 297:E1247-59.

Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harb Symp Quant Biol. 1986;51:263-73.

Nagel G, Wabitsch M, Galm C, Berg S, Brandstetter S, Fritz M, Klenk J, Peter R, Prokopchuk D, Steiner R, Stroth S, Wartha O, Weiland SK, Steinacker J: Determinants of obesity in the Ulm Research on Metabolism, Exercise and Lifestyle in Children (URMEL-ICE). Eur J Pediatr 2009, 10:1259-67.

Nelms K, O'Neill TJ, Li S, Hubbard SR, Gustafson TA, Paul WE. Alternative splicing, gene localization, and binding of SH2-B to the insulin receptor kinase domain. Mamm Genome. 1999;10:1160-7.

Ng MC, Hester JM, Wing MR, Li J, Xu J, Hicks PJ, Roh BH, Lu L, Divers J, Langefeld CD, Freedman BI, Palmer ND, Bowden DW. Genome-wide association of BMI in African Americans. Obesity (Silver Spring). 2012;20:622-7.

Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003;31:3812-4.

Nishi M, Werner ED, Oh BC, Frantz JD, Dhe-Paganon S, Hansen L, Lee J, Shoelson SE. Kinase activation through dimerization by human SH2-B. Mol Cell Biol. 2005;25:2607-21.

Nishina M, Kikuchi T, Yamazaki H, Kameda K, Hiura M, Uchiyama M. Relationship among systolic blood pressure, serum insulin and leptin, and visceral fat accumulation in obese children. Hypertens Res. 2003;26:281-8.

O'Brien KB, O'Shea JJ, Carter-Su C. SH2-B family members differentially regulate JAK family tyrosine kinases. J Biol Chem. 2002;277:8673-81.

O'Donovan MC, Oefner PJ, Roberts SC, Austin J, Hoogendoorn B, Guy C, Speight G, Upadhyaya M, Sommer SS, McGuffin P. Blind analysis of denaturing high-performance liquid chromatography as a tool for mutation detection. Genomics. 1998;52:44-9.

Ogden CL, Fryar CD, Carroll MD, Flegal KM. Mean body weight, height, and body mass index, United States 1960-2002. Adv Data 2004;347:1-17.

Okada Y, Kubo M, Ohmiya H, Takahashi A, Kumasaka N, Hosono N, Maeda S, Wen W, Dorajoo R, Go MJ, Zheng W, Kato N, Wu JY, Lu Q; GIANT consortium, Tsunoda T, Yamamoto K, Nakamura Y, Kamatani N, Tanaka T. Common variants at CDKAL1 and KLF9 are associated with body mass index in east Asian populations. Nat Genet. 2012;44:302-6.

Olds T, Maher C, Zumin S, Péneau S, Lioret S, Castetbon K, Bellisle, de Wilde J, Hohepa M, Maddison R, Lissner L, Sjöberg A, Zimmermann M, Aeberli I, Ogden C, Flegal K, Summerbell C. Evidence that the prevalence of childhood overweight is plateauing: data from nine countries. Int J Pediatr Obes. 2011;6:342-60.

Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. Proc Natl Acad Sci U S A. 1989;86:2766-70.

Pate RR, Stevens J, Webber LS, et al. Age-related change in physical activity in adolescent girls. J Adolesc Health 2009; 44: 275–82.

Perez-Iglesias R, Mata I, Amado JA, Berja A, Garcia-Unzueta MT, Martínez García O, Arranz MJ, Vazquez-Barquero JL, Crespo-Facorro B.Effect of FTO, SH2B1, LEP, and LEPR polymorphisms on weight gain associated with antipsychotic treatment. J Clin Psychopharmacol . 2010;6:661-6.

Perrone L, Marzuillo P, Grandone A, del Giudice EM. Chromosome 16p11.2 deletions: another piece in the genetic puzzle of childhood obesity. Ital JnPediatr. 2010;36:43.

Petersen TA, Rasmussen S, Madsen M. [BMI of Danish school children measured during the periods 1986/1987--1996/1997 compared to Danish measurement in 1971/1972]. Ugeskr Laeger 2002;164:5006-10.

Pietiläinen KH, Virtanen SM, Rissanen A, Rita H, Mäenpää J. Distribution and heritability of BMI in Finnish adolescents aged 16y and 17y: a study of 4884 twins and 2509 singletons. Int J Obes Relat Metab Disord. 1999;23:107-15.

Powell WT, Coulson RL, Crary FK, Wong SS, Ach RA, Tsang P, Yamada NA, Yasui DH, Lasalle JM. A Prader-Willi locus lncRNA cloud modulates diurnal genes and energy expenditure. Hum Mol Genet. 2013 Jun 13. [Epub ahead of print]

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559-75.

Pusch W, Wurmbach JH, Thiele H, Kostrzewa M. MALDI-TOF mass spectrometry-based SNP genotyping. Pharmacogenomics. 2002;3:537-48.

Raftogianis RB, Wood TC, Weinshilboum RM. Human phenol sulfotransferases SULT1A2 and SULT1A1: genetic polymorphisms, allozyme properties, and human liver genotype-phenotype correlations. Biochem Pharmacol. 1999;58:605-16.

Reinehr T, de Sousa G, Toschke AM, Andler W. Long-term follow-up of cardiovascular disease risk factors in children after an obesity intervention. Am J Clin Nutr. 2006;3:490-6.

Reinehr T, Temmesfeld M, Kersting M, de Sousa G, Toschke AM. Four-year follow-up of children and adolescents participating in an obesity intervention program. Int J Obes (Lond). 2007;7:1074-7

Reinehr T, Widhalm K, l'Allemand D, Wiegand S, Wabitsch M, Holl RW; APV-Wiss STudy Group and German Competence Net Obesity. Two-year follow-up in 21,784 overweight children and adolescents with lifestyle intervention. Obesity (Silver Spring). 2009;17:1196-9.

Ren D, Li M, Duan C, Rui L: Identification of SH2-B as a key regulator of leptin sensitivity, energy balance, and body weight in mice. Cell Metab 2005;2:95-104.

Ren D, Zhou Y, Morris D, Li M, Li Z, Rui L: Neuronal SH2B1 is essential for controlling energy and glucose homeostasis. J Clin Invest 2007;117:397-406.

Renström F, Payne F, Nordström A, Brito EC, Rolandsson O, Hallmans G, Barroso I, Nordström P, Franks PW; GIANT Consortium: Replication and extension of genome-wide association study results for obesity in 4923 adults from northern Sweden. Hum Mol Genet 2009, 18:1489-1496.

Revicki DA, Israel RG. Relationship between body mass indices and measures of body adiposity. Am J Public Health. 1986;76:992-4.

Richter DC, Ott F, Auch AF, Schmid R, Huson DH. MetaSim: a sequencing simulator for genomics and metagenomics. PLoS One. 2008;3:e3373.

Riedel H, Wang J, Hansen H, Yousaf N. PSM, an insulin-dependent, pro-rich, PH, SH2 domain containing partner of the insulin receptor. J Biochem. 1997;122:1105-13.

Robiou-du-Pont S, Bonnefond A, Yengo L, Vaillant E, Lobbens S, Durand E, Weill J, Lantieri O, Balkau B, Charpentier G, Marre M, Froguel P, Meyre D. Contribution of 24 obesity-associated genetic variants to insulin resistance, pancreatic beta-cell function and type 2 diabetes risk in the French population. Int J Obes (Lond). 2012 Oct 23. [Epub ahead of print]

Robiou-du-Pont S, Yengo L, Vaillant E, Lobbens S, Durand E, Horber F, Lantieri O, Marre M, Balkau B, Froguel P, Meyre D. Common variants near BDNF and SH2B1 show nominal evidence of association with snacking behavior in European populations. J Mol Med (Berl). 2013 May 3. [Epub ahead of print]

Rodríguez-Moran M, Guerrero-Romero F, Aradillas-García C, Violante R, Simental-Mendia LE, Monreal-Escalante E, Mendoza Ede L. Obesity and family history of diabetes as risk factors of impaired fasting glucose: implications for the early detection of prediabetes. Pediatr Diabetes. 2010;11:331-6.

Rolland-Cachera MF, Cole TJ, Sempé M, Tichet J, Rossignol C, Charraud A. Body mass index variations: centiles from birth to 87 years. Eur J Clin Nutr 1991;45:13-21.

Rosenblum CI, Vongs A, Tota MR, Varnerin JP, Frazier E, Cully DF, Morsy MA, Van der Ploeg LH. A rapid, quantitative functional assay for measuring leptin. Mol Cell Endocrinol. 1998;143:117-23.

Rui L, Carter-Su C. Identification of SH2-bbeta as a potent cytoplasmic activator of the tyrosine kinase Janus kinase 2. Proc Natl Acad Sci U S A. 1999;96:7172-7.

Rui L, Herrington J, Carter-Su C. SH2-B, a membrane-associated adapter, is phosphorylated on multiple serines/threonines in response to nerve growth factor by kinases within the MEK/ERK cascade. J Biol Chem. 1999;274:26485-92.

Rui L, Mathews LS, Hotta K, Gustafson TA, Carter-Su C. Identification of SH2-Bbeta as a substrate of the tyrosine kinase JAK2 involved in growth hormone signalling. Mol Cell Biol. 1997;17:6633-44.

Sallis JF. Age-related decline in physical activity: a synthesis of human and animal studies. Med Sci Sports Exerc. 2000;32: 1598–600.

Sandelin A, Wasserman WW, Lenhard B: ConSite: web-based prediction of regulatory elements using cross-species comparison. Nucleic Acids Res 2004 32:W249-252.

Sandholt CH, Vestmar MA, Bille DS, Borglykke A, Almind K, Hansen L, Sandbæk A, Lauritzen T, Witte D, Jørgensen T, Pedersen O, Hansen T. Studies of metabolic phenotypic correlates of 15 obesity associated gene variants. PLoS One. 2011;6:e23531.

Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A. 1977;74:5463-7.

Santos AR, Duarte CB. Validation of internal control genes for expression studies: effects of the neurotrophin BDNF on hippocampal neurons. J Neurosci Res. 2008;86:3684-92.

Schaefer E, Zaloszyc A, Lauer J, Durand M, Stutzmann F, Perdomo-Trujillo Y, Redin C, Bennouna Greene V, Toutain A, Perrin L, Gérard M, Caillard S, Bei X, Lewis RA, Christmann D, Letsch J, Kribs M, Mutter C, Muller J, Stoetzel C, Fischbach M, Marion V, Katsanis N, Dollfus H. Mutations

in SDCCAG8/NPHP10 Cause Bardet-Biedl Syndrome and Are Associated with Penetrant Renal Disease and Absent Polydactyly. Mol Syndromol. 2011;1:273-281.

Scherag A, Dina C, Hinney A, Vatin V, Scherag S, Vogel CI, Müller TD, Grallert H, Wichmann HE, Balkau B, Heude B, Jarvelin MR, Hartikainen AL, Levy-Marchal C, Weill J, Delplanque J, Körner A, Kiess W, Kovacs P, Rayner NW, Prokopenko I, McCarthy MI, Schäfer H, Jarick I, Boeing H, Fisher E, Reinehr T, Heinrich J, Rzehak P, Berdel D, Borte M, Biebermann H, Krude H, Rosskopf D, Rimmbach C, Rief W, Fromme T, Klingenspor M, Schürmann A, Schulz N, Nöthen MM, Mühleisen TW, Erbel R, Jöckel KH, Moebus S, Boes T, Illig T, Froguel P, Hebebrand J, Meyre D: Two new Loci for body-weight regulation identified in a joint analysis of genome-wide association studies for early-onset extreme obesity in French and german study groups. PloS Genet 2010, 6:e1000916.

Scherag A, Kleber M, Boes T, Kolbe AL, Ruth A, Grallert H, Illig T, Heid IM, Toschke AM, Grau K; NUGENOB Consortium, Sørensen TI, Hebebrand J, Hinney A, Reinehr T. SDCCAG8 obesity alleles and reduced weight loss after a lifestyle intervention in overweight children and adolescents. Obesity (Silver Spring). 2012;20:466-70.

Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods. 2010;7:575-6.

Scuteri A, Sanna S, Chen WM, Uda M, Albai G, Strait J, Najjar S, Nagaraja R, Orrú M, Usala G, Dei M, Lai S, Maschio A, Busonero F, Mulas A, Ehret GB, Fink AA, Weder AB, Cooper RS, Galan P, Chakravarti A, Schlessinger D, Cao A, Lakatta E, Abecasis GR. Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. PLoS Genet. 2007;3:e115.

Sheffield VC, Beck JS, Kwitek AE, Sandstrom DW, Stone EM. The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substitutions. Genomics. 1993;16:325-32.

Shi J, Long J, Gao YT, Lu W, Cai Q, Wen W, Zheng Y, Yu K, Xiang YB, Hu FB, Zheng W, Shu XO. Evaluation of genetic susceptibility loci for obesity in Chinese women. Am J Epidemiol. 2010;172:244-54.

Shinawi M, Liu P, Kang SH, Shen J, Belmont JW, Scott DA, Probst FJ, Craigen WJ, Graham BH, Pursley A, Clark G, Lee J, Proud M, Stocco A, Rodriguez DL, Kozel BA, Sparagana S, Roeder ER, McGrew SG, Kurczynski TW, Allison LJ, Amato S, Savage S, Patel A, Stankiewicz P, Beaudet AL, Cheung SW, Lupski JR. Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. J Med Genet. 2010;47:332-41.

Song W, Ren D, Li W, Jiang L, Cho KW, Huang P, Fan C, Song Y, Liu Y, Rui L. SH2B regulation of growth, metabolism, and longevity in both insects and mammals. Cell Metab. 2010;11:427-37.

Sørensen TI, Holst C, Stunkard AJ. Childhood body mass index--genetic and familial environmental influences assessed in a longitudinal adoption study. Int J Obes Relat Metab Disord. 1992;16:705-14.

Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, Lango Allen H, Lindgren CM, Luan J, Mägi R, Randall JC, Vedantam S, Winkler TW, Qi L, Workalemahu T, Heid IM, Steinthorsdottir V, Weedon MN, Wheeler E, Wood AR, Ferreira T, Weyant RJ, Segrè AV, Estrada K, Liang L, Nemesh J, Park JH, Gustafsson S, Kilpeläinen TO, Yang J, Bouatia-Naji N, Esko T, Feitosa MF, Kutalik Z, Mangino M, Raychaudhuri S, Scherag A, Smith AV, Welch R, Zhao JH, Aben KK, Absher DM, Amin N, Dixon AL, Fisher E, Glazer NL, Goddard ME, Heard-Costa NL, Hoesel V, Hottenga JJ, Johansson Å, Johnson T, Ketkar S, Lamina C, Li S, Moffatt MF, Myers RH, Narisu N, Perry JRB, Peters MJ, Preuss M, Ripatti S, Rivadeneira F, Sandholt C, Scott LJ, Timpson NJ, Tyrer JP, van Wingerden S, Watanabe RM, White CC, Wiklund F, Barlassina C, Chasman DI, Cooper MN, Jansson JO, Lawrence RW, Pellikka N, Prokopenko I, Shi J, Thiering E, Alavere H, Alibrandi MTS, Almgren P, Arnold AM, Aspelund T, Atwood LD, Balkau B, Balmforth AJ, Bennett AJ, Ben-Shlomo Y, Bergman RN, Bergmann S, Biebermann H, Blakemore AIF, Boes T, Bonnycastle LL, Bornstein SR, Brown MJ, Buchanan TA, Busonero F, Campbell H, Cappuccio FP, Cavalcanti-Proença C, Chen Y-DI, Chen C-M, Chines PS, Clarke R, Coin L, Connell J, Day INM, den Heijer M, Duan J, Ebrahim S, Elliott P, Elosua R, Eiriksdottir G, Erdos MR, Eriksson JG, Facheris MF, Felix SB, Fischer-Posovszky P, Folsom AR, Friedrich N, Freimer NB, Fu M, Gaget S, Gejman PV, Geus EJC, Gieger C, Gjesing AP, Goel A, Goyette P, Grallert H,

Gräßler J, Greenawalt DM, Groves CJ, Gudnason V, Guiducci C, Hartikainen AL, Hassanali N, Hall AS, Havulinna AS, Hayward C, Heath AC, Hengstenberg C, Hicks AA, Hinney A, Hofman A, Homuth G, Hui J, Igl W, Iribarren C, Isomaa B, Jacobs KB, Jarick I, Jewell E, John U, Jørgensen T, Jousilahti P, Jula A, Kaakinen M, Kajantie E, Kaplan LM, Kathiresan S, Kettunen J, Kinnunen L, Knowles JW, Kolcic I, König IR, Koskinen S, Kovacs P, Kuusisto J, Kraft P, Kvaløy K, Laitinen J, Lantieri O, Lanzani C, Launer LJ, Lecoeur C, Lehtimäki T, Lettre G, Liu J, Lokki ML, Lorentzon M, Luben RN, Ludwig B, MAGIC, Manunta P, Marek D, Marre M, McArdle WL, McCarthy A, McKnight B, Meitinger T, Melander O, Meyre D, Midthjell K, Montgomery GW, Morken MA, Morris AP, Mulic R, Ngwa JS, Nelis M, Neville MJ, Nyholt DR, O'Donnell CJ, O'Rahilly S, Ong KK, Oostra B, Paré G, Parker AN, Perola M, Pichler I, Pietiläinen KH, Platou CGP, Polasek O, Pouta A, Rafelt S, Raitakari O, Rayner NW, Ridderstråle M, Rief W, Ruokonen A, Robertson NR, Rzehak P, Salomaa V, Sanders AR, Sandhu MS, Sanna S, Saramies J, Savolainen MJ, Scherag S, Schipf S, Schreiber S, Schunkert H, Silander K, Sinisalo J, Siscovick DS, Smit JH, Soranzo N, Sovio U, Stephens J, Stringham HM, Surakka I, Swift AJ, Tammesoo ML, Tardif JC, Teder-Laving M, Teslovich TM, Thompson JR, Thomson B, Tönjes A, Tuomi T,van Meurs JBJ, van Ommen GJ, Vatin V, Viikari J, Visvikis-Siest S, Vitart V, Vogel CI, Voight BF, Waite LL, Wallaschofski H, Walters BG, Widen E, Wiegand S, Wild SH, Willemsen G, Witte DR, Witteman JC, Xu J, Zhang Q, Zgaga L, Ziegler A, Zitting P, Beilby JP, I. Farooqi S, Hebebrand J, Huikuri HV, James AL, Kähönen M, Levinson DF, Macciardi F, Nieminen MS, Ohlsson C, Palmer LJ, Ridker PM, Stumvoll M, Beckmann JS, Boeing H, Boerwinkle E, Boomsma DI, Caulfield MJ, Chanock SJ, Collins FS, Cupples LA, Smith GD, Erdmann J, Froguel P, Grönberg H, Gyllensten U, Hall P, Hansen T, Harris TB, Hattersley AT, Hayes RB, Heinrich J, Hu FB, Hveem K, Illig T, Jarvelin MR, Kaprio J, Karpe F, Khaw KT, Kiemeney LA, Krude H, Laakso M, Lawlor DA, Metspalu A, Munroe PB, Ouwehand WH, Pedersen O, Penninx BW, Peters A, Pramstaller PP, Quertermous T, Reinehr T, Rissanen A, Rudan I, Samani NJ, Schwarz PEH, Shuldiner AR, Spector TD, Tuomilehto J, Uda M, Uitterlinden A, Valle TT, Wabitsch M, Waeber G, Wareham NJ, Watkins H on behalf of the Procardis Consortium, Wilson JF, Wright AF, Zillikens CM, Chatterjee N, McCarroll SA, Purcell S, Schadt EE, Visscher PM, Assimes TL, Borecki IB, Deloukas P, Fox CS, Groop LC, Haritunians T, Hunter DJ, Kaplan RC, Mohlke KL, O'Connell JR, Peltonen L, Schlessinger D, Strachan DP, van Duijn CM, Wichmann HE, Frayling TM, Thorsteinsdottir U, Abecasis GR, Barroso I, Boehnke M, Stefansson K, North KE, McCarthy MI, Hirschhorn JN, Ingelsson E, Loos RJF. Association analyses of 249,796 individuals reveal eighteen new loci associated with body mass index. Nat Genet. 2010;42:937-48.

Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet. 1993;52:506-16.

Spink BC, Katz BH, Hussain MM, Pang S, Connor SP, Aldous KM, Gierthy JF, Spink DC. SULT1A1 catalyzes 2-methoxyestradiol sulfonation in MCF-7 breast cancer cells. Carcinogenesis. 2000;21:1947-57.

Stamatakis E, Primatesta P, Chinn S, Rona R, Falascheti E. Overweight and obesity trends from 1974 to 2003 in English children: what is the role of socioeconomic factors? Arch Dis Child 2005;90:999-1004.

Stankiewicz P, Lupski JR. Structural variation in the human genome and ist role in disease. Annu Rev Med. 2010;61:437-55.

Storm N, Darnhofer-Patel B, van den Boom D, Rodi CP. MALDI-TOF mass spectrometry-based SNP genotyping. Methods Mol Biol. 2003;212:241-62.

Strobel A, Issad T, Camoin L, Ozata M, Strosberg AD. A leptin missense mutation associated with hypogonadism and morbid obesity. Nat Genet. 1998;18:213-5.

Struwe C (in preparation) "Molekulargenetische Analyse des TUFM-Gens (Tu translation elongation factor mitochondrial gene) bei extrem adipösen Kindern und Jugendlichen. Medical PhD thesis at the University of Duisburg-Essen.

Stunkard AJ, Berkowitz RI, Stallings VA, Cater JR. Weights of parents and infants: is there a relationship? Int J Obes Relat Metab Disord. 1999;23:159-62.

Stunkard AJ, Foch TT, Hrubec Z. A twin study of human obesity. JAMA. 1986;256:51-4.

Stunkard AJ, Harris JR, Pedersen NL, McClearn GE. The body-mass index of twins who have been reared apart. N Engl J Med. 1990;322:1483-7.

Stutzmann F, Vatin V, Cauchi S, Morandi A, Jouret B, Landt O, Tounian P, Levy-Marchal C, Buzzetti R, Pinelli L, Balkau B, Horber F, Bougnères P, Froguel P, Meyre D. Non-synonymous polymorphisms in melanocortin-4 receptor protect against obesity: the two facets of a Janus obesity gene. Hum Mol Genet. 2007;16:1837-44.

Takeuchi F, Yamamoto K, Katsuya T, Nabika T, Sugiyama T, Fujioka A, Isono M, Ohnaka K, Fujisawa T, Nakashima E, Ikegami H, Nakamura J, Yamori Y, Yamaguchi S, Kobayashi S, Ogihara T, Takayanagi R, Kato N. Association of genetic variants for susceptibility to obesity with type 2 diabetes in Japanese individuals. Diabetologia. 2011;54:1350-9.

Taliani MR, Roberts SC, Dukek BA, Pruthi RK, Nichols WL, Heit JA. Sensitivity and specificity of denaturing high-pressure liquid chromatography for unknown protein C gene mutations. Genet Test. 2001;5:39-44.

Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N, Ladunga I, Ulitsky-Lazareva B, Muruganujan A, Rabkin S, Vandergriff JA, Doremieux O. PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. Nucleic Acids Res. 2003;31:334-41. Erratum in: Nucleic Acids Res. 2003;31:2024.

Thorleifsson G, Walters GB, Gudbjartsson DF, Steinthorsdottir V, Sulem P, Helgadottir A, Styrkarsdottir U, Gretarsdottir S, Thorlacius S, Jonsdottir I, Jonsdottir T, Olafsdottir EJ, Olafsdottir GH, Jonsson T, Jonsson F, Borch-Johnsen K, Hansen T, Andersen G, Jorgensen T, Lauritzen T, Aben KK, Verbeek AL, Roeleveld N, Kampman E, Yanek LR, Becker LC, Tryggvadottir L, Rafnar T, Becker DM, Gulcher J, Kiemeney LA, Pedersen O, Kong A, Thorsteinsdottir U, Stefansson K: Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. Nat Genet 2009, 41:18-24.

Troiano RP, Flegal KM. Overweight children and adolescents: description, epidemiology, and demographics. Pediatrics. 1998;101:497-504.

Truby H, Baic S, deLooy A, Fox KR, Livingstone MB, Logan CM, Macdonald IA, Morgan LM, Taylor MA, Millward DJ. Randomised controlled trial of four commercial weight loss programmes in the UK: initial findings from the BBC "diet trials". BMJ. 2006;332:1309-14.

Tseng YH, Butte AJ, Kokkotou E, Yechoor VK, Taniguchi CM, Kriauciunas KM, Cypess AM, Niinobe M, Yoshikawa K, Patti ME, Kahn CR. Prediction of preadipocyte differentiation by gene expression reveals role of insulin receptor substrates and necdin. Nat Cell Biol. 2005;7:601-11.

Valente L, Tiranti V, Marsano RM, Malfatti E, Fernandez-Vizarra E, Donnini C, Mereghetti P, De Gioia L, Burlina A, Castellan C, Comi GP, Savasta S, Ferrero I, Zeviani M. Infantile encephalopathy and defective mitochondrial DNA translation in patients with mutations of mitochondrial elongation factors EFG1 and EFTu. Am J Hum Genet. 2007;80:44-58.

Van Gaal LF, Vansant GA, De Leeuw IH. Upper body adiposity and the risk for atherosclerosis. J Am Coll Nutr. 1989;8:504-14.

Varela LM, Ortega A, Bermudez B, Lopez S, Pacheco YM, Villar J, Abia R, Muriana FJ. A high-fat meal promotes lipid-load and apolipoprotein B-48 receptor transcriptional activity in circulating monocytes. Am J Clin Nutr. 2011;93:918-25.

Västermark Å, Jacobsson JA, Johansson Å, Fredriksson R, Gyllensten U, Schiöth HB. Polymorphisms in sh2b1 and spns1 loci are associated with triglyceride levels in a healthy population in northern Sweden. J Genet. 2012;91:237-40.

Visscher TL, Seidell JC. Time trends (1993-1997) and seasonal variation in body mass index and waist circumference in the Netherlands. Int J Obes Relat Metab Disord. 2004;28:1309-16.

Volckmar AL*, Song JY*, Graniger J, Knoll N, Scherag A, Pütter C, Reinehr T, Wang H, Hebebrand J, Hinney A. Genetic variants in the APOB48R gene are associated with early-onset extreme obesity. (in preparation)

Volckmar AL, Bolze F, Jarick I, Knoll N, Scherag A, Reinehr T, Illig T, Grallert H, Wichmann HE, Wiegand S, Biebermann H, Krude H, Fischer-Posovszky P, Rief W, Wabitsch M, Klingenspor M, Hebebrand J, Hinney A. Mutation screen in the GWAS derived obesity gene SH2B1 including functional analyses of detected variants. BMC Med Genomics. 2012;5:65.

Volckmar AL, Knoll N, Hinney A. Genetische Prädikatoren der Adipositas. In: Ardelt-Gattinger, E, Ring-Dimitriou, S, Weghuber, D (eds). Der gesunde Adipöse. Hogrefe Verlag. 2013 (in press).

Volckmar AL, Pütter C, Song JY, Graninger J, Knoll N, Wolters B, Hebebbrand J, Scherag A, Reinehr T, Hinney A. Analyses of non-synonymous obesity risk alleles in SH2B1 (rs7498665) and APOB48R (rs180743) in obese children and adolescents undergoing a 1-year lifestyle intervention. Exp Clin Endocrinol Diabetes. 2013; 121:1-4.

Wabitsch M, Hauner H, Hertrampf M, Muche R, Hay B, Mayer H, Kratzer W, Debatin KM, Heinze E: Type II diabetes mellitus and impaired glucose regulation in Caucasian children and adolescents with obesity living in Germany. Int J Obes Relat Metab Disord. 2004, 28:307-313.

Walters RG, Jacquemont S, Valsesia A, de Smith AJ, Martinet D, Andersson J, Falchi M, Chen F, Andrieux J, Lobbens S, Delobel B, Stutzmann F, El-Sayed Moustafa JS, Chèvre JC, Lecoeur C, Vatin V, Bouquillon S, Buxton JL, Boute O, Holder-Espinasse M, Cuisset JM, Lemaitre MP, Ambresin AE, Brioschi A, Gaillard M, Giusti V, Fellmann F, Ferrarini A, Hadjikhani N, Campion D, Guilmatre A, Goldenberg A, Calmels N, Mandel JL, Le Caignec C, David A, Isidor B, Cordier MP, Dupuis-Girod S, Labalme A, Sanlaville D, Béri-Dexheimer M, Jonveaux P, Leheup B, Ounap K, Bochukova EG, Henning E, Keogh J, Ellis RJ, Macdermot KD, van Haelst MM, Vincent-Delorme C, Plessis G, Touraine R, Philippe A, Malan V, Mathieu-Dramard M, Chiesa J, Blaumeiser B, Kooy RF, Caiazzo R, Pigeyre M, Balkau B, Sladek R, Bergmann S, Mooser V, Waterworth D, Reymond A, Vollenweider P, Waeber G, Kurg A, Palta P, Esko T, Metspalu A, Nelis M, Elliott P, Hartikainen AL, McCarthy MI, Peltonen L, Carlsson L, Jacobson P, Sjöström L, Huang N, Hurles ME, O'Rahilly S, Farooqi IS, Männik K, Jarvelin MR, Pattou F, Meyre D, Walley AJ, Coin LJ, Blakemore AI, Froguel P, Beckmann JS: A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. Nature. 2010, 463:671-675.

Wang TC, Chiu H, Chang YJ, Hsu TY, Chiu IM, Chen L. The adaptor protein SH2B3 (Lnk) negatively regulates neurite outgrowth of PC12 cells and cortical neurons. PLoS One. 2011;6:e26433.

Watson CS, Turpenoff CM, Kelly WJ, Botwinick J. Age differences in resolving power and decision strategies in a weight discrimination task. J Gerontol. 1979;34:547-52.

Wen W, Cho YS, Zheng W, Dorajoo R, Kato N, Qi L, Chen CH, Delahanty RJ, Okada Y, Tabara Y, Gu D, Zhu D, Haiman CA, Mo Z, Gao YT, Saw SM, Go MJ, Takeuchi F, Chang LC, Kokubo Y, Liang J, Hao M, Le Marchand L, Zhang Y, Hu Y, Wong TY, Long J, Han BG, Kubo M, Yamamoto K, Su MH, Miki T, Henderson BE, Song H, Tan A, He J, Ng DP, Cai Q, Tsunoda T, Tsai FJ, Iwai N, Chen GK, Shi J, Xu J, Sim X, Xiang YB, Maeda S, Ong RT, Li C, Nakamura Y, Aung T, Kamatani N, Liu JJ, Lu W, Yokota M, Seielstad M, Fann CS; Genetic Investigation of ANthropometric Traits (GIANT) Consortium, Wu JY, Lee JY, Hu FB, Tanaka T, Tai ES, Shu XO. Meta-analysis identifies common variants associated with body mass index in east Asians. Nat Genet. 2012;44:307-11.

Whitmer RA, Gunderson EP, Quesenberry CP Jr, Zhou J, Yaffe K. Body mass index in midlife and risk of Alzheimer disease and vascular dementia. Curr Alzheimer Res. 2007;4:103-9.

WHO expert consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. The Lancet, 2004; 157-163.

WHO. Obesity: preventing and managing the global epidemic. Report of a WHO Consultation. WHO Technical Report Series 894. Geneva: World Health Organization, 2000.

WHO. Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee. WHO Technical Report Series 854. Geneva: World Health Organization, 1995.

Wichmann HE, Gieger C, Illig T; MONICA/KORA Study Group: KORA-gen—resource for population genetics, controls and a broad spectrum of disease phenotypes. Gesundheitswesen. 2005, 67(Suppl 1):S26-30.

Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, Heid IM, Berndt SI, Elliott AL, Jackson AU, Lamina C, Lettre G, Lim N, Lyon HN, McCarroll SA, Papadakis K, Qi L, Randall JC, Roccasecca RM, Sanna S, Scheet P, Weedon MN, Wheeler E, Zhao JH, Jacobs LC, Prokopenko I, Soranzo N, Tanaka T, Timpson NJ, Almgren P, Bennett A, Bergman RN, Bingham SA, Bonnycastle LL, Brown M, Burtt NP, Chines P, Coin L, Collins FS, Connell JM, Cooper C, Smith GD, Dennison EM, Deodhar P, Elliott P, Erdos MR, Estrada K, Evans DM, Gianniny L, Gieger C, Gillson CJ, Guiducci C, Hackett R, Hadley D, Hall AS, Havulinna AS, Hebebrand J, Hofman A, Isomaa B, Jacobs KB, Johnson T, Jousilahti P, Jovanovic Z, Khaw KT, Kraft P, Kuokkanen M, Kuusisto J, Laitinen J, Lakatta EG, Luan J, Luben RN, Mangino M, McArdle WL, Meitinger T, Mulas A, Munroe PB, Narisu N, Ness AR, Northstone K, O'Rahilly S, Purmann C, Rees MG, Ridderstråle M, Ring SM, Rivadeneira F, Ruokonen A, Sandhu MS, Saramies J, Scott LJ, Scuteri A, Silander

K, Sims MA, Song K, Stephens J, Stevens S, Stringham HM, Tung YC, Valle TT, Van Duijn CM, Vimaleswaran KS, Vollenweider P, Waeber G, Wallace C, Watanabe RM, Waterworth DM, Watkins N; Wellcome Trust Case Control Consortium, Witteman JC, Zeggini E, Zhai G, Zillikens MC, Altshuler D, Caulfield MJ, Chanock SJ, Farooqi IS, Ferrucci L, Guralnik JM, Hattersley AT, Hu FB, Jarvelin MR, Laakso M, Mooser V, Ong KK, Ouwehand WH, Salomaa V, Samani NJ, Spector TD, Tuomi T, Tuomilehto J, Uda M, Uitterlinden AG, Wareham NJ, Deloukas P, Frayling TM, Groop LC, Hayes RB, Hunter DJ, Mohlke KL, Peltonen L, Schlessinger D, Strachan DP, Wichmann HE, McCarthy MI, Boehnke M, Barroso I, Abecasis GR, Hirschhorn JN; Genetic Investigation of ANthropometric Traits Consortium. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat Genet. 2009;41:25-34.

Wolff OH. Obesity in childhood; a study of the birth weight, the height, and the onset of puberty. Q J Med. 1955;24:109-23.

Xiang Z, Litherland SA, Sorensen NB, Proneth B, Wood MS, Shaw AM, Millard WJ, Haskell-Luevano C: Pharmacological characterization of 40 human melanocortin-4 receptor polymorphisms with the endogenous proopiomelanocortin-derived agonists and the agouti-related protein (AGRP) antagonist. Biochemistry. 2006, 45:7277-88.

Xie T, Ho SL, Ma OC. High resolution single strand conformation polymorphism analysis using formamide and ethidium bromide staining. Mol Pathol. 1997;50:276-8.

Yang J, Manolio TA, Pasquale LR, Boerwinkle E, Caporaso N, Cunningham JM, de Andrade M, Feenstra B, Feingold E, Hayes MG, Hill WG, Landi MT, Alonso A, Lettre G, Lin P, Ling H, Lowe W, Mathias RA, Melbye M, Pugh E, Cornelis MC, Weir BS, Goddard ME, Visscher PM. Genome partitioning of genetic variation for complex traits using common SNPs. Nat Genet. 2011;43:519-25.

Yoganathan P, Karunakaran S, Ho MM, Clee SM: Nutritional regulation of genome-wide association obesity genes in a tissue-dependent manner. Nutr Metab (Lond). 2012,1:65.

Yoshiga D, Sato N, Torisu T, Mori H, Yoshida R, Nakamura S, Takaesu G, Kobayashi T, Yoshimura A. Adaptor protein SH2-B linking receptor-tyrosine kinase and Akt promotes adipocyte differentiation by regulating peroxisome proliferator-activated receptor gamma messenger ribonucleic acid levels. Mol Endocrinol. 2007;21:1120-31.

You FM, Huo N, Gu YQ, Luo MC, Ma Y, Hane D, Lazo GR, Dvorak J, Anderson OD. BatchPrimer3: a high throughput web application for PCR and sequencing primer design. BMC Bioinformatics. 2008;9:253.

Young EH, Wareham NJ, Farooqi S, Hinney A, Hebebrand J, Scherag A, O'Rahilly S, Barroso I, Sandhu MS. The V103I polymorphism of the MC4R gene and obesity: population based studies and meta-analysis of 29 563 individuals. Int J Obes (Lond). 2007;31:1437.

Yousaf N, Deng Y, Kang Y, Riedel H. Four PSM/SH2-B alternative splice variants and their differential roles in mitogenesis. J Biol Chem. 2001;276:40940-8.

Zhang M, Deng Y, Riedel H. PSM/SH2B1 splice variants: critical role in src catalytic activation and the resulting STAT3s-mediated mitogenic response. J Cell Biochem. 2008;104:105-18.

Zhang XH, Kangsamaksin T, Chao MS, Banerjee JK, Chasin LA. Exon inclusion is dependent on predictable exonic splicing enhancers. Mol Cell Biol. 2005;25:7323-32.

Zhang Y, Proenca R, Maffei M, Barone M, Leopold L, Friedman JM. Positional cloning of the mouse obese gene and its human homologue. Nature. 1994;372:425-32. Erratum in: Nature. 1995;374:479.

9. Appendix

Table S1: Sequences and PCR conditions of the primers used for the mutation screen of *SH2B1.*

Primer name	Sequence	GC [%]	T_{anneal} [°C]	MgCl2 [µmol/reaction]	Amplicon length [bp]
Fragment 2h inner F	5`-CCT CTT TGT GCC CCA CAG-3´	61	58	1	490
Fragment 2h inner R	3`-TCG GCT TGG AGG AGG TTG TA-5´	55			
Fragment 2h outer_F	5`- CAG GTT AGA CGC TGG GGA GC -3´	65	60	1	728
Fragment 2h outer_R	3`- GGC TTG GAG GAG GTT GTA GA -5´	55			
Fragment 2i_F	5`-ACA TGA CCT GTC CCT TGA GAG-3´	52	60	1	453
Fragment 2i_R	3`-CTC AGC AGC TCT TCC CTC TG-5´	60			
Fragment 2j_F	5`-GGG CTG TGA CCA CTC TGA G-3´	63	59	1	380
Fragment 2j_R	3`- TGG TAG GGG ACT GGT CAG T-5´	58			
Fragment 3_F	5`-TGG AAA CCA AAC ACC CAG AT-3´	45	58.7	1.5	217
Fragment 3_R	3`-CAC TTG CTT AAA TGC GCT GA-5´	45			
Fragment 4_F	5`-AGG CGT CGT CTC ATC TCT GT-3´	55	60.5	1.5	555
Fragment 4_R	3`-AGG GAT CAA AGA CCA ACA TT-5´	40			
Fragment 5_F	5`-CAG CAC CAT CTT CCC TGT CT-3´	55	65	1.5	285
Fragment 5_R	3`-TCT GGA GAG GGC ARG GCG-5´	71			
Fragment 6_F	5`-GGT AAA AGC ATC AGG GGT CA-3´	50	62	1	422
Fragment 6_R	3`-CAG GAG AGG TGG TCA GGA AG-5´	60			
Fragment 7_F	5`- CTT CCT GAC CAC CTC TCC TG -3´	60	63.7	1.5	423
Fragment 7_R	3`- CCC ACT GCC AAC TAT GCA C -5´	58			
Fragment 8_F	5`- GGA AAG GCT CTG TTC TGT GC -3´	55	63	1.5	416
Fragment 8_R	3`- TAC CCT GCT TCT CCA AGC -5´	56			
Fragment 9_F	5`- ACC CAT CCT GGC CTC GTC -3´	67	62	1.5	355
Fragment 9_R	3`- CCC CAC ATC CCA TCT CCT G -5´	63			
Fragment 10a inner_F	5`- CAC CTG CCT CCA CAA TCA -3´	56	60	1.2	358
Fragment 10a inner_R	3`- CCT AGT GGT GAC TGC TGC A -5´	58			
Fragment 10a outer_F	5`- AGT GTG GTC CTC CTC TCA CC -3´	60	62	1	554
Fragment 10a outer_R	3`- GTT GGC TCA CAC GAA GGA GT -5´	55			
Fragment 10b_F	5`- TCC CCG TGG TTG AGC TG-3´	65	62	1.5	476
Fragment 10b_R	3`- CTC GGG CTG GGG GTA GTT-5´	67			

Table S2: Sequences and PCR conditions of the primers used for the mutation screen of *APOB48R*.

Primer name	Sequence	GC [%]	T_{anneal} [°C]	MgCl2 [mM/reaction]	Amplicon length [bp]
Fragment 1 F	5`- AAG GTT GGC TTG TCT TGG TG-3´	50	60.3	1.5	231
Fragment 1 R	3`- ATC AGA AGG AGG AGG GGA AA-5´	50			
Fragment 2 F	5`- AAA GGG CTG GGA CTC TCC T-3´	58	65.7	1.5	377
Fragment 2 R	3`- GCA GCC TTG GCT GTC TCC-5´	67			
Fragment 3 F	5`-TGG GGA GCT CAG CTG TAG AA-3´	55	59.5	1.5	496
Fragment 3 R	3`-CAG GCC TTC TCC ACC ACT AC-5´	60			
Fragment 4 F	5`-AGA TGG AGC AGG GGG TCA G-3´	63	58.7	1	473
Fragment 4 R	3`- CTC CTC TTT GCC TGA GGT TG-5´	55			
Fragment 5 F	5`-GGA AGC CAG GAC AAT CTC A-3´	53	58.7	1	499
Fragment 5 R	3`-CAG GTC TAC CTG GCC CTC A-5´	63			
Fragment 6 F	5`-GTC CTG GGC ACT GAA AGA AC-3´	55	62.6	1.5	393
Fragment 6 R	3`-TTT GGT GAC GCT GTG TGT G-5´	53			
Fragment 7 F	5`-TCA GAT GGA GAG GCT GAA GG-3´	55	61.3	1.5	397
Fragment 7 R	3`-CCG CAT CCT CCT GAG TAT TT-5´	50			
Fragment 8 F	5`-CGG GGT CTG TAA AGC CTG A-3´	58	58.3	1	387
Fragment 8 R	3`-GCA TAG GCC TCT CCC AGT G-5´	63			
Fragment 9 F	5`- TGT GGA ACT GAG GAG GGA GA-3´	55	58.7	1.5	379
Fragment 9 R	3`- AGC CCT CCA GAC CAA ATT CT-5´	50			
Fragment 10 F	5`- GCT GGT GAA GCT TTG GAA GG-3´	55	62.8	1.5	374
Fragment 10 R	3`- CCC TCA CTG TCT TCC CTC TG-5´	60			
Fragment 11 F	5`- ATG GGA GCC ATG GTG GAG-3´	61	59.5	1.5	400
Fragment 11 R	3`- TCG GGA TTC CAC TCG TTC-5´	56			
Fragment 12 F	5`- AAC GAG TGG AAT CCC GAA G-3´	53	58.7	1.5	344
Fragment 12 R	3`- CTG AGT TGC CCT GTG CCT AC-5´	60			
Fragment 13 F	5`- GTA GGC ACA GGG CAA CTC AG-3´	60	64.7	1.5	394
Fragment 13 R	3`- GAG GTG GTC TCA GTG GGG TA-5´	60			
Fragment 14 F	5`- TAC CCC ACT GAG ACC ACC TC-3´	60	62.5	1.5	240
Fragment 14 R	3`- TCC CTC AGT CTT CCT TGC AG-5´	55			

T_{anneal}= primer specific annealing temperature

Table S3: Sequences and PCR conditions of the primers used for the mutation screen of *SULT1A2.*

Primer name	Sequence	GC [%]	T_{anneal} [°C]	MgCl2 [µmol/reaction]	Amplicon length [bp]
Fragment 1 F	5`- TTC CAC GCC AAC TTC AAC TA-3´	45	67	1	213
Fragment 1 R	3`-CTT GAT CCC CAA GTC CCT G-5´	58			
Fragment 2 F	5`- ACT TTG CAT TTT GGA ATG GT-3´	35	56#	2	209
Fragment 2 R	3`- ACA CAC ACA AAA AGA TAC TGA TAA CAT-5´	30			
Fragment 3 F	5`-CTG AGT GGC TTT GTG AGT GC-3´	55	65	1	378
Fragment 3 R	3`-GAG ATG GGA GGT GAG CAG G-5´	63			
Fragment 4 F	5`-CCT CAG CCT GCT CAC CTC-3´	67	65	1	443
Fragment 4 R	3`-GTG CTC TCA AAC TCC AAC C-5´	53			
Fragment 5 F	5`-CGT GCC TTG CTC CAG ATT G-3´	58	60*	2	426
Fragment 5 R	3`-ACC ACC CCT TAG CTC CAC A-5´	58			
Fragment 6 F	5`-GGA GAA GTT CAT GGT CGG AG-3´	55	61.4	2	252
Fragment 6 R	3`-CAG GAG TCA CAT GGA GGG AA-5´	55			
Fragment 7 F	5`-AGT ATC CGA GCC TCC ACT G-3´	58	67	1.5	320
Fragment 7 R	3`-AAA GCT GGA GTC TCA TCC CCA-5´	52			
Fragment 8 F	5`-TGG GAT GAG ACT CCA GCT TT-3´	50	65.4	1	479
Fragment 8 R	3`-CCT GTC CTC CAG TGA TCC TC-5´	60			

T_{anneal}= primer specific annealing temperature

Used PCR kit was Taq GOLD by Invitrogen with buffers and Taq according to manufacturer

* Used PCR kit was Optimase by Invitrogen with buffers and Taq according to manufacturer

Table S4: List of fragments from *SH2B1*, *APOB48R*, *and SULT1A2* that were screened for genetic variants with dHPLC.

Gene	Fragment	Fragment length [bp]	Temp 1 [°C]	Temp 2 [°C]
SH2B1	1	490	61	62
SH2B1	2	453	61	62.5
SH2B1	3	380	61	63
SH2B1	4	217	57	61
SH2B1	5	555	60.4	-
SH2B1	6	285	65	-
SH2B1	7	422	60.1	61.5
SH2B1	8	423	61	62
SH2B1	9	416	61	-
SH2B1	10	355	62	-
SH2B1	11	358	62.5	-
SH2B1	12	476	62.5	-
APOB48R	2	377	61	-
APOB48R	3	496	62	-
APOB48R	4	473	59	-
APOB48R	5	499	60	-
APOB48R	6	393	62.5	-
APOB48R	7	397	59	-
APOB48R	8	387	60	-
APOB48R	9	379	59	-
APOB48R	10	374	60	-
APOB48R	11	400	61.5	-
APOB48R	12	344	62	-
APOB48R	13	394	60	-
SULT1A2	3	409	60	-

Table S5: List of fragments from *SH2B1*, *APOB48R*, *and SULT1A2* that were screened for genetic variants with SSCP with used restriction enzymes is applicable.

Gene	Fragment	Fragment length [bp]	Restriction enzyme
APOB48R	1	231	-
APOB48R	14	240	-
SULT1A2	2	281	-
SULT1A2	4	453 (281+169)	HindIII
SULT1A2	5	319 (182+141)	NciI
SULT1A2	6	275	-

Table S6: Parameters of leptin receptor activity measured by STAT3 mediated luciferase response.

Genotype	Basal activity	SD	Emax	SD	EC50 [ng/ml]	SD
Control (pcDNA3)	8.1	2.9	102.5	15.4	16.6	8.6
WT SH2B1β	10.8	5.7	154.1	24.1	15	7.9
SH2B1β Thr484Ala	11.2	5.7	145.7	13.6	13.3	4.7
SH2B1β Thr656Ile	12.3	5.5	148.2	18.4	13.8	6.3
WT SH2B1γ	11.3	5.5	147.4	20.9	12.8	6.1
SH2B1γ Thr484Ala	10.4	5.2	137.1	17.9	13.2	5
SH2B1γ Pro674Ser	10.7	5.1	145.8	22.6	19.1	12.3

HEK293 cells (n=8 separate experiments) were co-transfected with LEPRb, a STAT3 responsive element and SH2B1 splice variants beta (left) and gamma (right) with and without the infrequent alleles at rs7498665 (Thr484Ala) and βThr656Ile/γPro674Ser. Depicted are basal activity without leptin treatment, maximal activation of the leptin receptor (Emax), and the half-maximal activation of the leptin receptor (EC50), of the leptin receptor co-transfected with clones of SH2B1 harbouring the different variants; all with standard deviation (SD; Volckmar et al. 2013).

Table S7: Distribution of coding variants in *SH2B1*, *APOB48R* and *SULT1A2* in the individuals from the mutation screen study group.

	SH2B1			*APOB48R*			*SULT1A2*				
Individual	**Thr175Ala**	**Thr484Ala**	**βThr656Ile/γPro674Ser**	**Ala345_Gly346insdelAlaGlyThrAla SerGlyGlyGluGluAlaGly**	**Ser323delSerGlyGlyGluGluAlaGlyThrAla**	**Pro419Ala**	**IL7Thr**	**Pro19Leu**	**Ser44Asn**	**Tyr62Phe**	**Ala164Val**	**Asn235Ile**
1	11	22	11	22	11	22	11	11	11	12	11	11
2	11	22	11	22	11	22	11	11	11	11	11	11
3	11	22	11	22	11	22	11	11	11	11	11	11
4	11	22	11	22	11	22	11	11	11	11	12	11
5	11	22	11	22	11	22	11	11	11	11	11	11
6	11	22	11	22	11	22	11	11	11	11	11	11
7	11	22	11	22	11	22	11	11	11	11	11	11
8	11	22	11	22	11	22	11	11	11	11	11	11
9	11	22	11	22	11	22	11	11	11	11	11	11
10	11	22	11	22	11	22	11	11	11	11	11	11
11	11	22	11	22	11	22	11	11	11	12	11	11
12	11	22	11	22	11	22	11	11	11	11	11	11
13	11	22	11	22	11	22	11	11	11	11	11	11
14	11	22	11	22	11	22	11	11	11	11	11	12
15	11	22	11	22	11	22	11	11	11	11	11	11
16	11	22	11	22	11	22	11	11	11	11	11	11
17	11	22	11	22	11	22	11	11	11	11	11	11
18	11	22	11	22	11	22	11	11	11	11	11	11
19	11	22	11	22	11	22	11	11	11	11	11	11
20	11	22	11	22	11	22	11	11	11	11	11	11
21	11	22	11	22	11	22	11	11	11	11	11	11
22	11	22	11	22	11	22	11	11	11	11	11	11
23	11	22	11	22	11	22	11	11	11	11	11	11
24	11	22	11	22	11	22	22	11	11	11	11	11
25	11	22	11	22	11	22	11	11	11	11	11	11
26	11	22	11	22	11	22	11	11	11	11	11	11
27	11	22	11	22	11	22	11	11	12	11	11	11
28	11	22	11	22	11	22	11	11	11	11	11	11
29	11	22	11	22	11	22	11	11	11	11	11	11
30	11	22	11	22	11	22	11	11	11	11	11	11
31	11	22	11	22	11	22	12	11	11	12	11	11
32	11	22	11	22	11	22	11	11	11	11	11	11
33	11	22	11	22	11	22	12	11	11	12	11	11
34	11	22	11	22	11	22	11	11	11	12	11	11
35	11	22	11	22	11	22	11	11	11	11	11	11
36	11	22	11	22	11	22	11	11	11	11	11	11
37	11	22	11	22	11	22	11	11	11	11	11	11

-Table S7 continued-

-Table S7 continued-

	SH2B1			*APOB48R*				*SULT1A2*				
Individual	**Thr175Ala**	**Thr484Ala**	**βThr656Ile/γPro674Ser**	**Ala345_Gly346insdelAlaGlyThrAla SerGlyGlyGluGluAlaGly**	**Ser323delSerGlyGlyGluGluAlaGlyThrAla**	**Pro419Ala**	**IL7Thr**	**Pro19Leu**	**Ser44Asn**	**Tyr62Phe**	**Ala164Val**	**Asn235Ile**
38	11	22	11	22	11	22	11	11	11	11	11	11
39	11	22	11	22	11	22	11	11	11	11	11	11
40	11	22	11	22	11	22	11	11	11	11	11	11
41	11	22	11	22	11	22	11	11	11	11	11	11
42	11	22	11	22	11	22	11	11	11	11	11	11
43	11	22	11	22	11	22	11	11	11	11	11	11
44	11	22	11	22	11	22	11	11	11	11	11	11
45	11	22	11	22	11	22	11	11	11	11	11	11
46	11	22	11	22	11	22	11	11	11	11	11	11
47	11	22	11	22	11	22	11	11	11	11	11	11
48	11	22	11	22	11	22	11	11	11	11	11	11
49	11	22	11	22	11	22	11	11	11	11	11	11
50	11	22	12	22	11	11	11	11	11	11	11	11
51	11	22	11	22	11	22	11	11	11	11	11	11
52	11	22	11	22	11	22	11	11	11	11	11	11
53	11	22	11	12	11	22	11	11	11	11	11	11
54	11	22	11	22	11	11	11	11	11	11	11	11
55	11	22	11	22	11	22	11	11	11	11	11	11
56	11	22	11	22	11	22	11	11	11	12	11	11
57	11	22	11	22	11	11	11	11	11	11	11	11
58	11	22	11	22	11	22	11	11	11	11	11	11
59	11	22	11	22	11	22	11	11	11	11	11	11
60	11	22	11	22	11	22	11	11	11	11	11	11
61	11	22	11	12	11	22	11	11	11	11	11	11
62	11	22	11	22	11	22	11	11	11	11	11	11
63	11	22	11	22	11	22	11	11	11	11	11	11
64	11	22	11	22	11	22	11	11	11	11	11	11
65	11	22	11	22	11	22	11	11	11	11	11	11
66	11	22	11	22	11	22	11	11	11	11	11	11
67	11	22	11	12	12	12	11	11	11	11	11	11
68	11	22	11	22	11	22	11	11	11	11	11	11
69	11	22	11	22	11	22	11	11	11	11	11	11
70	11	22	11	22	11	22	11	11	11	11	11	11
71	11	22	11	22	11	22	11	11	11	11	11	11
72	11	22	11	22	11	22	11	11	11	11	11	11
73	11	22	11	22	11	12	11	11	11	11	11	11
74	11	22	11	22	11	22	11	11	11	11	11	11
75	11	22	11	22	11	22	11	11	11	11	11	11
76	11	22	11	22	11	22	11	11	11	11	11	11

-Table S7 continued-

-Table S7 continued-

	SH2B1			*APOB48R*					*SULT1A2*			
Individual	**Thr175Ala**	**Thr484Ala**	**βThr656Ile/γPro674Ser**	**Ala345_Gly346insdelAlaGlyThrAla SerGlyGlyGluGluAlaGly**	**Ser323delSerGlyGlyGluGluAlaGlyThrAla**	**Pro419Ala**	**IL7Thr**	**Pro19Leu**	**Ser44Asn**	**Tyr62Phe**	**Ala164Val**	**Asn235Ile**
77	11	22	11	22	11	22	11	11	11	11	11	11
78	11	22	11	22	11	22	11	11	11	11	11	11
79	11	22	11	22	11	22	11	11	11	12	11	11
80	11	22	11	22	11	22	11	11	11	11	11	11
81	11	22	11	22	11	22	11	11	11	11	11	11
82	11	12	11	22	11	22	11	11	11	11	11	11
83	11	22	11	22	11	22	11	11	11	11	11	11
84	12	22	11	22	11	22	11	11	11	11	11	11
85	11	22	11	22	11	22	11	11	11	11	11	11
86	11	22	11	22	11	22	11	11	11	11	11	11
87	11	22	11	22	11	22	11	11	11	11	11	11
88	11	22	11	22	11	22	11	11	11	11	11	11
89	11	22	11	22	11	22	11	11	11	11	11	11
90	11	22	11	22	11	22	11	11	11	11	11	11
Del1	11	11	11	11	11	11	11	12	11	11	11	11
Del2	11	11	11	11	11	11	11	12	11	11	11	11
Del3	11	12	11	12	11	12	11	12	11	12	11	11
Del4	11	12	11	11	11	12	11	12	11	11	11	11
Del5	11	22	11	11	11	11	11	11	11	11	11	11

Del: Carriers of a heterozygous deletion of chr16p11.2 which does not harbor *SH2B1*, *APOB48R* and *SULT1A2*.

1 = the major allele, 2 = the minor allele at each position.

Table S8: Distribution of coding variants in *SH2B1*, *APOB48R* and *SULT1A2* in the individuals from the mutation screen study group. Variants are listed by gene and position. Nonsynonymous coding variants are marked in grey.

Gene	Position	rs-number	AA exchange	MAF in Obese Sample	MAF in CEU
SH2B1	g.2749C/A	rs181294111	Thr175Ala	0.005	NA
SH2B1	g.8164A/G	rs7498665	Thr484Ala	0.963	0.294
SH2B1	g.8250C/T	rs28433345	NA	0.026	0.367
SH2B1	g.8738A/G	rs62037368	NA	0.038	NA
SH2B1	g.8764C/T	rs62037369	NA	0.045	0.358
SH2B1	g.9483C/T	NA	βThr656Ile/γPro674Ser	0.005	NA
SH2B1	g.10182C/A	NA	NA	0.011	NA
APOB48R	c.57+50C/T	rs74949322	NA	0.005	0.05
APOB48R	c.66C/A/G/T	rs151233	Leu22=	0.026	0.142
APOB48R	c.510A/G	rs149271	Glu170=	0.963	0.342
APOB48R	c.961_996delACAGCCTCAGGCGGGGAGGAGGCCGGGACAGCCTC? *	NA	Thr321_Gly329delThrAlaSerGlyGlyGluGluAlaGly	0.962	NA
APOB48R	g.28507397_28507398insdelGGGACAGCCTCAGGAGGGGAGGAGGCC	rs3833080	Ala345_Gly346delinsAlaGlyThrAlaSerGlyGlyGluGluAlaGly	0.005	NA
APOB48R	c.1255C/G	rs180743	Pro419Ala	0.916	0.375
APOB48R	c.1659A/C/G/T	rs180744	Gln553=	0.942	0.350
APOB48R	c.1680C/T	rs151174	Gly560=	0.942	0.358
APOB48R	c.2058A/C/G/T	rs40831	Ala686=	0.953	0.389
APOB48R	c.3036G/A	rs61738759	Pro1012=	0.074	0.025
APOB48R	c.2955C/T	NA	Asp985=	0.005	NA
APOB48R	c.*218C/G	rs40833	NA	0.089	0.467
SULT1A1	g.5442G/A	rs56148728	NA	0.011	0.083
SULT1A1	c.57G/A	rs34513973	Pro19=	0.170	0.022
SULT1A1	c.153T/C	rs1126447	Thr51=	0.016	0.208
SULT1A1	c.161T/C	rs138013613	Ala54Thr	0.016	0
SULT1A1	c.162A/G	rs1126446	Thr51=	0.011	0.295
SULT1A1	c.373-132C/T	NA	NA	NA	NA
SULT1A1	c.373-147C/A	rs2925624	NA	NA	0.292
SULT1A1	c373-167G/T	NA	NA	NA	NA
SULT1A1	c.222+253G/C	NA	NA	NA	NA
SULT1A1	c.222+268G/A	rs4148410	NA	NA	0.059
SULT1A1	c.373-48A/G	rs2925623	NA	NA	0.283
SULT1A1	c.441G/A	rs1042005	Lys69=	NA	0.000
SULT1A1	c.451G/C	rs1042014	Glu73Gln	NA	0.000
SULT1A1	c.453G/C	rs1042011	Glu73Asp	NA	0.008

-Table S8 will be continued-

-Table S8 continued-

Gene	Position	rs-number	AA exchange	MAF in Obese Sample	MAF in CEU
SULT1A1	T/C	rs7499134	NA	NA	0.000
SULT1A1	A/G	rs9282862	NA	1.000	0.300
SULT1A1	g.17334G/A	NA	NA	1.000	NA
SULT1A1	g.17356G/C	rs3176926	Pro200=	1.000	0.325
SULT1A1	g.17413C/T	rs35497673	Thr219=	1.000	0.500
SULT1A1	g.17420T/C	NA	Phe222Leu	1.000	NA
SULT1A1	c.667G/A	rs1801030	Val223Met	1.000	0.000
SULT1A1	c.704A/C	rs35728980	Asn235Thr	1.000	0.210
SULT1A1	g.17487C/G	NA	Pro244Arg	1.000	NA
SULT1A1	c.775+3G/A	NA	NA	1.000	NA
SULT1A1	c.775+9T/C	rs34823028	NA	1.000	0.002
SULT1A1	c.775+28T/C	NA	NA	1.000	NA
SULT1A1	c.775+32T/C	NA	NA	1.000	NA
SULT1A1	c.776-54T/C	rs41278164	NA	1.000	0.500
SULT1A1	c.776-34T/C	rs9282864	NA	0.793	NA
SULT1A1	c.*14A/G	rs6839	NA	0.809	0.274
SULT1A2	c.20C/T	rs1136703	Ile7Thr	0.005	0.304
SULT1A2	c.24T/C	rs1690407	Ser8=	0.742	0.360
SULT1A2	c.56C/T/A/G	rs10797300	Pro19Leu	0.742	0.100
SULT1A2	c.131G/A	rs145008170	Ser44Asn	0.995	0.001
SULT1A2	c.148+ 34T/C	rs4149406	NA	0.005	0.352
SULT1A2	c.185A/T	rs4987024	Tyr62Phe	0.005	0.004
SULT1A2	g.241+39G/A	NA	NA	0.005	NA
SULT1A2	c.500-19T/C	rs3743963	NA	0.979	0.483
SULT1A2	c.491C/T	rs142241142	Ala164Val	0.005	0.002
SULT1A2	c.492T/C	rs139896537	Ala164=	0.084	0.002
SULT1A2	c.704A/C	rs1059491	Asn235Ile	0.005	0.317
SULT1A2	c.844A/G	rs75191166	Glu282Lys	0.042	NA
SULT1A2	c.*7T/C	rs710410	NA	0.021	0.258
SULT1A2	c.*14A/G	rs762634	NA	0.021	0.364
SULT1A2	g. 28603263T/C	rs145790611	NA	0.037	NA
TUFM	c.-55T/C	rs7187776	NA	0.074	0.362
TUFM	c.817+13T/C	rs4788099	NA	0.021	0.381
TUFM	c.248-18G/A	rs8061877	NA	0.037	0.254
TUFM	c.922+29C/G	rs61737565	NA	0.047	0.358
TUFM	c.3536C/G	NA	NA	0.005	NA

AA: Amino Acid, NA: not available, obese sample: 95 extremely obese children and adolescents from the initial mutation screen, APOB48R – Apolipoprotein B48 receptor, SH2B1 - SH2B adaptor protein 2 isoform 1, SULT1A1 - sulfotransferase family, cytosolic, 1A, member 1, SULT1A2 - sulfotransferase family, cytosolic, 1A, member 2, TUFM - Tu translation elongation factor, mitochondrial

* exact position of the deletion within the repetitive region can not be determined, hence the questionare in the nomenclature as recommended by Antonarakis (www.hgmd.cf.ac.uk/docs/mut_nom.html)

11. Acknowledgements:

I would like to thank my examiner Prof. Dr. Johannes Hebebrand and my direct supervisor Prof. Dr. Anke Hinney from the Department of Child- and Adolescent Psychiatry at the University of Duisburg-Essen for making this PhD's thesis possible. Thank you for analytical criticism and your suggestions for my work which improved it immensely. Thank you for giving me the opportunity to research freely, introduce my own ideas into the projects but always being supported and guided at the same time. I really appreciated your help and guidance in intellectual and practical ways.

I thank the Bundesministerium für Bildung und Forschung (NGFNplus 01GS0820) for financial aid which supported this thesis and all study participants whose participation enabled me to do this research.

Additional thanks go to our cooperation partners from the NGFNplus network "Molecular Mechanisms of Obesity" who offered samples, techniques and useful discussions which broadened my view and improved my research. Special thanks to Prof. Dr. Martin Klingenspor and Dr. Florian Bolze who taught me functional assays and Ivonne Jarick, PD. Dr. André Scherag and Carolin Pütter who helped me with anything statistical.

I also want to thank all members of the molecular genetic laboratory who helped me to stay sane and get my work done: Thank you Jitka, for your wit and encouragement. Thank you Sigi, for being there and helping with every (research) problem. Thank you Beate, for a different perspective on life. Thank you Susann, for helping out when things really got tough. Thank you again Anke, for your patience with me and everything you taught me. Special thanks to Nadja, who never let me down and helped intellectually, with her time and positive personality to make this thesis possible.

The last people I want to thank are my parents, Sebastian and my friends who kept me sane and were very patient when anything went wrong, took longer than expected or was more complicated than I thought. Thank you for your support and always being there for me!

12. Curriculum vitae

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.

For privacy reasons, the curriculum vitae is not included in the online version.

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.

For privacy reasons, the curriculum vitae is not included in the online version.

13. Eidesstattliche Erklärungen

Erklärung:

Hiermit erkläre ich, gem. § 6 Abs. (2) f) der Promotionsordnung der Fakultäten für Biologie, Chemie und Mathematik zur Erlangung der Dr. rer. nat., dass ich das Arbeitgebiet, dem das Thema *„In depth analysis of obesity candidate genes in the chromosomal region 16p11.2 identified in genome-wide association studies“* zuzuordnen ist, in Forschung und Lehre vertrete und den Antrag von Anna-Lena Volckmar befürworte und die Betreuung auch im Falle eines Weggangs, wenn nicht wichtige Gründe dem entgegenstehen, weiterführen werde.

Essen, den _____ _____

Unterschrift eines Mitglieds der Universität Duisburg-Essen

Erklärung:

Hiermit erkläre ich, gem. § 7 Abs. (2) c) + e) der Promotionsordnung Fakultäten für Biologie, Chemie und Mathematik zur Erlangung des Dr. rer. nat., dass ich die vorliegende Dissertation selbständig verfasst und mich keiner anderen als der angegebenen Hilfsmittel bedient habe.

Essen, den _____ _____

Unterschrift des/r Doktoranden/in

Erklärung:

Hiermit erkläre ich, gem. § 7 Abs. (2) d) + f) der Promotionsordnung der Fakultäten für Biologie, Chemie und Mathematik zur Erlangung des Dr. rer. nat., dass ich keine anderen Promotionen bzw. Promotionsversuche in der Vergangenheit durchgeführt habe und dass diese Arbeit von keiner anderen Fakultät/Fachbereich abgelehnt worden ist.

Essen, den _____ _____

Unterschrift des Doktoranden