Variation in Mitochondrial Genes in Obesity

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

Dr. rer. nat.

der Fakultät für

Biologie

an der

Universität Duisburg-Essen

vorgelegt von

Nadja Knoll
aus Plauen

August 2013

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

1. Gutachter: Prof. Anke Hinney
2. Gutachter: Prof. Bernhard Horsthemke

Vorsitzender des Prüfungsausschusses: Prof. Daniel Hoffmann

Tag der mündlichen Prüfung: 18.11.2013

Publications and Presentations

First authorship

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(1):75-7.

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(2):e55884.

Co-authorship

Föcker M, Timmesfeld N, Scherag S, Knoll N, Singmann P, Wang-Sattler R, Bühren K, Schwarte R, Egberts K, Fleischhaker C, Adamski J, Illig T, Suhre K, Albayrak O, Hinney A, Herpertz-Dahlmann B, Hebebrand J. Comparison of metabolic profiles of acutely ill and short-term weight recovered patients with anorexia nervosa reveals alterations of 33 out of 163 metabolites. J Psychiatr Res. 2012;46(12):1600-9.

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(8):888-97.

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, Pütter C, Song JY, Graniger 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(6):334-7.

Reviews/Book chapters

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(5):294-303. Review.

Hebebrand J, Hinney A, Knoll N, Volckmar AL, Scherag A. Molecular genetic aspects of weight regulation. Dtsch Arztebl Int 2013;110(19):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.

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).

Talks

"Gene set of nuclear encoded mitochondrial regulators is enriched for inherited variation in obesity" Knoll N, Jarick I, Volckmar AL, Klingenspor M, Illig T, Grallert H, Gieger C, Wichmann HE, Peters A, Hebebrand J, Scherag A, 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
- 4th Annual Meeting of NGFN-Plus and NGFN-Transfer in the Program of Medical Genome Research 2011, September 27-29 in Berlin
- 28. Jahrestagung der Deutschen Adipositasgesellschaft 2012, October 4-6 in Stuttgart

Posters

"Lack of association of a fatty *acid amide hydrolase (FAAH)* gene variant to weight loss in a one-year intervention for obese children and adolescents (Obeldicks)" Knoll N, Volckmar AL, Pütter C, Scherag A, Kleber M, Hebebrand J, Hinney A, Reinehr T.

- European Human Genetics Conference 2011, May 28-31 in Amsterdam
- 27. Jahrestagung der Deutschen Adipositasgesellschaft 2011, October 6-8 in Bochum

"Gene set of nuclear encoded mitochondrial regulators is enriched for inherited variation in obesity" Knoll N, Jarick I, Volckmar AL, Klingenspor M, Illig T, Grallert H, Gieger C, Wichmann HE, Peters A, Hebebrand J, Scherag A, Hinney A.

- Forschungstag des Universitätsklinikums Essen 2011, November 25 in Essen
- 19th European congress on Obesity 2012, May 9-12 in Lyon
- EMBO/EMBL Symposium Diabetes and Obesity 2012, September 13-16 in Heidelberg
- 28. Jahrestagung der Deutschen Adipositasgesellschaft 2012, October 4-6 in Stuttgart

"Common variation in mitochondrial DNA is not associated with obesity" Knoll N, Jarick I, Volckmar AL, Klingenspor M, Illig T, Grallert H, Gieger C, Peters A, Wiegand S, Biebermann H, Fischer-Posovszky P, Wabitsch M, Völzke H, Rosskopf D, Rimmbach C, Schreiber S, Jacobs G, Tittmann L, Franke A, Hebebrand J, Scherag A, Hinney A

- Forschungstag des Universitätsklinikums Essen 2012, November 23 in Essen
- 2nd Annual Meeting of the BIOME Core "Genetics and Cell Biology" 2012, November 29 in Essen
- 5th Annual Meeting of NGFN-Plus and NGFN-Transfer in the Program of Medical Genome Research 2012, December 11-13 in Heidelberg

Index

Publications and Presentations ..... **iii**

Index ..... **v**

Abbreviations and Units ..... **ix**

Index of Figures ..... **xiv**

Index of Tables ..... **xvi**

1 Introduction **1**

1.1 Obesity *1*

1.1.1 Definition and classification 1

1.1.2 Prevalence of obesity 3

1.1.3 Causes of obesity epidemics 4

1.2 Heritability of obesity *6*

1.3 Molecular genetic studies on obesity *7*

1.3.1 Molecular genetic approaches 7

1.3.2 Monogenic forms of obesity and major gene effects 9

1.3.3 Polygenic forms of obesity 10

1.3.4 Missing heritability 12

1.3.5 Pathway-based approaches and gene set enrichment analyses (GSEA) 12

1.4 Mitochondria *14*

1.4.1 Description and function 14

1.4.2 Mitochondrial DNA (mtDNA) 16

1.4.2.1 Origin of mtDNA 16

1.4.2.2 Amount of mtDNA molecules per cell 16

1.4.2.3 Characterization of sequence 17

1.4.2.4 Coding region 18

1.4.2.5 Control region 19

1.4.2.6 Transcription and replication 21

1.4.2.7 Mode of inheritance and haplogroups 22

1.4.3 Nuclear encoded mitochondrial genes 24

1.4.4 Nuclear-mitochondrial interactions 26

1.5 Mitochondrial/mtDNA alterations and body weight associated phenotypes 28

1.5.1 Variation in mitochondrial genes 28

1.5.1.1 Common polymorphisms in mtDNA 28

1.5.1.2 Mitochondrial diseases 28

1.5.2 Alterations in mitochondrial gene expression and function in obese individuals 29

2 Objectives 30

2.1 Analysis of mtDNA 30

2.2 Analysis of nuclear-encoded mitochondrial genes 30

3 Subjects and Methods 32

3.1 Subjects 32

3.1.1 Samples for early onset extreme obesity 33

3.1.1.1 Case-control sample 33

3.1.1.2 Family-based sample ('Trios') 33

3.1.2 Population-based samples 34

3.1.2.1 KORA 34

3.1.2.2 SHIP 34

3.1.2.3 POPGEN 34

3.2 Chemicals and buffers 37

3.3 Molecular genetic analyses 38

3.3.1 Genotyping 38

3.3.2 Variant detection by Sanger-re-sequencing 40

3.3.2.1 Complete mitochondrial DNA 40

3.3.2.2 Mitochondrial displacement loop (D-loop) 44

3.4 Haplogroup determination 46

3.5 Statistics 47

3.5.1 Association tests 47

3.5.1.1 Common SNPs of mtDNA 47

3.5.1.2 Variants of mtDNA detected by D-loop re-sequencing 47

3.5.1.3 Autosomal SNPs 48

3.5.2 Gene set enrichment analysis (GSEA) 49

3.5.2.1 Gene sets 49

3.5.2.2 Mapping of SNPs onto genes 51

3.5.2.3 Determination of gene-wise empirically corrected p-value P_g 51

3.5.2.4 Determination of gene set enrichment p-values P^{GSEA} 52

3.5.2.5 Meta-Analysis Gene set Enrichment of variant Associations (MAGENTA) 52

3.5.2.6 Determination of LD between best SNPs of gene set 1 53

4 Results 54

4.1 Variation in mitochondrial DNA 54

4.1.1 Association of common mitochondrial SNPs with obesity 55

4.1.2 Association of haplogroups with obesity 57

4.1.3 Re-sequencing of complete mtDNA 62

4.1.4 Re-sequencing of mitochondrial D-loop 65

4.2 Gene set enrichment analysis of nuclear encoded mitochondrial genes 74

4.2.1 Discovery 74

4.2.2 Confirmation 74

4.2.3 Meta-analysis 75

5 Discussion 82

5.1 Variation in mitochondrial DNA 82

5.1.1 Common variation in mtDNA in association with obesity 82

5.1.2 Evaluation of haplogroup assignment 83

5.1.3 D-loop variation in association with obesity 89

5.1.4 Limitations of mtDNA analysis 92

5.2 Gene set enrichment analysis on nuclear-encoded mitochondrial genes 93

5.2.1 Enrichment of association signals in gene set 1 93

5.2.2 Relevance of genes of gene set 1 pertaining to body weight 94

5.2.3 Evaluation of the method 95

5.2.4 Limitations of present GSEA 96

5.3 Study samples 97

6 Summary and Conclusion 99

7 References 101

8 Appendix 123

Danksagung xviii

Curriculum vitae xix

Ehrenwörtliche Erklärung xx

Abbreviations and Units

Abbreviations

A	adenosine
ADP	adenosine diphosphate
AgRP	agouti related peptide
AMPK	5' adenosine monophosphate-activated protein kinase
ATP	adenosine triphosphate
ATP6, 8	ATP synthase F0 subunit 6, 8
BDNF	brain derived neurotrophic factor
BMI	body mass index
BMI-SDS	body mass index standard deviation score
BS	binding site
C	cytosine
Ca	calcium
CaMKIV	calcium/calmodulin-dependent protein kinase type IV
cAMP	cAMP, cyclic adenosine monophosphate
CC	case-control
CE	control element
cGMP	cyclic guanosine monophosphate
CNV(s)	copy number variants
COI, II, III	subunits of OXPHOS complex IV
conc.	concentrated, concentration
COX	cytochrom-C-oxidase
CR	call rate
CREB	cAMP response element-binding protein
CRS	Cambridge Reference Sequence
CSB1,2, 3	conserved sequence block 1, 2, 3
Cyt b	subunit of OXPHOS complex III
Cyt C	cytochrom-C
DEGS1	German Health Interview and Examination Survey for Adults
D-loop	displacement loop
DNA	deoxyribonucleic acid
dNTP	desoxy nucleoside triphosphate
Drp1	dynamin-related protein 1
DZ	dizygotic
e.g.	for example
ECDF	empirical cumulative distribution functions
EDTA	ethylenediaminetetraacetic acid
EPIC	European Prospective Investigation into Cancer and Nutrition
ERRα	estrogene related receptor α
ESRRA	estrogen related receptor alpha
ESRRG	estrogen related receptor gamma
ETAS1, 2	extended termination associated sequence 1, 2
ETC	electron transport chain
F	forward primer

$FADH_2$	flavin adenine dinucleotide
Fe	iron
Fis1	fission 1
FTO	fat mass and obesity associated gene/protein
G	guanine
GABPA	GA-binding protein alpha subunit
GABPB1, 2	GA-binding protein beta subunit 1, 2
GNHIES98	German National Health Interview and Examination Survey 1998
GO	gene ontology
GSEA	gene set enrichment analysis
GWAS	genome-wide association study
H^+	protons
H_2O	water
HCl	hydrochloric acid
HeLa	Henrietta Lacks
HSP1, 2	heavy strand promoter 1, 2
H-strand	heavy strand
HV1, 2, 3	hypervariable region 1, 2, 3
HWE	Hardy-Weinberg-Equilibrium
i.e.	that is
ICS	inter cristae space
IM	inner membrane
IMS	inner membrane space
k.o.	knock out
KEGG	Kyoto Encyclopedia of Genes and Genomes
KIGGS	German Health Interview and Examination Survey for Children and Adolescents
KORA	Kooperative Gesundheitsförderung in der Region Augsburg
KS	Kolmogorov-Smirnov test
LD	linkage disequilibrium
LEP	leptin gene
LSP	light strand promoter
L-strand	light strand
MA	matrix of mitochondria
MAF	minor allele frequency
MAGENTA	Meta-Analysis Gen set Enrichment Analysis of variaNT Associations
MC4R	melanocortin 4 receptor protein/gene
MEF-2	myocyte enhancer factor-2
MEF2A	myocyte-specific enhancer factor 2A
MELAS	mitochondrial encephalopathy lactic acidosis and stroke-like syndrome
Mff	mitochondrial fission factor
Mfn1, 2	mitofusin 1, 2
$MgCl_2$	magnesium chloride
min	minimum
MIRAS	mitochondrial recessive ataxia syndrome
MNGIE	mitochondrial neurogastrointestinal encephalopathy
mRNA	messenger RNA

MS	mass spectroscopy
MSRA	methionine sulfoxide reductase A gene
Mt3, 4, 5	metallothionein 3, 4, 5
mtDNA	mtDNA, mitochondrial DNA
mtSSB	mitochondrial single stranded binding protein
mtTF1	mitochondrial transcription factor 1
MYC	myelocytomatosis viral oncogene homolog (avian)
MZ	monozygotic
n.d.	not defined
NAD^+	nicotinamide adenine dinucleotide – oxidized form
NADH	nicotinamide adenine dinucleotide – reduced form
NaOH	sodium hydroxide
ND1, …, 6	subunits 1 to 6 of OXPHOS complex I
NDUFB8	NADH dehydrogenase (ubiquinone) 1 beta subcomplex 8
NHANES	National Health and Nutrition Examination Survey
NRF1, 2	nuclear respiratory factor 1, 2
NRIP1	nuclear receptor-interacting protein 1
O_2	oxygene
ObRb	leptin receptor protein
O_H	heavy strand origin of replication
O_L	light strand origin of replication
OM	outer membrane
Opa1	optic atrophy gene 1
OR	odds ratio
OXPHOS	oxidative phosphorylation
PCR	polymerase chain reaction
PCSK1	prohormone-convertase-I/3
P_g	gene-wise corrected p-value
PGC-1b	peroxisome proliferator-activated receptor gamma coactivator 1 beta
PGC-1α	peroxisome proliferator-activated receptor gamma coactivator 1 alpha
p^{GSEA}	gen set enrichment p-value
PKA	protein kinase A
$p^{MAGENTA}$	gen set enrichment p-value derived by MAGENTA
Pol γ	polymerase γ
POLG1	polymerase γ
POLRMT	mitochondrial DNA-directed RNA polymerase
POMC	pro-opiomelanocortin
POPGEN	population genetic research project of the national genome research network
PPARA	peroxisome proliferator-activated receptor alpha
PPARD	peroxisome proliferator-activated receptor delta
PPARGC1A	peroxisome proliferator-activated receptor gamma coactivator 1 alpha
PPARGC1B	peroxisome proliferator-activated receptor gamma coactivator 1 beta
PPARα	peroxisome proliferator-activated receptor alpha
PPARγ	peroxisome-proliferator-activated receptor γ
PPARδ	peroxisome proliferator-activated receptor delta
Q	ubiquinone

QC	quality control
QH_2	ubiquinol
r	correlation coefficient
R	reverse primer
rCRS	revised Cambridge Reverence Sequence
RIP140	receptor-interacting protein 140
RNA	ribonucleic acid
rRNA	ribosomal RNA
SHIP	The Study of Health in Pomerania
SIRT1	Sirtuin 1
SNP(s)	single nucleotide polymorphism(s)
SP1	Specificity Protein 1
T	Thymidine
T2DM	type 2 diabetes mellitus
TAE	Tris-Acetat-EDTA-Puffer
TAS	termination associated sequence
TDT	transmission disequilibrium test
TFAM	mitochondrial transcription factor A
TFB1M	transcription factor B1, mitochondrial
TFB2M	transcription factor B2, mitochondrial
TNKS	tankyrase, TRF1-interacting ankyrin-related ADP-ribose polymerase gene
Tom/Tim	transporter outer membrane/transporter inner membrane
TRF1	telomeric repeat binding factor 1
Tris	tris(hydroxymethyl)-aminomethan
TRKB	tyrosine receptor kinase B
tRNA	transfer RNA
TYMP	thymidine phosphorylase
U	uridine
UTR	untranslated region
VIP	vasoactive intestinal peptide
vs.	versus
WHO	World Health Organization
WMW	Wilcoxon-Mann-Whitney test
YY1	transcriptional repressor protein YY1
α-MSH	α-melanocyte stimulating hormone

Units

°C	degree(s) Celsius
bp	base pair(s)
g	gram(s)
h	hours(s)
kg	kilogram(s)
L	liter(s)
m	meter(s)
M	molar (mol/L)
pH	$-\lg[H^+]$
s	second(s)
U	unit(s)
V	volt(s)
v/v	volume per volume
w/v	weight per volume
m	milli (10^{-3})
n	nano (10^{-9})
p	pico (10^{-12})
µ	micro (10^{-6})

Index of Figures

Figure 1.1 BMI percentile curves for girls aged 0-18 years 2

Figure 1.2 Hierarchical model of relationship among major determinants of energy balance 5

Figure 1.3 Leptinergic-melanocortinergic system 11

Figure 1.4 Schematic representation of mitochondrial TOM/TIM protein complex 14

Figure 1.5 Electron transport chain and oxidative phosphorylation 15

Figure 1.6 Coding region of human mtDNA 18

Figure 1.7 Control region (D-loop) of human mtDNA 19

Figure 1.8 Simplified haplogroup tree 24

Figure 1.9 Example of mitochondrial entrance sequence 25

Figure 1.10 Interplay of nuclear and mtDNA-encoded proteins 26

Figure 1.11 Nuclear-mitochondrial communication 27

Figure 3.1 Study samples and purpose in the present analysis 32

Figure 3.2 Coverage of sequencing primers for re-sequencing (Sanger) of complete mtDNA 40

Figure 3.3 Establishment of annealing temperature for fragment 1 and fragment 2 41

Figure 3.4 Selection of primers for the re-sequencing of the mtDNA D-Loop 44

Figure 3.5 Establishment of annealing temperature for D-loop PCR fragment 44

Figure 3.6 Description of GSEA procedure 49

Figure 4.1 Approach for analysis/detection of variation in mtDNA in association with obesity 55

Figure 4.2 Comparison of haplogroup distribution between discovery sample and individuals whose D-loop was re-sequenced (D-loop sample) 67

Figure 4.3 Detected length heteroplasmies 71

Figure 4.4 Comparison of mean number of D-loop deviations per individual between cases and controls 71

Figure 4.5 Empirical cumulative distribution functions (ECDF) of gene p-values Pg in all autosomal genes and in three gene sets of nuclear encoded mitochondrial genes (discovery) 75

Figure 4.6 Empirical cumulative distribution functions (ECDF) of gene p-values Pg in all autosomal genes and gene set 1 (confirmation) 76

Figure 5.1 Geographic distribution of frequencies of haplogroup X (A) and W (B) 83

Figure 5.2 Haplogroup I as part of the phylogenic mtDNA tree 86

Figure 5.3 Potential misclassification of haplogroups by HaploGrep due to lacking genotype information 87

Figure 5.4 Haplogroup W as part of the phylogenic mtDNA tree 88

Figure 8.1 Evaluation sheet of re-sequencing of mitochondrial D-loop 123

Index of Tables

Table 1.1 WHO Classification of adults according to BMI 1
Table 1.2 Functionally relevant regions of the D-loop 20
Table 1.3 Estimated haplogroup frequencies of West Europe in % 24
Table 3.1 Phenotypical characteristics of subjects 35
Table 3.2 Used chemicals in alphabetical order 37
Table 3.3 Components of 50 x TAE buffer for gel electrophoresis 37
Table 3.4 PCR product loading buffer for agaroses gels 37
Table 3.5 Quality control of SNPs 39
Table 3.6 Components of reaction mix for long range PCR fragments 42
Table 3.7 Long range PCR conditions for the amplification of both mtDNA fragments 42
Table 3.8 Self-selected primers for re-sequencing for complete mtDNA 43
Table 3.9 Components of reactions mixes for the D-Loop PCR product 45
Table 3.10 PCR conditions for the amplification of the D-Loop 45
Table 3.11 Nuclear-encoded regulators of mitochondrial genes (gene set 1) 50
Table 4.1 SNPs of mtDNA in association with obesity in discovery 56
Table 4.2 Nominally associated SNPs of mtDNA in discovery and follow-up in confirmation 57
Table 4.3 Haplogroup frequencies in discovery and confirmation 58
Table 4.4 Frequency of major haplogroups in cases and controls in discovery and confirmation 60
Table 4.5 Frequency of haplogroups in cases and controls in discovery and confirmation stratified by gender 61
Table 4.6 Variants detected by re-sequencing of complete mtDNA of each five lean and obese individuals 62
Table 4.7 Haplotype classification after complete re-sequencing using HaploGrep 65
Table 4.8 Comparison of haplogroup distribution between discovery and individuals whose D-loop was re-sequenced (D-loop sample) 66
Table 4.9 Detected variants (single nucleotide exchanges) by re-sequencing of mtDNA and frequencies in cases and controls 68
Table 4.10 Detected variants (complex nucleotide exchanges, insertions and deletions) by re-sequencing of mtDNA and frequencies in cases and controls 70

Table 4.11 Comparison of mean number of variants in functionally relevant D-loop locations between cases and control 73

Table 4.12 GSEA and MAGENTA for obesity in the CC sample (discovery) 77

Table 4.13 GSEA and MAGENTA for the gene set of 16 nuclear regulators of mitochondrial genes (confirmation) 78

Table 4.14 Best SNPs of nuclear regulators of mitochondrial genes (gene set 1) in each sample and linkage disequilibrium between best SNPs of the three different study samples 79

Table 4.15 MAGENTA for the gene set of 16 nuclear regulators of mitochondrial genes (meta-analysis) 80

Table 4.16 Best SNPs of nuclear regulators of mitochondrial genes (gene set 1) in meta-analysis and linkage disequilibrium to best SNP of family-based trios 81

Table 5.1 Distribution of haplogroup frequencies (in %) in present study samples and among West Europeans and Germans 85

Table 5.2 Mouse models (knockout, alterations in expression and mutations) for 9 of the 16 genes of gene set 1 showing a body weight associated phenotype 95

Table 8.1 Enrichment of association signals above 95th percentile in 705 family-based trios with initial genotype calling using eight batches 124

1 Introduction

1.1 Obesity

1.1.1 Definition and classification

Obesity is defined as an excessive accumulation of body fat, which increases the proportion of body fat mass to whole body mass, as a result of a longstanding positive energy balance (Hebebrand et al. 2001a). In general, the body mass index (BMI) calculated by body weight (in kg) over squared body height (in m^2) is used for classification of obesity. Due to a high correlation to body weight (r=0.8-0.9) and body fat mass (r=0.6-0.7) accompanied by a low correlation to body height (r=-0.2) in both men and women, the BMI represents a body height adjusted measure for body fat mass (Watson et al. 1979). According to World Health Organization (WHO) criteria, adult people are assigned to weight classes by use of the BMI, and obesity is defined as a BMI ≥ 30 kg/m^2 (Table 1.1; WHO 2000).

Table 1.1 WHO Classification of adults according to BMI

Classification	BMI [kg]
Underweight	<18.50
Normal weight	18.50-24.99
Overweight	≥25.00
Pre-obesity	25.00-29.99
Obesity	≥30.00
Class I	30.00-34.99
Class II	35.00-39.99
Class III	≥40.00

adapted from WHO (2000)

However, there are some limitations using the BMI as estimation for body fat mass. Individuals with a large muscle mass and hence a high body weight for their height might be misclassified as (pre-) obese (Lambert et al. 2012), while the body fat mass of individuals with a reduced lean body mass as for instance found in elderly can be underestimated (Gallagher et al. 1996). In addition, in children and adolescents, lean body mass fluctuates stronger than in adulthood due to different developmental stages, i.e. the BMI rises continuously during the first six months of life and decreases until the age of six to eight years because of increased linear growth; afterwards BMI rises again (Fig. 1.1; Rolland-Cachera et al. 1991). Hence, percentile curves (Fig. 1.1) which represent age and gender adjusted BMI distributions of a reference population are more adequate than the pure BMI to define weight classes (Hebebrand et al. 1994, Kromeyer-Hauschild et al. 2001). In

Germany, a meta-analysis of 17 studies including a total of 17,147 boys and 17,275 girls aged 0-18 years is primarily used as reference (Kromeyer-Hauschild et al. 2001). Currently, overweight and obesity are defined as the 90th and 97th age and gender adjusted BMI percentile, respectively, set according to the guidelines of the "Arbeitsgemeinschaft Adipositas im Kindes- und Jugendalter" (www.a-g-a.de). Formerly, the German National Survey I was used as reference, and overweight and obesity were defined as the 90th and 99th age and gender adjusted BMI percentile (Hebebrand et al. 1994). In addition, the standard deviation score (SDS) can be applied to quantify by which factor of a standard deviation the individual BMI deviates from the age and gender adjusted BMI median of the reference population (Kromeyer-Hauschild et al. 2001).

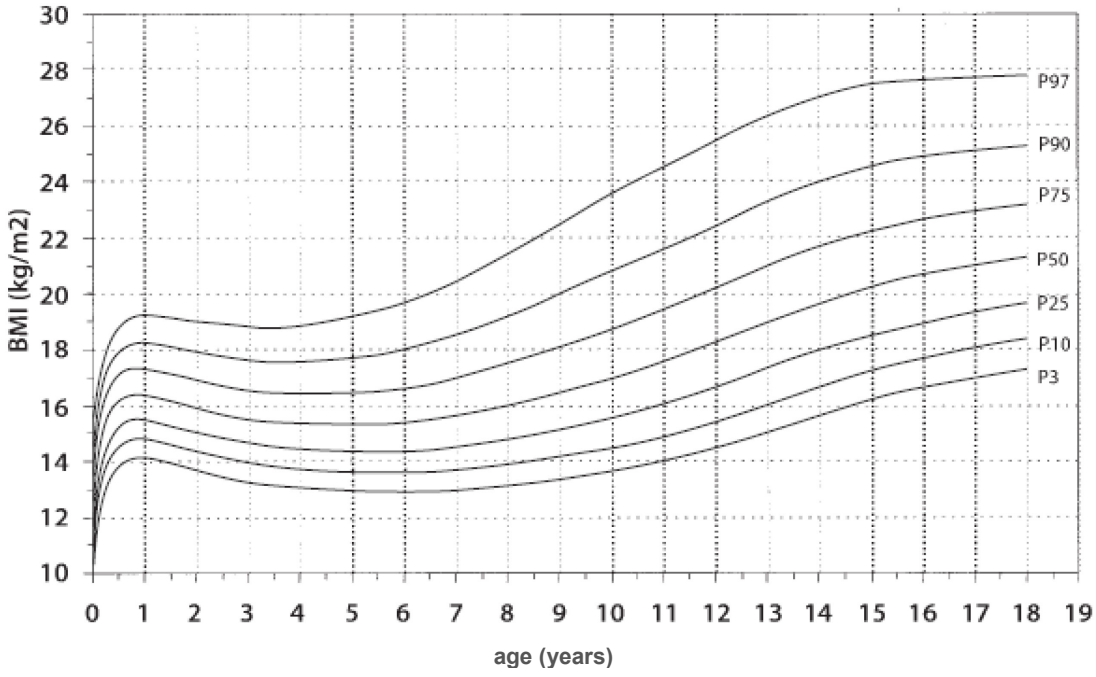

Figure 1.1 BMI percentile curves for girls aged 0-18 years

Curves derive from a meta-analysis of 17 studies (performed in Germany) including 17,275 girls aged 0-18 years. Data from the 17,147 boys included were used to establish similar BMI percentile curves. According to the guidelines of the "Arbeitsgemeinschaft Adipositas im Kindes- und Jugendalter" (www.a-g-a.de), the 90th and 97th age and gender adjusted BMI percentile are used to define overweight and obesity, respectively. *Figure from Kromeyer-Hauschild et al. (2001)*

1.1.2 Prevalence of obesity

According to a recent report by the WHO (2013), world-wide more than 1.4 billion adults of 20 years and older (35 %) were overweight (BMI ≥ 25 kg/m^2) in 2008. Among these, about 500 million, i.e. 11 % of adults and 7 % of the world's total population in 2008, were obese (BMI ≥ 30 kg/m ^{2}, UN World Population Prospects 2011, WHO 2013). Because of consequences in later life such as type two diabetes mellitus (T2DM), ischemic heart diseases and cancer, obesity has become a worldwide important health problem (WHO 2013).

Pertaining to results of the German Health Interview and Examination Survey for Adults (DEGS1) performed between 2008 and 2011, 67.1 % of men and 53.0 % of women were overweight (Mensink et al. 2013). These figures are comparable to those reported in the German National Health Interview and Examination Survey 1998 (GNHIES98, Bergmann and Mensink 1999). However, the prevalence of obesity has increased: 23.3 % of men and 23.9 % of women were obese in DEGS1 compared with 18.9 % and 22.5 % in GNHIES98 (Mensink et al. 2013). In the USA, data of the National Health and Nutrition Examination Survey (NHANES) show that obesity has been prevalent among 35.5 % of men and 35.8 % of women in 2009-2010 (Flegal et al. 2012). While between 1988-1994 and 1999-2000 age and gender dependent increases in the obesity prevalence of 3.4 % to 12.0 % were found (Flegal et al. 2002), the figures of 2009-2010 were similar to those in 2003-2008 indicating a leveling off effect of the obesity prevalence (Flegal et al. 2012).

In 2011, more than 40 million children under the age of five were overweight (WHO 2013). Childhood obesity is associated with a higher risk of obesity during adulthood (Power et al. 1997). During the last three to four decades, prevalence rates of obesity during childhood and adolescence rose especially in developed and newly industrializing countries. For instance, prevalence rates have increased 2.3-fold to 3.3-fold between 1971-74 and 1999 in the USA, 3.4-fold to 4.6-fold between 1985 and 1995 in Australia, 3.4-fold to 3.6-fold between 1974 and 1997 in Brazil and 1.1 to 1.4-fold between 1991 and 1997 in China (Ebbeling et al. 2002). In addition, the heaviest children from US surveys from the 1990s were significantly heavier than the heaviest children in 1960-1970, while the BMI distribution in the middle and lower ranges did not change (Troiano and Flegal 1998). Comparatively, this was also seen in obese children and adolescents treated as inpatients from six different German study groups over a period of 10 years (1985-1995). While the most pronounced increase was detected among the obese above the 9th decile (5 kg/m^2 and

2.5 kg/m^2 in boys and girls, respectively), in the lower BMI range a shift towards a higher BMI has not been seen (Barth et al. 1997).

According to representative data pertaining to overweight and obesity among 14,836 children and adolescents of 3 to 17 years of age from the German Health Interview and Examination Survey for Children and Adolescents (KiGGS), 15 % and 6.3 % were overweight and obese, respectively (Kurth and Schaffrath-Rosario 2007). Blüher et al. (2011) analyzed prevalence trends for overweight and obesity among 272,826 children and adolescents of 4 to 16 years in Germany. While there was an upward trend for both overweight and obesity prevalence rates between 1999 and 2003, a decrease was found between 2004 and 2008.

A stabilizing trend of the obesity prevalence was also seen among children and adolescents since the late 1990s or first years of the 2000s in other countries, such as England (Karlsen et al. 2013), Australia (Hardy et al. 2012), France (Salanave et al. 2009), the Netherlands (de Wilde et al. 2009) and Sweden (Sundblom et al. 2008). Interestingly, in some of these countries, the stabilizing trend was not seen among ethnic minorities (de Wilde et al. 2009, Karlsen et al. 2013).

1.1.3 Causes of obesity epidemics

The increase in the prevalence rates of obesity since the last three to four decades might be attributed to both environmental factors and a genetic predisposition (Bouchard 2007, Hebebrand and Hinney 2009). An excessive availability of cheap high-caloric food in combination with a more and more sedentary lifestyle form an “obesogenic” environment for humans especially in developed and newly industrializing countries, and a persisting positive energy balance seem to be the major causes for obesity (Hebebrand et al. 2001a, Ebbeling et al. 2002).

On the other hand, Neel hypothesized the existence of a “thrifty genotype” regarding the development of T2DM in the 1960s: a decreased energy expenditure accompanied by an increased energy storage capacity had been advantageous to survive when food was not as easy accessible and/or scarce. Consequently, genetic variation favoring this phenotype might have been enriched during human evolution (Neel 1962). However, a “thrifty genotype” in times of excessive availability of food, as found nowadays in many parts of the world, can no longer be considered a survival advantage, but it promotes the obesity epidemics. Bouchard (2007) derived a hierarchical model depicting the relationship between the major

determinants of energy balance (built environment, social environment, behavior and biology; Fig. 1.2).

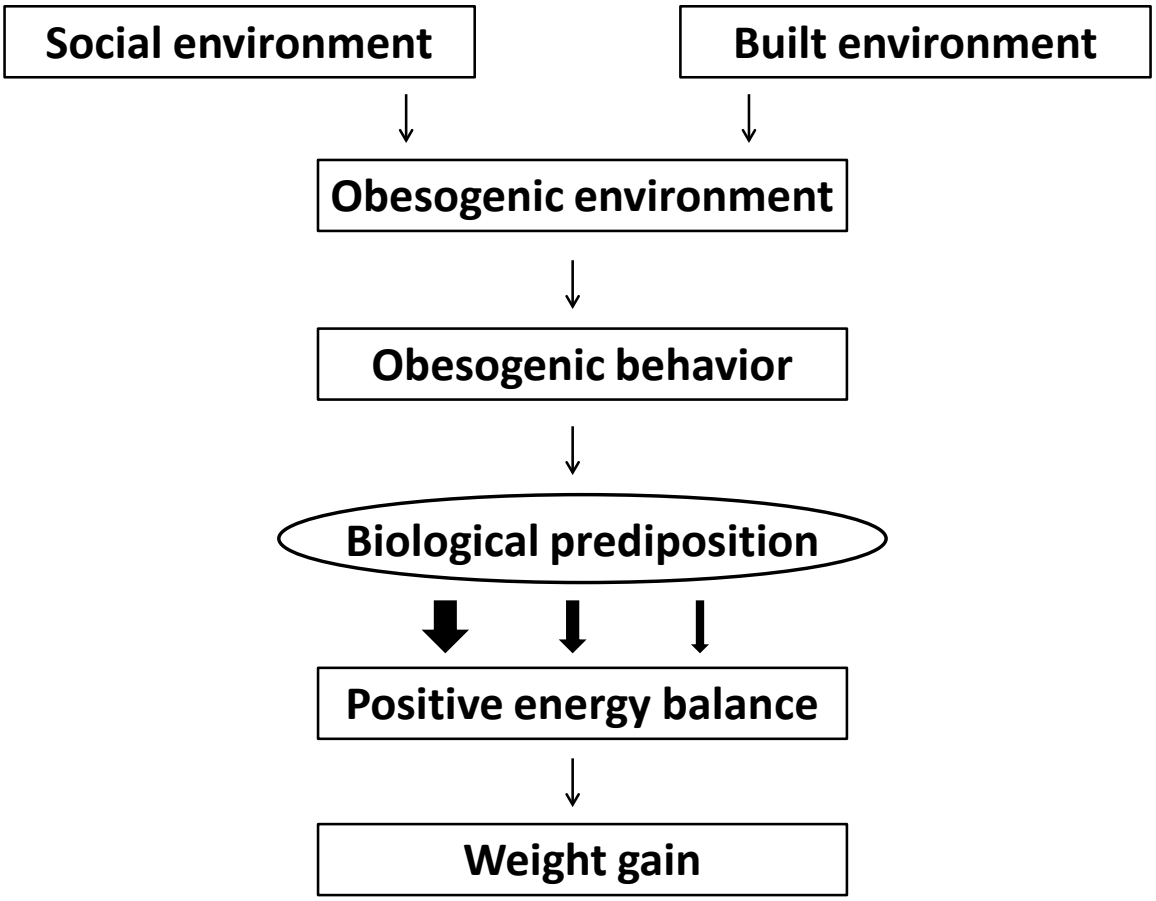

Figure 1.2 Hierarchical model of relationship among major determinants of energy balance

Factors in the built environment (e.g. reliance on the automobile) and the social environment (e.g. advertising, pressure to consume) create an obesogenic environment. This in turn is favorable for the adaptation of an obesogenic behavior (e.g. consumption of large portion sizes of meals, many hours spent watching TV). Both the environment and behavior have accelerated the increase in the obesity prevalence. However, a biological predisposition (e.g. genetic variation) – indicated by arrows of different sizes in the figure – is important for inter-individual differences in the predisposition of gaining weight and risk of becoming obese, particularly severe obese. *Adapted by permission from Macmillan Publishers Ltd: International Jornal of Obesity (Bouchard, 2007), © 2007.*

1.2 Heritability of obesity

Numerous twin, family and adoption studies have been performed to estimate the heritability of body weight variation. Of particular interest are twin pairs which were reared apart, as the contribution of a common environment which most twins experience, especially during childhood and adolescence, is reduced. Stunkard et al. (1990) investigated intra pair BMI correlations in a total of 673 monozygotic (MZ) and dizygotic (DZ) adult twin pairs which were reared together or apart from the Swedish Adoption/Twin Study of Aging. The intra pair BMI correlations were higher in the MZ twin pairs (between 0.66 and 0.74) compared with the DZ twin pairs (between 0.15 and 0.33), but were similar independent whether the twin pairs were reared together or apart. Among 34 reared apart and 38 reared together British MZ twin pairs, intra pair correlations of body fat of 0.61 and 0.75 were reported, respectively (Price and Gottesman 1991).

In family studies, correlations in BMI between siblings which were reared together ranged between 0.1 and 0.5, and were slightly lower than between parents and their offspring (~0.1-0.3; reviewed by Maes et al. 1997). Among 540 adult Danish adoptees, the relationship between the weight class (thin, median weight, overweight, and obese) of the adoptee and the BMI of the biological parents and siblings was strong, while BMI correlations between the adoptee and the adoptive parents or siblings were not existing (Sørensen et al. 1989, Stunkard et al. 1986). In a longitudinal adoption study of 840 Danish adoptees (partly the same as the above mentioned 540), 269 of which height and weight data at the ages 7 to 13 years were present, BMI correlations between the adoptees and their biological parents were higher than with their adoptive parents (0.16 to 0.17 vs. 0.03 to 0.10). Comparatively, BMI correlations between the adoptees and biological full siblings were 4.2-fold higher than BMI correlations between the adoptees and their adoptive siblings (0.59 vs. 0.14; Sørensen et al. 1992).

Considering all empirical study types, it is assumed that heritability estimates range from 0.4 to 0.7 (Hebebrand et al. 2010). Interestingly, some of these family and adoption studies have shown that correlations in BMI between mothers and their offspring are higher than between fathers and their offspring (Hebebrand et al. 2001a). For instance, Sørensen et al. (1998) reported (based on the above mentioned Danish adoptees) that correlation in BMI between the adoptee and the biological mother and father were 0.15 and 0.11, respectively. In 179 Italian families, the mother-offspring correlation in BMI was 0.37, while the father-

offspring's was 0.31 (Zonta et al. 1987). Price et al. (1987) reported that among 357 adult American adoptees correlations in BMI were 0.40/0.15 between biological mothers and daughters/sons, while correlations between biological fathers and daughters/sons were 0.18/0.08, respectively. Comparatively, in a study of young Swedish men, maternal half-brothers show two-fold higher correlations in BMI than paternal half-brothers do (r=0.21 vs. r=0.11, Magnusson and Rasmussen 2002).

1.3 Molecular genetic studies on obesity

1.3.1 Molecular genetic approaches

On a molecular genetic level, several approaches have been undertaken to elucidate genetic variation to be associated with obesity. There are basically two major approaches – candidate gene studies and genome-wide approaches – which have been and still are performed to get insight into the genetics of body weight regulation. In candidate gene studies, specific genes with a known role in metabolism or body weight regulation are investigated. These might be derived from animal models as for instance knockout (*k.o.*) models of the respective genes and are screened for the detection of mutations (Hinney et al. 2004). Later on, allele frequencies of detected variations are classically compared between cases of a certain trait (e.g. early onset (extreme) obesity) and controls, i.e. individuals which are ideally only different by the trait of investigation. Another option to follow-up detected variants are family-based studies. A family-based study typically consists of "trios", i.e. one index proband with the phenotype of interest (e.g. early onset (extreme) obesity) and both biological parents. Statistically, a transmission-disequilibrium-test (TDT) is performed (Spielman et al. 1993), which compares the transmission of each allele of the variant of interest, by only including heterozygous parents. A transmission of 50 % for each allele is expected by chance, and a significant deviation from the expected transmission (= transmission disequilibrium) indicates association of the variant with the investigated trait. One advantage of family-based studies compared with case-control (CC) studies is that they are not subject to population stratification. Both cases and controls of the investigated trait might be from different population subgroups and hence the allele frequencies *per se* might be different in cases and controls leading to false positive findings (Attia et al. 2009b). However, one disadvantage of family-based studies is that the parents might also carry the same phenotype as their offspring (Hebebrand et al. 2000), and thus might carry relevant genetic variants homozygously. This might decrease the power for association testing, as

only heterozygous parents are included in the TDT. Significantly associated variants may be either directly functionally relevant for the investigated trait (e.g. early onset (extreme) obesity), or a marker for a functionally relevant variant in high linkage disequilibrium (LD).

In contrast to candidate studies, genome-wide approaches for the detection of genes being associated with obesity are "hypothesis-free" (Kitsios and Zintzaras 2009, Day and Loos 2011). Formerly, some linkage analyses have been performed. In these studies, obesity predisposing genes were detected by genetic markers (e.g. microsatellites) which are theoretically inherited together with a nearby gene, i.e. in LD to the gene (Hebebrand et al. 2001b). However, linkage approaches as well as candidate studies majorly failed to be successful in the detection of robust loci or genes associated with obesity (Day and Loos 2011).

The most recent and more prosperous approaches regarding the identification of genes associated with obesity are genome-wide association studies (GWAS). These studies are mainly focused on single nucleotide polymorphisms (SNPs) and copy number variations (CNVs), which are evenly spread over the whole genome. More than 12 million common SNPs with a frequency of > 1 % have been identified so far (Attia et al. 2009a). Efficient genotyping technologies, the sequencing of the human genome, and the creation of SNP databases such as HapMap (www.hapmap.org; International HapMap Consortium 2003) helped this approach to be performed on a large scale (Day and Loos 2011). Recent GWAS arrays, such as the Affymetrix Genome-Wide Human SNP Array 6.0, cover each one million of SNPs and CNVs evenly distributed across the whole genome including sex chromosomes and mitochondrial DNA (mtDNA). SNPs on the array found to be associated with a trait can be either functionally relevant *per se*, or serve as a marker for a nearby SNP or genomic region which is in high LD with the SNP.

Since 2005, GWAS have contributed to a further elucidation of the genetic architecture of many complex traits, such as obesity, coronary heart disease or diabetes, and have revealed many gene loci which would have never been detected by classical candidate approaches as their impact on the investigated traits is still mainly unknown (Pearson and Manolio 2008, Kitsios and Zintzaras 2009). Nevertheless, as many variants are tested simultaneously for association, the risk of spurious association is high. Therefore, currently a p-value of $5x10^{-8}$ (correction for one million tests) is considered genome-wide significant and a confirmation of the initial findings in an independent study sample is required (Attia et al. 2009a, b, Manolio 2010).

1.3.2 Monogenic forms of obesity and major gene effects

Monogenic forms of obesity are characterized by mutations in a single gene that lead to (extreme) obesity. In case a gene has a clear impact on the development of obesity it is referred to as a major gene. Both monogenic forms and major gene effects regarding obesity are infrequent (Hinney et al. 2010). All current known monogenic forms of obesity are derived from mutations in genes belonging to the leptinergic-melanocortinergic system, which is involved in the central regulation of the energy balance (Fig. 1.3, Hinney et al. 2010, Hebebrand et al. 2013). In brief, the satiety hormone leptin is generated by adipocites. Leptin reaches the arcuate nucleus in the hypothalamus through the bloodstream. There, it induces the syntheses of pro-opiomelanocortin (POMC) which is cleaved to α-melanocyte stimulating hormone (α-MSH) by prohormone-convertase-I/3 (PCSK1). α-MSH is an agonist of the melanocortin-4-receptor (MC4R) located in higher order neurons of the hypothalamus. This in turn activates the brain derived neurotrophic factor (BDNF), the binding of which to tyrosine receptor kinase B (TRKB) signals a saturated – anorexigenic – condition. By contrast, in case of low levels or lack of leptin the agouti-related protein is expressed and acts as an antagonist of the MC4R, signaling a hunger stimulating – orexigenic – condition (Hebebrand et al. 2013). The identification of these genes as being involved in the development of early onset extreme obesity in a monogenic manner occurred through *k.o.* or transgenic mouse models (Hinney et al. 2010). Furthermore, large pedigree studies of consanguineous families based on extremely obese individuals were used (Hinney et al. 2010). Mutations in the leptin gene (*LEP*) have the strongest effects. Homozygous mutation carriers, which lack the synthesis of leptin despite a significantly elevated fat mass, are subject to a continuously increased orexigenic condition. This in turn leads to hyperphagia and results in early onset (extreme) obesity. Only 11 families with mutation carriers have been identified so far whose parents were, with one exception, consanguineous (Montague et al. 1997, Strobel et al. 1998, Mazen et al. 2009, Fischer-Posovszky et al. 2010). It was reported that in case of denying food to the severely hyperphagic children, they showed an aggressive behavior (Montague et al. 1997). After treatment with human recombinant leptin, these children developed a normal eating behavior. Sustained beneficial effects on fat mass, hyperinsulinemia, and hyperlipidemia were also observed (Farooqi et al. 2002). For instance, during the first 12 months of leptin therapy one of the children lost 16.4 kg of body weight, of which 95 % were fat (Farooqi et al. 1999).

As depicted in Fig. 1.3, MC4R plays a crucial role in the body weight regulation. More than 150 mutations in *MC4R* with impact on the receptor function have been identified so far (Hinney et al. 2013). The frequency of these mutations in extremely obese children and adolescents is 2.5 % (Hinney et al. 2003), while in obese adults, only 0.2 % are carriers of functionally relevant mutations (Hinney et al. 2006). Adult carriers are on average 15 kg (males) to 30 kg (females) heavier than relatives without mutation in *MC4R* (Dempfle et al. 2004). However, not every carrier of a functionally relevant mutation develops obesity (Hinney et al. 2006). In addition, the reported effect sizes regarding body weight are lower than those reported in other monogenic forms of obesity, which is why *MC4R* has rather major gene effects than monogenic (Hebebrand et al. 2013).

Interestingly, three of the genes with major gene or monogenic effects on obesity also contribute to polygenic forms of obesity (Fig. 1.3, *MC4R:* Geller et al. 2004, Stutzmann et al. 2007, Loos et al. 2008, *BDNF*: Thorleiffson et al. 2009, *POMC*: Speliotes et al. 2010).

1.3.3 Polygenic forms of obesity

In polygenic forms of obesity, many genes with small effect size contribute to obesity. In the latest population-based GWAS meta-analysis of ~250,000 individuals of European origin, 32 gene loci have been robustly associated with BMI (Speliotes et al. 2010). The effect sizes of these loci range from 0.06 kg/m^2 to 0.39 kg/m^2. The fat mass and obesity associated gene (*FTO*) had the largest effect size in the meta-analysis of Speliotes et al. (2010). It was first identified as a robustly BMI associated polygene in a GWAS for T2DM, with a subsequent replication in nearly 40,000 individuals (Frayling et al. 2007). Homozygous carriers of the risk allele were on average 2 to 3 kg heavier than non-carriers (Frayling et al. 2007, Hinney et al. 2007, Speliotes et al. 2010). According to bioinformatics analyses, the FTO protein has equal motifs as found in Fe(II) and 2-oxoglutarate-dependent oxygenases (Gerken et al. 2007). Moreover, it was shown that recombinant murine Fto is involved in demethylation of single stranded DNA (Gerken et al. 2007). Fischer et al. (2009) found *Fto k.o.* mice to be leaner with a reduced fat mass than littermates. Berulava and Horsthemke (2010) showed that increased FTO expression is associated with increased fat mass. Both studies are directionally consistent and point at a gain of function effect of *FTO* in obesity. In most human studies, *FTO* alleles predisposing to obesity are associated with increased food intake rather and not with energy expenditure (summarized by Müller et al. 2013).

A more recent GWAS meta-analysis of 14 studies with 5,530 obese cases and 8,318 lean controls detected two further SNPs associated with childhood and adolescent obesity (Bradfield et al. 2012). These were directionally consistent in the above mentioned meta-analysis on adult BMI (Speliotes et al. 2010). Another GWAS meta-analysis focusing on the extremes, i.e. individuals of the upper and lower 5th BMI percentiles are compared, yielded in the detection of seven new loci associated with different classes of obesity (Berndt et al. 2013).

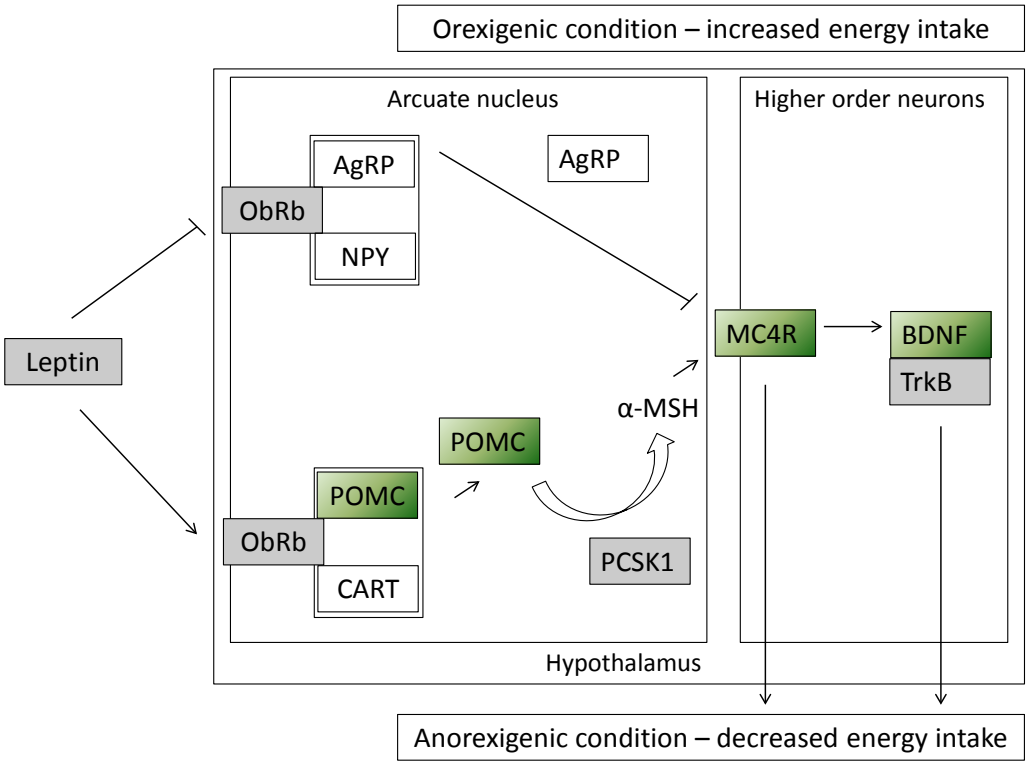

Figure 1.3 Leptinergic-melanocortinergic system

The satiety hormone leptin is generated by adipocytes in quantities depending on the size of the fat depot. It reaches the arcuate nucleus in the hypothalamus via the blood stream, and increased amounts of leptin are recognized by the leptin receptor (ObRb). The binding to the receptor stimulates the synthesis of pro-opiomelanocortin (POMC) which is cleaved to α-melanocyte stimulating hormone (α-MSH) by prohormone-convertase-I/3 (PCSK1). α-MSH is an agonist of the melanocortin-4-receptor (MC4R) being located in higher order neurons of the hypothalamus. This in turn activates the brain derived neurotrophic factor (BDNF), the binding of which to tyrosine receptor kinase B (TrkB) signals a saturated – anorexigenic – condition. By contrast, in case of low levels or lack of leptin the agouti-related protein (AgRP) is expressed and acts as an antagonist of the MC4R, signaling a hunger stimulating – orexigenic – condition.
Gray shaded boxes: mutations lead to monogenic forms of obesity; green shaded boxes: mutations leading to mono- and polygenic forms of obesity; *Figure and legend adapted from Hebebrand et al. (2013).*

1.3.4 Missing heritability

Though having a large effect on the body weight, all recently identified monogenic forms of obesity are infrequent. Thus, their contribution to the global obesity epidemics is certainly rather low. Considering the effect sizes of all 32 polygenic loci of the last population-based meta-analysis, only 1.5 % of the BMI variance can be explained (additive genetic effects). This is based on the polygenic nature of these genes and their small effect size. Even by increasing the sample size to 730,000 individuals to detect 250 further gene loci with similar effect sizes as those found in the 32 loci (i.e. 0.06 kg/m^2 to 0.39 kg/m^2), estimations predicted that still only 4.5 % of the variance in BMI would be explained (Speliotes et al. 2010). By contrast, empirical studies have shown heritability estimates for the variance of the BMI between 40 % and 70 %. The discrepancy between the empirical ascertained heritability and that explained by all variants identified in molecular genetic studies is coined as “missing heritability” (Hebebrand et al. 2010). Reasons for this phenomenon are among others that the empirically estimated high heritability might have been overestimated. On the other hand, there might be further infrequent monogenic variants with strong effects on the body weight, but due to their frequency they cannot be identified by GWAS (Hebebrand et al. 2010). The identified polygenic variants were assumed to have additive effects. Yang et al. (2011a) have shown that by including all 586,898 autosomal SNPs of a GWAS of 11,586 unrelated individuals, a total of 16.5 % of the variance in BMI would be explained. Moreover, non-additive genetic effects might be existent. According to simulations, such epistasis effects explain large parts of the missing heritability (Zuk et al. 2012).

1.3.5 Pathway-based approaches and gene set enrichment analyses (GSEA)

In GWAS, each SNP has been generally considered individually. However, genes and their gene products often interact in functionally relevant groups or pathways, which as a whole may contribute to a disease. Hence, for the discovery of further variants being associated with obesity, an extension of the single-locus-oriented GWAS to pathway-based approaches or gene set enrichment analyses (GSEA) might be an option for new genetic insight into the trait of interest (Liu et al. 2010). These approaches might thus help to diminish the gap between empirically ascertained and genetically explained heritability.

In a pathway-based approach or GSEA, gene association signals of pathways or gene sets which are biologically plausible for a given trait like obesity are compared with those of the genome-wide set of genes (Subramanian et al. 2005, Wang et al. 2007, Segrè et al. 2010).

Hence, these approaches concentrate on the combined effects across several loci; each locus might have a too small effect to be detected on its own, however in combination the impact could be striking enough (Wang et al. 2007, Torkamani et al. 2008, Segrè et al. 2010).

Pathway-based analyses or GSEA have been initially applied on gene expression microarray data (Mootha et al. 2003, Subramanian et al. 2005). For instance, Mootha et al. (2003) found PGC-1α responsive genes which are involved in the oxidative phosphorylation to be downregulated in skeletal muscle tissue of individuals with T2DM. On the gene expression level, each gene is represented by the maximum or median expression value of its transcript(s) and/or probe sets (Wang et al. 2007). On a SNP level, by contrast, each gene is represented by several common SNPs, of which only a few are causal or in LD with SNPs that contribute to a disease risk (Wang et al. 2007). Thus, Wang et al. (2007) suggested taking the maximum statistic for all SNPs in and near a gene to represent the significance of the gene and to use a permutation-based approach to adjust for multiple testing.

Liu et al. (2010) performed a pathway-based GWAS analysis of 963 pathways or gene sets in 1,000 US whites regarding obesity. The pathways or gene sets were generated using the public databases BioCarta, Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) and Ambion Gene Assist. The authors found the vasoactive intestinal peptide (VIP) pathway significantly associated with both BMI and body fat mass. This pathway was also shown to be significantly associated in a sub-sample of 3,552 individuals of the EPIC-Norfolk study (Evangelou et al. 2012).

Segrè et al. (2010) developed an extended GSEA method called MAGENTA (Meta-Analysis Gene set Enrichment Analysis of variaNT Associations). Using MAGENTA, large GWAS meta-analysis, of which only summary statistics are available, can be analyzed. For this purpose, for each gene, its best SNP p-value received from a meta-analysis has been taken forward and corrected for confounders such as gene size, variant number and LD properties using linear regression analysis. Segrè et al. (2010) did not find enrichment of association signals of common variants for T2DM and related glycemic traits in three gene sets of mitochondrial genes in a large GWAS meta-analysis using MAGENTA.

As mitochondria have a predominant role in the cellular energy generation, variation in mitochondrial genes might have an impact on body weight. Thus, the mitochondrial gene sets used by Segrè et al. (2010) have been tested for enrichment of obesity association signals in the present PhD thesis.

1.4 Mitochondria

1.4.1 Description and function

Mitochondria are oval to longish organelles with a constant thickness of ~1 µm and variable length. They are characterized by a double membrane. The outer membrane is permeable for intermediates of the metabolism, but not for correctly folded proteins. The protein complex TOM/TIM (transporter outer membrane/transporter inner membrane; Fig. 1.4) which connects the outer and inner membrane is responsible for the active transport of unfolded proteins synthesized in the cytosol (Löffler and Petrides 2003, pp.195ff and pp.287f, Bolender et al. 2008).

On the other hand, the inner membrane is impermeable for any substance. It consists of 70 % protein. In addition, the inner membrane is folded into cristae which enlarge its surface enormously (Löffler and Petrides 2003, p.195).

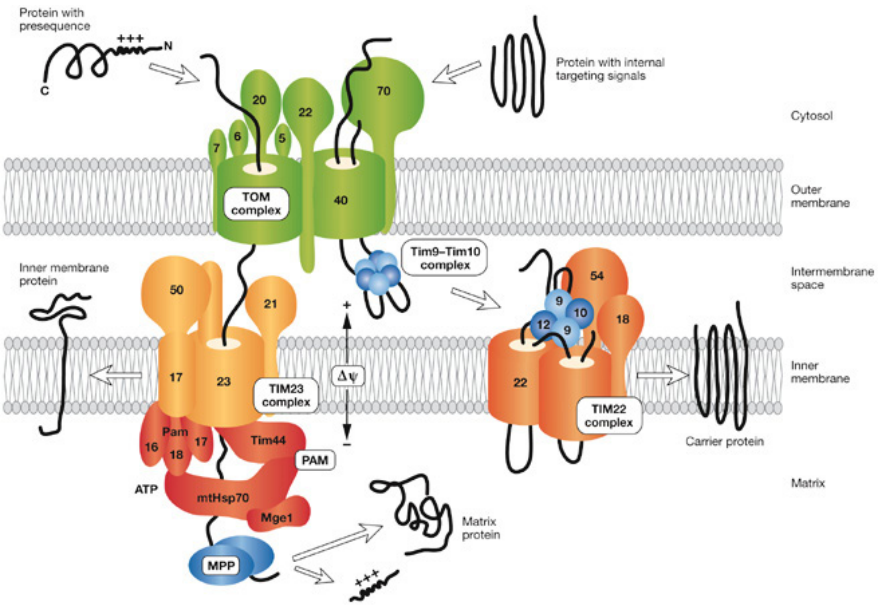

Figure 1.4 Schematic representation of mitochondrial TOM/TIM protein complex

Inner membrane proteins and matrix proteins synthesized in the cytosol are directed by a positively-charged presequece through the TOM complex, TIM23 complex and motor PAM to the matrix. The presequence is removed by the matrix processing peptidase (MPP). Inner membrane proteins are released laterally from TIM23 complex. Carrier proteins containing internal targeting signals, which are recognized by the receptor Tom70, are translocated by the TOM complex as well as Tim9-Tim10 chaperone of the intermembrane space and inserted into the inner membrane by TIM22 complex. The proton gradient at the inner membrane is the driving force for the passage through the mitochondrial membranes.Mge1, nucleotide exchange factor; MtHsp70, matrix heat shock protein 70; PAM, presequence translocase-associated motor; TIM, translocase (transporter) of inner membrane; TOM, translocase (transporter) of outer membrane. *Reprinted by permission from Macmillan Publishers Ltd: EMBO reports (Bolender et al., 2008), © 2008.*

Mitochondria are well known as the cellular power plants. Their main function is the generation of the energy equivalent adenosine triphosphate (ATP). For this purpose, the complexes I to IV of the electron transport chain (ETC) and complex V for the oxidative phosphorylation (OXPHOS, in the following "OXPHOS" is referred to as OXPHOS including ETC), which are located in the inner membrane, are necessary (Fig. 1.5). The reduction equivalent nicotinamide adenine dinucleotide (NADH), which is derived from the citric acid cycle, is converted to its oxidative state (NAD^+) by passing its electrons to complex I of the electron transport chain. This in turn induces a reduction cascade of the complexes II to IV finally to reduce molecular oxygen to water. Via the complexes I, III and IV, protons (H^+) are actively transported into the inter membrane space of the mitochondria and generate a proton gradient. This gradient is finally used for the OXPHOS of adenosine diphosphate (ADP) to ATP via the ATPase of complex V (Löffler and Petrides 2003, p.195 and p. 537). A resting adult requires 40 kg of ATP per 24 hours, which is mainly derived by the mitochondrial OXPHOS; intensive exercise might augment this requirement by 0.5 kg per minute (Stryer 1999, p. 470).

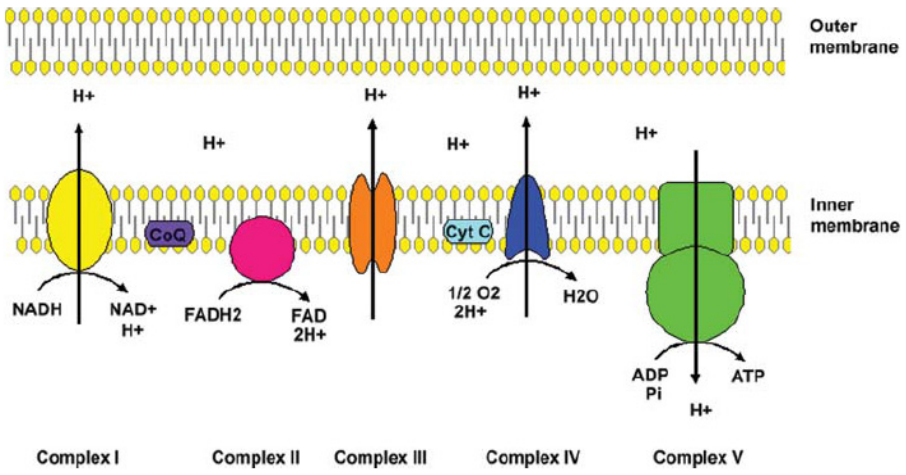

Figure 1.5 Electron transport chain and oxidative phosphorylation

The reduction equivalent nicotinamide adenine dinucleotide (NADH), which is derived from the citric acid cycle, is converted to its oxidative state (NAD^+) by passing its electrons to complex I (NADH-ubiquinone-oxidoreductase) of the electron transport chain. Subsequently, succinate is oxidized to fumerate while ubiquinone (Q) is reduced to ubiquinol (QH_2) at complex II (succinate-ubiquinone-oxidoreductase). FADH2 serves as reduction equivalent. At complex III (ubiquinol-cytochrome C-oxidoreductase), QH_2 is oxidized to Q, while cytochrome *C* (cyt *C*) is reduced. Thereafter, cyt *C* is oxidized and molecular oxygen (O_2) is reduced to water (H_2O) at complex IV (cytocrome *C*-oxidase). Finally, H+ influx is the driving force for the synthesis of ATP at complex V (*ATPase*). *Figure from St John et al. Mitochondrial DNA transmission, replication and inheritance: a journey from the gamete through the embryo and into offspring and embryonic stem cells, Hum Reprod Update, 2010, 16(5):488-509, by permission of Oxford University Press.*

Besides the generation of ATP via the OXPHOS, further processes as for instance the citric acid cycle and the fatty acid ß-oxidation can be found in mitochondria. Moreover, mitochondria can induce apoptosis by release of cyt *C* (Löffler and Petrides 2003, p. 214).

In addition, mitochondria are highly dynamic organelles which are able to fuse or to fission in order to regulate their morphology, distribution and activity (Chan 2006). Fusion and fission are generally counterbalanced processes, which can be shifted to one direction in order to regulate the energy demand; for instance, mitochondria tend to fuse and elongate during starvation, while fragmentation (fission) is predominant during excessive nutrient supply (Liesa and Shirihai 2013). The regulation of the mitochondrial dynamics is important for cellular health or proper function (Hales 2010). By fusion, mitochondria exchange mtDNA and share proteins which might enable the mitochondrion to restore its function; a fission process which follows fusions might lead to newly refurbished mitochondria (Lindinger et al. 2010). The fusion machinery of mammals consists of three proteins, mitofusin 1 (Mfn1) and Mfn2 as well as the optic atrophy gene 1 (Opa1), while dynamin-related protein 1 (Drp1), fission 1 protein (Fis1) and the mitochondrial fission factor (Mff) are part of the mammalian fission machinery (Chan 2006, Liesa and Shirihai 2013).

1.4.2 Mitochondrial DNA (mtDNA)

1.4.2.1 Origin of mtDNA

According to the endosymbiotic theory, mitochondria have been formerly autonomous prokaryotes which were taken up by another cell as an endosymbiont. This theory originated from the Russian biologist Constantin Mereschkovsky at the beginning of the 20th century and was extensively studied by Lynn Margulis during the last century (Campbell and Reece 2002, p. 549). Molecular genetic analysis of the smaller ribosomal subunit RNA revealed that mitochondria are descendents of the α-proteobacteria. The endosymbiotic theory is supported by the fact that mitochondria show similarities to bacteria as for instance: (1) the creation of new mitochondria occurs only by fission, (2) mitochondria are surrounded by a double membrane and (3) mitochondria contain their own DNA – a circular DNA which is distinct from the nuclear encoded eukaryotic DNA (Campbell and Reece 2002, p. 550).

1.4.2.2 Amount of mtDNA molecules per cell

Somatic cells have a varying number of ~1,000-10,000 copies of mtDNA molecules (Falkenberg et al. 2007). Robin and Wong (1988) reported the number of virtual mitochondria and mtDNA molecules to be constant in a specific cell type, but considerable

variation between different cell types was found. Each mitochondrion has one to three mtDNA molecules (Robin and Wong 1988, Wiesner et al. 1992). All copies of mtDNA of an organism or tissue are usually identical (homoplasmy). However, especially among individuals with pathogenic mtDNA mutations leading to mitochondrial disease, only a fraction of mtDNA might be affected (heteroplasmy, Falkenberg et al. 2007). As the generation of the energy equivalent ATP is the major function of mitochondria, tissues of high energy demand (e.g. brain, retinal, skeletal and cardiac muscle tissues) are predominantly affected by these mutations (Chial and Craig 2008). However, heteroplasmies at functionally neutral positions might also be present in individuals without obvious mitochondrial disease (Sondheimer et al. 2011).

1.4.2.3 Characterization of sequence

Human mtDNA was completely sequenced for the first time by Frederic Sanger and colleagues in Cambridge in 1981 (Anderson et al. 1981). This sequence is referred to as the Cambridge Reference Sequence (CRS). The CRS was mainly derived from one person of European descent (human placenta mtDNA preparation). However, some regions had to be determined using HeLa cell mtDNA (African origin), and a few nucleotides were ambiguous and thus assumed to be identical with the respective nucleotide found in bovine mtDNA, which was also sequenced at that time (Anderson et al. 1981). As the CRS differed at several sites from mtDNA sequences re-sequenced by others, reanalysis was done 18 years later using the original placental mtDNA sample (Andrews et al. 1999). As both sequencing errors and rare polymorphisms were harbored, the investigators suggested revising the CRS by (1) correcting the 10 sequencing errors, (2) retaining rare polymorphisms such as m.263A (instead of G) and a C tract of only five Cs (instead of six) between m.311 and m.315, and (3) retaining the original numbering by replacing the second C misleadingly found at m.3107 in the CRS by an N in the revised CRS (rCRS) in order to remain consistent with previous literature (Andrews et al. 1999).

Sequencing of human mtDNA revealed a length of 16,569 bp. Due to its unequal base distribution mtDNA consists of a heavy strand (H-strand) and a light strand (L-strand, 25 % of T, 31 % of C, 31 % of A and 13 % of G; Anderson et al. 1981). Moreover, the genetic code deviates from the universal genetic code, as in human or generally in vertebrate mtDNA (1) UGA encodes for tryptophan, (2) AGA and AGG are stop codons, (3) AUA encodes for

methionine, and (4) AUA and AUU are additional start codons (Anderson et al. 1981, Jukes and Osawa 1990).

1.4.2.4 Coding region

The human mtDNA coding region is located between m.577 and m.16025 (Fig. 1.6). It encodes for 13 messenger RNAs (mRNAs), 22 transfer RNAs (tRNAs) and two ribosomal RNAs (rRNAs); except for one mRNA and eight tRNAs all genes are encoded by the H-strand (Fig. 1.6, Taylor and Turnbull 2005). The 13 mRNAs are subunits of complex I, III, IV and V of the mitochondrial OXPHOS system. The coding region is very compact. Between the genes, only very short non-coding regions can be found, and nucleotide overlaps between the genes are present as well (Chial and Craig 2008).

The 22 tRNAs are sufficient to read all 64 possible codons. Eight tRNAs are able to read each four codons, which only differ in the last codon position. These tRNAs are characterized by a modified U (pseudouridine) in the first position of the anticodon. Moreover, 13 tRNAs are able to read each two codons by building G:U wobbles between the third codon and first anticodon position. Apart from leucine and serine, which have two, there is only one tRNA for each amino acid (Anderson et al. 1981).

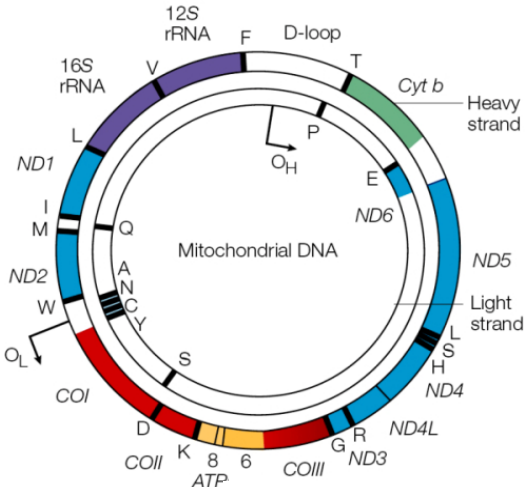

Figure 1.6 Coding region of human mtDNA

The coding region of human mtDNA encodes for subunits of complex I (=NADH-ubiquinone-oxidoreductase, *ND1*, *ND2*, *ND3*, *ND4*, *ND5*, *ND6* and *ND4L, blue*), complex III (=ubiquinol-cytochrome C-oxidoreductase, *Cyt b, green*), complex IV (cytocrome *c*-oxidase, *COI*, *COII* and *COIII, red*), complex V (ATPase, *ATP6* and *ATP8, yellow*), two rRNAs (12S and 16S, purple) and 22 tRNAs (*black line*) indicated as single letter code. The light strand origin of replication (O_L) is located in the coding region, while remaining function locations as e.g. the heavy strand origin of replication (O_H) are located in the control region (displacement (D)-loop). *Adapted by permission from Macmillan Publishers Ltd: Nature Reviews Genetics (Taylor and Turnbull, 2005), © 2005.*

1.4.2.5 Control region

The mtDNA control region is an 1122-bp-long non-coding region located between the sequences of $tRNA^{Phe}$ and $tRNA^{Pro}$ (Fig. 1.6). The control region harbors the origin of replication (O_H) of the H-strand. H-strand replication results in the formation of a displacement loop (D-loop) with the newly synthesized H-strand of ~680 bases; this DNA segment mostly does not grow to full length and is known as 7S DNA (Anderson et al. 1981). Because of this peculiarity, the control region is synonymously named D-loop. Subsequently, the term "D-loop" is used when referring to the control region of mtDNA. The D-loop has two hypervariable regions – HV1 (m.16024-m.16365) and HV2 (m.73-m.340) – characterized by a greater variability compared with the remaining mtDNA D-loop or coding region (Wilson et al. 1993; Fig. 1.7). Later, Lutz et al. (1998) suggested a third hypervariable region after having sequenced the D-loop in 200 German individuals. In that sample, the highest density of variants was obtained in HV1 (26 %, i.e. 88 variable positions in 342 bp), followed by HV2 (24 %) and HV3 (18 %), while variant rates of 7 % and 3 % were found in the sequence between the hypervariable regions (Lutz et al. 1998).

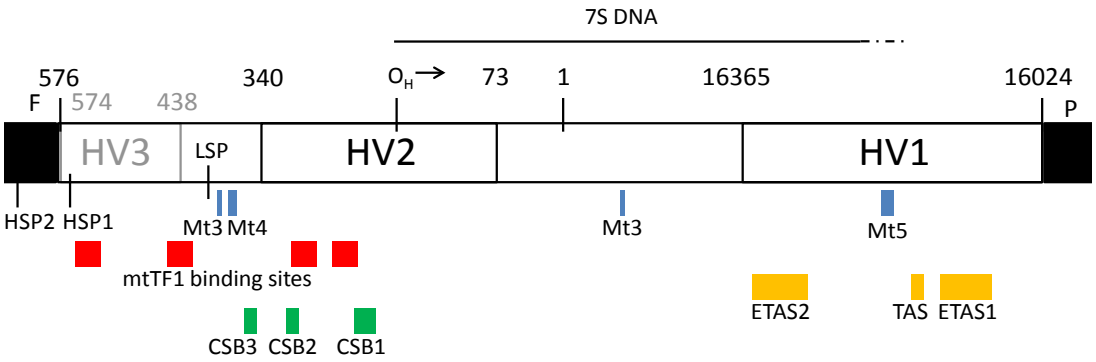

Figure 1.7 Control region (D-loop) of human mtDNA

The mtDNA control region is located between $tRNA^{Phe}$ (F) and $tRNA^{Pro}$ (P, cf. Fig. 1.6). Three hypervariable regions (HV1 to HV3) can be found characterized by a greater variant density. The D-loop harbors the origin of replication of the heavy strand (O_H), the light strand promoter (LSP) and the heavy strand promoter 1 (HSP1), while the second heavy strand promoter (HSP2) is located in $tRNA^{Phe}$. Premature termination of H-strand synthesis results in the formation of 7S DNA which is included in the formation of the displacement (D-loop). Moreover, mitochondrial transcription factor A (TFAM, formerly known as mtTF1) binding sites, several control element/transcription factor bindings sites (metallothionein 3 (Mt3) to Mt5), conserved sequence blocks (CSB1 to CSB3), as well (extended) termination associated sequences (TAS, ETAS1 and ETAS2) can be found in the D-loop. The function of each functionally relevant region is explained in Table 1.2

Table 1.2 Functionally relevant regions of the D-loop

Region	Description/Function	Location start	Location end	Reference
HV1a	Hypervariable regions	m.16024	m.16365	Wilson et al. 1993
HV1b		m.16024	m.16382	Meyer et al. 1999
HV2a		m.73	m.340	Wilson et al. 1993
HV2b		m.57	m.371	Meyer et al. 1999
HV3		m.438	m.574	Lutz et al. 1998
Mt5 (CE)	Intra- and interspecific control element (CE)	m.16194	m.16208	Ohno et al. 1991
Mt3 (L-strand CE)	- *cis* elements, first found in 5' region of nuclear encoded mitochondrial OXPHOS genes, later on also detected in mtDNA D-loop - potentially involved in coordinated expression of nuclear-encoded and mtDNA OXPHOS genes	m.16499	m.16506	Suzuki et al. 1991
Mt4 (L-strand CE)		m.371	m.379	Suzuki et al. 1991
Mt3 (H-strand CE)		m.384	m.391	Suzuki et al. 1991
mtTF1 BS	binding sites (BS) of mitochondrial transcription factor A (TFAM, formerly known as mtTF1)	m.233	m.260	Fisher et al. 1987
mtTF1 BS		m.276	m.303	Fisher et al. 1987
mtTF1 BS		m.418	m.445	Fisher et al. 1987
mtTF1 BS		m.525	m.552	Fisher et al. 1987
LSP (including mtTF1 BS)	L-strand and H-strand promoters	m.392	m.445	Chang and Clayton 1984
HSP1		m.545	m.567	Chang and Clayton 1984
HSP1 (including mtTF1 BS)		m.525	m.567	- -
HSP2		m.632	m.655	Lodeiro et al. 2012
CSB1	- Conserved sequence blocks	m.210	m.234	Sbisà et al. 1997
CSB2	- potentially involved in direction of transcription termination and H-strand primer formation	m.299	m.315	Sbisà et al. 1997, Pham et al. 2006
CSB3		m.346	m.363	Sbisà et al. 1997
ETAS1	- (extended) termination associated sequences - potentially involved in premature termination of H-strand synthesis	m.16081	m.16140	Sbisà et al. 1997
ETAS2		m.16294	m.16356	Sbisà et al. 1997
TAS		m.16157	m.16172	Roberti et al. 1998, Ingman and Gyllensten 2001

[a] LSP region alone also indicated at m.392 to m.445 (Chang and Clayton 1984)

[b] HSP2 was indicated at mitomap as m.645 (Montoya et al. 1982, 1983, Yoza et al. 1984), however, more recent investigation mapped start of HSP2 at m.644 (Zollo et al. 2012, Lodeiro et al. 2012); m.632 to m.655 was selected as HSP2 region, as Lodeiro et al. (2012) randomized these 24 nucleotides around the transcription start and did not detect transcription *in vitro*, thus, these 24 nucleotides might be important HSP2 control elements (e.g. transcription factor binding sites)

CE, control element; CSB, conserved sequence block; ETAS, extended termination associated sequence, HSP, heavy strand promoter; HV, hypervariable region; LSP, light strand promoter; mtTF1 BS, mitochondrial transcription factor A (TFAM) binding side (TFAM, formerly known as mtTF1), OXPHOS, oxidative phosphorylation; TAS, termination associated sequence

Besides O_H mapping at m.191 (Anderson et al. 1981), two promoters, the light strand promoter (LSP) and the heavy strand promoter 1 (HSP1) can be found in the D-loop (Fig. 1.7). Binding sites of the mitochondrial transcription factor A (TFAM, formerly known as mtTF1) are located near LSP and HSP1, as well as in HV2 (Fig. 1.7, Fisher et al. 1987). Moreover, several control elements (i.e. *cis* elements) are located in the D-loop: Mt3 on L- and H-strand, as well as Mt4 and Mt5 on L-strand. These elements are also present in the 5' region of nuclear-encoded mitochondrial OXPHOS genes and hence might be involved in the coordinated expression of nuclear and mtDNA encoded OXPHOS genes (Suzuki et al. 1991).

In addition, the D-loop has three conserved sequence blocks (CSB1 to CSB3) which are moderately to strongly conserved when comparing the mtDNA control region among different mammalian orders (Sbisà et al. 1997). They are potentially involved in the termination of LSP transcription and direction to primer formation for H-strand replication (Sbisà et al. 1997, Pham et al. 2006). The termination associated sequence (TAS) and both extended TAS (ETAS1 and ETAS2; Fig. 1.7) are conserved among different mammalian orders as well (Sbisà et al. 1997). They are potentially involved in the premature termination of H-strand synthesis resulting in the 7S DNA and hence D-loop formation (Anderson et al. 1981, Sbisà et al. 1997, Roberti et al. 1998).

1.4.2.6 Transcription and replication

Transcription of mtDNA L-strand is initiated in the light strand promoter (LSP), the start of which was mapped to m.407A. For H-strand transcription two promoters are present – heavy strand promoter 1 (HSP1) at m.561A and HSP2 at m.644A. The latter is not located in the D-loop, but in the adjacent $tRNA^{Phe}$ (Fig. 1.7, Zollo et al. 2012). Transcription from LSP and HSP2 produce polycistronic precursor RNAs which contain the entire genetic information of the respective strand, and excision of tRNAs enables the production of mature mRNA and rRNA molecules. By contrast, HSP1 transcription terminates at the 3' end of the 16S rRNA gene (Falkenberg et al. 2007). The mitochondrial DNA-directed RNA polymerase (POLRMT) requires the mitochondrial transcription factor A (TFAM) and either TFB1M or TFB2M (mitochondrial transcription factor B), as POLRMT is not able to interact with the mtDNA promoter sequence itself (Falkenberg et al. 2007). *In vitro*, however, only POLRMT and TFB2M were necessary for transcription of mtDNA fragments (Shutt et al. 2011, Zollo et al. 2012). The addition of TFAM at low concentrations leads to further activation of

transcription, while excess TFAM concentrations had inhibitory effects (Zollo et al. 2012, Lodeiro et al. 2012).

Replication of mtDNA occurs continuously without depending on cell cycle phase (Falkenberg et al. 2007). It is coupled with the LSP transcription, as during LSP transcription RNA primers for H-strand mtDNA replication starting at O_H are produced. The molecular mechanisms which direct the switch between LSP transcription and primer formation have not been completely identified yet. The three CSBs located near LSP might be involved in this process, as transition from RNA primer to DNA have been mapped near or within CSB2 (Falkenberg et al. 2007). Pham et al. (2006) have shown *in vitro* that mutation of CSB2 and, to a smaller extent, mutation of CSB3 decreased premature transcription termination. Considering these data, one can conclude that the primer formation and hence the H-strand synthesis is realized through sequence specific DNA elements (Pham et al. 2006, Falkenberg et al. 2007). After approximately two thirds of the H-strand (leading strand) has been synthesized, L-strand (lagging strand) replication starts at its origin of replication (O_L) in the opposite direction (strand-asymmetric model of replication, Falkenberg et al. 2007). Performing *in vitro* studies, Wanrooij et al. (2008) suggested that compared to H-strand replication POLRMT acts as primase in mammalian mitochondria by synthesis of 25 to 75 base-long RNA primers for lagging strand replication. Besides this classical strand-asymmetric replication model, advanced electrophoresis methods led to the proposition of further replication mechanisms of mtDNA, as for instance a bidirectional coupled leading and lagging strand synthesis at the same initiation site (strand-coupled model, Pham et al. 2006, Kasiviswanathan et al. 2012).

Furthermore, mtDNA is organized in nucleoids, which have been isolated by immunoaffinity purification to characterize the protein content. Twenty one proteins were identified, of which many are involved in mtDNA transcription and replication, e.g. TFAM, mitochondrial single-stranded DNA-binding proteins (mtSSB), which mediate the unwinding of mtDNA and stabilize mtDNA during replication, TWINKLE (=helicase) or mtDNA polymerase γ (POL γ; Falkenberg et al. 2007, St John et al. 2010, Campbell et al. 2012).

1.4.2.7 Mode of inheritance and haplogroups

Human or generally mammalian DNA is exclusively maternally inherited. During spermatogenesis, spermatogonia are labeled by ubiquitin (St John et al. 2010). After fertilization, these masked sperm mitochondria and their mtDNA are eliminated from the

oocyte cytoplasm by proteolysis at the 4 to 8-cell-stage in humans (St John et al. 2000). Consequently, the mtDNA from the population present in the oocyte just prior to fertilization which predominantly consists of identical copies is inherited to the offspring (St John et al. 2010).

mtDNA is characterized by a at least 5- to 15-fold higher mutation rate compared with nuclear encoded DNA, partly due to its exposition to reactive oxygen species generated by the ETC and also limited repair mechanisms (Payne et al. 2013). Mutation rate was estimated to $200x10^{-9}$ bp^{-1} $year^{-1}$ in HV1 and HV2 compared with $0.5x10^{-9}$ bp^{-1} $year^{-1}$ in the nuclear-encoded DNA (Scally and Durbin 2012). The high mutation rate results in high mtDNA diversity (Payne et al. 2013).

mtDNA is particularly suitable for the investigation of evolutionary aspects, as it is exclusively maternally inherited without any recombination events, and its high mutation rate enables to differentiate between most recently separated populations (Torroni et al. 1994). Based on the assumption that all human mtDNA haplotypes can be traced back to one common matrilineal ancestor living in Africa ~200,000 years ago, mtDNA variation evolved as consecutive accumulation of mutations along maternally inherited lineages. These accumulated mutations can be represented in a tree which reflects the phylogenetic relationship of mtDNA variants (van Oven and Kayser 2009). Major European and Asian haplogroups are branches of the L3 African haplotype, of which haplogroup M (predominantly found in the Indian subcontinent and farther east) and N have evolved as the first non-African founder nodes (Fernandes et al. 2012). All European haplogroups are derived from haplogroup N (Fig. 1.8, Ruiz-Pesini et al. 2007, www.mitomap.org). Fernandes et al. (2012) tried to locate the geographical site of the first steps of modern human outside of Africa by comparing Southwest Asian samples (Arabian Peninsula) of three minor haplogroups of western Eurasia (N1, N2 and X) branching directly from haplogroup N with 300 European samples. The authors found a relict distribution of these haplogroups in the European samples, and thus suggested that modern humans spread from the Gulf Oasis region to the Near East and Europe 55,000 to 24,000 years ago.

Pertaining to Europe, haplogroup H is the most frequent (41 % in West Europe, 48.6 % in Germany) followed by haplogroup U (18 % in West Europe, 13.5 % in Germany; www.mitomap.org; Pliss et al. 2006, Table 1.3). The rCRS (Anderson et al. 1981, Andrews et al. 1999) has a sub-haplogroup of H, H2a2a (Fig. 1.8, www.phylotree.org; built 11).

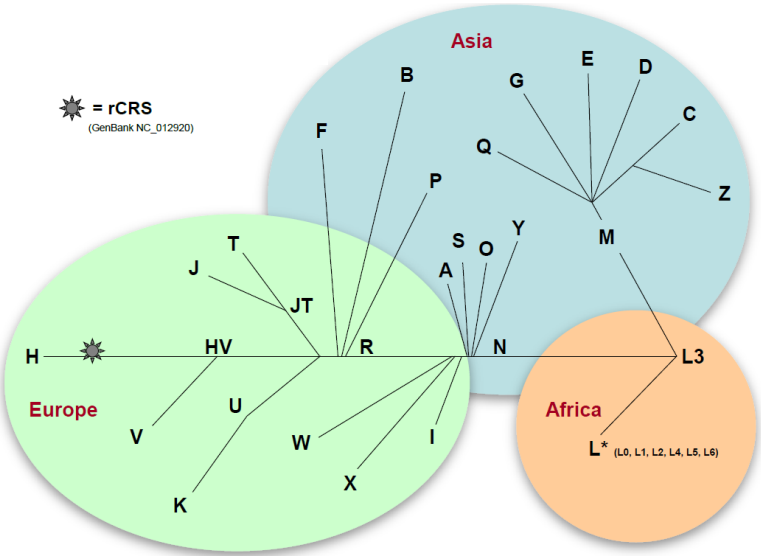

Figure 1.8 Simplified haplogroup tree

Major European and Asian haplogroups are branches from the L3 African haplotype. Haplogroups M and N are the first non-African founder nodes. All European haplogroups are derived from haplogroup N. For each major haplogroup, sub-haplogroups are existed. The rCRS (indicated as star) belongs to haplogroup H, the most frequent European major haplogroup. *Figure from www.mitomap.org, 2012 (Ruiz-Pesini et al. 2007). Figure content licensed by a Creative Commons Attribution 3.0 license.*

Table 1.3 Estimated haplogroup frequencies of West Europe in %

	D	H	I	J	K	L	M	N	R	T	U	V	W	X	n. d.
West Europe [a]	1	41	2	9	5	1	1	1	-	8	18	7	2	2	3
Germany [b]	0.6	48.6 [c]	1.8	8.4	7.5	1.2	-[d]	0.6	0.3	9.0	13.5	4.5	2.7	1.2	-

[a] Estimations based on means from published frequencies (bearing in mind that sometimes not all haplogroups have been typed), compiled in 2009 for Mitomap only for illustrative purpose (www.mitomap.org, Ruiz-Pesini et al. 2007); n. d., not defined

[b] data from Pliss et al. (2006) which are based on a total of n=333 German individuals

[c] haplogroups HV and preHV included

[d] all individuals of haplogroup M belonged to haplogroup D which directly branches off of M

1.4.3 Nuclear encoded mitochondrial genes

Besides those 13 protein encoding genes on the mtDNA, about 1,500 genes are estimated to be necessary to realize mitochondrial biogenesis (Lopez et al. 2000, Bar-Yaacov et al. 2012). Whereas for most of the mitochondrial functions (e.g. apoptosis, citric acid cycle or nucleotide biosynthesis) only proteins encoded by nuclear genes are involved, for

OXPHOS and the mitochondrial translation machinery proteins encoded by nuclear DNA and mtDNA are required (Bar-Yaacov et al. 2012).

Up to date, the largest investigation on the identification of mitochondrial proteins was done by Pagliarini et al. (2008). The authors isolated mitochondria from 14 different mouse tissues and performed mass spectroscopy (MS). Subsequently, the received MS data were compared with six pre-existent genome-scale datasets of mitochondrial localization (such as presence of mitochondrial target sequence (Fig. 1.9), yeast homology, transcriptional co-expression with known mitochondrial genes or ancestry with *Rickettsia prowazekii*, the closest still living bacterial relative of human mitochondria; Calvo et al. 2006). Further validation was achieved by green fluorescent protein tagging and a literature check of prior experimentally identified mitochondrial proteins. This resulted in an inventory of 1,098 mouse mitochondrial proteins called MitoCarta (www.broad.mit.edu/publications/MitoCarta/), from which 1,012 human homologs can be derived (including the 13 from the mtDNA and 31 encoded on sex chromosomes; Pagliarini et al. 2008). The 1,012 proteins/genes represent ~85 % of all assumed mitochondrial proteins/genes (Pagliarini et al. 2008, Segrè et al. 2010).

In addition to characteristic entrance sequences (Fig. 1.9), the identified MitoCarta genes are significantly shorter (UTRs and coding region), and more highly expressed in comparison with all mouse genes. Their promoters are enriched for sequence motifs, as for instance transcription factor binding sites, and tend to have CpG islands, while TATA boxes are not present. The latter feature might explain their higher expression levels, as housekeeping genes generally lack these boxes (Pagliarini et al. 2008).

Figure 1.9 Example of mitochondrial entrance sequence

Entrance sequence of nuclear encoded mitochondrial genes is recognized by receptors being located on the outer mitochondrial membrane, and only proteins carrying such sequences are imported. Single letter amino acid code is used (red, hydrophobic proteins; yellow, basic proteins; green, serine and threonine). Entrance sequences typically have a length of 15 to 35 amino acids and are enriched by positively charged amino acids (serine and threonine). A consensus sequence has not yet been found. *Figure adapted from Stryer (1999, p. 975)*.

1.4.4 Nuclear-mitochondrial interactions

For most mitochondrial functions, only nuclear-encoded proteins are necessary, however, the mitochondrial translation machinery and OXPHOS require both nuclear- and mtDNA-encoded proteins (Fig. 1.10, Bar-Yaacov et al. 2012). Thus a communication between both compartments (nucleus and mitochondria) as well as a coordinated expression of the two genomes is indispensable (Ryan and Hoogenraad 2007).

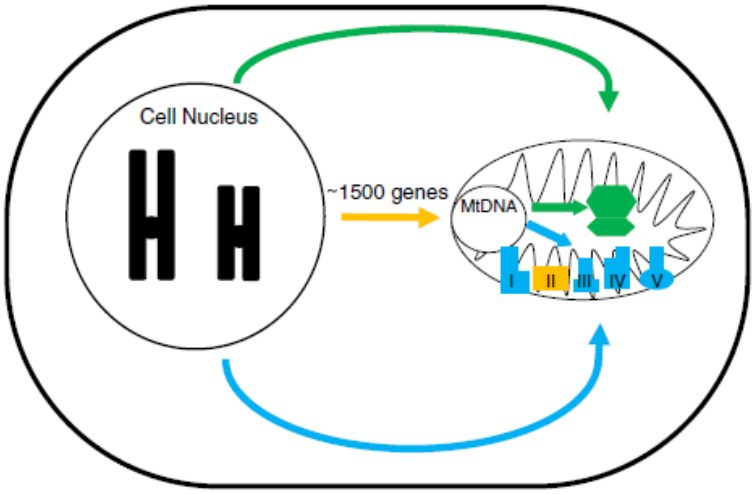

Figure 1.10 Interplay of nuclear and mtDNA-encoded proteins

For most mitochondrial functions only nuclear-encoded proteins are necessary (*yellow*). For OXPHOS (complex I, III, IV and V, *blue*) and the mitochondrial translation machinery (*green*) both nuclear-encoded and mtDNA-encoded genes are required. *Reprinted from Biochimica et Biophysica Acta (BBA) – Gene Regulatory Mechanisms, 1819, Bar-Yaacov et al., Mitochondrial-nuclear co-evolution and its effects on OXPHOS activity and regulation, pp. 1107-1111, © 2012, with permission of Elsevier.*

Several transcription factors and coactivators are involved in coordinating nuclear-mitochondrial interactions. The peroxisome-proliferator-activated receptor coactivator-1α (PGC-1α) can be considered the key regulator/coordinator of mitochondrial biogenesis. External stimuli such as energy deprivation during exercise, fasting or cold have been found to increase PGC1α transcription in the nucleus through upstream cascades and transcription factors as for instance the cAMP response element-binding protein (CREB). PGC1α activates the key transcription factors (e.g. ERRα, NRF1, NRF2, MEF-2 and PPARα) which in turn enhance transcription of nuclear genes responsible for several mitochondrial functions. For instance, the nuclear respiratory factor 1 (NRF1) activates TFAM which regulates mtDNA

transcription and regulation (Fig. 1.11, Ryan and Hoogenraad 2007, Scarpulla 2008). As mentioned above, nuclear-encoded mitochondrial genes are enriched for sequence motifs such as transcription factor binding sites (Pagliarini et al. 2008). Those genes belonging to one pathway in mitochondria (e.g. OXPHOS or β-oxidation) share recognition elements of a few transcriptions factors. This might enable a collective gene expression of all genes of the respective pathway (Moyes et al. 1998). Interestingly, sequence recognition sites of some transcription factors such as metallothionein 3 (Mt3) or Mt4 have been found in nuclear-encoded OXPHOS genes and in the D-loop. These might be involved in the coordinated expression of both nuclear- and mtDNA encoded OXPHOS genes (Suzuki et al. 1991).

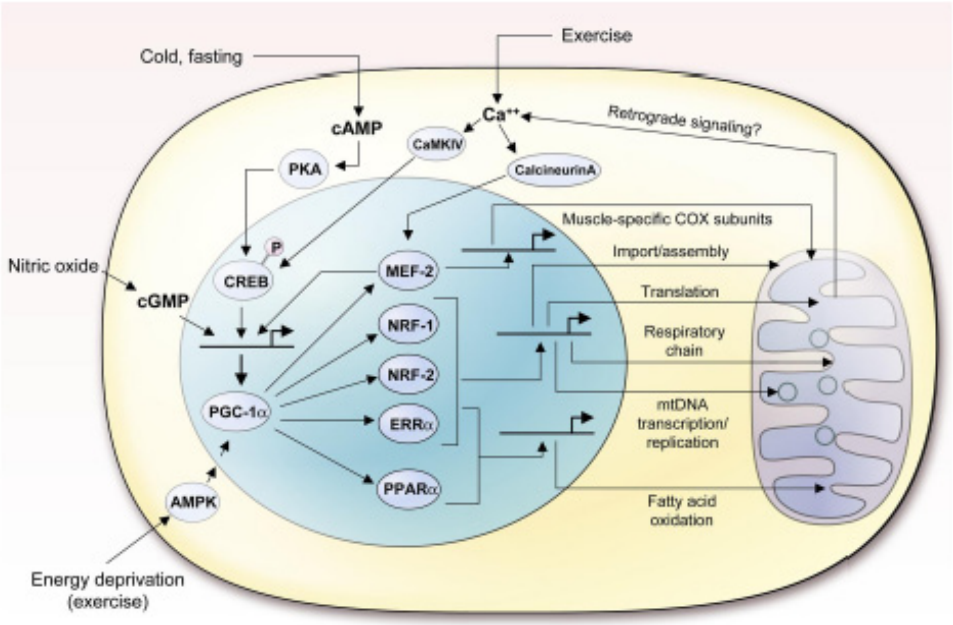

Figure 1.11 Nuclear-mitochondrial communication

Nuclear activity of peroxisome-proliferator-activated receptor coactivator-1α (PGC-1α) is a central regulator or coordinator of mitochondrial biogenesis, which is activated by different stimuli via different cascades.

AMPK, 5' adenosine monophosphate-activated protein kinase; Ca, calcium; CaMKIV, calcium/calmodulin-dependent protein kinase type IV; cAMP, cyclic adenosine monophosphate; cGMP, cyclic guanosine monophosphate; COX, cytochrom-c-oxidase; CREB, cAMP response element-binding protein; ERRα, estrogene related receptor α; MEF-2, myocyte enhancer factor-2; mtDNA, mitochondrial DNA; NRF1, nuclear respiratory factor 1; NRF2, nuclear respiratory factor 2; PKA, protein kinase A; PPARγ, peroxisome-proliferator-activated receptor γ.

Figure from Scarpulla (2008).

1.5 Mitochondrial/mtDNA alterations and body weight associated phenotypes

1.5.1 Variation in mitochondrial genes

1.5.1.1 Common polymorphisms in mtDNA

Pertaining to common polymorphisms, the variant allele m.16189C of the mtDNA D-loop polymorphism m.16189T/C accompanied by an uninterrupted poly-C tract at m.16184 to m.16193 was found to be nominally associated with leanness in 161 Australian mothers and their 20-year-old offspring (Parker et al. 2005). Saxena et al. (2006) did not find association between common SNPs of the mtDNA coding region and BMI in adults of European descent. The study was primarily focused on investigating T2DM in association with mtDNA SNPs using a cohort of each 3,304 diabetic and non-diabetic adults. In both groups, a mean BMI above 25 kg/m^2 was present. The subsample used for association testing of BMI was not described in more detail. Hence, the power of this study regarding BMI was elusive.

Although first GWAS on BMI or obesity were published in 2007 (e.g. Frayling et al. 2007, Hinney et al. 2007), mtDNA SNPs have not been investigated until recently in association with body weight and/or related traits. In 2011, Yang et al. described the mitochondrial haplogroup X to be associated with a lower BMI in a sample of 2,286 adult unrelated Caucasians (Yang et al. 2011b). However, the study lacked confirmation in an independent sample. Grant et al. (2012) performed a GWAS on both European-American and African-American case-control samples of obese and lean children, and did not find any mtDNA variant or heteroplasmy being associated with increased BMI.

1.5.1.2 Mitochondrial diseases

Infrequent mutations in mtDNA with non-synonymous amino acid exchanges or tRNA nucleotide exchanges are predominantly associated with severe phenotypes like blindness, muscle weakness or movement disorders (Li et al. 2012), and the function of tissues with high energy demand (nervous system and skeletal muscle) are most pronouncedly affected (Chial and Craig 2008). Only a few reports referred to the body weight of the patients. For instance, patients with mitochondrial encephalopathy lactic acidosis and stroke-like syndrome (MELAS), which is caused by a transition in $tRNA^{Leu(UUR)}$ (m.3243A/G), are characterized by a lean to normal weight phenotype (Suomalainen et al. 2011). This transition was also associated with the cyclic vomiting syndrome (Salpietro et al. 2003). Horváth et al. (2009) reported of a very lean girl (BMI < 1^{st} percentile) with clinical symptoms

of mitochondrial neurogastrointestinal encephalopathy (MNGIE) syndrome-like gastrointestinal dysmotility and cachexia. Comparatively, MNGIE patients harboring mutations in the thymidine phosphorylase (*TYMP*) gene and being characterized by abnormal mitochondria and OXPHOS defects are very lean as well (Suomalainen et al. 2011). By contrast, patients with mitochondrial recessive ataxia syndrome (MIRAS) were found to be obese (Suomalainen et al. 2011). These patients harbor mutations in the polymerase γ (*POLG1*). Thus, mtDNA replication might lead to multiple deletions (Hamosh et al. 2005).

1.5.2 Alterations in mitochondrial gene expression and function in obese individuals

As cellular power plants, mitochondria are key organs of energy metabolism, and thus, the assumption of a relationship between altered body weight and disturbed mitochondrial function might suggest itself. Although there is a body of publications on alterations of mitochondrial function and T2DM or insulin resistance, the number of studies focusing only on obesity in combination with mitochondria is rather small.

In adult obese individuals, smaller mitochondria were found in skeletal muscle compared with lean individuals (Kelley et al. 2002). Moreover, these individuals were characterized by a reduced skeletal muscle complex I activity. Niemann et al. (2011) investigated the right atrial cardiomyocytes of 60 male cardiac surgery patients without pre-existing T2DM, which were assigned to four groups of each 15 patients (young/normal weight, young/obese, old/normal weight and old/obese). The division by age occurred as mitochondria function diminishes by age. Young patients were <55 years, while old were >70 years, and body weight status was defined by BMI (Table 1.1). The authors found that mRNA expression of NRF1 and TFAM as well as mRNA expression of ND6, a mtDNA encoded subunit of complex I, were significantly reduced in old patients (both weight statuses) and young obese patients. In addition, a nuclear-encoded protein of complex I – NADH dehydrogenase (ubiquinone) 1 beta subcomplex 8 (NDUFB8) – was slightly reduced in both groups of old patients and young obese. Comparatively to Kelley et al. (2002), complex I activity was reduced in the obese as well as in normal weight old individuals (Niemann et al. 2011).

2 Objectives

Due to the central role of mitochondrial in energy metabolism as the cellular power plants, the overall objective of the present thesis was to assess whether variation in mitochondrial genes are associated with obesity. Within this scope, both variation in mtDNA and nuclear-encoded mitochondrial genes were addressed.

2.1 Analysis of mtDNA

Based on the empirical observation that correlations in BMI between mothers and their offspring were higher compared with fathers and their offspring and the greater correlations in BMI between maternal half-brothers than paternal ones (Hebebrand et al. 2001a, Magnusson and Rasmussen 2002), one specific objective was to assess whether variations in the exclusively maternally inherited mtDNA contributes to this parental effect.

To address this question, two basic approaches were performed:

1) an association study of common mtDNA SNPs and haplogroups in two case-control samples for obesity in children and adolescents (discovery) and adults (confirmation) with subsequent re-sequencing of complete mtDNA of ten individuals of predominantly haplogroup W; this haplogroup was nominally over-represented in the controls of the discovery and thus potentially protective against increased body weight, and the re-sequencing was done to detect functionally relevant variants with impact on body weight explaining its initial association
2) re-sequencing of mtDNA D-loop in (extremely) obese children and adolescents (cases) and lean adult controls to detect further variations potentially associated with obesity

The experiments have been started in August 2010. At that time, a well powered GWAS on mtDNA SNPs pertaining to obesity or BMI had not been published.

2.2 Analysis of nuclear-encoded mitochondrial genes

Regarding to the latest GWAS meta-analysis of 250,000 individuals newly identified and confirmed loci only explain 1.5 % of the BMI variance which is in large discrepancy to the empirical estimated heritability of 40 % to 70 %. Due to their polygenic nature and small effect sizes, a three-fold increase in sample size would not lead to a substantial increase in explained variance (Hebebrand et al. 2010, Speliotes et al. 2010). Pathway-based approaches or gene set enrichment analysis (GSEA) might reveal further insight into the

genetics of obesity. Based on the observations of a reduced mitochondrial function in obese individuals (Kelley et al. 2002, Niemann et al. 2011), one specific objective was to assess whether nuclear-encoded mitochondrial genes are enriched for modest association signals that collectively have an impact on mitochondrial function and potentially influence body weight.

To address this question, a GSEA was performed using three gene sets of nuclear encoded mitochondrial genes. The three gene sets – (1) 16 nuclear regulators of mitochondrial genes, (2) 91 OXPHOS genes, and (3) 966 nuclear-encoded human mitochondrial genes listed in the MitoCarta compendium (Pagliarini et al. 2008) – were from Segrè et al. (2010) who used these in meta-analysis GSEA for T2DM and related traits.

3 Subjects and Methods

3.1 Subjects

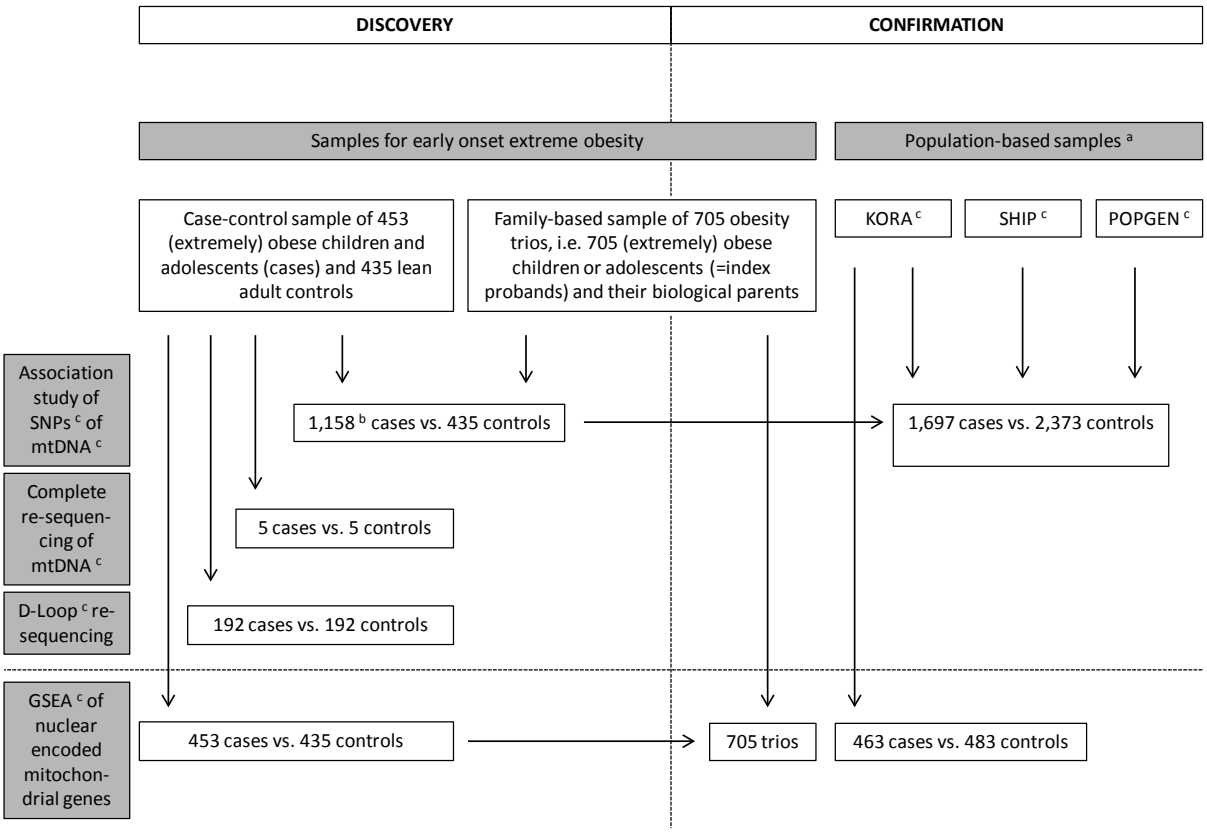

Figure 3.1 Study samples and purpose in the present analysis

[a] only individuals with a BMI > 30 kg/m^2 (cases) and < 25 kg/m^2 (controls) were included into the analyses

[b] 453 and 705 (extremely) obese children and adolescents from the case-control and family-based sample, respectively

[c] D-loop, displacement loop; GSEA, gene set enrichment analysis; mtDNA, mitochondrial DNA; KORA, Kooperative Gesundheitsförderung in der Region Augsburg; POPGEN, population genetic research project of the national genome research network; SHIP, The Study of Health in Pomerania; SNPs, single nucleotide polymorphisms

Two types of samples for early onset extreme obesity, i.e. a case-control (CC) and a family-based sample ("trios"), as well as three population-based adult samples (KORA, SHIP and POPGEN) were used for the present analyses (Fig. 3.1; Table 3.1). All samples had been recruited for various health related purposes including genetic association studies (Scherag et al. 2010, Rückert et al. 2011, Völzke et al. 2006, Nöthlings and Krawczak 2012). Thus, for the present analysis, both genotype and phenotype data were already present.

All participants (in case of minors their parents) gave written informed consent. The studies were approved by the Ethics Committees of the Universities of Marburg, Essen,

Greifswald and Kiel, as well as the Bavarian Medical Association, and conducted in accordance with the Declaration of Helsinki.

3.1.1 Samples for early onset extreme obesity

3.1.1.1 Case-control sample

The case-control (CC) sample consisted of 453 (extremely) obese children and adolescents and 435 lean or normal weight adult controls. The mean body mass index (BMI) of the cases (mean age: 14.37 ± 3.75 years) was 33.15 ± 6.68 kg/m^2 (Table 3.1). 84.4 % of the cases were extremely obese (BMI ≥ 99 th percentile) using age- and sex-specific percentile criteria for the German population from the National Nutrition Survey I (Hebebrand et al. 1994). The recruitment of the cases occurred in hospitals which are specialized for the treatment of extreme obesity in children and adolescents (Murnau, Berchtesgaden, Gießen, Ulm and Marburg). The lean or normal weight healthy adult controls (mean age: 26.08 ± 5.75 years) had a mean BMI of 18.31 ± 1.11 kg/m^2 (Table 3.1). These individuals were recruited at the University of Marburg and interviewed regarding their course of body weight development. At the age of 15 years, i.e. a similar age to that of the cases, 78 % of the lean and normal weight controls reported that their body weight was below the average body weight (Hinney et al. 2007). The CC sample was taken as the discovery sample for the gene set enrichment analysis (GSEA, Fig. 3.1). To increase the power for association testing of the mtDNA SNPs, the sample was enlarged by 705 (extremely) obese children and adolescents cases (i.e., the index probands from the family-based trios, cf. 3.1.1.2). The mean BMI of the 1158 cases was 32.45 ± 6.19 kg/m^2 (Table 3.1).

The complete mtDNA was re-sequenced from five lean (all of haplogroup W) and five obese individuals (three of haplogroup W, each one of haplogroup H and HV) from the CC sample (Fig. 3.1). In addition, a case-control sample of each 192 cases and controls was used for variant detection in the mitochondrial D-loop (Fig. 3.1). Most of these individuals (except for 14 cases and six controls) were from the CC sample. Both mean BMI and mean age was similar to the full CC sample (Table 3.1).

3.1.1.2 Family-based sample ('Trios')

The family-based sample comprised 705 obesity trios, each consisting of one (extremely) obese child or adolescent (index proband) and both biological parents. Similar to the cases of the CC sample, the families were recruited in hospitals specialized for the inpatient treatment of extreme obesity in children and adolescents (Bad Orb, Murnau, Wien, Ulm,

Gießen, Berchtesgaden, and Marburg). The mean age and BMI of the index probands was 13.44 ± 3.01 years and 32.02 ± 5.82 kg/m^2, respectively (Table 3.1). According to reference data from the German National Nutrition Survey I (Hebebrand et al. 1994), 83.8 % of the index probands were extremely obese (BMI ≥ 99th percentile). The parents (mean age: 42.54 ± 6.02 years) had a mean BMI of 30.28 ± 6.33 kg/m^2 (Table 3.1). The index probands were independent of the cases of the CC sample. The family-based sample was used as a confirmation sample for the GSEA (Fig. 3.1).

3.1.2 Population-based samples

3.1.2.1 KORA

KORA is an epidemiological study group of the region of Augsburg (Kooperative Gesundheitsförderung in der Region Augsburg) which includes about 18,000 individuals between 25 and 74 years of age at recruitment. The 1,743 adult participants (53 % females), of which genome-wide SNP data on the Affymetrix Genome-Wide Human SNP Array 6.0 are available, are a sub-sample of the total KORA follow-up 4 sample (Rückert et al. 2011, Table 3.1). 463 of these individuals categorized as obese cases (mean BMI: 33.55 ± 3.66 kg/m^2, mean age: 55.84 ± 8.54 years) and 483 individuals as normal weight or lean controls were included in the present study (mean BMI: 22.84 ± 1.48 kg/m^2, mean age: 50.55 ± 8.64 years, Table 3.1).

3.1.2.2 SHIP

SHIP (The Study of Health in Pomerania) is a cross-sectional population-based health survey in Northeast Germany comprising 4,310 individuals between 20 and 79 years of age at recruitment (Völzke et al. 2006). For 4,073 of these individuals (51 % females, Table 3.1) GWAS data (Affymetrix Genome-Wide Human SNP Array 6.0) are available. 1,045 were obese (mean BMI: 33.56 ± 3.30 kg/m^2, mean age: 54.76 ± 14.28 years) and 1,379 were normal weight and lean individuals (mean BMI: 26.26 ± 4.20 kg/m^2, mean age: 42.43 ± 16.08 years) and included in the present analysis (Table 3.1).

3.1.2.3 POPGEN

POPGEN is a population-based genetic research project founded in 2003 at the University Clinic of Schleswig-Holstein for the research of genetic risk factors for complex diseases (Nöthlings and Krawczak 2012). Specific patient groups as well as a random population-based sample of 1,317 individuals (19 to 77 years of age, 45 % females) were recruited. For

1,198 individuals of the population-based sample GWAS data (Affymetrix Genome-Wide Human SNP Array 6.0) were available. Amang these, 738 subjects (46 % females) were recruited via the local population registry and 460 (41 % females) as blood donors. BMI of the individuals recruited via the local population registry was estimated by self-report, while the BMI of the blood donors was measured. The cases and controls of the sample were used for confirmation of mtDNA SNP analysis (Fig. 3.1). Because of the different modes of recruitment and determination of BMI, statistical analysis was done stratified for both POPGEN sub-samples (Table 3.1).

A large case-control sample was created from all three population-based samples which consisted of 1,697 cases (mean BMI: 33.57 ± 3.47 kg/m^2, mean age: 55.08 ± 12.87 years) and 2,373 controls (mean BMI: 22.57 ± 1.69 kg/m^2, mean age: 46.15 ± 15.37 years, Table 3.1). This sample was used for confirmation of mtDNA SNP analysis. In addition, KORA as a case-control sample was used for confirmation of GSEA (Fig. 3.1).

Table 3.1 Phenotypical characteristics of subjects

Sample	**Description**	**Status**	**n total** {% female}	**Age [years]** {female} (male)	**BMI [kg/m^2]** {female} (male)	**BMI SDS** [a] {female} (male)
Samples for early onset extreme obesity				**mean ± SD**	**mean ± SD**	**mean ± SD**
Case-control (CC) sample	(extremely) obese children and adolescents	cases	453 {58.00}	14.37 ± 3.75 {14.50 ± 3.67} (14.18 ± 3.85)	33.15 ± 6.68 {33.18 ± 6.84} (33.11 ± 6.48)	4.55 ± 2.16 {4.66 ± 2.18} (4.40 ± 2.12)
	Lean or normal weight subjects	controls	435 {61.00}	26.08 ± 5.75 {26.54 ± 6.37} (25.38 ± 4.57)	18.31 ± 1.11 {17.58 ± 0.95} (18.86 ± 0.94)	-1.45 ± 0.34 {-1.35 ± 0.30} (-1.60 ± 0.36)
Reduced CC sample (D-loop sample) [b]	(extremely) obese children and adolescents	cases	192 {52.60}	13.87 ± 3.05 {13.79 ± 3.11} (13.95 ± 2.99)	33.00 ± 7.24 {33.37 ± 8.20} (32.59 ± 6.01)	4.49 ± 2.38 {4.75 ± 2.67} (4.19 ± 1.98)
	Lean or normal weight subjects	controls	192 {45.31}	25.50 ± 3.95 {24.93 ± 3.82} (25.97 ± 4.01)	18.37 ± 1.09 {17.61 ± 0.71} (19.01 ± 0.93)	-1.46 ± 0.33 {-1.31 ± 0.24} (-1.59 ± 0.34)
Family-based sample (Trios)	(extremely) obese children and adolescents	cases	705 {54.89}	13.44 ± 3.01 {13.54 ± 3.04} (13.31 ± 2.98)	32.02 ± 5.82 {32.36 ± 6.04} (31.60 ± 5.51)	4.23 ± 1.96 {4.50 ± 2.03} (3.91 ± 1.81)
	parents	parents	1410 {50.00}	42.54 ± 6.02 {40.89 ± 5.44} (44.21 ± 6.12)	30.28 ± 6.33 {30.23 ± 7.12} (30.34 ± 5.42)	1.65 ± 1.84 {1.75 ± 1.88} (1.55 ± 1.78)
Enlargement of CC by 705 cases from Trios	(extremely) obese children and adolescents	cases	1158 {55.81}	13.79 ± 3.35 {13.91 ± 3.34} (13.65 ± 3.36)	32.45 ± 6.19 {32.66 ± 6.39} (32.18 ± 5.93)	4.35 ± 2.04 {4.55 ± 2.10} (4.10 ± 1.95)

Table 3.1 is continued on the next page

Table 3.1 Phenotypical characteristics of subjects – *continued*

Sample	Description	Status	n total {% female}	Age [years] {female} (male)	BMI [kg/m^2] {female} (male)	BMI SDS [a] {female} (male)
Population-based samples				**mean ± SD**	**mean ± SD**	**mean ± SD**
KORA	population-based	population-based	1743 {51.06}	53.87 ± 8.86 {53.63 ± 8.80} (54.12 ± 8.91)	27.75 ± 4.56 {27.51 ± 5.08} (28.00 ± 3.94)	0.68 ± 1.28 {0.71 ± 1.30} (0.65 ± 1.27)
As case-control sample	obese subjects (BMI ≥ 30)	cases	463 {53.13}	55.84 ± 8.54 {56.46 ± 8.36} (55.14 ± 8.70)	33.55 ± 3.66 {34.04 ± 3.75} (33.00 ± 3.48)	2.31 ± 1.10 {2.37 ± 1.00} (2.24 ± 1.19)
	normal weight subjects (BMI < 25)	controls	483 {66.25}	50.55 ± 8.64 {49.83 ± 8.17} (51.96 ± 9.38)	22.84 ± 1.48 {22.63 ± 1.51} (23.26 ± 1.32)	-0.61 ± 0.45 {-0.49 ± 0.41} (-0.83 ± 0.42)
SHIP	population-based	population-based	4073 {50.75}	49.73 ± 16.27 {48.6 ± 16.04} (50.89 ± 16.44)	27.31 ± 4.77 {26.95 ± 5.35} (27.69 ± 4.04)	0.75 ± 1.33 {0.76 ± 1.36} (0.75 ± 1.29)
As case-control sample	obese subjects (BMI ≥ 30)	cases	1045 {51.39}	54.76 ± 14.28 {54.78 ± 14.31} (54.73 ± 14.27)	33.56 ± 3.3 {34.15 ± 3.6} (32.94 ± 2.82)	2.47 ± 1.03 {2.55 ± 1.05} (2.39 ± 1.00)
	normal weight subjects (BMI < 25)	controls	1379 {62.73}	42.43 ± 16.08 {41.49 ± 15.06} (44.01 ± 17.56)	22.43 ± 1.76 {22.14 ± 1.82} (22.93 ± 1.53)	-0.46 ± 0.56 {-0.35 ± 0.52} (-0.63 ± 0.57)
POPGEN (from local registry)	population-based	population-based	738 {45.53}	61.06 ± 10.57 {61.76 ± 10.7} (60.48 ± 10.44)	26.26 ± 4.2 {25.85 ± 4.64} (26.6 ± 3.76)	0.21 ± 1.22 {0.19 ± 1.19} (0.22 ± 1.24)
As case-control sample	obese subjects (BMI ≥ 30)	cases	111 {47.75}	61.5 ± 10.49 {63.06 ± 9.52} (60.07 ± 11.19)	33.46 ± 3.81 {33.93 ± 3.96} (33.02 ± 3.65)	2.26 ± 1.18 {2.25 ± 1.02} (2.28 ± 1.32)
	normal weight subjects (BMI < 25)	controls	306 {52.61}	59.9 ± 11.76 {59.53 ± 12.14} (60.32 ± 11.36)	22.67 ± 1.69 {22.28 ± 1.79} (23.11 ± 1.45)	-0.79 ± 0.54 {-0.7 ± 0.55} (-0.89 ± 0.52)
POPGEN (blood donors)	population-based	population-based	460 {40.87}	43 ± 12.66 {40.02 ± 11.82} (45.07 ± 12.82)	26.23 ± 4.53 {25.84 ± 5.05} (26.51 ± 4.11)	0.51 ± 1.42 {0.66 ± 1.43} (0.4 ± 1.4)
As case-control sample	obese subjects (BMI ≥ 30)	cases	78 {42.31}	45.22 ± 11.34 {41.73 ± 9.71} (47.78 ± 11.85)	33.86 ± 4.09 {34.74 ± 3.84} (33.21 ± 4.18)	2.75 ± 1.58 {2.99 ± 1.32} (2.57 ± 1.74)
	normal weight subjects (BMI < 25)	controls	205 {47.32}	39.89 ± 12.44 {39.07 ± 12.48} (40.62 ± 12.41)	22.64 ± 1.56 {22.21 ± 1.59} (23.03 ± 1.43)	-0.48 ± 0.57 {-0.3 ± 0.56} (-0.64 ± 0.53)
All cases and all controls from the population-based samples	obese subjects (BMI ≥ 30)	cases	1697 {51.21}	55.08 ± 12.87 {55.27 ± 12.89} (54.87 ± 12.86)	33.57 ± 3.47 {34.13 ± 3.67} (32.98 ± 3.15)	2.42 ± 1.10 {2.49 ± 1.06} (2.36 ± 1.14)
	normal weight subjects (BMI < 25)	controls	2373 {60.81}	46.15 ± 15.37 {45.23 ± 14.71} (47.57 ± 16.26)	22.57 ± 1.69 {22.27 ± 1.75} (23.03 ± 1.47)	-0.53 ± 0.55 {-0.42 ± 0.52} (-0.71 ± 0.54)

[a] BMI SDS calculation based on reference data of the National Nutrition Survey I (Hebebrand et. al. 1994).[b] All individuals except for 14 cases and six controls are from CC sample, as analyses of Affymetrix Genome-Wide Human SNP Array 6.0 failed for these 20 individuals.

3.2 Chemicals and buffers

Table 3.2 Used chemicals in alphabetical order

Chemicals	Producer	Method
Bromophenol blue	Merck KGaG, Darmstadt, Germany	Gel electrophoresis
DirectLoad™ Wide Range DNA Marker (50-10,000 bp)	Sigma-Aldrich Chemie GmbH, Munich, Germany	Gel electrophoresis
dNTPs (100 mM)	Sigma-Aldrich Chemie GmbH, Munich, Germany	PCR
EDTA	Karl Roth GmbH, Karlsruhe, Germany	Gel electrophoresis
Ethidium bromide	Karl Roth GmbH, Karlsruhe, Germany	Gel electrophoresis
Expand Long Template PCR System	Roche Diagnostics Deutschland GmbH, Roche Applied Science, Mannheim, Germany	PCR
Ficoll Type 400	Sigma-Aldrich Chemie GmbH, Munich, Germany	Gel electrophoresis
Glacial acetic acid	Karl Roth GmbH, Karlsruhe, Germany	Gel electrophoresis
HCl	Karl Roth GmbH, Karlsruhe, Germany	Gel electrophoresis
LongAmp® *Taq* PCR Kit	New England BioLabs® Inc., Ipswich, MA, USA	PCR
$MgCl_2$ (25 mM)	Sigma-Aldrich Chemie GmbH, Munich, Germany	PCR
NaOH	Karl Roth GmbH, Karlsruhe, Germany	Gel electrophoresis
Oligonucleotides (Primers)	Sigma-Aldrich Chemie GmbH, Munich, Germany	PCR
PCR buffer 10x	Sigma-Aldrich Chemie GmbH, Munich, Germany	PCR
peqGOLD Universal Agaroses	PEQLAB Biotechnologie GMBH, Erlangen, Germany	Gel electrophoresis
Taq DNA Polymerase	Sigma-Aldrich Chemie GmbH, Munich, Germany	PCR
Tris	Karl Roth GmbH, Karlsruhe, Germany	Gel electrophoresis
Tris-HCl	Karl Roth GmbH, Karlsruhe, Germany	Gel electrophoresis
Xylene cyanol	Karl Roth GmbH, Karlsruhe, Germany	Gel electrophoresis

For 50 x TAE buffer production, Tris, Tris HCl and EDTA were diluted in 500 mL H_2O (bidest.). Glacial acetic acid was added, and H_2O (bidest.) filled-up to 1 L (Table 3.3). For production of the loading buffer, all ingredients (Table 3.4) were diluted in 50 mL H_2O (bidest.), and H_2O (bidest.) was filled-up to 100 mL. Afterwards, pH was adjusted to 8.3 with NaOH or HCl in both buffers.

Table 3.3 Components of 50 x TAE buffer for gel electrophoresis

	Final volume 1 L	Final concentration [M]
Tris	230 g	2.0
Tris HCl	15 g	
glacial acetic acid	57.1 mL	2.0
EDTA	100 mL of 0.5 M EDTA	0.05

Table 3.4 PCR product loading buffer for agaroses gels

	Final volume 100 mL	Final concentration [mM]
Bromophenol blue	0.25 g	3.6
Xylene cyanol	0.25 g	4.6
Ficoll 400	15 g	0.4
EDTA	4.46 g	12

3.3 Molecular genetic analyses

3.3.1 Genotyping

All individuals (except for 14 cases and six controls whose D-loop was re-sequenced) were genotyped by the Affymetrix Genome-Wide Human SNP Array 6.0. This SNP array has genome-wide approximate 910,000 SNPs, of which approximately 870,000 SNPs are autosomal and 119 of mtDNA. Only individuals with a genome-wide SNP call rate (CR) ≥ 95 % were included.

The following quality control (QC) criteria were applied on the 119 SNP of the mtDNA for each study sample separately (Table 3.5): (1) the sample call-rate (CR) per SNP had to be ≥ 95 % and (2) the minor allele frequency (MAF) ≥ 1 %. Only ~40 SNPs passed these criteria due to the fact that most of the mitochondrial SNPs had a very low MAF or were even monomorphic (96 % of all excluded SNPs). As a third QC criterion (3), the cluster graphs of the 40 SNPs were checked by two independent raters. Only those SNPs showing a clear separation of both alleles were finally included in the analyses (CC: 40, Trios: 35, KORA: 37, SHIP: 32 and POPGEN: 35; Table 3.5).

For GSEA, all autosomal SNPs having passed QC were included. Four QC criteria were applied for the ~870,000 autosomal SNPs (Table 3.5): (1) sample CR per SNP had to be ≥ 95 %; (2) the MAF had to be ≥ 1 % in the CC sample and in KORA, and ≥ 5% in the set of all parents of the family-based sample; (3) the two-sided exact p-value of the test for Hardy-Weinberg-Equilibrium (HWE, Wigginton et al. 2005) in the whole KORA sample, in the parents of the family-based sample and in the controls of the case-controls GWAS sample, respectively, had to be ≥ 0.001, and (4) after setting all Mendelian inconsisten t calls to "missing" in the family-based sample, at least one major allele and one minor allele transmission at each SNP was claimed. 703,015 / 641,991 / 659,502 autosomal SNPs passed this QC in the CC / family-based / population-based sample, respectively (Table 3.5).

Genotype calling of the family-based trios initially was performed in eight batches of a varying number of individuals between n=26 and n=432, while for both the CC and the KORA sample one batch was used each. In the trios, gene set 1 was initially enriched for association signals above the 95th percentile (Table 8.1 in the appendix). Association results can be influenced by batch effects (Miclaus et al. 2010). In case the cluster graph of a SNP does not show clear allele separation, the SNP might be false-positively associated with the trait of investigation (Browning and Yu 2009). Therefore, I checked cluster graphs of 2,986

SNPs (i.e. all SNPs initially involved in GSEA of gene set 1 of CC, trios and KORA) for clear allele separation. In the trios, 14 % of all SNPs (nine of the 16 best SNPs were involved) did not show clear allele separation compared with only 2-3 % in CC sample and KORA (only one best SNP was involved in KORA). Consequently, genotype calling of the trios was repeated using one batch and GSEA was recalculated (Jarick, Scherag) leading to non-enrichment of gene set 1.

Table 3.5 Quality control of SNPs

Sample	**CC** **n=888**	**TRIOS** **n=2,115** **705 trios**	**KORA** **n=1,743**	**SHIP** **n=4,073**	**POPGEN** **n=1,198**
Mitochondrial SNPs					
number of mtDNA SNPs genotyped [a]	119	115	115	115	115
1st SNP-QC criterion	**sample call-rate per SNP ≥ 95 %**				
number of SNPs failing 1st SNP-QC criterion	3	6	6	2	11
% of SNPs failing 1st SNP-QC criterion	2.52	5.22	5.22	1.74	9.57
2nd SNP-QC criterion	in sample	in sample	**MAF ≥ 1 %** in sample	in sample	in sample
number of SNPs failing 2nd SNP-QC criterion	76	68	72	71	71
% of SNPs failing 2nd SNP-QC criterion	63.9	59.1	62.6	61.7	61.7
3rd SNP-QC criterion	**clear separation of the 2 alleles in the cluster graph of each SNP**				
number of SNPs failing 3rd SNP-QC criterion	1	7	7	17	7
% of SNPs failing 3rd SNP-QC criterion	0.84	6.09	6.09	14.8	6.09
number of SNPs left after SNP-QC [b]	40	35	37	32	35
% of SNPs left after SNP-QC	33.6	30.4	32.2	27.8	30.4
Autosomal SNPs					
number of autosomal SNPs genotyped [a]	869,224	868,257	868,278		
1st SNP-QC criterion	**sample call-rate per SNP ≥ 95 %**				
number of SNPs failing 1st SNP-QC criterion	33,616	36,507	79,032		
% of SNPs failing 1st SNP-QC criterion	3.87	4.20	9.10		
2nd SNP-QC criterion	**MAF ≥ 1 %** in sample	**MAF ≥ 5 %** in parents	**MAF ≥ 1 %** in sample		
number of SNPs failing 2nd SNP-QC criterion	134,847	192,243	120,498		
% of SNPs failing 2nd SNP-QC criterion	15.51	22.14	13.88		
3rd SNP-QC criterion	**two-tailed exact p-value in test for HWE ≥ 0.001** in controls	in parents	in sample		
number of SNPs failing 3rd SNP-QC criterion	4,563	13,050	40,538		
% of SNPs failing 3rd SNP-QC criterion	0.52	1.50	4.67		
4th SNP-QC criterion	none	**at least 1 minor and 1 major allele transmission**	none		
number of SNPs failing 4th SNP-QC criterion	0	67,682	0		
% of SNPs failing 4th SNP-QC criterion	0.00	7.80	0.00		
number of SNPs left after SNP-QC [b, c]	703,015	641,991	659,502		
% of SNPs left after SNP-QC	80.88	73.94	75.96		

[a] This number is differing due to the fact that each sample has been analyzed at different points of time and for each moment of analysis the latest Affymetrix annotation file has been used. [b] There are SNPs failing more than one SNP-QC criterion. [c] after setting all Mendelian inconsistent calls to "missing"

3.3.2 Variant detection by Sanger-re-sequencing

3.3.2.1 Complete mitochondrial DNA

Complete mtDNA of five lean (all haplogroup W) and five obese individuals (three haplogroup W, each one H and HV) was re-sequenced (Sanger re-sequencing) by *Seqlab, Sequence Laboratories, Göttingen*. The company provided the sequences of two primer pairs to amplify the complete mtDNA with two long range polymerase chain reaction (PCR) products of 10 kb each. For sequencing of the PCR products, 48 (forward and reverse) primers were selected. The distance between the starts of the forward and reverse primers was approximately 700 bp, respectively, in order to guarantee a bi-directional sequencing in a major part of the complete mtDNA (Fig. 3.2). The exact starting positions and sequences of the primers are kept a secret by *SeqLab*.

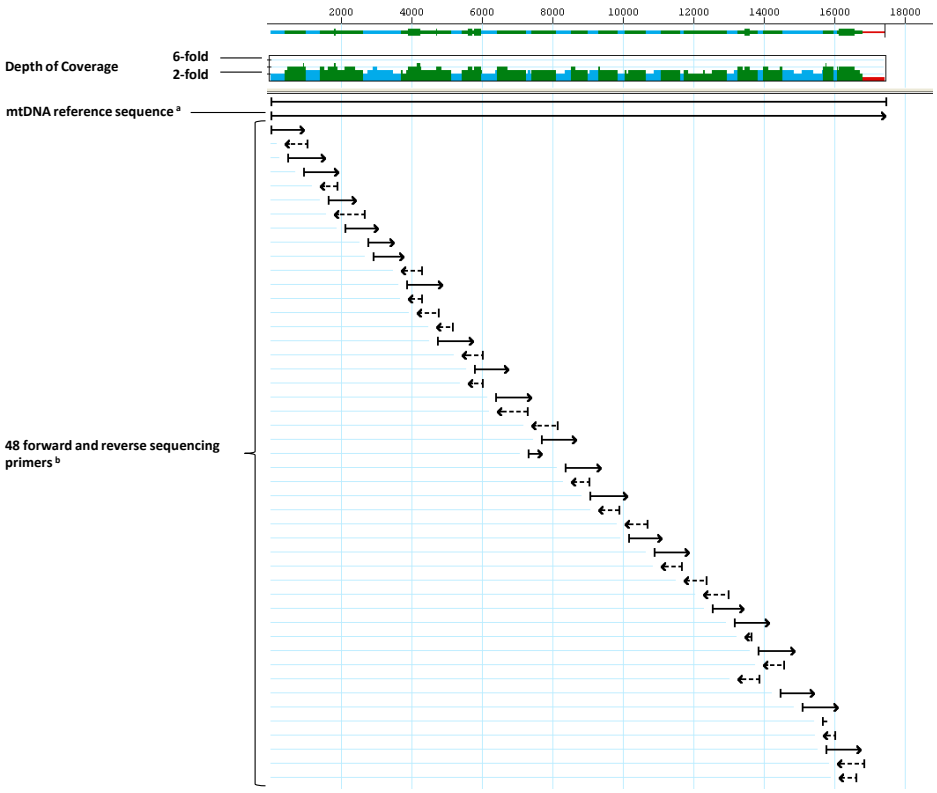

Figure 3.2 Coverage of sequencing primers for re-sequencing (Sanger) of complete mtDNA

This strategy view from the analysis tool *Seqman Pro (v.10.1.0 (174), 419, DNASTAR, Inc., Madison (WI), USA)* is of one individual whose mtDNA was completely re-sequenced The distance between the starts of the forward and reverse primers was approximately 700 bp, respectively, so that a bi-directionally sequencing of a major part of the complete mtDNA could be guaranteed. The exact starting positions and/or sequences of the primers are a secret of SeqLab.

[a] revised Cambridge Reverence Sequence (rCRS) enlarged by adding the last 129 nucleotides from the end of the rCRS to its start (3'-end) and the first 660 nucleotides to its 5'-end

[b] In the present example, only 47 primers are shown as one reverse primer failed to work. However, the remaining primers still covered the whole mtDNA.

Preparation of both PCR fragments for re-sequencing was performed in our laboratory. For both PCR fragments, the annealing temperature was established using a temperature gradient between 57 °C and 65 °C (fragment 1) and 58 °C and 68 °C (fragment 2; Fig. 3.3). The amplification of fragment 1 occurred with a kit from *Roche Diagnostics Deutschland GmbH* (Expand Long Template PCR System), and that of fragment 2 with the LongAmp® *Taq* PCR kit from *New England BioLabs® Inc*. These two commercial kits were necessary, as during the establishment, fragment 1 showed strong by-products with the mix from *New England BioLabs®*, and for fragment 2, the amplification was either weak or absent with the mix of *Roche*. PCR components and conditions were chosen according to the manual of the respective producer (Table 3.6 and 3.7). For visualization, 20 µL of the long range PCR product was loaded on a 0.8 % (w/v) agaroses gel containing ethydiumbromide (0.0025 % (v/v)). Primer annealing temperatures of 57 °C and 65 °C have been chosen for fragment 1 and 2, respectively (Fig. 3.3).

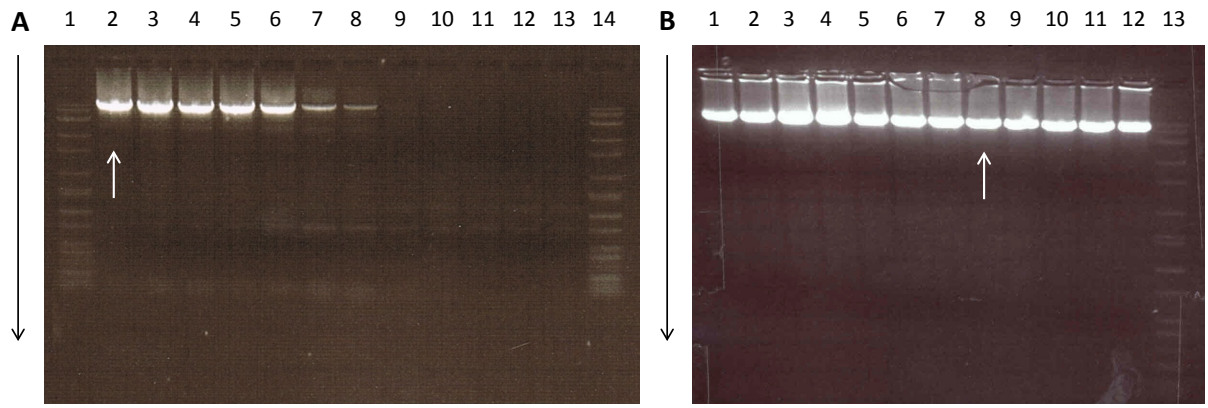

Figure 3.3 Establishment of annealing temperature for fragment 1 and fragment 2

A lane 1 and 14: DNA marker (50 to 10,000 bp, *DirectLoad™ Wide Range DNA Marker, Sigma-Aldrich*, lane 2-13: PCR product with annealing temperature gradient from 57 °C to 65 °C; 57 °C (lane 2) was chosen as annealing temperature

B lane 13: DNA marker (cf. **A**), lane 1-12: PCR product with annealing temperature gradient from 58 °C to 68 °C, 65 °C (lane 8) was chosen as annealing temperature

0.8 % (w/v) agaroses gel in TAE buffer, containing ethydiumbromide (0.0025 % (v/v)); 20 µL PCR product and 4 µL loading buffer; electrophoresis conditions: 4 h running time and 80 V; black arrows indicate direction of electrophoresis

Table 3.6 Components of reaction mix for long range PCR fragments

	Fragment 1		Fragment 2	
	Volume [µL]	**Final Conc.**	**Volume [µL]**	**Final Conc.**
Total volume	50		50	
DNA (100 ng/µL)	2	4 ng/µL	2	4 ng/µL
H_2O	38.55		21	
Reaction mix	5 [a]	2.75 mM $MgCl_2$	25 [c]	1 x
Forward primer (25 pmol/µL)	0.6	0.3 µM	1	0.5 µM
Reverse primer (25 pmol/µL)	0.6	0.3 µM	1	0.5 µM
dNTP (10 mM)	2.5	500 µM		
Taq [b]	0.75	3.75 U		

[a] Expand long template buffer 2, 10 x conc. with 27.5 mM $MgCl_2$ from *Roche Diagnostics Deutschland GmbH*

[b] Expand long template enzyme mix from *Roche Diagnostics Deutschland GmbH*

[c] LongAmp® *Taq* from *New England BioLabs® Inc.*

Table 3.7 Long range PCR conditions for the amplification of both mtDNA fragments

	Fragment 1 [a]			Fragment 2 [a]		
	Time [s]	**Temperature [°C]**		**Time [s]**	**Temperature [°C]**	
Denaturation	120	94		30	94	
Denaturation	10	94	10 x	30	94	30 x
Primer annealing	30	57		30	65	
Extension	480	68		540	65	
Denaturation	15	94	25 x			
Primer annealing	30	57				
Extension	480 (+20 s each new cycle)	68				
Final extension	420	68		600	65	
Storage	∞	4		∞	4	

[a] Conditions were chosen according to the producers' manuals (fragment 1: "Expand Long Template PCR System" from *Roche Diagnostics Deutschland GmbH*; fragment 2: LongAmp® *Taq* from *New England BioLabs® Inc.*)

Evaluation of electropherograms received by *Seqlab* was done with Seqman Pro (v.10.1.0 (174), 419, DNASTAR, Inc., Madison (WI), USA) and Microsoft® Office Excel® 2007 (Microsoft Coop., Redmond (WA)) in our laboratory. As reference, the rCRS (Anderson et al. 1981, Andrews et al. 1999) was copied from PubMed (www.ncbi.nlm.nih.gov/nuccore/251831106). This sequence was enlarged by adding the last 129 nucleotides from the end of the rCRS to its start (3'-end) and the first 660 nucleotides to its 5'-end. The enlargement was done due to the circularity of mtDNA. The enlarged reference was copied into an Excel® sheet (cf. Fig. 8.1 in appendix), so as to find 10 nucleotides per cell and 60 nucleotides per line, which facilitated the counting to find a

certain nucleotide position. All (up to 48) obtained electropherograms of one individual were aligned to the enlarged reference sequence using Seqman Pro (Fig. 3.2). The alignment was manually checked for deviations from the reference sequence, as an automatic search was not accurate enough to distinguish between real deviations and false positive ones due to a low sequence quality. All deviations were noted in the prepared Excel® sheet.

The sequencing primers of *SeqLab* were selected to obtain an at least two-fold coverage of a major part of the complete mtDNA (Fig. 3.3). However, several gaps or low sequence qualities were seen in the alignments during evaluation. A total of eight sequencing gaps were present in seven of the 10 individuals (Table 3.8). Hence, eight additional primer pairs were selected using primer3 (v.0.4.0; http://frodo.wi.mit.edu/). As parts of the mtDNA sequence are present in the nuclear DNA as well, mtDNA fragment 1 and 2 were taken as PCR template for re-sequencing. The re-sequencing (Sanger) of these gaps was performed by *LGC Genomics, Berlin*. Obtained electropherograms were added to the alignment of the respective individual and again manually checked for deviations as described above.

Table 3.8 Self-selected primers for re-sequencing for complete mtDNA

	Start of gap	End of gap	Sequences of primers	Length [bp]	No. of individuals re-sequenced
Fragment 1					
Gap 1	3541	3781	F: 5‘-TCTTCACCAAAGAGCCCCTA-3‘ R: 5‘-GGGTCATGATGGCAGGAGTA-3‘	377	2
Gap 2	7201	7981	F: 5‘-TCTTCCCACAACACTTTCTCG-3‘ R: 5‘-TTGTCAACGTCAAGGAGTCG-3‘	828	4
Gap 3	7981	8461	F: 5‘-CGACTACGGCGGACTAATCT-3‘ R: 5‘-TTTTATGGGCTTTGGTGAGG-3‘	577	2
Gap 4	8761	8941	F: 5‘-CAACACTAAAGGACGAACCTGA-3‘ R: 5‘-CTAGGGCTATTGGTTGAATG-3‘	287	4
Gap 5	9901	10141	F: 5‘-TCCGCCAACTAATATTTCACTTT-3‘ R: 5‘-GGGTGGATTTTTCTATGTAGCC-3‘	300	3
Fragment 2					
Gap 6	1681	1861	F: 5‘-TGACCGCTCTGAGCTAAACC-3‘ R: 5‘-TTGGCTCTCCTTGCAAAGTT-3‘	234	1
Gap 7	12721	13081	F: 5‘-CCCAAACATTAATCAGTTCTTCAA-3‘ R: 5‘-GGTGGAAGCGGATGAGTAAG-3‘	461	5
Gap 8	13081	13681	F: 5‘-ACCCCAGTCTCAGCCCTACT-3‘ R: 5‘-CAGGCGTTTAATGGGGTTTA-3‘	651	4

F, forward primer; R, reverse primer; primers were chosen with primer3 (v.0.4.0; http://frodo.wi.mit.edu/)

3.3.2.2 Mitochondrial displacement loop (D-loop)

For re-sequencing of the mtDNA D-Loop, primers were chosen according to Cardoso et al. (2012) who sequenced the D-loop of individuals of European origin (Fig. 3.4). However, instead of the heavy strand primer H616, which was used by Cardoso et al. (2012), a primer located out of the D-Loop was selected (H715) using primer3 (v.0.4.0; http://frodo.wi.mit.edu/). This extension was done to include the sequence of the heavy strand promoter 2 (HSP2), the start of which is located at m.644A in $tRNA^{Phe}$ (Zollo et al. 2012, Lodeiro et al. 2012). The D-Loop was amplified using the outer primers, light strand primer 15988 (L15988) and H715, yielding a PCR product of 1319 bp (Fig. 3.4). A primer annealing temperature of 65 °C was chosen (Fig. 3.5). The components and conditions of this PCR can be found in Table 3.9 and Table 3.10.

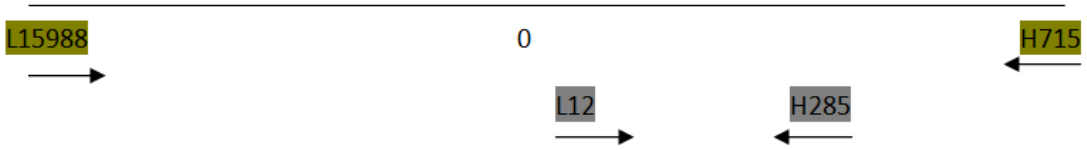

Figure 3.4 Selection of primers for the re-sequencing of the mtDNA D-Loop

L15988, 5'-AAGTCTTTAACTCCACCATTAGC-3'; L12, 5'-ACATCACGATGGATCACAGGTC-3' and H285, 5'-GGGGTTTGGTGGAAATTTTTTG-3' were chosen according to Cardoso et al. (2012), and H715, 5'-TGGAACGGGGATGCTTGCAT-3' was selected using primer3 (v.0.4.0; http://frodo.wi.mit.edu/), to include the heavy strand primer 2 (HSP2) starting position at m.644A (Zollo et al. 2012, Lodeiro et al. 2012) in $tRNA^{Phe}$.

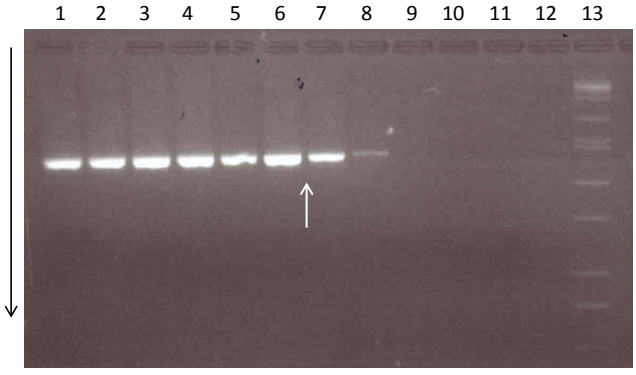

Figure 3.5 Establishment of annealing temperature for D-loop PCR fragment

lane 1 and 12: PCR product with annealing temperature gradient from 58 °C to 72 °C; 65 °C (lane 6-7) was chosen as annealing temperature; lane 13: DNA marker (50 to 10,000 bp, *DirectLoad™ Wide Range DNA Marker, Sigma-Aldrich*); 2.5 % (w/v) agarose gel in TAE buffer, containing ethydiumbromide (0.0025 % (v/v)); 20 µL PCR product and 4 µL loading buffer; electrophoresis conditions: 3 h running time and 120 V; black arrow indicates direction of electrophoresis

Table 3.9 Components of reactions mixes for the D-Loop PCR product

	Volume [µL]	Final Conc.
Total volume	25	
DNA (20 ng/µL)	1	0.8 ng/µL
H_2O	18.65	
PCR buffer 10x	2.5	1 x
$MgCl_2$ (25 mM)	2.0	2 mM
Forward primer (25 µM)	0.25	0.25 µM
Reverse primer (25 µM)	0.25	0.25 µM
dNTP (10 mM)	0.25	100 µM
***Taq* DNA polymerase**	0.1	

Table 3.10 PCR conditions for the amplification of the D-Loop

	Time [s]	Temperature [°C]	
Denaturation	300	95	
Denaturation	45	95	35 x
Primer annealing	60	65	
Extension	90	72	
Final extension	600	72	
Storage	∞	4	

The re-sequencing of the D-loop was done using four primers (L15988, L12, H285 and H715) and performed by *LGC Genomics, Berlin*. The four primers were necessary as (1) the length of the PCR product was longer than 1100 bp, which is the maximum length for Sanger re-sequencing according to *LGC Genomics, Berlin*, and (2) the D-loop contains several poly-C tracts (at m.16184 to m.16193, m.303 to m.315, and m.568 to m.573) at which length heteroplasmy might be induced at a certain length of uninterrupted Cs (Cardoso et al. 2012). Length heteroplasmy resulted in un-analyzable electropherogram peaks. If a length heteroplasmy was induced at m.16189 as an example, the re-sequencing with L15988 was stopped at this position, and the remaining re-sequencing of the direction of L15988 was done by L12 starting downstream of the heteroplasmic location.

As for the complete re-sequencing of mtDNA, evaluation of electropherograms was performed using *Seqman Pro* and *Excel*® (2007). The extended D-Loop sequence amplified for re-sequencing was copied from the rCRS into *Excel*® (ten nucleotides per cell and 60 per line, Fig. 8.1 in appendix). The four sequences of each individual were assembled with the extended D-loop from the rCRS using *Seqman Pro*. Again, all deviations from the reference were noted in the prepared Excel sheet. Afterwards, for each variation found, the number of lean and obese individuals harboring that variation was count. Evaluation of electropherograms, notation of variation and counting were performed by two independent raters to minimize the rate of mistakes.

3.4 Haplogroup determination

Haplogroups were determined in all individuals of the CC sample, the 705 index patients of the family-based trios and the population-based samples. This determination was performed using HaploGrep, a web application which is freely accessible at http://haplogrep.uibk.ac.at (Kloss-Brandstätter et al. 2011). HaploGrep applies the latest version of Phylotree (www.phylotree.org), an updated comprehensive phylogeny of the global human mtDNA variation based on variation in both coding and control region (van Oven and Kayser 2009). The haplogroup analysis of the discovery CC sample started in June 2011; at that time, the latest version was Phylotree build 11. Haplogroup determination of the confirmation sample was also performed using Phylotree build 11 for comparative results.

A text file from the available mtDNA SNPs having passed QC was created and imported into HaploGrep. From the up-loaded genotype data, HaploGrep determined each individual's most likely haplogroup. In addition, a rank or quality value in % was calculated to estimate the reliability of the determined haplogroup. Only those haplogroups with a quality value≥ 90 % were included in statistical analysis, as this threshold indicates a quite reliable haplogroup assignment according to the user manual of HaploGrep. All received haplogroups were assigned to major haplogroups (A, B, D, H, J, K, L, M, N, P, R, S, T, U, V, W, X and Z) for association testing (cf. Fig. 1.8). HV was assigned to H, and JT to J.

The haplogroup of the 10 individuals, of which complete mtDNA was re-sequenced, was re-determined using all detected variants as well as the information of the 40 SNPs found on the SNP array having passed QC, and only detected D-loop variants.

For reasons of comparability, haplogroups were re-determined using the detected D-loop variants from the 364 individuals of the CC sample, of which SNP array-based data were available.

3.5 Statistics

All statistical analyses were performed by Dipl.-Math. Ivonne Jarick from the Institute for Medical Biometry and Epidemiology of the Philipps-University of Marburg and PD Dr. André Scherag from the Institute for Medical Informatics, Biometry and Epidemiology of the University Clinic Essen.

3.5.1 Association tests

3.5.1.1 Common SNPs of mtDNA

In the discovery sample (Fig. 3.1, Table 3.1) Fisher's two-sided exact test was used for association testing of all single mtDNA SNPs and major haplogroups. The analyses were done in all individuals, as well as stratified by gender. A p-value below 0.05 was considered nominally significant. Odds ratios and 95 % confidence intervals were calculated.

For independent confirmation, all nominally significant SNPs and haplogroups were followed-up in the confirmation sample (Fig. 3.1, Table 3.1), and those SNPs and major haplogroups which were discovered in either females or males, were only followed-up in the respective sub-group.

3.5.1.2 Variants of mtDNA detected by D-loop re-sequencing

Most of the individuals (except for 14 cases and six controls) whose D-loop was re-sequenced (D-loop sample) were from the CC sample, which was part of the discovery sample. Hence, haplogroup data based on the 40 SNPs of the SNP array were present in the majority of individuals of the D-loop sample. In a first step, the haplogroup distribution between D-loop and discovery sample was compared using a χ^2 test. Only individuals with a HaploGrep's quality value ≥ 90 % were included in the analysis.

The frequencies of all detected variants were compared between cases and controls using Fisher's two sided exact test. Odds ratios as well as 95 % confidence intervals were calculated. In addition, the absolute number of all variants detected per individual was compared between cases and controls using a two-sided t-test.

Finally, the mean absolute number of variants per case was compared with the mean absolute number of variants per control in 23 functionally relevant locations of the D-loop – i.e. potential protein binding sites during transcription and replication of mtDNA (Fig. 1.7, Table 1.2) – using a two-sided t-test. These locations have been summarized in the mitomap data base (www.mitomap.org, *Last Edited: Aug 18, 2009*). A literature check was performed to confirm indicated start and end sites of all suggested locations, and to check whether

further locations potentially involved in the transcription and replication process of mtDNA exist. Only for HSP2, which was listed at mitomap at only m.645 (Montoya et al. 1982, 1983, Yoza et al. 1984), an extended area of m.632 to m.655 was selected. This was done as Lodeiro et al. (2012) randomized these 24 nucleotides around the transcription start recently mapped at m.644 (Zollo et al. 2012, Lodeiro et al. 2012) and did not detect transcription *in vitro*. Hence, these 24 nucleotides might be important HSP2 control elements (e.g. transcription factor binding sites).

For all comparisons between cases and controls, a p-value below 0.05 was considered nominally significant.

3.5.1.3 Autosomal SNPs

For GSEA, all autosomal SNPs having passed QC were included in the analysis. In the CC sample and in KORA analyzed as a CC (Fig. 3.1, Table 3.1), the Cochran-Armitage trend test was applied to each autosomal SNP for an additive mode of inheritance. By contrast, in the family-based trios, a TDT (Spielman et al. 1993) was calculated for each SNP using PLINK v1.07 (http://pngu.mgh.harvard.edu/purcell/plink/; Purcell et al. 2007) under the assumption of an additive allelic model of inheritance. All families with missing genotypes were excluded from TDT analysis of the respective SNP. Due to genotyping failures and/or Mendelian inconsistencies, for just 0.37 % of all SNPs more than 5 % of the trios were excluded from the TDT.

3.5.2 Gene set enrichment analysis (GSEA)

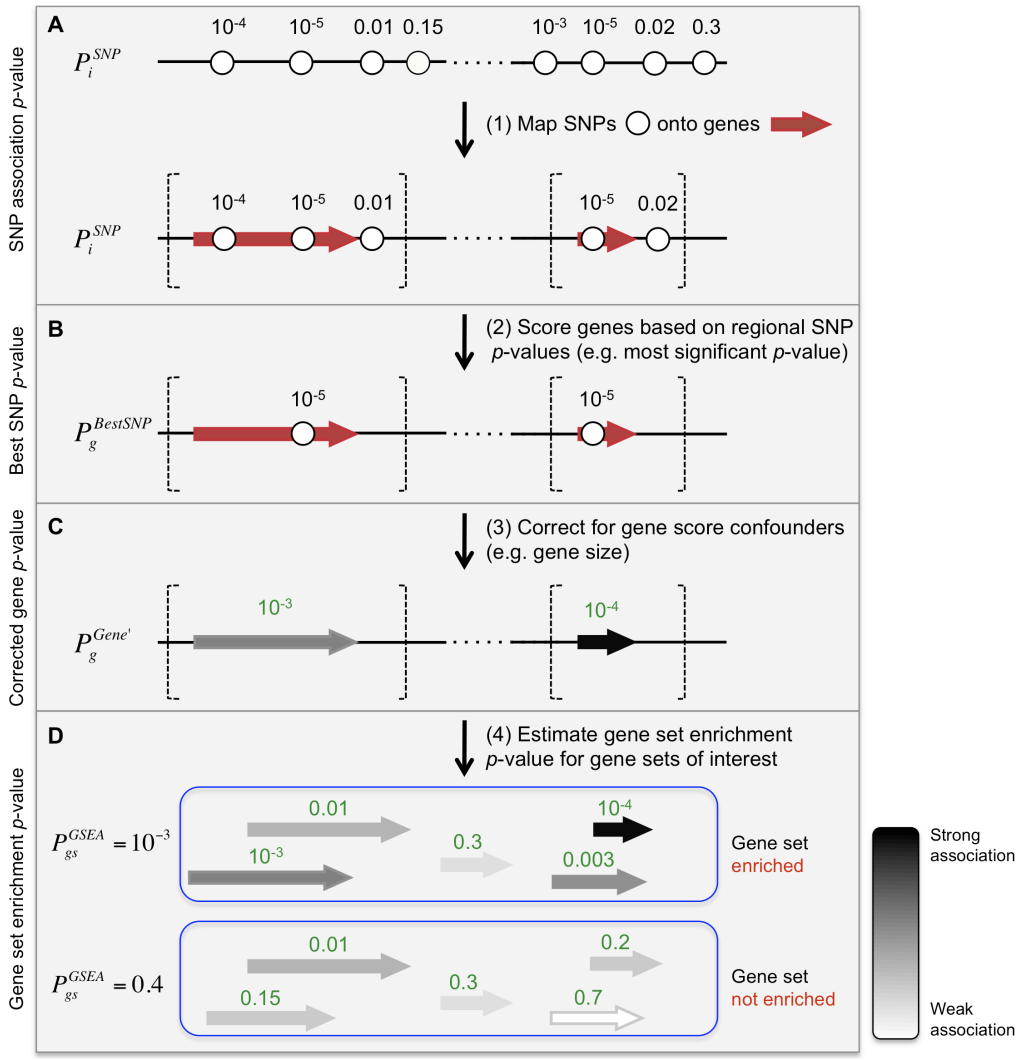

Figure 3.6 Description of GSEA procedure

In a GSEA, first of all, all analyzed autosomal SNPs of a study sample are mapped to their corresponding gene (**A**). Afterwards, a test statistic (e.g. p-value) that aggregates the SNP information is derived for each gene (**B**, **C**), and the distribution of test statistics between the gene set chosen pertaining to a certain biological function (*here:* mitochondrial genes) is compared with the gene set of all autosomal genes (**D**). *Figure from Segrè et al. (2010).*

3.5.2.1 Gene sets

In a large GWA meta-analysis, Segrè et al. 2010 tested three mitochondrial gene sets for enrichment of association signals of common variation in T2DM and related glycemic traits. These three gene sets were tested for obesity in the present PhD thesis. The first gene set consisted of 16 autosomal nuclear regulators of mitochondrial genes (Table 3.11), which were selected by Segrè et al. (2010) based on literature (Goffart and Wiesner 2003, Kelly and Scapulla 2004, Finck and Kelly 2006, Giguère 2008, Yu and Auwerx 2009 and Wan et al.

2010). A literature re-check at the beginning of the analysis (Dec. 2010) revealed that the most important regulators are included. In order to be comparable with the results of the analysis in a phenotype similar to obesity, the list of genes had been applied as described by Segrè et al. (2010).

Table 3.11 Nuclear-encoded regulators of mitochondrial genes (gene set 1)

Gene ID	Full name of gene	Chromosome
ESRRA	Estrogen related receptor alpha	11
ESRRG	Estrogen related receptor gamma	1
GABPA	GA-binding protein alpha subunit	21
GABPB1	GA-binding protein beta subunit 1	15
GABPB2	GA-binding protein beta subunit 2	1
MEF2A	Myocyte-specific enhancer factor 2A	15
MYC	Myelocytomatosis viral oncogene homolog (avian)	8
NRF1	Nuclear respiratory factor 1	7
NRIP1	Nuclear receptor-interacting protein 1	21
PPARA	Peroxisome proliferator-activated receptor alpha	22
PPARD	Peroxisome proliferator-activated receptor delta	6
PPARGC1A	Peroxisome proliferator-activated receptor gamma coactivator 1 alpha	4
PPARGC1B	Peroxisome proliferator-activated receptor gamma coactivator 1 beta	5
SIRT1	Sirtuin 1	10
SP1	Specificity protein 1	12
YY1	Transcriptional repressor protein YY1	14

The second gene set comprised 91 autosomal genes involved in the OXPHOS. This gene set was created by Mootha et al. (2003) for a GSEA of expression profiles of different mitochondrial gene sets (Mootha et al. 2003, Segrè et al. 2010). Finally, the third gene set consisted of a list of 966 autosomal nuclear-encoded human mitochondrial genes which was derived from the MitoCarta compendium (Pagliarini et al. 2008). The proteins belonging to the genes of the latter gene set were identified in 14 mouse tissues using various approaches (mass spectrometry, green fluorescent protein tagging and integrated analysis of 7 genome-scale data sets). The corresponding human gene homologs can be found at http://www.broadinstitute.org/pubs/MitoCarta/human.mitocarta.html. According to the authors, the 1098 identified mitochondrial genes of the MitoCarta compendium, from which 1012 human homologs can be derived, represent ~85 % of all assumed mitochondrial genes (Pagliarini et al. 2008, Segrè et al. 2010). Among the 1,012 mitochondrial genes, 13 are from the mtDNA, and 31 are located on the sex chromosomes. These 44 genes were not included in the analysis of Segrè et al. (2010), as SNPs of these genes are usually not analyzed in large scale GWAS studies and different association tests would have to be applied for these non-

autosomal genes. Two further genes were removed, as they were absent in the human gene list used for the analyses of Segrè et al. (2010). Again, for reasons of comparability, the same 966 genes were analyzed in the present GSEA.

3.5.2.2 Mapping of SNPs onto genes

A list of all human gene transcripts (n=26,914 for the hg18 March 2006 version) was downloaded from the UCSC Genome Browser (http://genome.uscs.edu/). Genes with transcripts on separate chromosomes or genes with transcripts on one chromosome, whose distance is more than 1 Mb, were excluded. Hence, 17,680 autosomal genes were followed-up. To the most extreme transcript start and end site of each gene 110 kb upstream and 40 kb downstream were added (Fig 3.6A). These boundaries represent the 99^{th} percentiles of cis-eQTLs distances from transcript start and end sites of adjacent genes (Segrè et al. 2010 according to Veyrieras et al. 2008). SNPs (with their corresponding p-values determined by Cochran-Armitage test or TDT, cf. 3.5.1.3) were mapped onto genes within the extended boundaries (Fig. 3.6A). The 55 genes without SNPs within the gene and the extended boundaries were excluded from further analyses.

3.5.2.3 Determination of gene-wise empirically corrected p-value P_g

Each gene was assigned a gene-wise empirically corrected p-value P_g. Therefore, the lowest observed p-value $P_{g;min}$ of each gene was determined and stored (Fig. 3.6B). Afterwards, in CC and KORA-CC, 10,000 permutations of the genotype data were performed using PLINK v1.07 (Purcell et al. 2007). The null distribution was generated by flipping the affection status for all SNPs in each permutation. P_g was eventually calculated as the fraction of permutations whose minimal p-value per gene was equal to or smaller than $P_{g;min}$. For genes with $P_g \leq 0.01$ (0.001), the procedure was repeated with 100,000 (1,000,000) permutations. This was done to achieve maximal accuracy. In the family-based trios, P_g was calculated by randomly flipping the parentally transmitted allele for each family and each permutation.

A gene with an identical selected SNP as found already in one of the gene sets was excluded due to the physical clustering of the gene. This exclusion was done to avoid significant gene set enrichment based on identical association signals (Segrè et al. 2010).

3.5.2.4 Determination of gene set enrichment p-values P^{GSEA}

Several tests were applied to test the alternative hypothesis that P_g in one of the three gene sets skewed towards high ranks in comparison with the set of all autosomal genes. First of all, the leading-edge-fraction-comparison test as proposed by Segrè et al. (2010) was used. The 50^{th}, 75^{th} and 95^{th} percentile of the set of all autosomal P_g were chosen as cut-offs (i.e. leading edge fraction). With this test, the fraction of genes with P_g below a certain cut-off (i.e. above a certain percentile) of the gene set of interest was compared with that of the full set of genes. The null distribution of these fractions is generated by randomly sampling 10,000 gene sets from the genome which are identical in size compared with the three gene sets, respectively. The corresponding GSEA p-values ($P^{GSEA,95}$, $P^{GSEA,75}$ and $P^{GSEA,50}$) were generated by division of the number of samplings with equal or larger leading edge fraction as observed in the respective gene set for a given cut-off by the number of samplings generated.

As proposed by Segrè et al. (2010), three alternative one-sided GSEA tests were used (Wilcoxon-Mann-Whitney test ($P^{GSEA,WMW}$), Kolmogorov-Smirnov test ($P^{GSEA,KS}$) and t-test ($P^{GSEA,t}$)) to test the robustness of the leading-edge-fraction-comparison test. Due to the fact that three gene sets were tested for enrichment of association signals, enrichment was considered significant in case of $p^{GSEA} < 0.017$.

3.5.2.5 Meta-Analysis Gene set Enrichment of variant Associations (MAGENTA)

For the performance of a gene set enrichment analysis in a meta-analysis in which individual genotypes are not available, Segrè et al. (2010) used the MAGENTA software available at http://www.broadinstitute.org/mpg/magenta/. MAGENTA is based on a linear regression-based approach which accounts for physical gene size, the number of SNPs and their genetic properties (LD between SNPs, number of recombination hotspots and genetic distance of the gene) to determine P_g (Fig.3.6C). This approach was performed, as in case of missing individual genotype data, the above mentioned permutation procedures could not be used for the determination of P_g.

Within the present PhD thesis, a meta-analysis was performed of all three samples (CC, trios and KORA-CC). The p-values of each SNP from three samples were meta-analyzed with the inverse variance method under the assumption of a fixed effect model adopted from Kazeem and Farrell (2005). Therefore, the METAL software package (www.sph.umich.edu/csg/abecasis/metal; Willer et al. 2010) was used, and MAGENTA was

applied to the p-values obtained. As mentioned above, the leading edge fraction tests for the 95th, the 75th and the 50th percentile cut-off were performed ($p^{MAGENTA}$). The exact Wilcoxon-Mann-Whitney test included in the software was applied as alternative test.

MAGENTA was also applied to each of the three samples individually in order to be comparable to the permutation-based GSEA approach. As for GSEA testing, enrichment was considered significant at $p^{MAGENTA} < 0.017$.

3.5.2.6 Determination of LD between best SNPs of gene set 1

The first gene set was enriched for association signals. It turned out that the best SNP of each gene was mostly different between the samples. The parents of the family-based trios served as reference sample to calculate the LD between these best SNPs using HaploView 4.2 (Barrett et al. 2005).

4 Results

4.1 Variation in mitochondrial DNA

First of all an association study of 35/40 mtDNA SNPs of the Affymetrix Genome-Wide Human SNP array was performed in 1,158 (extremely) obese children and adolescents and 435 lean adult controls. The minor allele A of m.8994G/A (rs28358887), and haplogroup W were nominally overrepresented in the lean controls. m.8994G/A is a synonymous SNP in *ATP6.* Its minor allele was present in all individuals of haplogroup W. Thus, mtDNA of each five cases (three of haplogroup W, each one of H and HV) and controls (all of haplogroup W) was completely re-sequenced in order to detect causal variants explaining the initial associations. In the confirmation analysis, which was only completed after having started re-sequencing of complete mtDNA, the initial associations were not confirmed. Although all eight individuals of haplogroup W were identical regarding the genotype information found on the SNP array, by re-sequencing both in the D-loop and coding region inter-individual variation was detected. Thus, the array provides only limited information pertaining to mtDNA variants, and the detection of further variants will require re-sequencing. The D-loop was chosen for re-sequencing in each 192 cases and controls, mainly from the CC sample, as (1) the inter-individual variability was greater in the D-loop compared with the coding region among individuals of haplogroup W, (2) the D-loop was not sufficiently covered, and (3) the D-loop is an important control region of mtDNA transcription and replication; thus variation in this region might have an impact on these processes and in further consequence on body weight. Lastly, (4) re-sequencing of the whole mtDNA (~16,569 bp) for a meaningful sample size of ~400 individuals would have been too extensive in the present PhD thesis, whereas the D-loop has a length of only ~1,100 bp. Fig. 4.1 sketches the approach of investigating whether variation in mtDNA is associated with obesity.

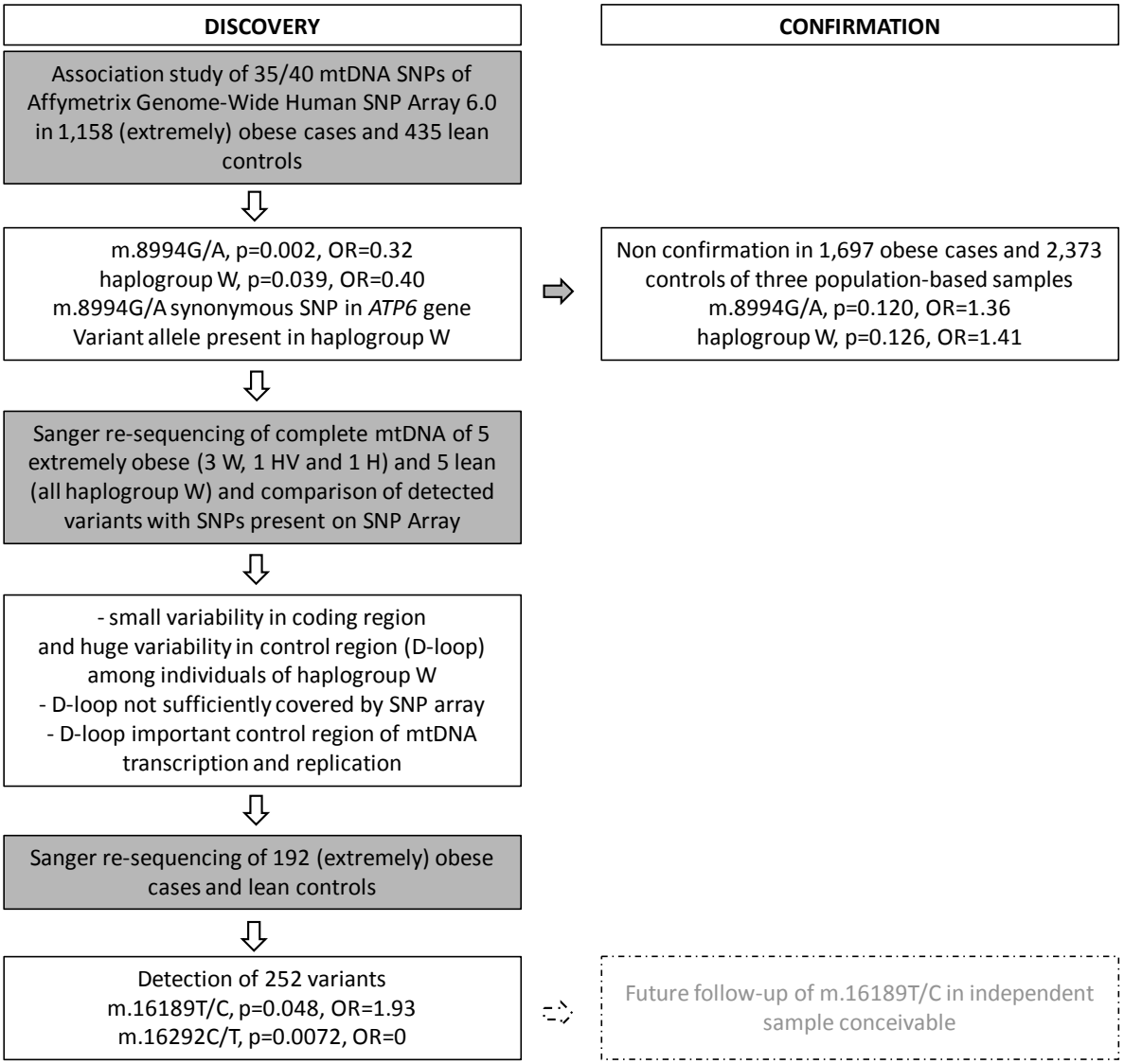

Figure 4.1 Approach for analysis/detection of variation in mtDNA in association with obesity

OR, odds ratio

4.1.1 Association of common mitochondrial SNPs with obesity

In the discovery, 35 mtDNA SNPs (+ 5 SNPs only in the CC sample) were analyzed in association with obesity (Table 4.1). One SNP, m.8994G/A, was nominally associated (p=0.002), and its minor allele A was more frequent among the controls (3.92 % vs. 1.30 %, Table 4.1). m.8994G/A is synonymous and located in *ATP6*. Analysis stratified by gender showed m.8994G/A to be associated in both males and females. In addition, one and three further SNPs were nominally significant only in males and females, respectively (Table 4.2).

Table 4.1 SNPs of mtDNA in association with obesity in discovery

Position on mtDNA according to rCRS	rs number	Minor allele [a]	Major allele [a]	MAF Cases [%] [b]	MAF Controls [%] [b]	Odds ratio for minor allele	Confidence Interval	p-value [c]
				n=1,158	n=435			
m.1438	rs2001030	**A**	G	2.94	2.76	1.06	0.53-2.28	1.000
m.1700	rs2854126 [d]	C	**T**	1.12	1.62	0.69	0.17-2.54	0.573
m.1811	rs28358576	G	**A**	13.04	13.56	0.96	0.69-1.35	0.803
m.1888	rs28358577 [d]	A	**G**	11.04	8.28	1.37	0.86-2.22	0.175
m.2706	rs2854128	**A**	G	46.28	45.73	1.02	0.81-1.28	0.865
m.3010	rs3928306	A	**G**	24.37	26.68	0.89	0.68-1.15	0.362
m.3197	rs2854131	C	**T**	8.23	7.64	1.08	0.71-1.69	0.757
m.4580	rs28357975	A	**G**	3.46	3.70	0.93	0.50-1.81	0.879
m.4769	rs3021086	**A**	G	3.11	1.84	1.71	0.77-4.30	0.229
m.7028	rs2015062	**C**	T	45.76	44.76	1.04	0.83-1.31	0.734
m.8994	rs28358887	A	**G**	1.30	3.92	0.32	0.15-0.69	**0.002**
m.9055	rs28358268 [d]	A	**G**	6.62	7.60	0.86	0.50-1.49	0.603
m.9123	rs28358270	A	**G**	1.56	1.61	0.96	0.38-2.75	1.000
m.9698	rs9743	C	**T**	6.65	8.35	0.78	0.51-1.22	0.272
m.10238	rs28358275	C	**T**	3.11	2.76	1.13	0.57-2.41	0.869
m.10463	rs28358279	C	**T**	10.45	8.99	1.18	0.80-1.77	0.454
m.10550	rs28358280	G	**A**	6.13	7.16	0.85	0.54-1.36	0.490
m.11251	rs3915952	G	**A**	19.43	21.43	0.88	0.67-1.17	0.398
m.11299	rs28358285	C	**T**	6.07	6.93	0.87	0.55-1.40	0.562
m.11467	rs2853493	G	**A**	21.56	21.43	1.01	0.77-1.33	1.000
m.11674	rs28358286	T	**C**	1.39	2.99	0.46	0.20-1.04	0.055
m.11719	rs2853495	A	**G**	49.70	50.34	0.97	0.78-1.22	0.822
m.11812	rs3088053 [d]	G	**A**	6.74	6.53	1.04	0.59-1.83	1.000
m.11914	rs2853496	A	**G**	2.34	1.15	2.06	0.77-6.89	0.162
m.12007	rs2853497	A	**G**	1.64	2.07	0.79	0.34-1.99	0.527
m.12308	rs2853498	G	**A**	21.78	21.61	1.01	0.77-1.34	1.000
m.12612	rs28359172	G	**A**	8.89	11.72	0.74	0.51-1.07	0.105
m.12705	rs2854122	T	**C**	7.81	7.14	1.10	0.71-1.74	0.750
m.13368	rs3899498	A	**G**	10.11	8.51	1.21	0.81-1.84	0.392
m.13617	rs2853503	C	**T**	8.12	7.37	1.11	0.72-1.74	0.677
m.13708	rs28359178	A	**G**	10.13	12.47	0.79	0.56-1.14	0.203
m.14470	rs3135030	C	**T**	2.60	2.07	1.26	0.58-3.04	0.716
m.14905	rs28357682	A	**G**	10.19	8.28	1.26	0.84-1.91	0.295
m.15043	rs28357684	A	**G**	4.24	2.99	1.44	0.76-2.91	0.309
m.15218	rs2853506	G	**A**	4.40	3.22	1.39	0.75-2.74	0.322
m.15326	rs2853508	**A**	G	1.47	0.69	2.14	0.62-11.5	0.312
m.15452	rs3088309	A	**C**	18.70	19.82	0.93	0.70-1.25	0.616
m.15607	rs28357372	G	**A**	10.19	8.51	1.22	0.82-1.85	0.343
m.15924	rs2853510	G	**A**	4.23	3.69	1.15	0.64-2.20	0.672
m.16140	rs3134562 [d]	C	**T**	6.25	6.96	0.89	0.50-1.57	0.686

[a] Reference allele of rCRS marked in bold

[b] MAF, minor allele frequency

[c] Fisher's exact test, two-sided, p-values below 0.05 are highlighted in bold

[d] SNPs that did not pass quality control in family-based trios

All nominally associated SNPs were followed-up in the confirmation sample (Fig. 3.1). However, association could not be confirmed for any of these SNPs. For most of these SNPs the direction of effect has even changed (Table 4.2).

Table 4.2 Nominally associated SNPs of mtDNA in discovery and follow-up in confirmation

	Discovery					Confirmation				
SNP [a]	**MAF Cases [%]** [b]	**MAF Controls [%]** [b]	**Odds Ratio**	**Confidence Interval** [c]	**p-value** [d]	**MAF Cases [%]** [b]	**MAF Controls [%]** [b]	**Odds Ratio**	**Confidence Interval** [c]	**p-value** [d]
All	**n=1,158**	**n=435**				**n=1,697**	**n=2,373**			
m.8994G/A	1.30	3.92	0.32	0.15-0.69	**0.002**	3.24	2.41	1.36	0.92-2.02	0.120
Males	**n=508**	**n=171**				**n=828**	**n=930**			
m.8994G/A	0.79	2.94	0.26	0.05-1.23	**0.048**	2.78	2.16	1.30	0.68-2.51	0.441
m.11674C/T	0.59	3.51	0.16	0.03-0.78	**0.010**	2.06	1.63	1.27	0.63-2.56	0.593
Females	**n=650**	**n=264**				**n=869**	**n=1,443**			
m.4769A/G	3.38	0.76	4.60	1.12-40.6	**0.022**	2.99	2.36	1.28	0.76-2.15	0.348
m.8994G/A	1.69	4.55	0.36	0.14-0.91	**0.019**	3.68	2.57	1.45	0.87-2.41	0.132
m.12612A/G	8.00	2.88	0.58	0.36-0.95	**0.023**	10.60	9.01	1.20	0.90-1.59	0.216
m.13708G/A	9.12	13.69	0.61	0.38-0.98	**0.040**	10.87	12.81	1.08	0.83-1.42	0.238

[a] mtDNA position according to rCRS

[b] MAF, minor allele frequency

[c] 95 % confidence of odds ratio for minor allele

[d] Fisher's exact test, 2-sided, p-values below 0.05 are highlighted in bold

4.1.2 Association of haplogroups with obesity

A total of 80 haplogroups with a quality value of at least 90 % were identified with HaploGrep (Table 4.3). These haplogroups were assigned to 18 major haplogroups for association analysis (Table 4.3). The HaploGrep's quality 90 % was exceeded by 96 % of individuals, and only these individuals were included for association testing.

In the discovery, haplogroup W was nominally associated with obesity (p=0.034). Comparable to the minor allele A of m.8994G/A, which was present in all individuals of haplogroup W, haplogroup W was more frequent in the controls (2.84 % vs. 1.17 %, Table 4.4. Stratified by gender, haplogroup W remained nominally associated only in the males (p=0.012), and haplogroup J became nominally significant in the females (p=0.032, Table 4.5). None of the findings from the discovery was confirmed in the population-based adults, and the direction of effect has changed (Table 4.5).

Table 4.3 Haplogroup frequencies in discovery and confirmation

		Discovery		Confirmation		
Haplogroup [a]	Total n=5,430	Cases n=1,114	Controls n=422	Cases n=1,623	Controls n=2,271	Major Haplogroup
A2i	5	0	0	2	3	**A**
B4c1b	1	0	1	0	0	
B4c1b2b	2	0	0	2	0	**B**
B5b1a'b	19	0	0	8	11	
D4j5	3	0	0	2	1	**D**
H	1171	247	96	344	484	
H1	871	184	69	268	350	
H18	7	4	0	2	1	
H1h	17	6	0	3	8	
H2	8	1	0	2	5	
H21	1	0	1	0	0	
H27	47	12	5	10	20	**H**
H2a2	44	19	4	8	13	
H2a5	8	2	1	1	4	
H4	75	15	7	16	37	
H6a1b	5	0	0	2	3	
HV	109	23	9	28	49	
HV1	8	3	0	1	4	
J	131	24	8	48	51	
J1	353	59	35	107	152	
J1b1a	41	12	6	10	13	**J**
J1c1b1a	5	2	1	0	2	
J1c3c	16	5	1	3	7	
JT	1	0	0	1	0	
K	198	63	28	46	61	
K1a1	22	4	1	6	11	**K**
K1a1b	21	4	2	4	11	
K1a4c	1	0	0	1	0	
L0f2a1	1	0	0	0	1	
L0k	1	0	0	0	1	**L**
L2a1c	1	1	0	0	0	
L3e2	1	0	0	1	0	
M	9	3	1	1	4	
M20	9	2	0	3	4	
M34	12	3	0	2	7	**M**
M3a	1	1	0	0	0	
M5	2	2	0	0	0	
M8a	1	1	0	0	0	
N	3	0	0	1	2	
N1	13	4	2	5	2	
N1a'e'I	19	2	1	4	12	
N1b	3	0	0	0	3	**N**
N1c	4	2	0	2	0	
N1e'I	134	26	8	50	50	
N2	2	1	0	1	0	
N9a1	3	2	0	1	0	

Table 4.3 is continued on the next page

Table 4.3 Haplogroup frequencies in discovery and confirmation – *continued*

Haplogroup [a]	Total n=5,430	Discovery Cases n=1,114	Discovery Controls n=422	Confirmation Cases n=1,623	Confirmation Controls n=2,271	Major Haplogroup
P	2	1	0	0	1	**P**
R	16	3	1	4	8	
R0a1a1	4	0	0	2	2	**R**
R0a2b	1	1	0	0	0	
S3	1	0	0	1	0	**S**
T	407	81	8	122	196	
T2	108	27	27	21	33	**T**
T2d	7	2	0	1	4	
T2f1	41	6	1	14	20	
U	18	5	1	4	8	
U2'3'4'7'8'9	299	60	22	89	128	
U2b	7	2	0	1	4	
U2d	2	0	0	1	1	
U2e1a1a	3	1	0	1	1	
U3a1	29	10	2	9	8	
U4b1a2	5	0	0	3	2	
U4b1a3	6	2	1	1	2	
U5	301	55	17	89	140	
U5a1	121	24	4	44	49	**U**
U5a1a	6	1	4	1	0	
U5a1a1	60	12	3	13	32	
U5a1b1c	3	2	1	0	0	
U5a2b2	6	0	0	3	3	
U5b1a	1	0	1	0	0	
U5b1c1	3	0	1	2	0	
U5b2a1a2	1	0	1	0	0	
U8	186	5	4	86	91	
U8b	2	1	1	0	0	
V	186	37	16	59	74	**V**
W	93	13	12	32	36	**W**
W3a	20	0	0	12	8	
X	39	19	4	7	9	**X**
X2b'd	36	6	2	6	22	
Z	1	0	0	0	1	**Z**

[a] Haplogroups determined by HaploGrep (Kloss-Brandstätter et al. 2011), only individuals with HaploGrep's quality ≥ 90 % were included

Table 4.4 Frequency of major haplogroups in cases and controls in discovery and confirmation

	Discovery					Confirmation				
Haplo-group [a]	Frequency Cases [%]	Frequency Controls [%]	Odds Ratio	Confidence Interval [b]	p-value [c]	Frequency Cases [%]	Frequency Controls [%]	Odds Ratio	Confidence Interval [b]	p-value [c]
	n=1,114	n=422				n=1,623	n=2,271			
A	0	0	-	-	-	0.12	0.13	0.93	0.08-8.15	1.000
B	0.00	0.24	0.00	0.00-14.65	0.275	0.62	0.48	1.27	0.48-3.31	0.659
D	0	0	-	-	-	0.12	0.04	2.80	0.15-165	0.575
H	46.23	45.73	1.00	0.80-1.26	0.909	42.21	43.06	0.97	0.85-1.10	0.599
J	9.16	12.09	0.73	0.50-1.06	0.104	10.41	9.91	1.06	0.85-1.31	0.628
K	6.37	7.35	0.85	0.54-1.36	0.492	3.51	3.65	0.96	0.67-1.37	0.862
L	0.09	0	Inf	0.01-Inf	1.000	0.06	0.09	0.70	0.01-13.4	1.000
M	1.08	0.24	4.54	0.67-195	0.129	0.37	0.66	0.56	0.18-1.53	0.271
N	3.32	2.61	1.27	0.63-2.79	0.516	3.94	3.04	1.31	0.91-1.88	0.129
P	0.09	0.00	Inf	0.01-Inf	1.000	0.00	0.04	0.00	0.00-54.5	1.000
R	0.36	0.24	1.50	0.15-74.2	1.000	0.37	0.44	0.84	0.25-2.55	0.804
S	0	0	-	-	-	0.06	0.00	Inf	0.04-Inf	0.417
T	10.41	8.53	1.23	0.83-1.88	0.293	9.74	11.14	0.86	0.69-1.07	0.169
U	16.16	14.93	1.09	0.79-1.51	0.584	21.32	20.70	1.04	0.88-1.21	0.661
V	3.32	3.79	0.86	0.46-1.68	0.641	3.64	3.26	1.12	0.78-1.61	0.532
W	1.17	2.84	0.40	0.17-0.97	**0.034**	2.71	1.94	1.41	0.90-2.20	0.126
X	2.24	1.42	1.58	0.63-4.73	0.416	0.80	1.37	0.58	0.28-1.15	0.124
Z	0	0	-	-	-	0.00	0.04	0.00	0.00-54.5	1.000

[a] only individuals with HaploGrep's quality ≥ 90 % were included (~96 % of all individuals)

[b] 95 % confidence interval for odds ratio

[c] Fisher's exact test, 2-sided, p-values below 0.05 are highlighted in bold

Table 4.5 Frequency of haplogroups in cases and controls in discovery and confirmation stratified by gender

	Discovery					Confirmation				
Haplo-group [a]	Frequency Cases [%]	Frequency Controls [%]	Odds Ratio	Confidence Interval [b]	p-value [c]	Frequency Cases [%]	Frequency Controls [%]	Odds Ratio	Confidence Interval [b]	p-value [c]
males	**n=491**	**n=163**				**n=793**	**n=890**			
A						0.25	0.22	1.12	0.08-15.5	1.000
B	0.00	0.00	0.00	0.00-Inf	-	0.38	0.22	1.69	0.19-20.2	0.671
D						0.13	0.11	1.12	0.01-88.2	1.000
H	46.03	46.01	1.03	0.71-1.48	1.000	44.14	42.81	1.06	0.87-1.28	0.588
J	10.39	10.43	1.01	0.55-1.93	1.000	9.84	10.90	0.89	0.64-1.24	0.522
K	4.89	6.75	0.72	0.33-1.67	0.421	2.90	3.48	0.83	0.46-1.48	0.580
L	0.00	0	0.00	0.00-Inf	-	0.00	0.11	0.00	0.00-43.8	1.000
M	0.41	0.00	Inf	0.06-Inf	1.000	0.63	0.56	1.12	0.26-4.90	1.000
N	3.87	1.84	2.17	0.63-11.6	0.315	3.28	3.71	0.88	0.50-1.53	0.691
P	0.00	0.00	0.00	0.00-Inf	-	0.00	0.00	0.00	0.00-Inf	-
R	0.61	0.61	1.01	0.08-53.3	1.000	0.63	0.45	1.41	0.30-7.11	0.742
S						0.00	0.00	0.00	0.00-Inf	-
T	10.59	9.82	1.10	0.60-2.14	0.883	10.21	11.12	0.91	0.66-1.26	0.581
U	15.48	17.18	0.90	0.55-1.50	0.622	21.31	19.89	1.09	0.86-1.39	0.506
V	4.48	3.68	1.24	0.48-3.82	0.824	3.15	2.92	1.08	0.59-1.97	0.887
W	0.41	3.07	0.13	0.01-0.81	**0.012**	2.14	1.69	1.28	0.60-2.77	0.592
X	2.85	0.61	4.81	0.72-204	0.132	1.01	1.80	0.56	0.21-1.39	0.217
Z						0.00	0.00	0.00	0.00-Inf	-
females	**n=623**	**n=259**				**n=830**	**n=1,381**			
A						0.00	0.07	0.00	0.00-64.7	1.000
B	0.00	0.39	0.00	0.00-15.8	0.294	0.84	0.65	1.29	0.41-3.92	0.612
D						0.12	0.00	Inf	0.04-Inf	0.375
H	46.39	45.56	0.99	0.74-1.34	0.824	40.36	43.23	0.89	0.75-1.06	0.197
J	8.19	13.13	0.58	0.36-0.94	**0.032**	10.96	9.27	1.20	0.89-1.61	0.211
K	7.54	7.72	0.95	0.54-1.73	1.000	4.10	3.77	1.09	0.68-1.73	0.734
L	0.16	0	Inf	0.01-Inf	1.000	0.12	0.07	1.66	0.02-1306	1.000
M	1.61	0.39	4.11	0.58-179	0.190	0.12	0.72	0.17	0.00-1.16	0.062
N	2.89	3.09	0.91	0.37-2.45	0.830	4.58	2.61	1.79	1.09-2.93	**0.014**
P	0.16	0.00	Inf	0.01-Inf	1.000	0.00	0.07	0.00	0.00-64.7	1.000
R	0.16	0.00	Inf	0.01-Inf	1.000	0.12	0.43	0.28	0.01-2.28	0.267
S						0.12	0.00	Inf	0.04-Inf	0.375
T	10.27	7.72	1.33	0.78-2.38	0.259	9.28	11.15	0.81	0.60-1.09	0.173
U	16.69	13.51	1.25	0.81-1.94	0.265	21.33	21.22	1.00	0.81-1.24	0.957
V	2.41	3.86	0.60	0.25-1.52	0.266	4.10	3.48	1.18	0.73-1.89	0.486
W	1.77	2.70	0.63	0.22-1.95	0.433	3.25	2.10	1.56	0.88-2.76	0.123
X	1.77	1.93	0.89	0.28-3.31	1.000	0.60	1.09	0.55	0.16-1.60	0.354
Z						0.00	0.07	0.00	0.00-64.7	1.000

[a] only individuals with HaploGrep's quality ≥ 90 % were included (~96 % of all individuals)

[b] 95 % confidence interval for odds ratio

[c] Fisher's exact test, 2-sided, p-values below 0.05 are highlighted in bold

4.1.3 Re-sequencing of complete mtDNA

By re-sequencing of the complete mtDNA of 10 individuals, a total of 65 different variants (i.e. deviations of rCRS) were detected which were not present on the SNP array (Table 4.6). Among these, 33 belong to the mitochondrial D-loop and 32 to the coding region. All of these variations have been reported previously (www.mitomap.org, *last edited on Apr 23, 2013*), except for the deletion mt.8270_8278DelCACCCCCTC in the non-coding region. However, a similar one displaced by two bases, mt.8272_8280DelCCCCCTCTA, had been already described (Cann and Wilson 1983). In addition, genotype information of all 40 SNPs from the SNP array was re-confirmed.

Table 4.6 Variants detected by re-sequencing of complete mtDNA of each five lean and obese individuals

			Individual [a] (Haplogroup [b])									
Position [c]	**Reference allele [d]**	**Variant Allele**	**1 (W)**	**2 (W)**	**3 (W)**	**4 (H)**	**5 (HV)**	**6 (W)**	**7 (W)**	**8 (W)**	**9 (W)**	**10 (W)**
m.16093	T	C				C						
m.16104	C	A	A									
m.16140	T	C										
m.16192	C	T			T							
m.16213	G	A						A				
m.16221	C	T				T						
m.16223	C	T	T	T	T			T	T	T	T	T
m.16261	C	T									T	
m.16286	C	T										T
m.16292	C	T	T	T	T				T	T	T	
m.16295	C	T								T		
m.16301	C	T						T				
m.16311	T	C	C				C					
m.16324	T	C									C	
m.16325	T	C			C							
m.16362	T	C		C					C			
m.16519	T	C	C	C	C	C		C	C	C	C	C
m.73	A	G	G	G	G			G	G	G	G	G
m.94	G	A						A				
m.119	T	C								C	C	
m.143	G	A										A
m.152	T	C					C					
m.189	A	G	G	G	G				G	G	G	G
m.192	T	C										C
m.194	C	T	T	T	T				T			T
m.195	T	C	C	C	C				C	C	C	C
m.196	T	C										C
m.199	T	C	C									
m.204	T	C	C	C	C			C	C	C	C	C
m.207	G	A	A	A	A			A	A	A	A	A
m.263	A	G	G	G	G	G	G	G	G	G	G	G
m.309						309.1C			309.1C		309.1C	
m.309							309.1CC					
m.315			315.1C	315.1C	315.1C	315.1C	315.1C	315.1C	315.1C	315.1C	315.1C	315.1C

Table 4.6 is continued on the next page

Table 4.6 Variants detected by re-sequencing of complete mtDNA of each five lean and obese individuals – *continued*

Position [c]	Reference Allele [d]	Variant Allele	Individual [a] (Haplogroup [b])									
			1 (W)	2 (W)	3 (W)	4 (H)	5 (HV)	6 (W)	7 (W)	8 (W)	9 (W)	10 (W)
m.709	G	A	A	A	A			A	A	A	A	A
m.750	A	G	G	G	G	G	G	G	G	G	G	G
m.960			960.1C									
m.1243	T	C	C	C	C			C	C	C	C	C
m.1406	T	C	C									
m.1438	A	G	G	G	G	G	G	G	G	G	G	G
m.1700	T	C										
m.1811	A	G										
m.1888	G	A										
m.2706	A	G	G	G	G		G	G	G	G	G	G
m.3010	G	A										
m.3197	T	C										
m.3505	A	G	G	G	G			G	G	G	G	G
m.3531	G	A										A
m.4093	A	G			G							
m.4363	T	C		C					C			
m.4580	G	A										
m.4769	A	G	G	G	G	G	G	G	G	G	G	G
m.4833	A	G								G		
m.5046	G	A	A	A	A			A	A	A	A	A
m.5460	G	A	A	A	A			A	A	A	A	A
m.6528	C	T		T					T			
m.7028	C	T	T	T	T		T	T	T	T	T	T
m.7864	C	T						T		T	T	
m.8251	G	A	A	A	A			A	A	A	A	A
m.8270										8270_8278 DelCACCCCCTC		
m.8610	T	C			C							
m.8614	T	C			C							
m.8860	A	G	G	G	G	G	G	G	G	G	G	G
m.8994	G	A	A	A	A		A	A	A	A	A	A
m.9055	G	A										
m.9123	G	A										
m.9275	A	G							G			
m.9698	T	C										
m.10097	A	G		G					G			
m.10238	T	C										
m.10310	G	A									A	
m.10410	T	C		C					C			
m.10463	T	C										
m.10550	A	G										
m.11227	C	A									A	
m.11251	A	G										
m.11299	T	C										
m.11467	A	G										
m.11674	C	T	T	T	T			T	T	T	T	T
m.11719	G	A	A	A	A			A	A	A	A	A

Table 4.6 is continued on the next page

Table 4.6 Variants detected by re-sequencing of complete mtDNA of each five lean and obese individuals – *continued*

Position [c]	Reference Allele [d]	Variant Allele	Individual [a] (Haplogroup [b])									
			1 (W)	2 (W)	3 (W)	4 (H)	5 (HV)	6 (W)	7 (W)	8 (W)	9 (W)	10 (W)
m.11812	A	G										
m.11914	G	A										
m.11947	A	G	G	G	G			G	G	G	G	G
m.12007	G	A										
m.12308	A	G										
m.12414	T	C	C	C	C			C	C	C	C	C
m.12612	A	G										
m.12705	C	T	T	T	T			T	T	T	T	T
m.12923	G	T	T									
m.13368	G	A										
m.13617	T	C										
m.13708	G	A										
m.14148	A	G									G	
m.14470	T	A				A						
m.14602	A	G				G						
m.14766	C	T	T	T	T			T	T	T	T	T
m.14905	G	A										
m.15043	G	A										
m.15218	A	G										
m.15326	A	G	G	G	G	G	G	G	G	G	G	G
m.15452	C	A										
m.15607	A	G										
m.15775	A	G		G					G			
m.15884	G	C	C	C	C			C	C	C	C	C
m.15924	A	G										

[a] Individual 1-5 extremely obese children and adolescent, individual 6-10

[b] Haplogroup determined using HaploGrep (Kloss-Brandstätter et al. 2011), only individuals with HaploGrep's Quality ≥ 90 % were included

[c] grey shaded positions represent the 40 SNPs present on the Affymetrix Genome-Wide Human SNP Array 6.0

[d] according to rCRS; only deviations from the reference are shown

Considering only those eight individuals of haplogroup W, which were identical regarding genotype information from the SNP array, 29 and 30 variants were detected in the D-loop and coding region, respectively. In the D-loop, seven variants were detected in all individuals, while 15 were detected in only one individual. By contrast, in the coding region, 12 variants were detected in all individuals and 12 only in one individual (Table 4.6).

Finally, using all detected variants as well as only the detected D-loop variants, haplogroup was re-determined with HaploGrep (Table 4.7). Although haplogroups changed in all individuals, the assigned major haplogroup, which was used for association analysis,

remained the same apart from one exception. In individual 6, by using only D-loop variants for haplogroup determination, haplogroup turned out to be N9b1c1 and hence the assigned major haplogroup would be N. However, due to a HaploGrep's quality of only 80.5 %, the individual would not have been considered for major haplogroup assignment or association analysis. HaploGrep's quality values were slightly lower by including all detected variants and lowest by using only D-loop variants compared with the haplogroup determination using SNP array-based genotype information (Table 4.7).

Table 4.7 Haplotype classification after complete re-sequencing using HaploGrep

Individual [a]	Genotype information based on SNP array		Assigned major haplo-group	All detected variants by re-sequencing		Assigned major haplo-group	Only detected D-Loop variants		Assigned major haplo-group
	Haplogroup	Quality [b]		Haplogroup	Quality [b]		Haplogroup	Quality [b]	
1	W	100	W	W3b	99.6	W	W3a1+199	98.3	W
2	W	100	W	W5a1a	100	W	W5a	96.8	W
3	W	100	W	W6a	100	W	W6+16192	100	W
4	H	100	H	H10e2	100	H	H10e	75.5	H
5	HV	91.9	H	HV9+152	100	H	H2a+152 16311	63.9	H
6	W	100	W	W1	93.2	W	N9b1c1	80.5	N
7	W	100	W	W5a1a1	100	W	W5a	100.0	W
8	W	100	W	W1+119	97.1	W	W1+119	95.2	W
9	W	100	W	W1c	96.5	W	W1+119	91.4	W
10	W	100	W	W4a	97.6	W	W4a	96.0	W

[a] Individuals 1-5 extremely obese children and adolescents, individuals 6-10 lean adults

[b] a quality of at least 90 % indicates a quite reliable haplogroup assignment according to the user manual of HaploGrep

4.1.4 Re-sequencing of mitochondrial D-loop

D-loop re-sequencing (Sanger) was performed in 192 cases and 192 controls (D-loop sample) predominantly from the CC sample. In one case and one control, 9 and 4 clearly visible point heteroplasmies, respectively, were found. D-loop re-sequencing was repeated for both individuals using DNA from the stock solution of the respective individuals, and the initial results were confirmed. As point heteroplasmies at more than one position in one individual may be existing, but are rather infrequent (Budowle et al. 2002), in particular in blood cells (Calloway et al. 2000), contamination with DNA from another individual during DNA isolation might also be conceivable (Andréasson et al. 2006). Therefore, both individuals were excluded from further analyses.

Among the 191 extremely obese children and adolescents and 191 controls followed-up for all subsequent considerations, n=178 and n=186 were derived from the CC sample,

respectively. The major haplogroup distribution based on the 40 SNPs from the SNP array did not differ between the D-loop sample and the complete discovery sample (p=0.966; Table 4.8 and Fig. 4.2). Haplogroup of these 364 individuals was re-assigned using only the detected D-loop variants. The necessary HaploGrep's quality threshold of 90 % was reached in only 90 individuals (25 % vs. 96 % in SNP array-based assignment). However, among these 90 individuals, 91 % were assigned to the same major haplogroup as compared to the SNP-based assignment. Taking the D-loop variant derived assignment of all 364 individuals into consideration irrespective of quality value, accordance of the major haplogroup still has been found in 87 % of the individuals.

Table 4.8 Comparison of haplogroup distribution between discovery and individuals whose D-loop was re-sequenced (D-loop sample)

	Discovery sample			D-loop sample [a]		
Major Haplogroup [b]	Frequency cases [%]	Frequency controls [%]	p-value [c]	Frequency cases [%]	Frequency controls [%]	p-value [c]
	n=1,114	n=422		n=169	n=183	
B	0.00	0.24	0.275	0.00	0.55	1.000
H	46.23	45.73	0.909	44.97	47.54	0.669
J	9.16	12.09	0.104	13.02	8.20	0.165
K	6.37	7.35	0.492	5.92	8.20	0.534
L	0.09	0.00	1.000	0.00	0.00	-
M	1.08	0.24	0.129	0.59	0.00	0.480
N	3.32	2.61	0.516	2.37	3.28	0.752
P	0.09	0.00	1.000	0.00	0.00	-
R	0.36	0.24	1.000	0.00	0.55	1.000
T	10.41	8.53	0.293	12.43	10.38	0.615
U	16.16	14.93	0.584	14.79	14.75	1.000
V	3.32	3.79	0.641	4.14	3.28	0.780
W	1.17	2.84	**0.039**	0.00	2.73	0.062
X	2.24	1.42	0.416	1.78	0.55	0.354

[a] most of individuals of D-loop sample were derived from the case-control sample

[b] Haplogroup was determined using HaploGrep (Kloss-Brandstätter et al. 2011), only individuals with HaploGrep's Quality ≥ 90 % were included

[c] Fisher's exact test, two-sided

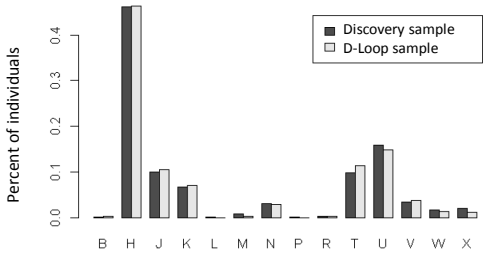

Figure 4.2 Comparison of haplogroup distribution between discovery sample and individuals whose D-loop was re-sequenced (D-loop sample)

Haplogroup was determined using HaploGrep (Kloss-Brandstätter et al. 2011) and SNP array-based genotype information of the 40 SNPs; only individuals with HaploGrep's Quality ≥ 90 % were included, χ^2 test was performed to compare haplogroup distribution between Discovery sample and individuals whose D-loop was re-sequenced (p=0.966, most of individuals of D-loop sample were derived from case-control sample)

In both cases and controls a total of 252 variants, i.e. deviations from the rCRS, were detected in the 1319-bp-long fragment. Four of these variants (m.576_577InsCA, m.628C/A, m.634T/C and 678T/C) were not located in the actual D-loop region between m.576 and m.16024 (Anderson et al. 1981). All detected variants were found at 233 different positions, of which 221 were nucleotide positions and 12 were spaces between nucleotides due to insertions (Table 4.9 and Table 4.10). Altogether, 223 single nucleotide exchanges at 213 positions were found, as at 10 positions tri-allelic single nucleotide exchanges were present (Table 4.9). Moreover, three complex nucleotide exchanges (m.16183A/CC, m.16183A/CCC and m.16189T/CC), 20 insertions of one to six nucleotides, and eight deletions of one to three nucleotides were detected (Table 4.10). Most of the variants detected have been previously described (www.mitomap.org, *last edited on Apr 23, 2013*). Only five (eight) single nucleotide exchanges and six (seven) complex nucleotide exchanges, insertions or deletions in the D-Loop (in the whole fragment) have not been reported previously (Table 4.9 and Table 4.10).

Four cases and nine controls had one point heteroplasmy each, which could be clearly identified by Sanger re-sequencing. These point heteroplasmies were found at 12 positions, as one was found twice (Table 4.9).

Table 4.9 Detected variants (single nucleotide exchanges) by re-sequencing of mtDNA and frequencies in cases and controls

Detected variants	Frequency cases [%] n=191	Frequency controls [%] n=191	Detected variants	Frequency cases [%] n=191	Frequency controls [%] n=191
m.63T/C	0.00	0.52	**m.310T/C**	0.52	0.52
m.64C/T	0.52	1.05	**m.316G/A**	0.52	0.00
m.72T/C	4.71	2.62	**m.318T/C**	0.52	0.00
m.73A/G	53.40	53.40	**m.319T/C**	1.57	1.57
m.73A/T [a]	0.00	0.52	**m.321T/C**	0.00	0.52
m.93A/G	1.57	0.52	**m.340C/T**	0.52	0.00
m.94G/A	0.00	0.52	**m.357A/G**	0.00	0.52
m.95A/C	0.52	0.00	**m.372T/C**	0.00	0.52
m.114C/T	0.00	0.52	**m.385A/G**	1.57	0.00
m.118G/C [a]	0.00	0.52	**m.408T/A**	0.52	1.05
m.119T/C	0.00	1.05	**m.449T/C**	0.52	0.00
m.143G/A	0.52	0.52	**m.455T/C**	0.52	0.00
m.146T/C	9.95	10.21	**m.456C/T**	6.28	4.71
m.150C/T	9.42	7.85	**m.458C/T**	0.52	1.05
m.151C/T	0.52	0.26	**m.462C/T**	8.90	6.28
m.152T/C	25.65	23.04	**m.469C/T**	0.00	0.52
m.153A/G	1.05	0.00	**m.477T/C**	4.71	3.66
m.182C/T	0.00	0.52	**m.480T/C**	0.00	0.52
m.183A/G	0.00	0.52	**m.482T/C**	0.52	0.00
m.185G/A	6.81	4.71	**m.489T/C**	12.57	7.85
m.188A/G	1.57	3.14	**m.497C/T**	2.09	4.19
m.189A/G	0.52	3.66	**m.499G/A**	5.24	3.66
m.193A/G	0.00	0.52	**m.508A/G**	1.05	1.05
m.194C/T	0.52	1.57	**m.509C/T**	0.52	0.00
m.195T/C	18.06	18.59	**m.513G/A**	2.09	1.57
m.196T/C	0.00	0.52	**m.533A/G**	0.00	1.05
m.198C/T	0.52	0.52	**m.535C/T**	0.00	0.52
m.199T/C	2.62	3.14	**m.549T/C**	0.26	0.00
m.200A/G	1.05	0.52	**m.550C/T** [a]	0.00	0.52
m.203G/A	1.57	0.00	**m.564G/A** [a]	0.52	0.00
m.204T/C	2.62	7.33	**m.567A/C**	0.52	0.00
m.207G/A	2.62	6.28	**m.568C/T** [a]	0.52	0.00
m.210A/G	0.52	0.00	**m.574A/C** [a]	0.52	0.00
m.215A/G	1.57	1.57	**m.628C/A** [a]	0.00	0.52
m.217T/C	1.05	0.52	**m.634T/C**	0.00	0.52
m.225G/A	1.05	0.00	**m.678T/C**	0.52	0.00
m.226T/C	1.05	0.00	**m.16051A/G**	2.09	2.62
m.227A/T	0.52	0.00	**m.16063T/C**	0.52	0.00
m.228G/A	5.24	4.19	**m.16067C/T**	0.52	0.00
m.234A/G	0.52	0.00	**m.16069C/T**	11.52	7.85
m.235A/G	1.57	0.00	**m.16082C/T**	0.52	0.00
m.236T/C	0.00	0.52	**m.16086T/C**	1.57	1.05
m.239T/C	2.09	3.14	**m.16092T/C**	1.05	0.52
m.240A/T [a]	0.00	0.52	**m.16093T/C**	5.76	5.24
m.242C/T	1.57	2.09	**m.16104C/T**	0.00	1.05
m.246T/C	0.52	0.00	**m.16104C/A**	0.52	0.00
m.247G/A	1.05	0.52	**m.16111C/T**	0.00	0.52
m.250T/C	2.09	2.09	**m.16114C/T**	1.05	0.26
m.257A/G	1.05	0.52	**m.16126T/C**	23.04	19.37
m.259A/G	0.52	0.00	**m.16129G/A**	4.19	5.24
m.260G/A	0.52	0.00	**m.16129G/C**	1.05	0.52
m.262C/T	0.52	0.00	**m.16134C/T**	0.52	2.09
m.263A/G	97.91	99.48	**m.16145G/A**	4.19	3.66
m.282T/C	0.00	2.09	**m.16146A/G**	0.00	1.05
m.285C/T	1.05	0.52	**m.16147C/A**	0.00	0.52
m.295C/T	10.99	7.85	**m.16147C/T**	0.00	0.52
m.295C/A	0.00	1.05	**m.16148C/T**	0.52	0.52

Table 4.9 is continued on the next page

Table 4.9 Detected variants (single nucleotide exchanges) by re-sequencing of mtDNA and frequencies in cases and controls – *continued*

Detected variants	Frequency cases [%] n=191	Frequency controls [%] n=191	Detected variants	Frequency cases [%] n=191	Frequency controls [%] n=191
m.16153G/A	1.05	0.00	**m.16265A/G**	2.09	0.52
m.16154T/C	0.00	0.52	**m.16266C/T**	1.57	1.57
m.16160A/G	0.52	0.00	**m.16269A/G**	0.52	0.00
m.16162A/G	2.62	4.71	**m.16270C/T**	6.81	4.71
m.16163A/G	2.62	2.62	**m.16271T/C**	0.00	1.05
m.16164A/G	0.00	0.52	**m.16274G/A**	0.52	0.00
m.16168C/T	0.52	0.00	**m.16278C/T**	3.14	1.57
m.16170A/G	0.52	0.00	**m.16284A/G**	0.52	0.00
m.16171A/T	0.26	0.00	**m.16286C/T**	0.52	0.52
m.16172T/C	4.19	3.14	**m.16288T/C**	0.52	0.00
m.16174C/T	1.05	0.52	**m.16289A/G**	0.52	0.00
m.16176C/T	0.26	0.52	**m.16290C/T**	0.00	0.52
m.16176C/G	0.52	0.52	**m.16291C/T**	2.62	0.52
m.16179C/T	0.52	0.00	**m.16292C/T** [b]	**0.00**	**4.19**
m.16180A/G	0.00	0.52	**m.16293A/T**	0.00	0.52
m.16182A/G	0.52	0.52	**m.16293A/G**	2.09	2.62
m.16182A/C	3.14	0.52	**m.16294C/T**	10.99	11.52
m.16183A/C	3.66	4.19	**m.16295C/T**	0.00	0.52
m.16185C/T	0.52	0.00	**m.16296C/T**	5.24	7.33
m.16186C/T	3.14	2.09	**m.16297T/C**	0.00	0.52
m.16187C/T	0.00	0.52	**m.16298T/C**	6.81	4.71
m.16188C/A	0.00	0.52	**m.16300A/G**	0.00	0.26
m.16188C/G	0.00	0.52	**m.16301C/T**	0.00	0.26
m.16189T/C [b]	**16.75**	**9.42**	**m.16304T/C**	10.47	12.04
m.16192C/T	4.71	3.14	**m.16309A/G**	0.00	0.52
m.16193C/T	1.57	2.09	**m.16311T/C**	13.61	15.97
m.16201C/T	0.52	0.00	**m.16316A/G**	0.00	1.05
m.16209T/C	2.62	4.19	**m.16318A/T**	0.52	0.52
m.16213G/A	0.00	1.57	**m.16319G/A**	0.00	0.52
m.16215A/G	0.52	0.00	**m.16320C/T**	1.05	1.57
m.16216A/G	0.52	0.00	**m.16324T/C**	0.52	1.57
m.16217T/C	0.00	0.52	**m.16325T/C**	0.00	1.57
m.16218C/T	0.52	0.52	**m.16327C/T**	0.00	0.52
m.16219A/G	0.00	1.57	**m.16335A/G**	0.00	0.79
m.16221C/T	1.05	0.52	**m.16342T/C**	0.52	1.57
m.16222C/T	3.14	1.57	**m.16343A/G**	0.52	1.05
m.16223C/T	5.76	7.85	**m.16344C/T**	0.52	0.00
m.16224T/C	5.24	7.85	**m.16353C/T**	0.00	0.52
m.16230A/G	0.00	0.52	**m.16354C/T**	0.52	0.00
m.16231T/C	1.57	1.57	**m.16355C/T**	1.57	1.57
m.16234C/T	1.05	2.09	**m.16356T/C**	6.28	5.76
m.16235A/G	0.52	0.00	**m.16357T/C**	0.52	0.00
m.16239C/T	0.00	0.52	**m.16360C/T**	0.00	0.52
m.16240A/G	0.00	0.52	**m.16362T/C**	7.85	8.90
m.16243T/C	0.52	0.52	**m.16366C/T**	0.52	1.05
m.16245C/T	0.52	1.57	**m.16390G/A**	2.09	2.62
m.16247A/G	0.00	0.52	**m.16391G/A**	2.09	2.09
m.16248C/T	0.52	0.52	**m.16398G/A**	1.05	0.00
m.16249T/C	1.05	0.52	**m.16399A/G**	3.66	2.88
m.16256C/T	4.19	3.14	**m.16482A/G**	1.05	2.62
m.16258A/C	0.52	0.00	**m.16497A/G**	0.00	1.05
m.16258A/G	0.00	0.52	**m.16519T/C**	66.49	70.16
m.16260C/T	0.52	0.52	**m.16526G/A**	2.62	1.05
m.16261C/T	4.19	4.19	**m.16527C/T**	0.52	0.00
m.16263T/C	1.57	1.05			

[a] variant has not been described previously based on www.mitomap.org, *last edited on Apr 23, 2013*

[b] nominal p<0.05, Fisher's exact test, two-sided

grey shaded frequency: point heteroplasmy was detected in either one case or one control

Table 4.10 Detected variants (complex nucleotide exchanges, insertions and deletions) by re-sequencing of mtDNA and frequencies in cases and controls

Detected Variants	Frequency cases [%] n=191	Frequency controls [%] n=191
m.16183A/CC	2.62	1.05
m.16183A/CCC [a]	0.52	0.00
m.16189T/CC [a]	1.05	2.62
m.43_44insG [a]	1.05	0.00
m.297_298insC [a]	0.00	0.52
m.309_310insC	34.03	39.79
m.309_310insCC	14.14	12.57
m.310_311insTC	0.52	1.57
m.315_316insC	95.29	97.91
m.315_316insCC	1.05	0.00
m.315_316insCCC	0.52	0.00
m.451_452insT	0.52	0.52
m.514_515insAC	5.24	8.38
m.514_515insACAC	3.14	1.05
m.514_515insACACAC	0.52	0.52
m.567_568insC	1.05	0.00
m.567_568insCCC	0.52	1.05
m.567_568insCCCC	0.52	0.00
m.567_568insCCCCC	0.52	3.14
m.576_577insCA [a]	0.52	0.00
m.16193_16194insC	0.52	0.00
m.310delT	1.05	0.00
m.311_313delCCC	0.52	0.00
m.498delC	2.62	1.05
m.513_514delGC [a]	0.00	0.52
m.515_516delAC	10.99	10.99
m.568delC [a]	0.52	0.00
m.16189delT	0.52	0.00
m.16257delC	0.52	0.00

[a] variant has not been described previously based on www.mitomap.org, *last edited on Apr 23, 2013*

In addition, length heteroplasmies, i.e. mixtures of various lengths of a certain mtDNA region in one individual, were found at four locations: in individuals with (1) an uninterrupted poly-C tract at m.16184 to m.16193 of at least 10 Cs through m.16189T/C(C), (2) alterations at the CCCCCCCTCCCCC-tract at m.303 to m.315, (3) insertions of AC or ACAC at an AC-repeat at m.515 to m.524 and (4) insertions of three to five Cs at a poly-C tract at m.568 to m.573 (Fig. 4.3).

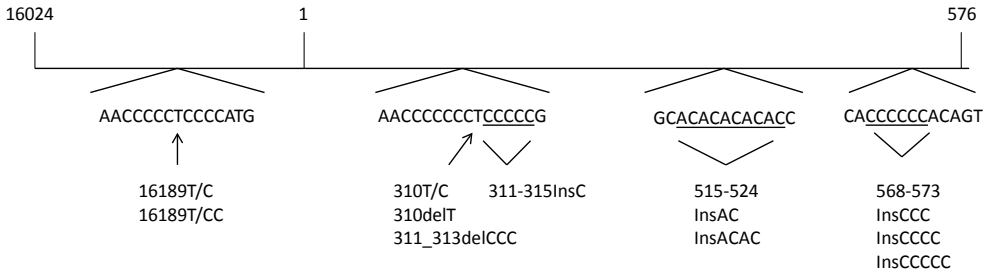

Figure 4.3 Detected length heteroplasmies

Length heteroplasmies, i.e. mixtures of various lengths of a certain mtDNA region in one individual, were found at four locations in the D-loop (m.16024 o m.576), predominantly by alterations of these as indicated at poly-C tracts. Numbering according to rCRS.

The frequency of each detected variant ranged from 0.26 % (n=1) to 98.69 % (n=377). Three variants were found in more than two thirds of all individuals (m.263A/G, m.315_316insC and m.16519T/C; Table 4.9 and Table 4.10). In both cases and controls, on average 8.3 variants per individual were detected (p=0.989; Fig. 4.4).

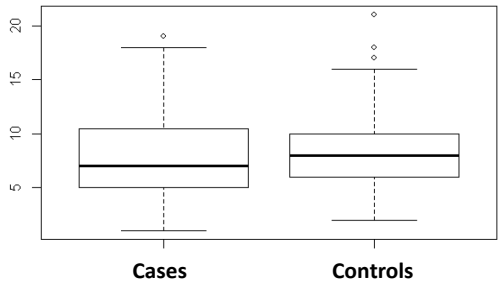

Figure 4.4 Comparison of mean number of D-loop deviations per individual between cases and controls

Comparing the frequencies of each variant between cases and controls, two variants were nominally associated with obesity (Table 4.9). m.16292C/T was only found in eight controls (p=0.0072), but five of these eight controls belong to haplogroup W, the frequency of which was significantly higher in the controls in the initial CC sample and borderline significant (p=0.062, Table 4.8) among the D-loop sample. The second nominally associated variant, m.16189T/C, was more frequent in the cases (17 % vs. 9 %; p=0.048, Table 4.9). The transition at m.16189 might create an uninterrupted poly-C tract of 10 Cs (m.16184 to m.16193; Fig. 4.3), in case of no further transition in any of the other Cs in this C-tract. Comparing only individuals with the uninterrupted C-tract, which was done by Parker et al. (2005) for BMI in an Australian cohort, a tendency of overrepresentation in the cases remained (15 % vs. 10 %).

Finally, the mean number of variants per case was not different compared with the mean number of variants per control in 23 functionally relevant locations of the D-loop (Table 4.11).

Table 4.11 Comparison of mean number of variants in functionally relevant D-loop locations between cases and control

Location [a]	Location start	Location end	Number of detected variants in location	Mean number of variants per case	Mean number of variants per control	p-value [b]
HV1a	m.16024	m.16365	126	2.39	2.39	0.977
HV1b	m.16024	m.16382	127	2.39	2.40	0.953
HV2a	m.73	m.340	69	4.25	4.33	0.686
HV2b	m.57	m.371	73	4.30	4.38	0.707
HV3	m.438	m.574	35	0.75	0.64	0.232
Mt5 (CE)	m.16194	m.16208	1	0.01	0	0.319
Mt3 (L-strand CE)	m.16499	m.16506	0	0	0	NaN
Mt4 (L-strand CE)	m.371	m.379	1	0	0.01	0.319
Mt3 (H-strand CE)	m.384	m.391	1	0.02	0	0.083
mtTF1 BS	m.233	m.260	12	0.12	0.09	0.537
mtTF1 BS	m.276	m.303	5	0.12	0.12	1.000
mtTF1 BS	m.418	m.445	0	0	0	NaN
mtTF1 BS	m.525	m.552	4	0.01	0.02	0.178
LSP (including mtTF1 BS)	m.392	m.445	1	0.01	0.01	0.563
HSP1	m.545	m.567	4	0.02	0.01	0.414
HSP1 (including mtTF1 BS)	m.525	m.567	6	0.02	0.02	0.738
HSP2	m.632	m.655	1	0	0.01	0.319
CSB1	m.210	m.234	8	0.12	0.06	0.101
CSB2	m.299	m.315	9	1.48	1.52	0.368
CSB3	m.346	m.363	1	0.01	0	0.319
ETAS1	m.16081	m.16140	12	0.39	0.36	0.578
ETAS2	m.16294	m.16356	25	0.59	0.73	0.103
TAS	m.16157	m.16172	8	0.12	0.11	0.878

[a] description/function of location in 1.4.2.5 of *Introduction*

[b] t-test, 2-sided

CE, control element; CSB, conserved sequence block; ETAS, extended termination associated sequence, HSP, heavy strand promoter; HV, hypervariable region; LSP, light strand promoter; mtTF1 BS, mitochondrial transcription factor A (TFAM) binding side (TFAM formerly known as mtTF1); NaN, not a number; TAS, termination associated sequence

4.2 Gene set enrichment analysis of nuclear encoded mitochondrial genes

The GSEA was done in cooperation with Dipl.-Math. Ivonne Jarick and PD Dr. André Scherag who performed statistical analysis. My contribution was (1) the check of the cluster graphs for clear allele separation of all ~1,000 SNPs per sample from gene set 1 (p. 38f); (2) a literature re-check of the literature-based gene set 1 at the beginning of the analysis (p. 49f). (3) Together with Ivonne Jarick, I decided to use the population-based sample KORA as a CC sample to increase power of the sample (p. 34, p. 98). I suggested (4) to meta-analyze data-sets of all three sample (p. 52f) and (5) to calculate the LD between best SNPs of the same gene of gene set 1 to conclude whether signals were independent (p.53).

4.2.1 Discovery

All 16 and 91 genes of gene sets 1 and 2, respectively, had genotyped SNPs within the genes and their extended boundaries. Due to physical clustering, two genes of the 91 genes of the second gene set were removed. All genes but one from gene set 3 had SNPs within the genes and their extended boundaries and 85 had to be removed due to physical clustering. In the reference gene set of all autosomal genes, 55 of them had no genotyped SNPs within the gene and their extended gene boundaries. 7,445 were removed due to physical clustering. Thus, the effective number of genes for gene sets 1, 2, 3 and all autosomal genes was 16, 89, 880 and 10,180, respectively. Genes of gene set 1, 2 and 3 were covered by 0.1 %, 0.4 % and 4.9 % of the 703,015 autosomal SNPs included in the GSEA. The reference gene set of all autosomal genes was covered by 73.0 % (Table 4.12).

Only the first gene set of 16 nuclear regulators of mitochondrial genes was enriched for obesity association signals ($P^{GSEA,WMW}$=0.0075, $P^{GSEA,KS}$=0.0195, $P^{GSEA,t}$=0.0053). This enrichment was above the 50th percentile of the set of all autosomal gene-wise p-values ($P^{GSEA,50}$=0.0103, Fig. 4.5 and Table 4.12). The enrichment remained significant after Bonferroni correction for the three gene sets tested except for the Kolmogorov-Smirnov-Test. Applying MAGENTA, enrichment was found above the 50th percentile as well ($p^{MAGENTA,50}$=0.0099). In addition, enrichment was also found above the 75th percentile ($p^{MAGENTA,75}$=0.0074), which only tended to be significant using the permutation GSEA approach ($p^{GSEA,75}$=0.0796; Table 4.12).

4.2.2 Confirmation

In the family-based GWAS sample, enrichment of gene set 1 was not confirmed ($P^{GSEA,50}$=0.5991, $P^{GSEA,WMW}$=0.7879, $P^{GSEA,KS}$=0.7930, $P^{GSEA,t}$=0.7588; Fig. 4.6A and Table 4.13).

In KORA-CC, by contrast, the initial finding of enrichment of gene set 1 above the 50th percentile was confirmed ($P^{GSEA,50}$=0.0398, $P^{GSEA,WMW}$=0.0260, $P^{GSEA,KS}$=0.0431, $P^{GSEA,t}$=0.0211, and $p^{MAGENTA,50}$=0.0405; Fig. 4.6B and Table 4.13).

The SNP with minimal P_g was different for most genes of gene set 1 in the three analyzed samples, and LD between best SNPs of one gene was generally quite low (Table 4.14 and Table 4.16).

4.2.3 Meta-analysis

Applying MAGENTA for the meta-analysis of all three samples, significant enrichment for gene set 1 was found ($P^{MAGENTA,WMW}$=0.0357). This enrichment was found above the 75th percentile ($P^{MAGENTA,75}$=0.0251; Fig. 4.6C and Table 4.15). Hence, the effect of enrichment of association signals in gene set 1 remained stable in the meta-analysis.

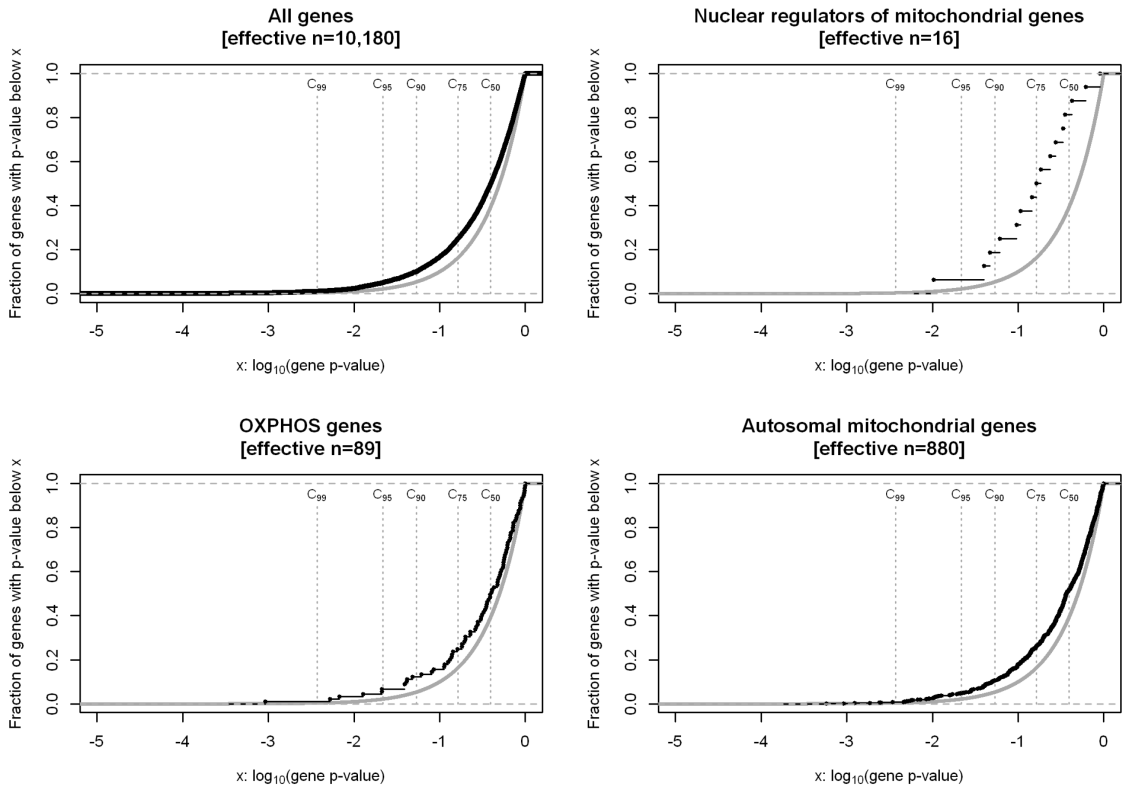

Figure 4.5 Empirical cumulative distribution functions (ECDF) of gene p-values Pg in all autosomal genes and in three gene sets of nuclear encoded mitochondrial genes (discovery)

The grey line represents the ECDF of the uniform distribution (null hypotheses of no association) in each panel and the black line represents the ECDF of the respective gene set. The CC sample of 453 extremely obese children and adolescents and 435 lean adult controls was used as the discovery sample. P_g, gene-wise corrected p-value; *Figure from Knoll et al. (2013).*

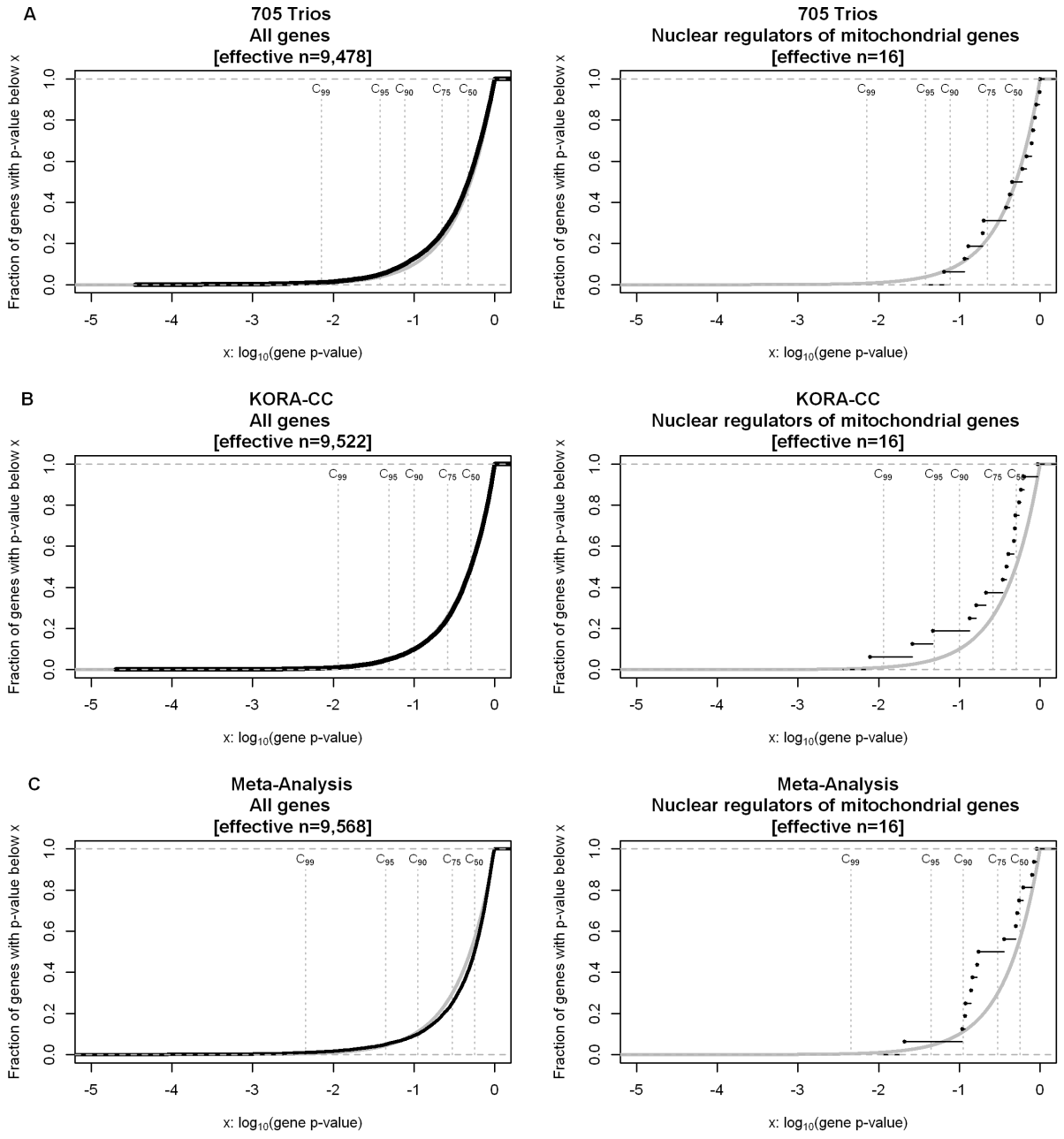

Figure 4.6 Empirical cumulative distribution functions (ECDF) of gene p-values Pg in all autosomal genes and gene set 1 (confirmation)

The initial finding of enrichment of gene set 1 was independently confirmed in the family-based sample of 705 trios (**A**) and in 463 obese cases and 483 normal weight or lean controls of the KORA-CC sample (**B**). Moreover, a meta-analysis of the discovery and both confirmation samples has been performed (**C**). The grey line represents the ECDF of the uniform distribution (null hypotheses of no association) in each panel and the black line represents the ECDF of the respective gene set. P_g, gene-wise corrected p-value; *Figure from Knoll et al. (2013).*

Table 4.12 GSEA and MAGENTA for obesity in the CC sample (discovery)

Gene set	total number of genes	effective number of genes	number of SNPs involved	% of all autosomal SNPs involved [a]	$P^{GSEA,WMW}$ Wilcoxon-Mann-Whitney test	$P^{GSEA,KS}$ Kolmogorov-Smirnov-test	$P^{GSEA,t}$ t-test	$P^{GSEA,95}$ 95th percentile cut-off test [b]	$P^{GSEA,75}$ 75th percentile cut-off test [c]	$P^{GSEA,50}$ 50th percentile cut-off test [d]	$P^{MAGENTA,WMW}$ Wilcoxon-Mann-Whitney test [e]	$P^{MAGENTA,95}$ 95th percentile cut-off test [b]	$P^{MAGENTA,75}$ 75th percentile cut-off test [c]	$P^{MAGENTA,50}$ 50th percentile cut-off test [d]
1) Nuclear regulators of mitochondrial genes	16	16	1,014	0.1	**0.0075**	**0.0195**	**0.0053**	0.5644	0.0796	**0.0103**	**0.0043**	0.575	**0.0074**	**0.0099**
2) OXPHOS genes	91	89	2,781	0.4	0.6225	0.8586	0.6374	0.2873	0.5643	0.5834	0.8447	0.6565	0.7495	0.7369
3) Nuclear-encoded mitochondrial genes	966	880	35,223	4.9	0.3841	0.2502	0.4104	0.6437	0.1905	0.1196	0.8969	0.5287	0.7372	0.7577
All autosomal genes [f]	17,680	10,180	521,469	73.0	-	-	-	-	-	-	-	-	-	-

[a] 703,015 autosomal SNPs were included

[b] cut-off = 0.0216

[c] cut-off = 0.1631

[d] cut-off = 0.3951

[e] exact GSEA Wilcoxon-Mann-Whitney test

[f] reference gene set

GSEA and MAGENTA p-values below 0.05 are highlighted in bold

Table adapted from Knoll et al. (2013).

Table 4.13 GSEA and MAGENTA for the gene set of 16 nuclear regulators of mitochondrial genes (confirmation)

Gene set	total number of genes	effective number of genes	number of SNPs involved	% of all autosomal SNPs involved [a]	$P^{GSEA,WMW}$ Wilcoxon-Mann-Whitney test	$P^{GSEA,KS}$ Kolmogorov-Smirnov-test	$P^{GSEA,t}$ t-test	$P^{GSEA,95}$ 95th percentile cut-off test [b]	$P^{GSEA,75}$ 75th percentile cut-off test [c]	$P^{GSEA,50}$ 50th percentile cut-off test [d]	$P^{MAGENTA,WMW}$ Wilcoxon-Mann-Whitney test [e]	$P^{MAGENTA,95}$ 95th percentile cut-off test [b]	$P^{MAGENTA,75}$ 75th percentile cut-off test [c]	$P^{MAGENTA,50}$ 50th percentile cut-off test [d]
Family-based sample [f]														
Gene set 1	16	16	919	0.1	0.7879	0.7930	0.7588	1.0000	0.3711	0.5991	0.6817	1	0.6024	0.7683
All autosomal genes [g]	17,680	9,478	471,858	73.5	-	-	-	-	-	-	-	-	-	-
KORA-CC [h]														
Gene set 1	16	16	933	0.1	**0.0260**	**0.0431**	**0.0211**	**0.0432**	0.1939	**0.0398**	**0.0083**	0.1918	0.1888	**0.0405**
All autosomal genes	17,680	9522	483,270	73.3	-	-	-	-	-	-	-	-	-	-

[a] 641,991 and 659,502 autosomal SNPs were included in family-based sample and KORA-CC, respectively

[b] Trios: cut-off = 0.0382, KORA-CC: cut-off = 0.0486

[c] Trios: cut-off = 0.2216, KORA-CC: cut-off = 0.2611

[d] Trios: cut-off = 0.4687, KORA-CC: cut-off = 0.5085

[e] exact GSEA Wilcoxon-Mann-Whitney test

[f] 705 trios

[g] reference gene set

[h] BMI ≥ 30 (cases) vs. BMI < 25 (controls)

GSEA and MAGENTA p-values below 0.05 are highlighted in bold

Table adapted from Knoll et al. (2013)

Table 4.14 Best SNPs of nuclear regulators of mitochondrial genes (gene set 1) in each sample and linkage disequilibrium between best SNPs of the three different study samples

	Discovery					Confirmation (1)				Confirmation (2)				
Sample	453 cases (children/adolescents) and 435 controls (adults)					705 family-based trios (index: child/adolescent)				463 cases and 483 controls (KORA-CC) [a] (adults)				
Gene ID	**Number of SNPs in gene region**	**Best SNP in gene [b]**	**Best SNP p-value $P_{g;\,min}$**	**Gene p-value P_g**	**LD: best SNP in Trios – best SNP in CC $[r^2]$ [c]**	**Number of SNPs in gene region**	**Best SNP in gene [b]**	**Best SNP p-value $P_{g;\,min}$**	**Gene p-value P_g**	**Number of SNPs in gene region**	**Best SNP in gene [b]**	**Best SNP p-value $P_{g;\,min}$**	**Gene p-value P_g**	**LD: best SNP in Trios – best SNP in CC $[r^2]$ [c]**
ESRRA	12	rs2429455[§]	0.0065	0.0616	0.403	10	rs1059440[§]	0.0719	0.3778	11	rs11231740[§]	0.0681	0.3853	0.129
ESRRG	335	rs2185226[#]	0.0035	0.4240	0.001	316	rs12033461[#]	0.0155	0.8585	320	rs11577585[#]	0.0185	0.9301	0
GABPA	36	rs2051180[§]	0.0022	0.0400	0.945	35	rs11087972[§]	0.1216	0.7873	32	rs7284014[§]	0.0012	0.0261	0.024
GABPB1	37	rs4775886[§]	0.0205	0.3542	0	32	rs12910368[#]	0.3374	0.9953	35	rs16963477[&]	0.0071	0.1336	0
GABPB2	19	rs3754210[§]	0.2138	0.9071	0.072	**19**	**rs4970989[§]**	**0.0052**	**0.0644**	19	rs267738[§]	0.0539	0.4848	NA
MEF2A	59	rs7173943[§]	0.0057	0.1643	0.001	60	rs4313794[§]	0.0185	0.4216	59	rs7175248[§]	0.0156	0.4036	0.065
MYC	60	rs11990827[§]	0.0094	0.2750	0.001	55	rs4395860[&]	0.0422	0.6785	59	rs13252644[§]	0.0194	0.4720	0
NRF1	51	rs2693737[&]	0.0104	0.2379	0.013	52	rs9792084[§]	0.1758	0.9805	47	rs11771549[§]	0.0418	0.6346	0.012
NRIP1	47	rs2776043[#]	0.0052	0.1075	0.005	46	rs17274722[&]	0.0506	0.6005	**43**	**rs10482862[&]**	**0.0003**	**0.0078**	**0.004**
PPARA	55	rs3744749[§]	0.0358	0.6169	0.009	46	rs12170325[§]	0.0084	0.1999	48	rs4253754[#]	0.0084	0.2128	0.008
PPARD	40	rs9658085[#]	0.0097	0.1854	0.016	22	rs2894401[&]	0.1644	0.8156	35	rs2267666[#]	0.0201	0.3453	0.134
PPARGC1A	75	rs17574213*	0.0034	0.1456	0.001	71	rs10517032[§]	0.0641	0.8988	67	rs17576576[§]	0.0163	0.4964	0.009
PPARGC1B	110	rs10069462[§]	0.0081	0.3352	0.017	102	rs7713955[§]	0.0141	0.4494	104	rs10065816[§]	0.0166	0.5462	0.145
SIRT1	**16**	**rs7895833[d,§]**	**0.0011**	**0.0104**	**0.304**	16	rs10509291[§]	0.0190	0.1282	15	rs16924888[d,§]	0.0053	0.0471	0.01
SP1	16	rs4759082[§]	0.0126	0.0964	same SNP	11	rs4759082[§]	0.0436	0.1949	14	rs2016266[§]	0.0207	0.1624	0.209
YY1	25	rs8007801[§]	0.0034	0.0470	0.011	26	rs9291[&]	0.0079	0.1156	25	rs2766692[§]	0.0732	0.5759	0.015

[a] BMI ≥ 30 (cases) vs. BMI < 25 (controls); [b] Location of SNP: *, exonic; [#], intronic; [§], upstream of gene and [&], downstream of gene; [c] Linkage Disequilibrium (LD) was calculated in the parents of the family-based GWAS sample by use of HaploView 4.2; [d] LD between rs7895833 and rs16924888: r^2=0.581; best gene of each sample is indicated in bold letter. *ESRRA*, Estrogen related receptor alpha; *ESRRG*, Estrogen related receptor gamma; *GABPA*, GA-binding protein alpha subunit; *GABPB1*, GA-binding protein beta subunit 1; *GABPB2*, GA-binding protein beta subunit 2; *MEF2A*, Myocyte-specific enhancer factor 2A; *MYC*, Myelocytomatosis viral oncogene homolog (avian); *NRF1*, Nuclear respiratory factor 1; *NRIP1*, Nuclear receptor-interacting protein 1; *PPARA*, Peroxisome proliferator-activated receptor alpha; *PPARD*, Peroxisome proliferator-activated receptor delta; *PPARGC1A*, Peroxisome proliferator-activated receptor gamma coactivator 1 alpha; *PPARGC1B*, Peroxisome proliferator-activated receptor gamma coactivator 1 beta; *SIRT1*, Sirtuin 1; *SP1*, Specificity protein 1; *YY1*, Transcriptional repressor protein YY. *Table adapted from Knoll et al. (2013)*.

Table 4.15 MAGENTA for the gene set of 16 nuclear regulators of mitochondrial genes (meta-analysis)

Gene set	total number of genes	effective number of genes	number of SNPs involved	% of all autosomal SNPs (728,844) involved [a]	$P^{MAGENTA,WMW}$ Wilcoxon-Mann-Whitney test [b]	$P^{MAGENTA,95}$ 95th percentile cut-off test [c]	$P^{MAGENTA,75}$ 75th percentile cut-off test [d]	$P^{MAGENTA,50}$ 50th percentile cut-off test [e]
Gene set 1	16	16	1,036	0.1	**0.0357**	0.5587	**0.0251**	0.1052
All autosomal genes [f]	17,680	9,568	559,705	76.8	-	-	-	-

[a] for all SNPs that are available in at least one sample

[b] exact GSEA Wilcoxon-Mann-Whitney test

[c] cut-off = 0.0443

[d] cut-off = 0.2969

[e] cut-off = 0.5619

[f] reference gene set

MAGENTA p-values below 5 % are highlighted in bold

Table adapted from Knoll et al. (2013).

Table 4.16 Best SNPs of nuclear regulators of mitochondrial genes (gene set 1) in meta-analysis and linkage disequilibrium to best SNP of family-based trios

Gene ID	Number of SNPs in gene region	Best SNP in gene [b]	Best SNP p-value $P_{g;\,min}$	Gene p-value P_g	LD: best SNP in Trios – best SNP in CC [r^2] [c]
ESRRA	12	rs4930702[&]	0.0161	0.1662	0.004
ESRRG	349	rs7531250[#]	0.0090	0.8458	0.003
GABPA	37	rs2829866[§]	0.0048	0.1191	0.206
GABPB1	38	rs16963477[&]	0.0046	0.1091	0
GABPB2	22	rs7526955[§]	0.0666	0.6166	0.243
MEF2A	65	rs7173943[§]	0.0039	0.1462	0.001
MYC	62	rs12155669[&]	0.0923	0.9125	0.384
NRF1	56	rs11771549[§]	0.0756	0.8021	0.012
NRIP1	49	rs10482862[&]	0.0047	0.1730	0.004
PPARA	55	rs4253655[#]	0.0219	0.4980	0.021
PPARD	40	rs9658085[#]	0.0181	0.3595	0.016
PPARGC1A	78	rs7682906[§]	0.0182	0.5493	0.089
PPARGC1B	114	rs11746690[#]	0.0096	0.5180	0.016
SIRT1	**17**	**rs17712705[§]**	**0.0019**	**0.0209**	**0.146**
SP1	16	rs4759082[§]	0.0101	0.1173	same SNP
YY1	26	rs9291[&]	0.0085	0.1391	same SNP

[a] Location of SNP: *, exonic; [#], intronic; [§], upstream of gene and [&], downstream of gene; [b] SNP-wise p-values of the meta-analysis were derived by application of the METAL software [c] Linkage Disequilibrium (LD) was calculated in the parents of the family-based GWAS sample by use of HaploView 4.2; best gene is indicated in bold letter.

ESRRA, Estrogen related receptor alpha; *ESRRG*, Estrogen related receptor gamma; *GABPA*, GA-binding protein alpha subunit; *GABPB1*, GA-binding protein beta subunit 1; *GABPB2*, GA-binding protein beta subunit 2; *MEF2A*, Myocyte-specific enhancer factor 2A; *MYC*, Myelocytomatosis viral oncogene homolog (avian); *NRF1*, Nuclear respiratory factor 1; *NRIP1*, Nuclear receptor-interacting protein 1; *PPARA*, Peroxisome proliferator-activated receptor alpha; *PPARD*, Peroxisome proliferator-activated receptor delta; *PPARGC1A*, Peroxisome proliferator-activated receptor gamma coactivator 1 alpha; *PPARGC1B*, Peroxisome proliferator-activated receptor gamma coactivator 1 beta; *SIRT1*, Sirtuin 1; *SP1*, Specificity protein 1; *YY1*, Transcriptional repressor protein YY.

Table adapted from Knoll et al. (2013).

5 Discussion

Based on the central role of mitochondria in energy metabolism as cellular power plants, the overall objective of the present thesis was to assess whether variation in mitochondrial genes is associated with obesity. Therefore, variation in both mtDNA and nuclear-encoded mitochondrial genes has been investigated.

5.1 Variation in mitochondrial DNA

mtDNA was specifically analyzed, as empirical studies have shown greater correlations in BMI between mothers and their offspring than between fathers and their offspring as well as greater correlations in BMI between maternal half-brothers compared with paternal half-brothers (Hebebrand et al. 2001a, Magnusson and Rasmussen 2002), and variation in the exclusively maternally inherited mtDNA might contribute to this parental effect. Therefore, common SNPs of the mtDNA coding region as well as D-loop variants have been addressed in the present analysis.

5.1.1 Common variation in mtDNA in association with obesity

Analysis of up to 40 array-based common mtDNA SNPs did not lead to any robust association of either a single mtDNA SNP or a haplogroup. Although in the discovery, a total of five SNPs and two haplogroups (J, W) reached nominally significant association, in both the whole sample and stratified by gender, independent confirmation of any of these SNPs was not possible. Moreover, for both haplogroups and all but one SNP the direction of effect was opposite between discovery and confirmation samples, underscoring that the initial associations might have been spurious.

These results are consistent with those of Grant et al. (2012) who did not find association between common mtDNA SNPs and obesity in neither 3,580 European American nor 3054 African American children. Yang et al. (2011b), by contrast, found haplogroup X to be associated with lower BMI in 2,286 unrelated population-based adult Caucasians, but the study lacked a confirmation in an independent sample. In the present study, haplogroup X was more frequent in the extremely obese cases in the discovery. In the population-based adults, however, though not significant, the direction of effect was consistent with that reported by Yang et al. (2011b).

Haplogroup W and m.8994G/A initially showed the strongest associations in all individuals and stratified by gender. However, results were not independent of each other,

as all individuals of haplogroup W carry the minor allele A of m.8994G/A. This association might have resulted from the composition of the discovery CC sample. As indicated in Fig. 5.1, some (infrequent) haplogroups as for instance haplogroups X and W, seem to be unequally distributed in Germany (Fernandes et al. 2012). The cases of the discovery sample were recruited in different parts of Germany/Austria (i.e. Bad Orb, Berchtesgaden, Gießen, Marburg, Murnau, Ulm and Vienna). The controls, however, were recruited at the University of Marburg. Students coming from or near Marburg thus might have been overrepresented in the controls. Hence, it is possible that the initial association might be rather conditioned by different regional haplogroup frequency distributions, than by the trait of investigation.

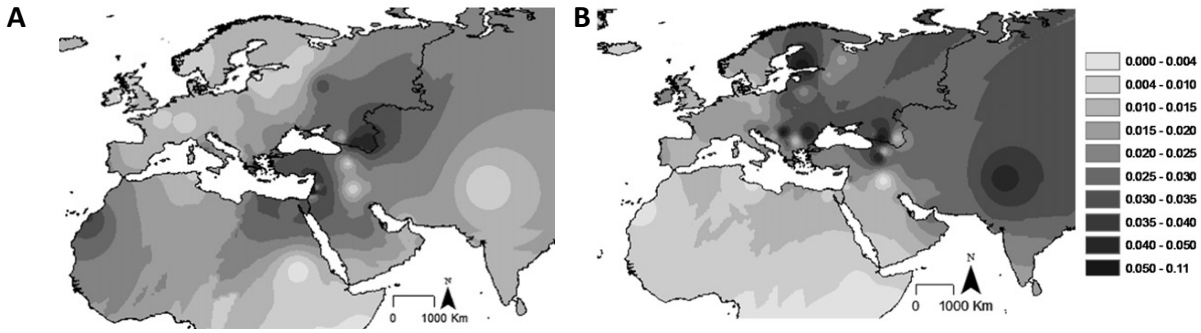

Figure 5.1 Geographic distribution of frequencies of haplogroup X (A) and W (B)

HV1 data (literature based pertaining to European individuals) were used for visualization of the geographical distribution of haplogroup X and W. Distribution plot created by ‘‘Spatial Analyst Extension of ArcView version 3.2. *Reprinted from The American Journal of Human Genetics, 90, Fernandes et al., The Arabian Cradle: Mitochondrial Relicts of the First Steps along the Southern Route out of Africa, pp.347-355, © 2012, with permission from Elsevier.*

On the other hand, it might be that the initial findings were not confirmed, as mtDNA SNPs and/or haplogroups associated with early onset obesity (discovery sample) might not be associated with adult obesity (confirmation sample). For instance, autosomal SNPs between *TNKS* (tankyrase, TRF1-interacting ankyrin-related ADP-ribose polymerase gene) and *MSRA* (methionine sulfoxide reductase A gene) were found to be associated with obesity only in children/adolescents (Scherag et al. 2010).

5.1.2 Evaluation of haplogroup assignment

Haplogroups were determined using the genotype information of up to 40 SNPs from the SNP array and HaploGrep. HaploGrep is a web application (Kloss-Brandstätter et al. 2011)

based on Phylotree built 11, a comprehensive phylogeny of global human mtDNA variation (van Oven and Kayser 2009).

As expected, haplogroup H was the most prevalent haplogroup (41 % to 44 %; Table 5.1), followed by U (14 % to 22 %), T (9 % to 12 %), J (8 % to 11 %), V (2.6 % to 3.4 %), and K (1 % to 8 %) considering each discovery sample (CC sample and index patients of trios) and the three population-based samples separately (Table 5.1). The percentages found of these haplogroups were mainly similar to those reported among West Europeans (www.mitomap.org) or Germans (Pliss et al. 2006), except for haplogroup K and U in SHIP (Table 5.1). Haplogroup K was only found at 0.6 % (vs. 6.2 % to 7.8 % in the other samples, 5 % and 7.5 % among West Europeans and Germans, respectively), while haplogroup U was found at 22.2 % (vs. 14.4 % to 16.3 % in the other samples, 18 % and 13.5 % among West Europeans and Germans, respectively). An explanation for these deviations might be the different number of SNPs available for haplogroup assignment. For SHIP, only 32 SNPs passed QC, while at least 35 from the other samples could be used. As depicted in the simplified haplogroup tree (Fig. 1.8), haplogroup K is a side branch of U (van Oven and Kayser 2009). To reach K, m.10550G is necessary (Achilli et al. 2005, González et al. 2006, Behar et al. 2006, 2008, van Oven and Kayser 2009). This SNP failed QC only in SHIP, and consequently, all individuals which would have been assigned to K by m.10550G, “remained” in U.

Table 5.1 Distribution of haplogroup frequencies (in %) in present study samples and among West Europeans and Germans

Haplogroup	CC sample [a]	Trios [a]	KORA [a, b]	SHIP [a, b]	POPGEN [a, b]	West Europeans [c]	Germans [d]
Number of SNP used for haplogroup determination	40	35	37	32	35	-	-
A	-	-	-	-	0.6	-	-
B	0.1	-	1.9	-	-	-	-
C	-	-	-	-	-	-	-
D	-	-	0.1	<0.05	0.1	1	0.6
F	-	-	-	-	-	0	-
G	-	-	-	-	-	-	-
H	44.4	44.5	41.6	41.2	41.2	41	48.6 [e]
I	-	-	-	-	-	2	1.8
J	10.6	8.4	8.9	9.6	11.1	9	8.4
K	6.5	6.2	6.3	0.6	7.8	5	7.5
L	-	0.1	0.1	<0.05	0.1	1	1.2
M	1.0	0.6	0.7	0.4	0.2	1	- [f]
N	2.6	3.5	2.5	3.3	3.0	1	0.6
P	0.1	-	-	<0.05	-	-	-
R	0.1	0.6	0.3	0.3	0.3	0	0.3
S	-	-	-	-	0.1	-	-
T	9.3	9.8	11.2	10.6	11.9	8	9.0
U	14.4	16.3	16.2	22.2	15.5	18	13.5
V	3.4	3.3	3.3	3.2	2.6	7	4.5
W	1.7	1.4	1.9	2.2	2.2	2	2.7
X	1.5	2.6	0.4	1.5	0.4	2	1.2
Z	-	-	-	<0.05	-	0	-
n. d.	4.3	2.7	4.8	4.8	2.8	3	-

[a] n. d., not defined represented by all individuals with a HaploGrep's quality <90 %

[b] whole population-based sample

[c] Estimations based on means from published frequencies (bearing in mind that sometimes not all haplogroups have been typed), compiled in 2009 for Mitomap only for illustrative purpose (www.mitomap.org, Ruiz-Pesini et al. 2007)

[d] data from Pliss et al. (2006) which are based on a total of n=333 German individuals

[e] including haplogroups HV and preHV

[f] all individuals of haplogroup M belonged to haplogroup D which directly branches off of M

Regarding haplogroups I, W and X occurring with a frequency of on average 2 % among West Europeans (www.mitomap.org) or Germans (Pliss et al. 2006), only W was detected at a similar and stable level in the samples of the present study (1.4 % to 2.2 %), while X had a larger deviation (0.4 % to 2.6 %), and I was not existent (Table 5.1). The absence of haplogroup I can be explained by the fact that neither SNPs at m.10034, m.16129, m.16391

leading to haplogroup I (Finnilä et al. 2001, Palanichamy et al. 2004, Derenko et al. 2007) nor any at a position leading to a sub-haplogroup of I were present on the SNP array (Fig. 5.2). Individuals of haplogroup I might have "remained" in N1, from where haplogroup I branches off, which would also explain the on average 2 % higher occurrence of haplogroup N (2.5 % to 3.5 %, Table 5.1) compared with the 1 % and 0.6 % found among West European and Germans, respectively. The low occurrence of haplogroup X in KORA and POPGEN might be explained by the absence of m.12705 (Fig. 5.3A), which lacked QC in these two samples.

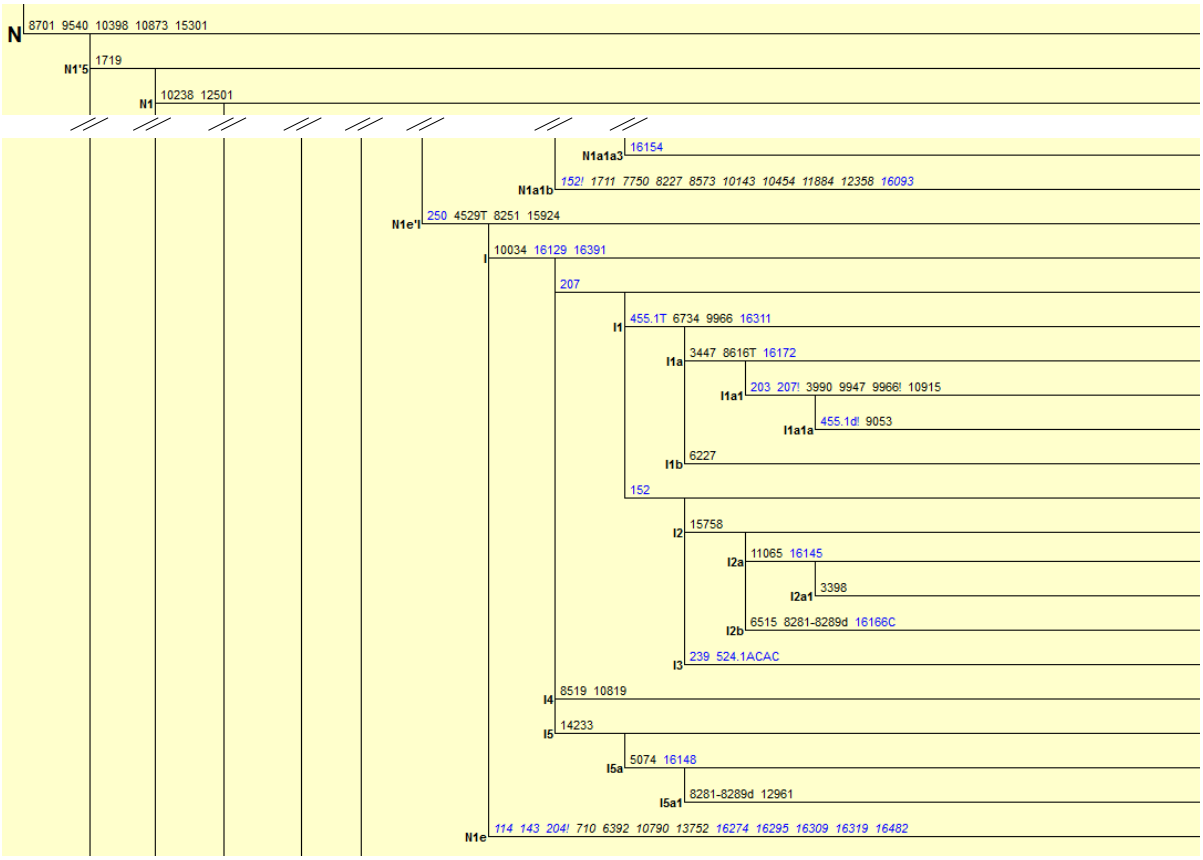

Figure 5.2 Haplogroup I as part of the phylogenic mtDNA tree

Screen shot from Phylotree built 11 (van Oven and Kayser 2009) for demonstration of haplogroup I as a side branch of haplogroup N1. Mutations/polymorphisms are transitions if not otherwise indicated by the respective nucleotide. Nucleotide positions (blue, control region; black, coding region) are relative to the rCRS (Anderson et al. 1981, Andrews et al. 1999). Italic nucleotide positions are preliminary and are likely to be further refined as additional sequences become available.

In addition, several non-European haplogroups were found; however, only in very small percentages, except for haplogroup B in KORA (1.9 %, Table 5.1). According to HaploGrep, most of these individuals had haplogroup B5b1a'b, which is defined by transitions at m.11146 and m.14470 (www.phylotree.org, van Oven and Kayser 2009). Only m.14470 was present on the SNP array, which also defines haplogroup X. Thus, as m.12705 was not available in KORA, a misclassification to B5b1a'b might have occurred (Fig. 5.3).

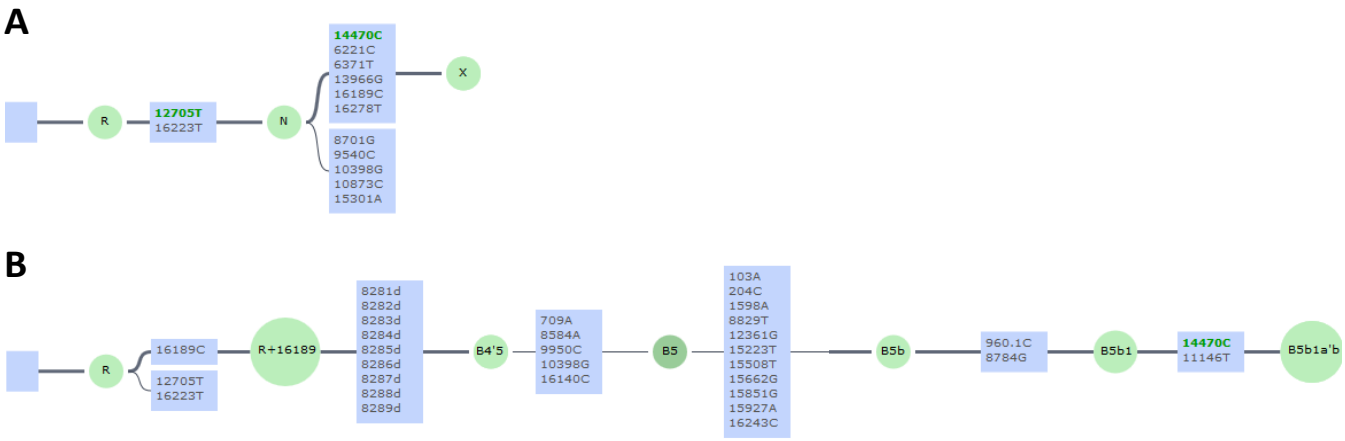

Figure 5.3 Potential misclassification of haplogroups by HaploGrep due to lacking genotype information

Screen shot of the web application HaploGrep (Kloss-Brandstätter et al. 2011) is based on Phylotree built 11 (van Oven and Kayser 2009) using one individual of the CC sample and KORA each. All deviations from the rCRS leading to a certain haplogroup (in green circles) are listed (in blue boxes). Deviations from rCRS marked in green were found in the individual. In the CC sample and KORA, 40 and 37 SNPs passed QC, respectively. The intersection was 35 SNPs. Both individuals were identical in genotype regarding these 35 SNPs. However, in the individual of the CC sample, m.12705 was available, and hence the presence of m.14470T/C resulted in assignment of haplogroup X (**A**) as HaploGrep's first choice, while the lack of m.12705 accompanied by the presence of m.14470T/C in the second individual has led to assignment of haplogroup B5b1a'b (**B**) as HaploGrep's first choice.

From eight individuals of haplogroup W (considering the 40 SNPs of the SNP array), the complete mtDNA was available to re-determine each individual's haplogroup by HaploGrep (Table 4.7, p. 65). Due to the greater amount of variants available after re-sequencing, haplogroups branching off from W (Fig. 5.4) could be determined. Though identical pertaining to SNP array based genotype information, different sub-haplogroups among these

individuals were detected. Nevertheless, this would not have had an impact on association testing, as for this purpose these sub-haplogroups would have also been assigned to major haplogroup W (Table 4.7, p. 65).

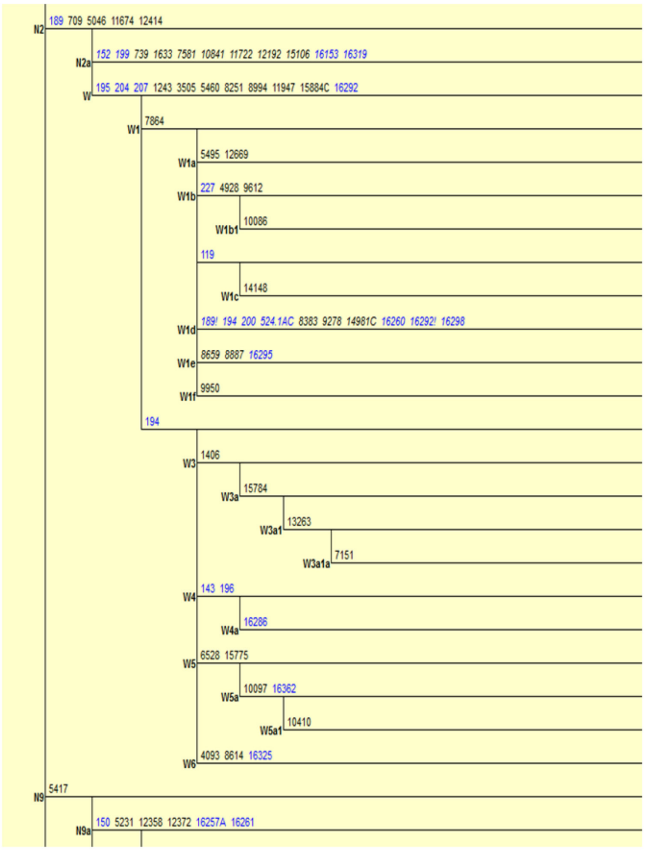

Figure 5.4 Haplogroup W as part of the phylogenic mtDNA tree

Screen shot from Phylotree built 11 (van Oven and Kayser 2009) to demonstrate sub-haplogroups of haplogroup W. Mutations/polymorphisms are transitions if not otherwise indicated by the respective nucleotide (e.g. 14981C is an A to C exchange at m.14981). Nucleotide positions (blue, control region; black, coding region) are relative to the rCRS (Anderson et al. 1981, Andrews et al. 1999). Italic nucleotide positions are preliminary and are likely to be further refined as additional sequences become available. Nucleotide positions with an exclamation mark refer to back transitions to the rCRS.

Moreover, a high accordance was found between the SNP array based major haplogroup assignment and the re-assignment based on the detected D-loop variants in the D-loop sample, even though most individuals had a HaploGrep's quality value below the required 90 %. Among the 10 individuals whose mtDNA was completely re-sequenced, quality values were lower using only D-loop variants compared with the use of all detected variants. There is no explanation for the low HaploGrep's quality values received by haplogroup determination using only D-loop data in the present analyses. Kloss-Brandstätter et al. (2011) – the authors of HaploGrep – provided mitochondrial test data of 60 European individuals. Of 28 of these individuals only D-loop data were available, of four data of the entire genome and of 28 D-loop data and selected coding region SNPs. In contrast to the results of the present analysis, quality values of haplogroup determination were even highest among the 28 individuals of which only D-loop data were available (~97 %), and lowest among the 28 of which D-loop data and selected coding region SNPs were available (~93 %).

Bandelt et al. (2012) recently compared available programs (mtDNA manager, MitoTool, HaploGrep) for automated haplogroup assignment from D-loop data, within which HaploGrep was evaluated as the most sophisticated. Regarding the results of the present study (major accordance of haplogroup distribution of present samples between the one expected for West Europeans and Germans, as well as equal assignment of one major haplogroup by use of either SNP array based data or detected variants by re-sequencing), HaploGrep can be considered a valid tool for the assignment of (major) haplogroups. Nevertheless, due to a limited number of available SNPs on the SNP array having passed QC, some haplogroups could not be detected or correctly assigned. Consequently, this might have slightly masked the association testing of some haplogroups with obesity.

5.1.3 D-loop variation in association with obesity

D-loop re-sequencing (Sanger) was initially performed in 192 cases and 192 controls (D-loop sample). Only 191 cases and 191 controls were used for analyses, as in one case and one control, 9 and 4 clearly visible point heteroplasmies, respectively, were found. Point heteroplasmies at more than one position in one individual may exist, but are rather infrequent, especially in blood cells (Calloway et al. 2000; Budowle et al. 2002). DNA of the mothers of these two individuals was not available to check whether these heteroplasmies might have been inherited. Contamination with DNA of another individual during DNA

isolation might also explain the occurrence of several point heteroplasmies (Andréasson et al. 2006), which is why these two individuals were excluded from the D-loop sample for further analyses.

The mean number of sequence deviations per individual (8.3) was comparable to that reported in a study, in which the D-loop of 200 unrelated German individuals was re-sequenced (~7.6 estimated from bar graphs; Lutz et al. 1998).

The mean number of variants per individual did not differ between cases and controls; neither did the mean number of variants in any of the 23 presumably functionally relevant locations of the mitochondrial D-loop. These regions are majorly involved in the transcription and replication of mtDNA. Thus, an accumulation of variants in these locations possibly having an impact on transcription and replication of mtDNA and potently energy metabolism or in the whole D-loop was not detected.

Among the 252 detected variants, frequencies of only two variants (m.16292C/T and m.16189T/C) differed nominally between cases and controls. m.16292C/T was found in only eight lean controls, of which five were of haplogroup W. As the initial association of haplogroup W could not be confirmed independently, this variant was not followed-up in a larger sample. m.16189T/C was more frequent in the (extremely) obese cases (17 % vs. 9 %). Considering all individuals with an uninterrupted poly-C tract in this region, as it was done by Parker et al. (2005), the tendency of an overrepresentation in the (extremely) obese cases was still found (15 % vs. 10 %). The uninterrupted poly-C tract was achieved by m.16189T/C, m.16189delT or m.16189T/CC without a further transition in any of the Cs between m.16184 and m.16193. In contrast to the results of the present study, Parker et al. (2005) found the uninterrupted poly-C tract to be associated with leanness in Australian mothers and their 20-year-old offspring.

m.16189T/C has also been associated with obesity-related traits like coronary artery disease among 1,963 Austrian individuals (Mueller et al. 2011) or insulin resistance among 251 English men (64 years of age; Poulton et al. 1998). Moreover, this variant has been associated with T2DM in 5,459 Asians (Park et al. 2008) and in 932 British individuals (Poulton et al. 2002). However, in a further British cohort (n=2,557) and a European meta-analysis (n=4,587), this variant has not been associated with T2DM (Chinnery et al. 2005). Whether the trend of m.16189T/C and/or the uninterrupted poly-C tract at m.16184 to m.16193 is a valid association with obesity could only be analyzed by a follow-up of this variant in further study groups.

The number of clearly identifiable point heteroplasmies was twofold higher in the lean individuals compared with the obese (9 vs. 4). At time of blood collection for DNA isolation, the lean controls were on average 12 years older than the obese cases. Sondheimer et al. (2011) found that point heteroplasmies accumulate significantly with progressing age. Hence, the difference seen between the lean and obese might be rather attributed to their distinct ages than to the trait of investigation (obesity). Altogether, in 3.4 % of all individuals of the present study, one clearly identifiable point heteroplasmy was found. Calloway et al. (2000) found point heteroplasmies in the HV2 in 2.4 % of individuals analyzing blood samples.

All four detected length heteroplasmies are known (Bendall and Sykes 1995, Szibor et al. 2007 and Forster et al. 2010). Sanger re-sequencing without prior cloning to separate the different alleles as performed in the present study is not an appropriate method to determine the exact pattern (number of varying Cs and their respective percentage) of length heteroplasmy, which is why a statistical comparison of their frequencies between cases and controls was not performed.

The most prevalent length heteroplasmy was found in HVI at m.16184 to m.16193. All individuals of the present study with an uninterrupted poly-C tract at m.16184 to m.16193, but none of the individuals with m.16189T/C and any further transition in this C tract showed length heteroplasmy. An uninterrupted poly-C tract is an unstable sequence which might result in strand slippage during the replication process and might generate C-tracts in variable length (length heteroplasmy); further transitions at this C tract, however, are preventive against slippage (Chinnery et al. 2005). The predominant length of this heteroplasmic tract was reported to be 10 to 12 Cs; shorter tracts of up to eight and longer of up to 14 Cs have been also described in some individuals (Bendall and Sykes 1995). This tract, which is nearly identical in and distinct between maternal lineages (Bendall and Sykes 1995), is located near the TAS element (m.16157 to m.16172). This element is potentially involved in the premature termination of H-strand synthesis forming the 7S DNA and hence the D-loop (Anderson et al. 1981, Sbisà et al. 1997, Roberti et al. 1998). Proteins were shown to bind at this sequence and the binding capacity and consequently transcription/replication of mtDNA might be influenced by the presence of an uninterrupted C tract (Poulton et al. 2002). Liou et al. (2010) reported that mtDNA content in leucocytes of healthy adult Taiwanese (n=837; mean BMI=24.5 kg/m^2) was lowest among individuals with an uninterrupted C tract compared with individuals with the wild-type or an otherwise

interrupted poly-C tract. However, mean BMI did not differ between the three groups. Whether this is also the case among European subjects warrants further investigation.

5.1.4 Limitations of mtDNA analysis

Using SNP array based genotype data of up to 40 SNPs of mtDNA, the initial analysis in association with obesity was primarily focused on common mtDNA variation. Infrequent variants of the coding region can have a strong impact on mitochondrial function and potentially body weight. For instance, MELAS (mitochondrial encephalopathy lactic acidosis and stroke-like syndrome) patients harboring a heteroplasmic m.3243A/G transition in $tRNA^{Leu(UUR)}$ have been described as normal weight to lean (Suomalainen et al. 2011). For screening such infrequent variants, re-sequencing of the complete mtDNA among a study sample at least as large in size as the D-loop sample would be required, which would have been too time consuming to fit in the time frame of the present thesis.

Moreover, by the availability of just up to 40 array-based SNPs having passed QC, not all common mtDNA variants of the coding region could be tagged. Some of them might have been evolutionally recently occurred and thus could be present in several different haplogroups. Comparatively, the D-loop polymorphism m.16189T/C detected by re-sequencing can be detected in ~15 % of Europeans (Bendall and Sykes 1995) and is present in several European (Poulton et al. 2002) but also non-European haplogroups (www.phylotree.org).

For association analysis of haplogroups, only major haplogroups were taken into consideration. A refined analysis of sub-haplogroups (e.g. W1 to W6) has not been performed for two reasons. First, with the limited number of available SNPs from the array, determination of sub-haplogroups was not possible for each individual. In addition, some major haplogroups are present at relative low frequencies *a priori*. Association analysis of their sub-haplogroups would only be sensitive in a larger study sample to maintain an adequate power. Nevertheless, in the sub-haplogroups, biologically relevant variants for obesity might be present which consequently would have been masked by comparing only the respective major haplogroups.

Simulations of Salas and Amigo (2010) suggested that only 10 SNPs are necessary for a given population as for instance Europeans to discriminate between 95 % of the maximum haplogroup diversity. Genotyping of specific SNPs that unambiguously differentiate between

haplogroups e.g. by multiplex approaches might be conceivable as a further approach in the detection of mtDNA variants potentially associated with obesity.

By re-sequencing the D-loop, both common and infrequent variants were detected. For the infrequent variants, which were only detected once or twice in either cases or controls, more individuals would have to be genotyped for these variants to assess their association with obesity.

5.2 Gene set enrichment analysis on nuclear-encoded mitochondrial genes

Despite an identification of several BMI associated loci during the last years, there is still a large discrepancy between the empirically assumed heritability of 40 % to 70 % and the currently genetically explained BMI variance of about 1.5 % (Hebebrand et al. 2013). As GSEA or pathway-based approaches concentrate on the combined effects of several loci, new insight into the genetic impact on obesity might be revealed. Based on the observations of a reduced mitochondrial function in obese individuals (Kelley et al. 2002, Niemann et al. 2011), a GSEA on nuclear-encoded mitochondrial genes was performed. This was done to assess whether these genes are enriched for modest association signals that collectively might have an impact on mitochondrial function and potentially influence body weight.

5.2.1 Enrichment of association signals in gene set 1

GSEA and MAGENTA procedure as well as the three mitochondrial gene sets were adopted from Segrè et al. (2010). Segrè et al. (2010) did not find enrichment in association with T2DM and related glycemic traits. In the present analysis, enrichment for obesity association signals in gene set 1 (16 regulators of nuclear-encoded mitochondrial genes) was detected in two independent CC samples (total n=1,834). Enrichment for obesity association signals was found for the 50^{th} percentile, i.e. modest association signals with gene adjusted p-values P_g between ~0.2 and 0.5. Moreover, enrichment was also present above the 75^{th} percentile in the discovery sample and in the meta-analysis using MAGENTA, but in none of the confirmation samples individually.

None of these 16 genes has been found among the 32 BMI loci reported in the latest and largest meta-analysis so far (Speliotes et al. 2010), thus in a single locus-oriented approach none of these genes revealed significant association with obesity. Hence, the results of the present analysis support the hypothesis that a GSEA may detect the combined association effects of several loci (Subramanian et al. 2005, Wang et al. 2007, Torkamani et al. 2008, Liu et al. 2010, and Evangelou et al. 2012). In addition, in both CC samples different SNPs and/or

genes with lowest p-values $P_{g;\,min}$ were found. Consequently, it was not possible to identify one specific weight associated candidate gene. This finding also underscores the idea of GSEA or pathway-based approaches that the combined effect of several loci leads to an association with the trait of investigation, rather than an individual gene of a gene set (Knoll et al. 2013).

Enrichment was not detected in the family-based sample of 705 obesity trios. This sample might have been too small for a confirmation. In case the effect was basically driven by lean and normal weight subjects, the variant frequencies might have been very low in the predominantly obese trio parents, because assortative mating is an observed phenomenon among obese individuals (Hebebrand et al. 2000, Katzmarzyk et al. 2002). This again might result in a reduced power of the trio sample. Finally, genetic heterogeneity including both locus and allelic heterogeneity might have been present (Li et al. 2006).

5.2.2 Relevance of genes of gene set 1 pertaining to body weight

The 16 genes of gene set 1 are transcription factors and/or co-activators. They are involved in the network of nuclear-mitochondrial interactions by regulating the transcription of mitochondrial functional pathways in response to external stimuli such as cold or energy deprivation (Fig. 1.11, Scarpulla 2008). Mouse models for nine of the 16 genes are existent showing that knockout (*k.o.*) or alterations in the expression of these genes are related to body weight or related traits (Table 5.2, Knoll et al. 2013). *PGC-1b k.o.* mice, for instance, have a reduced body weight and fat mass compared with controls (Lelliott et al. 2006). By contrast, female *PGC-1α k.o*. mice show increased body fat, and after short-term starvation of 24 hours *PGC-1α k.o*. mice of both genders developed hepatic steatosis (Leone et al. 2005). Taken together, these nine mouse models underscore the enrichment of modest association signals for obesity found in gene set 1. Importance as potential candidate genes for obesity and related traits can be ascribed to these 16 regulators of mitochondrial genes. Similarly, in a large meta-analysis of 123,564 individuals (Speliotes et al. 2010), a recent GSEA revealed enrichment of association signals for a gene set of 547 obesity-susceptibility candidate genes, which were derived from animal models, Mendelian syndromes, linkage/genetic association studies or expression studies (Vimaleswaren et al. 2012).

Table 5.2 Mouse models (knockout, alterations in expression and mutations) for 9 of the 16 genes of gene set 1 showing a body weight associated phenotype

Gene	Body weight associated phenotype	Reference
ESRRA	ERR$\alpha^{-/-}$ mice have reduced body weight and fat mass and are resistant to a high-fat diet-induced obesity compared with wild type littermates	Luo et al. 2003
GABPA	= NRF2 targeted knock-out of *Nrf2* in mice leads to - 20 % lower body weight after *ad libitum* diet - lower adipose tissue mass - smaller adipocytes and protects against weight gain and obesity otherwise induced by a high fat diet in comparison with wild type littermates	Pi et al. 2010
MYC	Transgenic mice overexpressing c-*myc* in the liver show - decreased body weight increase - decreased fat accumulation in adipose tissue on a three-month-high-fat diet in comparison with control mice	Riu et al. 2003
NRIP1	Formerly known as RIP140 Knock out mice are 15-20 % lighter than wild-type or heterozygous littermates	White et al. 2000
PPARA	*PPARα*-null mice on two different backgrounds (Sv/129 or C57BL/6N) were not obese, but showed - hepatic accumulation of fat - larger gonadal adipose stores in comparison with wild type controls	Akiyama et al. 2001
PPARD	*PPARδ*-null mice are smaller and have smaller gonadal fat sores in comparison with wild type controls	Peters et al. 2000
PPARGC1A	female PGC-1$\alpha^{-/-}$ mice show increased body fat and after 24-h-starvation PGC-1$\alpha^{-/-}$ mice develop hepatic steatosis	Leone et al. 2005
PPARGC1B	*PGC-1b* knock out mice have reduced body weight and fat mass	Lelliott et al. 2006
SIRT1	Sirt1 transgenic (*knock in*) mice are lighter and have white adipose tissue per body weight in comparison with wild type littermates	Bordone et al. 2007

ESRRA, Estrogen related receptor alpha; *GABPA*, GA-binding protein alpha subunit; *MYC*, Myelocytomatosis viral oncogene homolog (avian); *NRIP1*, Nuclear receptor-interacting protein 1; *PPARA*, Peroxisome proliferator-activated receptor alpha; *PPARD*, Peroxisome proliferator-activated receptor delta; *PPARGC1A*, Peroxisome proliferator-activated receptor gamma coactivator 1 alpha; *PPARGC1B*, Peroxisome proliferator-activated receptor gamma coactivator 1 beta; *SIRT1*, Sirtuin 1. *Table adapted from Knoll et al. (2013).*

5.2.3 Evaluation of the method

Several statistical tests recommended for GSEA or pathway-based approaches (leading-edge fraction test, Wilcoxon-Mann-Whitney-test, Kolmogorov Smirnov test, t-test) were performed to evaluate robustness of GSEA and demonstrate independence from method choice for the significant enrichment of association signals in gene set 1. All tests, except for the Kolmogorov Smirnov test which was only nominally significant in the discovery for gene set 1, were similar regarding significance within a tested sample and gene set.

MAGENTA software was initially designed for meta-analysis of GWAS samples (Segrè et al. 2010). In the present study, MAGENTA was also applied to each sample individually to guarantee robustness of the results within a sample and comparability between different samples and the meta-analysis. Indeed, regression-corrected gene p-values P_g achieved by MAGENTA highly correlated with permutation-based gene p-values P_g within a sample ($r=0.95$, $p<2x10^{-16}$), which was comparable with the results of Segrè et al. (2010; $r=0.95$, $p<1x10^{-30}$).

The 50th, 75th and 95th percentile of the set of all autosomal gene-wise p-values P_g were chosen as cut-offs. Simulations have shown that for weak and modest effects of SNPs, the 75th and 95th percentile, respectively, yielded the optimal power to detect gene set enrichment by analyzing GWAS data at a gene set level (Segrè et al. 2010). The conditions used in these simulations based on a sample size of n=10,000 individuals (Segrè et al. 2010). Given the relatively small sample size of n=888 individuals in the discovery, and the small effect sizes of the 32 BMI loci of the latest and largest meta-analysis, the 50th percentile was also selected as cut-off in the present analysis. Indeed, enrichment was more solidly found above this percentile rather than above the 75th.

As LDs between best SNPs of each gene of gene set 1 were rather low among different samples, association signals seemed to be independent. The gene-based approach which considers large parts of common variation within a gene might thus be a useful approach to handle multiple ancestral mutations (Pennisi 1998, Neale and Sham 2004).

5.2.4 Limitations of present GSEA

Although several tests have been performed to guarantee maximum robustness of the results, the present analysis still has some limitations. First of all, only autosomal mitochondrial genes have been included in the present analysis. There are 1,012 unique mitochondrial genes existing according to MitoCarta compendium (Pagliarini et al. 2008). Among these, 13 are protein coding genes of mtDNA (1.3 % of all mitochondrial genes), 30 are X- and one is X/Y-chromosomal (3.1 % of all mitochondrial genes, Pagliarini et al. 2008). These genes were not included in the meta-analysis of Segrè et al. (2010), because most GWAS only focused on autosomal SNPs. Apart from aspects of comparability to the results of Segré et al. (2010), analyses were primarily constrained to autosomal SNPs because association analyses for autosomal and X-chromosomal SNPs strongly differ. In fact, multiple alternatives including different underlying ideas on X inactivation for X-chromosomal SNP

analyses are existing (Loley et al. 2011). For the autosomal SNPs, the Cochran-Armitage trend test for CC association testing was used, as this test could be regarded as standard test for CC GWAS (Wellek and Ziegler 2012). Pertaining to mtDNA SNPs Fisher's exact test was used, as even for small minor allele frequencies which are present among mtDNA SNPs, reliable results were achieved. Based on these differences between association tests for autosomal, X-chromosomal and mtDNA SNPs, it was decided to limit the present GSEA to autosomal SNPs. As both mtDNA and sex-chromosomal genes together represent less than 5 % of all mitochondrial genes, the impact of these genes on the enrichment analysis might however be small.

Furthermore, as the present GSEA is only based on GWAS data and thus common variants, infrequent variants with a potentially stronger effect on obesity or body weight are hardly addressed.

Finally, despite hints of a reduced mitochondrial function as well as lower expression levels of two mitochondrial regulators NRF1 and TFAM (Kelley et al. 2002, Niemann et al. 2011), from the enrichment found in the present study, it is not deducible how and to what extent the genes of gene set 1 are involved in the observed alterations of mitochondrial function among obese individuals. It is uncertain, whether the decreased mitochondrial function and expression of regulators as reported above contributed to obesity or were a consequence of obesity. Due to the enrichment of association signals in the 16 mitochondrial regulators, results of the present study might indicate that a certain genetic predisposition for regulation of mitochondrial function with impact on body weight might exist. However, at this stage, the statement is rather speculative and further investigation is warranted.

5.3 Study samples

For discovery of both mtDNA SNP analysis and GSEA, a CC sample of extremely obese children and adolescents vs. lean adult controls was used for association testing. The use of lean adults as controls can be justified by the fact that misclassification is reduced, as lean children still might become overweight or obese in adulthood (Hinney et al. 2007). Moreover, at the age of 15 years, i.e. at a similar age to that of the cases, 78 % of the lean and normal weight controls reported that their body weight was below the average body weight, underscoring the power of the sample (Hinney et al. 2007). Common variants contributing to BMI have general small effect sizes making large sample sizes necessary to

detect them. Therefore, the CC sample was enlarged by the 705 (extremely) obese index probands of the family-based approach as further cases for mtDNA SNP analysis to maximally increase the power.

However, the sample size of the present discovery sample (n=1,593) is still slightly smaller than the 2nd discovery stage sample used in a meta-analysis for (early onset) obesity (1,181 overweight/obese cases vs. 1,960 controls plus the 705 obesity trios; Scherag et al. 2010). As in that slightly larger sample, initially detected best SNPs apart from SNPs of *FTO*, *MC4R,* and *TMEM18* mainly reached p-values between 0.01 and 0.04, it was decided to follow-up nominally associated mtDNA SNPs in the present analysis in order not to miss loci/haplogroups potentially associated with obesity due to a relatively small sample size.

For independent confirmation of the initial findings, population-based samples were converted into CC samples by categorizing all individuals with a BMI ≥ 30 kg/m^2 as obese cases and those with a BMI < 25 kg/m^2 as normal weight and lean controls. The remaining individuals were excluded from the analysis. According to simulations, genetic markers with an effect in the extremes of a trait are detected more solidly within a CC design compared with a linear regression design despite smaller sample sizes (Pütter et al. 2011). This is based on the fact that linear regression results are mainly influenced by the majority of individuals with a moderate trait and only little by the few individuals of the extremes (Pütter et al. 2011).

Moreover, individuals of all samples (except for 20 from the D-loop sample) used in the present analyses were genotyped with the same SNP array (Affymetrix Human Genome-Wide SNP Array 6.0). Although from the initially available 119 (115) mtDNA SNPs only up to 40 passed all QC criteria, the SNPs used for analyses had a high concordance between the samples. This was in particular of importance for haplogroup assignment.

6 Summary and Conclusion

In 2008, more than 35 % of the world's adult population were overweight (BMI ≥ 25 kg/m^2) and 11 % were obese (BMI ≥ 30 kg/m^2, WHO 2013). Although in the last five to ten years a stabilizing trend in prevalence had been observed (Blüher et al. 2011, Flegal et al. 2012), obesity still is a major global health problem, because of health consequences in later life such as type 2 diabetes mellitus (T2DM, WHO 2013). Both environmental and genetic factors were shown to contribute to the global obesity epidemic (Bouchard 2007). In empirical studies, the heritability of the BMI variance was estimated to 40 % to 70 % (Hebebrand et al. 2013). Interestingly, larger correlations in BMI between mothers and their offspring than between fathers and their offspring were found (Hebebrand et al. 2001a).

Mitochondria are well known as cellular power plants and contain an exclusively maternally inherited circular DNA (mtDNA) of 16,569 bp with 37 genes of which 13 are protein coding subunits of the oxidative phosphorylation system (OXPHOS). On the other hand, ~1,000 to 1,500 nuclear-encoded genes are required to maintain mitochondrial biogenesis (Bar-Yaacov et al. 2012). Alterations in mitochondrial function were found in obese individuals (Kelley et al. 2002).

Both variation of mtDNA and nuclear-encoded mitochondrial genes have been analyzed in association with obesity within the present PhD thesis, because of (1) the central role of mitochondria in the energy metabolism, (2) hints of altered mitochondrial function in obese individuals, and (3) the parental effect of correlations in BMI to which genetic variation in the exclusively maternally inherited mtDNA might contribute.

For analysis of variation in mtDNA, first of all, an association study of up to 40 array-based SNPs of mtDNA (all but one of the SNPs located in mtDNA coding region) was performed in a case-control (CC) sample of 1,158 (extremely) obese cases and 435 lean adult controls (discovery). SNPs were analyzed as single SNPs and as haplogroups determined by HaploGrep. Analysis was done (a) in all individuals and (b) stratified by gender. For independent confirmation, nominally associated SNPs were followed-up among adults of three population-based samples analyzed as CC sample of 1,697 obese cases and 2,373 normal weight controls. In addition, the mtDNA control region (D-loop) of each 192 cases and controls was screened for variants by re-sequencing (Sanger). Fisher's two-sided exact test was used for association testing. Five SNPs (m.4769A/G, m.8994G/A, m.11674C/T, m.12612A/G and m.13708G/A) and two haplogroups (W, J) were initially found to be

nominally associated with obesity (a) in the whole sample (m.8994G/A, W) or (b) stratified by gender, but none of these SNPs or haplogroups could be confirmed in the independent sample. By re-sequencing, 252 variants were detected. Frequencies of two of these variants differed nominally between cases and controls (m.16189T/C, m.16292C/T). m.16189C was more frequent among the cases (17 % vs. 9 %; p=0.048), while 16292T was only found in eight controls.

For analysis of nuclear-encoded mitochondrial genes, a gene set enrichment analysis (GSEA) was performed using three gene sets previously investigated in association with T2DM: (1) 16 nuclear-encoded regulators of mitochondrial genes, (2) 91 OXPHOS genes, and (3) 966 nuclear-encoded human mitochondrial genes (Segrè et al. 2010). For discovery, GSEA was performed in a CC sample of 453 (extremely) obese cases and 435 lean adult. Independent confirmation occurred in a family-based sample of 705 obesity trios and an adult population-based analyzed as a CC sample (463 obese cases and 483 normal weight controls, KORA-CC). A meta-analysis of all three samples was performed. The distribution of association signals (i.e. gene-wise corrected p-values P_g) between a gene set and the gene set of all genes was compared using a leading-edge-fraction-comparison test with cut-offs between the 50^{th} and the 95^{th} percentile in the set of all P_g and alternative tests (e.g. Wilcoxon-Mann-Whitney-test). In the discovery, gene set 1 was significantly enriched for modest to weak association signals above the 50^{th} percentile ($p^{GSEA,50}$=0.0103). This enrichment was not confirmed in the trios, but in KORA-CC. In the meta-analysis, enrichment was not detected above the 50^{th} percentile, but above the 75^{th}.

In conclusion, analysis of up to 40 array-based common mtDNA SNPs did not lead to robust association of either a single mtDNA SNP or a haplogroup. Pertaining to D-loop variants, m.16189T/C seems to be promising to be followed-up in a further sample for independent confirmation of the initial association. The results of variant m.16292C/T, by contrast, might be rather spurious as five of the eight controls carrying the variant allele T were of haplogroup W, association of which could not be independently confirmed. Regarding nuclear-encoded mitochondrial genes, GSEA revealed that modest to weak association signals for obesity might be enriched in the gene set of 16 nuclear-encoded regulators of mitochondrial genes. Although the impact is small, the results of the present thesis contribute to elucidate the heritability of the BMI variance on a molecular genetic level.

7 References

Achilli A, Rengo C, Battaglia V, Pala M, Olivieri A, Fornarino S, Magri C, Scozzari R, Babudri N, Santachiara-Benerecetti AS, Bandelt HJ, Semino O, Torroni A. Saami and Berbers--an unexpected mitochondrial DNA link. Am J Hum Genet. 2005;76(5):883-6.

Akiyama TE, Nicol CJ, Fievet C, Staels B, Ward JM, Auwerx J, Lee SS, Gonzalez FJ, Peters JM. Peroxisome proliferator-activated receptor-alpha regulates lipid homeostasis, but is not associated with obesity: studies with congenic mouse lines. J Biol Chem. 2001;276(42):39088-93.

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG. Sequence and organization of the human mitochondrial genome. Nature. 1981;290(5806):457-65.

Andréasson H, Nilsson M, Budowle B, Frisk S, Allen M. Quantification of mtDNA mixtures in forensic evidence material using pyrosequencing. Int J Legal Med.2006;120(6):383-90.

Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet. 1999;23(2):147.

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. 2009a;301(1):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. 2009b;301(2):191-7.

Bandelt HJ, van Oven M, Salas A. Haplogrouping mitochondrial DNA sequences in Legal Medicine/Forensic Genetics. Int J Legal Med. 2012;126(6):901-16.

Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005;21(2):263-5.

Barth N, Ziegler A, Himmelmann GW, Coners H, Wabitsch M, Hennighausen K, Mayer H, Remschmidt H, Schäfer H, Hebebrand J. Significant weight gains in a clinical sample of obese children and adolescents between 1985 and 1995. Int J Obes Relat Metab Disord. 1997;21(2):122-6.

Bar-Yaacov D, Blumberg A, Mishmar D. Mitochondrial-nuclear co-evolution and its effects on OXPHOS activity and regulation. Biochim Biophys Acta. 2012;1819(9-10):1107-11.

Behar DM, Metspalu E, Kivisild T, Achilli A, Hadid Y, Tzur S, Pereira L, Amorim A, Quintana-Murci L, Majamaa K, Herrnstadt C, Howell N, Balanovsky O, Kutuev I, Pshenichnov A, Gurwitz D, Bonne-Tamir B, Torroni A, Villems R, Skorecki K. The matrilineal ancestry of Ashkenazi Jewry: portrait of a recent founder event. Am J Hum Genet. 2006;78(3):487-97.

Behar DM, Metspalu E, Kivisild T, Rosset S, Tzur S, Hadid Y, Yudkovsky G, Rosengarten D, Pereira L, Amorim A, Kutuev I, Gurwitz D, Bonne-Tamir B, Villems R, Skorecki K. Counting the founders: the matrilineal genetic ancestry of the Jewish Diaspora. PLoS One. 2008;3(4):e2062.

Bendall KE, Sykes BC. Length heteroplasmy in the first hypervariable segment of the human mtDNA control region. Am J Hum Genet. 1995;57(2):248-56.

Bergmann KE, Mensink GB. [Anthropometric data and obesity]. Gesundheitswesen. 1999;61 Spec No:S115-20. German.

Berndt SI, Gustafsson S, Mägi R, Ganna A, Wheeler E, Feitosa MF, Justice AE, Monda KL, Croteau-Chonka DC, Day FR, Esko T, Fall T, Ferreira T, Gentilini D, Jackson AU, Luan J, Randall JC, Vedantam S, Willer CJ, Winkler TW, Wood AR, Workalemahu T, Hu YJ, Lee SH, Liang L, Lin DY, Min JL, Neale BM, Thorleifsson G, Yang J, Albrecht E, Amin N, Bragg-Gresham JL, Cadby G, den Heijer M, Eklund N, Fischer K, Goel A, Hottenga JJ, Huffman JE, Jarick I, Johansson A, Johnson T, Kanoni S, Kleber ME, König IR, Kristiansson K, Kutalik Z, Lamina C, Lecoeur C, Li G, Mangino M, McArdle WL, Medina-Gomez C, Müller-Nurasyid M, Ngwa JS, Nolte IM, Paternoster L, Pechlivanis S, Perola M, Peters MJ, Preuss M, Rose LM, Shi J, Shungin D, Smith AV, Strawbridge RJ, Surakka I, Teumer A, Trip MD, Tyrer J, Van Vliet-Ostaptchouk JV, Vandenput L, Waite LL, Zhao JH, Absher D, Asselbergs FW, Atalay M, Attwood AP, Balmforth AJ, Basart H, Beilby J, Bonnycastle LL, Brambilla P, Bruinenberg M, Campbell H, Chasman DI, Chines PS, Collins FS, Connell JM, Cookson WO, de Faire U, de Vegt F, Dei M, Dimitriou M, Edkins S, Estrada K, Evans DM, Farrall M, Ferrario MM, Ferrières J, Franke L, Frau F, Gejman PV, Grallert H, Grönberg H, Gudnason V, Hall AS, Hall P, Hartikainen AL, Hayward C, Heard-Costa NL, Heath AC, Hebebrand J, Homuth G, Hu FB, Hunt SE, Hyppönen E, Iribarren C, Jacobs KB, Jansson JO, Jula A, Kähönen M, Kathiresan S, Kee F, Khaw KT, Kivimäki M, Koenig W, Kraja AT, Kumari M, Kuulasmaa K, Kuusisto J, Laitinen JH, Lakka TA, Langenberg C, Launer LJ, Lind L, Lindström J, Liu J, Liuzzi A, Lokki ML, Lorentzon M, Madden PA, Magnusson PK, Manunta P, Marek D, März W, Leach IM, McKnight B, Medland SE, Mihailov E, Milani L, Montgomery GW, Mooser V, Mühleisen TW, Munroe PB, Musk AW, Narisu N, Navis G, Nicholson G, Nohr EA, Ong KK, Oostra BA,

Palmer CN, Palotie A, Peden JF, Pedersen N, Peters A, Polasek O, Pouta A, Pramstaller PP, Prokopenko I, Pütter C, Radhakrishnan A, Raitakari O, Rendon A, Rivadeneira F, Rudan I, Saaristo TE, Sambrook JG, Sanders AR, Sanna S, Saramies J, Schipf S, Schreiber S, Schunkert H, Shin SY, Signorini S, Sinisalo J, Skrobek B, Soranzo N, Stančáková A, Stark K, Stephens JC, Stirrups K, Stolk RP, Stumvoll M, Swift AJ, Theodoraki EV, Thorand B, Tregouet DA, Tremoli E, Van der Klauw MM, van Meurs JB, Vermeulen SH, Viikari J, Virtamo J, Vitart V, Waeber G, Wang Z, Widén E, Wild SH, Willemsen G, Winkelmann BR, Witteman JC, Wolffenbuttel BH, Wong A, Wright AF, Zillikens MC, Amouyel P, Boehm BO, Boerwinkle E, Boomsma DI, Caulfield MJ, Chanock SJ, Cupples LA, Cusi D, Dedoussis GV, Erdmann J, Eriksson JG, Franks PW, Froguel P, Gieger C, Gyllensten U, Hamsten A, Harris TB, Hengstenberg C, Hicks AA, Hingorani A, Hinney A, Hofman A, Hovingh KG, Hveem K, Illig T, Jarvelin MR, Jöckel KH, Keinanen-Kiukaanniemi SM, Kiemeney LA, Kuh D, Laakso M, Lehtimäki T, Levinson DF, Martin NG, Metspalu A, Morris AD, Nieminen MS, Njølstad I, Ohlsson C, Oldehinkel AJ, Ouwehand WH, Palmer LJ, Penninx B, Power C, Province MA, Psaty BM, Qi L, Rauramaa R, Ridker PM, Ripatti S, Salomaa V, Samani NJ, Snieder H, Sørensen TI, Spector TD, Stefansson K, Tönjes A, Tuomilehto J, Uitterlinden AG, Uusitupa M, van der Harst P, Vollenweider P, Wallaschofski H, Wareham NJ, Watkins H, Wichmann HE, Wilson JF, Abecasis GR, Assimes TL, Barroso I, Boehnke M, Borecki IB, Deloukas P, Fox CS, Frayling T, Groop LC, Haritunian T, Heid IM, Hunter D, Kaplan RC, Karpe F, Moffatt MF, Mohlke KL, O'Connell JR, Pawitan Y, Schadt EE, Schlessinger D, Steinthorsdottir V, Strachan DP, Thorsteinsdottir U, van Duijn CM, Visscher PM, Di Blasio AM, Hirschhorn JN, Lindgren CM, Morris AP, Meyre D, Scherag A, McCarthy MI, Speliotes EK, North KE, Loos RJ, Ingelsson E. Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. Nat Genet. 2013;45(5):501-12.

Berulava T, Horsthemke B. The obesity-associated SNPs in intron 1 of the FTO gene affect primary transcript levels. Eur J Hum Genet. 2010;18(9):1054-6.

Blüher S, Meigen C, Gausche R, Keller E, Pfäffle R, Sabin M, Werther G, Odeh R, Kiess W. Age-specific stabilization in obesity prevalence in German children: a cross-sectional study from 1999 to 2008. Int J Pediatr Obes. 2011;6(2-2):e199-206.

Bolender N, Sickmann A, Wagner R, Meisinger C, Pfanner N. Multiple pathways for sorting mitochondrial precursor proteins. EMBO Rep. 2008;9(1):42-9. Review.

Bordone L, Cohen D, Robinson A, Motta MC, van Veen E, Czopik A, Steele AD, Crowe H, Marmor S, Luo J, Gu W, Guarente L. SIRT1 transgenic mice show phenotypes resembling calorie restriction. Aging Cell. 2007;6(6):759-67.

Bouchard C. The biological predisposition to obesity: beyond the thrifty genotype scenario. Int J Obes (Lond). 2007;31(9):1337-9.

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(5):526-31.

Browning BL, Yu Z. Simultaneous genotype calling and haplotype phasing improves genotype accuracy and reduces false-positive associations for genome-wide association studies. Am J Hum Genet. 2009;85(6):847-61.

Budowle B, Allard MW, Wilson MR. Critique of interpretation of high levels of heteroplasmy in the human mitochondrial DNA hypervariable region I from hair. Forensic Sci Int. 2002;126(1):30-3.

Calloway CD, Reynolds RL, Herrin GL Jr, Anderson WW. The frequency of heteroplasmy in the HVII region of mtDNA differs across tissue types and increases with age. Am J Hum Genet. 2000;66(4):1384-97.

Calvo S, Jain M, Xie X, Sheth SA, Chang B, Goldberger OA, Spinazzola A, Zeviani M, Carr SA, Mootha VK. Systematic identification of human mitochondrial disease genes through integrative genomics. Nat Genet. 2006;38(5):576-82.

Campbell CT, Kolesar JE, Kaufman BA. Mitochondrial transcription factor A regulates mitochondrial transcription initiation, DNA packaging, and genome copy number. Biochim Biophys Acta. 2012;1819(9-10):921-9.

Campbell NA, Reece JB. Biology. 6th edition, Pearson Education, Inc., publishing as Benjamin Cummings, San Francisco 2002.

Cann RL, Wilson AC. Length mutations in human mitochondrial DNA. Genetics.1983;104(4):699-711.

Cardoso S, Villanueva-Millán MJ, Valverde L, Odriozola A, Aznar JM, Piñeiro-Hermida S, de Pancorbo MM. Mitochondrial DNA control region variation in an autochthonous Basque population sample from the Basque Country. Forensic Sci Int Genet. 2012;6(4):e106-8.

Chan DC. Mitochondrial fusion and fission in mammals. Annu Rev Cell Dev Biol. 2006;22:79-99. Review.

Chang DD, Clayton DA. Precise identification of individual promoters for transcription of each strand of human mitochondrial DNA. Cell. 1984;36(3):635-43.

Chial H, Craig J. mtDNA and mitochondrial disease. Nature Education 2008;1(1).

Chinnery PF, Elliott HR, Patel S, Lambert C, Keers SM, Durham SE, McCarthy MI, Hitman GA, Hattersley AT, Walker M. Role of the mitochondrial DNA 16184-16193 poly-C tract in type 2 diabetes. Lancet. 2005;366(9497):1650-1.

Day FR, Loos RJ. Developments in obesity genetics in the era of genome-wide association studies. J Nutrigenet Nutrigenomics. 2011;4(4):222-38. Review.

de Wilde JA, van Dommelen P, Middelkoop BJ, Verkerk PH. Trends in overweight and obesity prevalence in Dutch, Turkish, Moroccan and Surinamese South Asian children in the Netherlands. Arch Dis Child. 2009;94(10):795-800.

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(10):795-800.

Derenko M, Malyarchuk B, Grzybowski T, Denisova G, Dambueva I, Perkova M, Dorzhu C, Luzina F, Lee HK, Vanecek T, Villems R, Zakharov I. Phylogeographic analysis of mitochondrial DNA in northern Asian populations. Am J Hum Genet. 2007;81(5):1025-41.

Ebbeling CB, Pawlak DB, Ludwig DS. Childhood obesity: public-health crisis, common sense cure. Lancet. 2002;360(9331):473-82. Review.

Evangelou M, Rendon A, Ouwehand WH, Wernisch L, Dudbridge F. Comparison of methods for competitive tests of pathway analysis. PLoS One. 2012;7(7):e41018.

Falkenberg M, Larsson NG, Gustafsson CM. DNA replication and transcription in mammalian mitochondria. Annu Rev Biochem. 2007;76:679-99. Review.

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(12):879-84.

Farooqi IS, Matarese G, Lord GM, Keogh JM, Lawrence E, Agwu C, Sanna V, Jebb SA, Perna F, Fontana S, Lechler RI, DePaoli AM, O'Rahilly S. Beneficial effects of leptin on obesity, T cell hyporesponsiveness, and neuroendocrine/metabolic dysfunction of human congenital leptin deficiency. J Clin Invest. 2002;110(8):1093-103.

Fernandes V, Alshamali F, Alves M, Costa MD, Pereira JB, Silva NM, Cherni L, Harich N, Cerny V, Soares P, Richards MB, Pereira L. The Arabian cradle: mitochondrial relicts of the first steps along the southern route out of Africa. Am J Hum Genet. 2012;90(2):347-55.

Finck BN and Kelly DP. PGC-1 coactivators: inducible regulators of energy metabolism in health and disease. J Clin Invest. 2006;116: 615–622. Review.

Finnilä S, Lehtonen MS, Majamaa K. Phylogenetic network for European mtDNA. Am J Hum Genet. 2001;68(6):1475-84.

Fischer J, Koch L, Emmerling C, Vierkotten J, Peters T, Brüning JC, Rüther U. Inactivation of the Fto gene protects from obesity. Nature. 2009;458(7240):894-8.

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 withoutaffecting T cell responsiveness. J Clin Endocrinol Metab. 2010;95(6):2836-40.

Fisher RP, Topper JN, Clayton DA. Promoter selection in human mitochondria involves binding of a transcription factor to orientation-independent upstream regulatory elements. Cell. 1987;50(2):247-58

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(5):491-7.

Flegal KM, Carroll MD, Ogden CL, Johnson CL. Prevalence and trends in obesity among US adults, 1999-2000. JAMA. 2002;288(14):1723-7.

Forster L, Forster P, Gurney SM, Spencer M, Huang C, Röhl A, Brinkmann B. Evaluating length heteroplasmy in the human mitochondrial DNA control region. Int J Legal Med. 2010;124(2):133-42. Erratum in: Int J Legal Med. 2010;124(4):319.

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(5826):889-94.

Gallagher D, Visser M, Sepúlveda D, Pierson RN, Harris T, Heymsfield SB. How useful is body mass index for comparison of body fatness across age, sex, and ethnic groups? Am J Epidemiol. 1996;143(3):228-39.

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(3):572-81.

Gerken T, Girard CA, Tung YC, Webby CJ, Saudek V, Hewitson KS, Yeo GS, McDonough MA, Cunliffe S, McNeill LA, Galvanovskis J, Rorsman P, Robins P, Prieur X, Coll AP, Ma M, Jovanovic Z, Farooqi IS, Sedgwick B, Barroso I, Lindahl T, Ponting CP, Ashcroft FM, O'Rahilly S, Schofield CJ. The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. Science. 2007;318(5855):1469-72.

Giguère V. Transcriptional control of energy homeostasis by the estrogen-related receptors. Endocr Rev. 2008;29, 677–696. Review.

Goffart S and Wiesner RJ. Regulation and co-ordination of nuclear gene expression during mitochondrial biogenesis. Exp Physiol. 2003;88: 33–40. Review.

González AM, García O, Larruga JM, Cabrera VM. The mitochondrial lineage U8a reveals a Paleolithic settlement in the Basque country. BMC Genomics. 2006;7:124.

Grant SF, Glessner JT, Bradfield JP, Zhao J, Tirone JE, Berkowitz RI, Hakonarson H, Sondheimer N. Lack of relationship between mitochondrial heteroplasmy or variation and childhood obesity. Int J Obes (Lond). 2012;36(1):80-3.

Hales KG. Mitochondrial Fusion and Division. Nature Education 2010;3(9):12.

Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Res. 2005;33(Database issue):D514-7.

Hardy LL, Cosgrove C, King L, Venugopal K, Baur LA, Gill T. Shifting curves? Trends in thinness and obesity among Australian youth, 1985 to 2010. Pediatr Obes. 2012;7(2):92-100.

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. Aktuelle Ernährungsmedizin 1994;19: 259-265.

Hebebrand J, Wulftange H, Goerg T, Ziegler A, Hinney A, Barth N, Mayer H, Remschmidt H. Epidemic obesity: are genetic factors involved via increased rates of assortative mating? Int J Obes Relat Metab Disord. 2000;24(3):345-53.

Hebebrand J, Hinney A, Oeffner F. Molekulare Grundlagen der Adipositas. In: Ganten D, Ruckpaul K (Hrsg) Molkularmedizinische Grundlagen von Endokrinopathien. Springer-Verlag, Berlin Heidelberg 2001a; pp.387-426.

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

Hebebrand J, Hinney A. Environmental and genetic risk factors in obesity. Child Adolesc Psychiatr Clin N Am. 2009;18(1):83-94.

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(5):294-303.

Hebebrand J, Hinney A, Knoll N, Volckmar AL, Scherag A. Molecular genetic aspects of weight regulation. Dtsch Arztebl Int. 2013;110(19):338-44.

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(9):4258-67.

Hinney A, Friedel S, Remschmidt H, Hebebrand J. Genetic risk factors in eating disorders. Am J Pharmacogenomics. 2004;4(4):209-23. Review.

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(5):1761-9.

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(12):e1361.

Hinney A, Vogel CI, Hebebrand J. From monogenic to polygenic obesity: recent advances. Eur Child Adolesc Psychiatry. 2010;19(3):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.

Horváth R, Bender A, Abicht A, Holinski-Feder E, Czermin B, Trips T, Schneiderat P, Lochmüller H, Klopstock T. Heteroplasmic mutation in the anticodon-stem of mitochondrial tRNA(Val) causing MNGIE-like gastrointestinal dysmotility and cachexia. J Neurol. 2009;256(5):810-5.

Ingman M, Gyllensten U. Analysis of the complete human mtDNA genome: methodology and inferences for human evolution. J Hered. 2001;92(6):454-61.

International HapMap Consortium. The International HapMap Project. Nature. 2003;426(6968):789-96.

Jukes TH, Osawa S. The genetic code in mitochondria and chloroplasts. Experientia 1990;46(11-12):1117-26. Review.

Karlsen S, Morris S, Kinra S, Vallejo-Torres L, Viner RM. Ethnic variations in overweight and obesity among children over time: findings from analyses of the Health Surveys for England 1998-2009. Pediatr Obes. 2013 Apr 2. doi: 10.1111/j.2047-6310.2013.00159.x. [Epub ahead of print]

Kasiviswanathan R, Collins TR, Copeland WC. The interface of transcription and DNA replication in the mitochondria. Biochim Biophys Acta. 2012;1819(9-10):970-8.

Katzmarzyk PT, Hebebrand J, Bouchard C. Spousal resemblance in the Canadian population: implications for the obesity epidemic. Int J Obes Relat Metab Disord. 2002;26(2):241-6.

Kazeem GR, Farrall M. Integrating case-control and TDT studies. Ann Hum Genet 2005; 69: 329–335.

Kelley DE, He J, Menshikova EV, Ritov VB. Dysfunction of mitochondria in human skeletal muscle in type 2 diabetes. Diabetes. 2002;51(10):2944-50.

Kelly DP, Scarpulla RC Transcriptional regulatory circuits controlling mitochondrial biogenesis and function. Genes Dev 2004;18: 357–368. Review.

Kitsios GD, Zintzaras E. Genome-wide association studies: hypothesis-"free" or "engaged"? Transl Res. 2009;154(4):161-4.

Kloss-Brandstätter A, Pacher D, Schönherr S, Weissensteiner H, Binna R, Specht G, Kronenberg F. HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. Hum Mutat. 2011 Jan;32(1):25-32.

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(2):e55884.

Kromeyer-Hauschild K, Wabitsch M, Geller F, Ziegler A, Geiß HC, Hesse V, v. Hippel, Jaeger U, Johnsen D, Kiess W, Korte W, Kunze D, Menner K, Müller M, Niemann-Pilatus A, Remer Th, Schaefer F, Wittchen HU, Zabransky S, Zellner K, Hebebrand J. Perzentile für den Body Mass Index für das Kindes- und Jugendalter unter Heranziehung verschiedener deutscher Stichproben. Monatsschrift Kinderheilkunde 2001;149(8): 807-818.

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(5-6):736-43. German.

Lambert BS, Oliver JM, Katts GR, Green JS, Martin SE, Crouse SF. DEXA or BMI: clinical considerations for evaluating obesity in collegiate division I-A American football athletes. Clin J Sport Med. 2012;22(5):436-8.

Lelliott CJ, Medina-Gomez G, Petrovic N, Kis A, Feldmann HM, Bjursell M, Parker N, Curtis K, Campbell M, Hu P, Zhang D, Litwin SE, Zaha VG, Fountain KT, Boudina S, Jimenez-Linan M, Blount M, Lopez M, Meirhaeghe A, Bohlooly-Y M, Storlien L, Strömstedt M, Snaith M, Oresic M, Abel ED, Cannon B, Vidal-Puig A. Ablation of PGC-1beta results in defective mitochondrial activity, thermogenesis, hepatic function, and cardiac performance. PLoS Biol. 2006;4(11):e369.

Leone TC, Lehman JJ, Finck BN, Schaeffer PJ, Wende AR, Boudina S, Courtois M, Wozniak DF, Sambandam N, Bernal-Mizrachi C, Chen Z, Holloszy JO, Medeiros DM, Schmidt RE, Saffitz JE, Abel ED, Semenkovich CF, Kelly DP. PGC-1alpha deficiency causes multi-system energy metabolic derangements: muscle dysfunction, abnormal weight control and hepatic steatosis. PLoS Biol. 2005;3(4):e101.

Li H, Liu D, Lu J, Bai Y. Physiology and pathophysiology of mitochondrial DNA. Adv Exp Med Biol. 2012;942:39-51. Review.

Li R, Tsaih SW, Shockley K, Stylianou IM, Wergedal J, Paigen B, Churchill GA. Structural model analysis of multiple quantitative traits. PLoS Genet. 2006;2(7):e114.

Liesa M, Shirihai OS. Mitochondrial dynamics in the regulation of nutrient utilization and energy expenditure. Cell Metab. 2013;17(4):491-506.

Lindinger A, Peterli R, Peters T, Kern B, von Flüe M, Calame M, Hoch M, Eberle AN, Lindinger PW. Mitochondrial DNA content in human omental adipose tissue. Obes Surg. 2010;20(1):84-92.

Liou CW, Lin TK, Chen JB, Tiao MM, Weng SW, Chen SD, Chuang YC, Chuang JH, Wang PW. Association between a common mitochondrial DNA D-loop polycytosine variant and alteration of mitochondrial copy number in human peripheral blood cells. J Med Genet. 2010;47(11):723-8.

Liu YJ, Guo YF, Zhang LS, Pei YF, Yu N, Yu P, Papasian CJ, Deng HW. Biological pathway-based genome-wide association analysis identified the vasoactive intestinal peptide (VIP) pathway important for obesity. Obesity (Silver Spring). 2010;18(12):2339-46.

Lodeiro MF, Uchida A, Bestwick M, Moustafa IM, Arnold JJ, Shadel GS, Cameron CE. Transcription from the second heavy-strand promoter of human mtDNA is repressed by transcription factor A in vitro. Proc Natl Acad Sci U S A. 2012;109(17):6513-8.

Löffler G, Petrides P. Biochemie und Pathobiochemie. 7. Auflage, Springer-Verlag Berlin Heidelberg New York, 2003.

Loley C, Ziegler A, König IR. Association tests for X-chromosomal markers—a comparison of different test statistics. Hum Hered. 2011;71(1):23-36.

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(6):768-75.

Lopez MF, Kristal BS, Chernokalskaya E, Lazarev A, Shestopalov AI, Bogdanova A, Robinson M. High-throughput profiling of the mitochondrial proteome using affinity fractionation and automation. Electrophoresis. 2000;21(16):3427-40.

Luo J, Sladek R, Carrier J, Bader JA, Richard D, Giguère V. Reduced fat mass in mice lacking orphan nuclear receptor estrogen-related receptor alpha. Mol Cell Biol. 2003;23(22):7947-56.

Lutz S, Weisser HJ, Heizmann J, Pollak S. Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany. Int J Legal Med. 1998;111(2):67-77.

Maes HH, Neale MC, Eaves LJ. Genetic and environmental factors in relative body weight and human adiposity. Behav Genet. 1997;27(4):325-51. Review.

Magnusson PK, Rasmussen F. Familial resemblance of body mass index and familial risk of high and low body mass index. A study of young men in Sweden. Int J Obes Relat Metab Disord. 2002;26(9):1225-31.

Manolio TA. Genomewide association studies and assessment of the risk of disease. N Engl J Med. 2010;363(2):166-76. Review.

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(4):305-8.

Mensink GB, Schienkiewitz A, Haftenberger M, Lampert T, Ziese T, Scheidt-Nave C. [Overweight and obesity in Germany : Results of the German Health Interview and

Examination Survey for Adults (DEGS1)]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2013;56(5-6):786-94.

Meyer S, Weiss G, von Haeseler A. Pattern of nucleotide substitution and rate heterogeneity in the hypervariable regions I and II of human mtDNA. Genetics. 1999;152(3):1103-10.

Miclaus K, Wolfinger R, Vega S, Chierici M, Furlanello C, Lambert C, Hong H, Zhang L, Yin S, Goodsaid F. Batch effects in the BRLMM genotype calling algorithm influence GWAS results for the Affymetrix 500K array. Pharmacogenomics J. 2010;10(4):336-46.

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(6636):903-8.

Montoya J, Christianson T, Levens D, Rabinowitz M, Attardi G. Identification of initiation sites for heavy-strand and light-strand transcription in human mitochondrial DNA. Proc Natl Acad Sci U S A. 1982;79(23):7195-9.

Montoya J, Gaines GL, Attardi G. The pattern of transcription of the human mitochondrial rRNA genes reveals two overlapping transcription units. Cell. 1983;34(1):151-9.

Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstråle M, Laurila E, Houstis N, Daly MJ, Patterson N, Mesirov JP, Golub TR, Tamayo P, Spiegelman B, Lander ES, Hirschhorn JN, Altshuler D, Groop LC. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet. 2003;34(3):267-73.

Moyes CD, Battersby BJ, Leary SC. Regulation of muscle mitochondrial design. J Exp Biol. 1998;201(Pt 3):299-307. Review.

Mueller EE, Eder W, Ebner S, Schwaiger E, Santic D, Kreindl T, Stanger O, Paulweber B, Iglseder B, Oberkofler H, Maier R, Mayr JA, Krempler F, Weitgasser R, Patsch W, Sperl W, Kofler B. The mitochondrial T16189C polymorphism is associated with coronary artery disease in Middle European populations. PLoS One. 2011;6(1):e16455.

Müller TD, Tschöp MH, Hofmann S. Emerging function of fat mass and obesity-associated protein (fto). PLoS Genet. 2013;9(1):e1003223.

Neale BM, Sham PC. The future of association studies: gene-based analysis and replication. Am J Hum Genet 2004;75: 353–62.

Neel JV. Diabetes mellitus: a "thrifty" genotype rendered detrimental by "progress"? Am J Hum Genet. 1962;14:353-62.

Niemann B, Chen Y, Teschner M, Li L, Silber RE, Rohrbach S. Obesity induces signs of premature cardiac aging in younger patients: the role of mitochondria. J Am Coll Cardiol. 2011;57(5):577-85.

Nöthlings U, Krawczak M. [PopGen. A population-based biobank with prospective follow-up of a control group]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2012;55(6-7):831-5.

Ohno K, Tanaka M, Suzuki H, Ohbayashi T, Ikebe S, Ino H, Kumar S, Takahashi A, Ozawa T. Identification of a possible control element, Mt5, in the major noncoding region of mitochondrial DNA by intraspecific nucleotide conservation. Biochem Int. 1991;24(2):263-72.

Pagliarini DJ, Calvo SE, Chang B, Sheth SA, Vafai SB, Ong SE, Walford GA, Sugiana C, Boneh A, Chen WK, Hill DE, Vidal M, Evans JG, Thorburn DR, Carr SA, Mootha VK. A mitochondrial protein compendium elucidates complex I disease biology. Cell. 2008;134(1):112-23.

Palanichamy MG, Sun C, Agrawal S, Bandelt HJ, Kong QP, Khan F, Wang CY, Chaudhuri TK, Palla V, Zhang YP. Phylogeny of mitochondrial DNA macrohaplogroup N in India, based on complete sequencing: implications for the peopling of South Asia. Am J Hum Genet. 2004;75(6):966-78.

Park KS, Chan JC, Chuang LM, Suzuki S, Araki E, Nanjo K, Ji L, Ng M, Nishi M, Furuta H, Shirotani T, Ahn BY, Chung SS, Min HK, Lee SW, Kim JH, Cho YM, Lee HK; Study Group of Molecular Diabetology in Asia. A mitochondrial DNA variant at position 16189 is associated with type 2 diabetes mellitus in Asians. Diabetologia. 2008;51(4):602-8.

Parker E, Phillips DI, Cockington RA, Cull C, Poulton J. A common mitochondrial DNA variant is associated with thinness in mothers and their 20-yr-old offspring. Am J Physiol Endocrinol Metab. 2005;289(6):E1110-4.

Payne BA, Wilson IJ, Yu-Wai-Man P, Coxhead J, Deehan D, Horvath R, Taylor RW, Samuels DC, Santibanez-Koref M, Chinnery PF. Universal heteroplasmy of human mitochondrial DNA. Hum Mol Genet. 2013;22(2):384-90.

Pearson TA, Manolio TA. How to interpret a genome-wide association study. JAMA. 2008;299(11):1335-44.

Pennisi E. A closer look at SNPs suggests difficulties. Science 1998;281: 1787–1789.

Peters JM, Lee SS, Li W, Ward JM, Gavrilova O, Everett C, Reitman ML, Hudson LD, Gonzalez FJ. Growth, adipose, brain, and skin alterations resulting from targeted disruption of the

mouse peroxisome proliferator-activated receptor beta(delta). Mol Cell Biol. 2000;20(14):5119-28.

Pham XH, Farge G, Shi Y, Gaspari M, Gustafsson CM, Falkenberg M. Conserved sequence box II directs transcription termination and primer formation in mitochondria. J Biol Chem. 2006;281(34):24647-52.

Pi J, Leung L, Xue P, Wang W, Hou Y, Liu D, Yehuda-Shnaidman E, Lee C, Lau J, Kurtz TW, Chan JY. Deficiency in the nuclear factor E2-related factor-2 transcription factor results in impaired adipogenesis and protects against diet-induced obesity. J Biol Chem. 2010;285(12):9292-300.

Pliss L, Tambets K, Loogväli EL, Pronina N, Lazdins M, Krumina A, Baumanis V, Villems R. Mitochondrial DNA portrait of Latvians: towards the understanding of the genetic structure of Baltic-speaking populations. Ann Hum Genet. 2006;70(Pt 4):439-58.

Poulton J, Brown MS, Cooper A, Marchington DR, Phillips DI. A common mitochondrial DNA variant is associated with insulin resistance in adult life. Diabetologia. 1998;41(1):54-8.

Poulton J, Luan J, Macaulay V, Hennings S, Mitchell J, Wareham NJ. Type 2 diabetes is associated with a common mitochondrial variant: evidence from a population-based case-control study. Hum Mol Genet. 2002;11(13):1581-3.

Power C, Lake JK, Cole TJ. Body mass index and height from childhood to adulthood in the 1958 British born cohort. Am J Clin Nutr. 1997;66(5):1094-101.

Price RA, Cadoret RJ, Stunkard AJ, Troughton E. Genetic contributions to human fatness: an adoption study. Am J Psychiatry. 1987;144(8):1003-8.

Price RA, Gottesman II. Body fat in identical twins reared apart: roles for genes and environment. Behav Genet. 1991;21(1):1-7.

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(3):559-75.

Pütter C, Pechlivanis S, Nöthen MM, Jöckel KH, Wichmann HE, Scherag A. Missing heritability in the tails of quantitative traits? A simulation study on the impact of slightly altered true genetic models. Hum Hered. 2011;72(3):173-81.

Riu E, Ferre T, Hidalgo A, Mas A, Franckhauser S, Otaegui P, Bosch F. Overexpression of c-myc in the liver prevents obesity and insulin resistance. FASEB J. 2003;17(12):1715-7.

Roberti M, Musicco C, Polosa PL, Milella F, Gadaleta MN, Cantatore P. Multiple protein-binding sites in the TAS-region of human and rat mitochondrial DNA. Biochem Biophys Res Commun. 1998;243(1):36-40.

Robin ED, Wong R. Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. J Cell Physiol. 1988;136(3):507-13.

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.

Rückert IM, Heier M, Rathmann W, Baumeister SE, Döring A, Meisinger C. Association between markers of fatty liver disease and impaired glucose regulation in men and women from the general population: the KORA-F4-study. PLoS One. 2011;6(8):e22932.

Ruiz-Pesini E, Lott MT, Procaccio V, Poole JC, Brandon MC, Mishmar D, Yi C, Kreuziger J, Baldi P, Wallace DC. An enhanced MITOMAP with a global mtDNA mutational phylogeny. Nucleic Acids Res 2007, 35 (Database issue):D823-D828.

Ryan MT, Hoogenraad NJ. Mitochondrial-nuclear communications. Annu Rev Biochem. 2007;76:701-22. Review.

Salanave B, Peneau S, Rolland-Cachera MF, Hercberg S, Castetbon K. Stabilization of overweight prevalence in French children between 2000 and 2007. Int J Pediatr Obes. 2009;4(2):66-72.

Salas A, Amigo J. A reduced number of mtSNPs saturates mitochondrial DNA haplotype diversity of worldwide population groups. PLoS One. 2010;5(5):e10218.

Salpietro CD, Briuglia S, Merlino MV, Di Bella C, Rigoli L. A mitochondrial DNA mutation (A3243G mtDNA) in a family with cyclic vomiting. Eur J Pediatr. 2003;162(10):727-8.

Saxena R, de Bakker PI, Singer K, Mootha V, Burtt N, Hirschhorn JN, Gaudet D, Isomaa B, Daly MJ, Groop L, Ardlie KG, Altshuler D. Comprehensive association testing of common mitochondrial DNA variation in metabolic disease. Am J Hum Genet. 2006;79(1):54-61.

Sbisà E, Tanzariello F, Reyes A, Pesole G, Saccone C. Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. Gene. 1997;205(1-2):125-40.

Scally A, Durbin R. Revising the human mutation rate: implications for understanding human evolution. Nat Rev Genet. 2012;13(10):745-53.

Scarpulla RC. Transcriptional paradigms in mammalian mitochondrial biogenesis and function. Physiol Rev. 2008;88(2):611-38.

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(4):e1000916.

Segrè AV; DIAGRAM Consortium; MAGIC investigators, Groop L, Mootha VK, Daly MJ, Altshuler D. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. PLoS Genet. 2010;6(8).

Shutt TE, Bestwick M, Shadel GS. The core human mitochondrial transcription initiation complex: It only takes two to tango. Transcription. 2011;2(2):55-59.

Sondheimer N, Glatz CE, Tirone JE, Deardorff MA, Krieger AM, Hakonarson H. Neutral mitochondrial heteroplasmy and the influence of aging. Hum Mol Genet. 2011;20(8):1653-9.

Sørensen TI, Price RA, Stunkard AJ, Schulsinger F. Genetics of obesity in adult adoptees and their biological siblings. BMJ. 1989;298(6666):87-90.

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(9):705-14.

Sørensen TI, Holst C, Stunkard AJ. Adoption study of environmental modifications of the genetic influences on obesity. Int J Obes Relat Metab Disord. 1998;22(1):73-81.

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, Stringham HM, 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 A, Johnson T, Ketkar S, Lamina C, Li S, Moffatt MF, Myers RH, Narisu N, Perry JR, 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 MT, Almgren P, Arnold AM, Aspelund T, Atwood LD, Balkau B, Balmforth AJ, Bennett AJ, Ben-Shlomo Y, Bergman RN, Bergmann S, Biebermann H, Blakemore AI, Boes T, Bonnycastle LL, Bornstein SR, Brown MJ, Buchanan TA, Busonero F, Campbell H, Cappuccio FP, Cavalcanti-Proença C, Chen YD, Chen CM, Chines PS, Clarke R, Coin L, Connell J, Day IN, 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 EJ, 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, Martin NG, 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 CG, 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, Surakka I, Swift AJ, Tammesoo ML, Tardif JC, Teder-Laving M, Teslovich TM, Thompson JR, Thomson B, Tönjes A, Tuomi T, van Meurs JB, van Ommen GJ, Vatin V, Viikari J, Visvikis-Siest S, Vitart V, Vogel CI, Voight BF, Waite LL, Wallaschofski H, Walters GB, Widen E, Wiegand S, Wild SH, Willemsen G, Witte DR, Witteman JC, Xu J, Zhang Q, Zgaga L, Ziegler A, Zitting P, Beilby JP, Farooqi IS, 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 PE, Shuldiner AR, Spector TD, Tuomilehto J, Uda M, Uitterlinden A, Valle TT, Wabitsch M, Waeber G, Wareham NJ, Watkins H; Procardis Consortium, Wilson JF, Wright AF, Zillikens MC, 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 RJ. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet. 2010;42(11):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: 506–516.

St John J, Sakkas D, Dimitriadi K, Barnes A, Maclin V, Ramey J, Barratt C, De Jonge C. Failure of elimination of paternal mitochondrial DNA in abnormal embryos. Lancet. 2000;355(9199):200.

St John JC, Facucho-Oliveira J, Jiang Y, Kelly R, Salah R. Mitochondrial DNA transmission, replication and inheritance: a journey from the gamete through the embryo and into offspring and embryonic stem cells. Hum Reprod Update. 2010;16(5):488-509. Review.

Strobel A, Issad T, Camoin L, Ozata M, Strosberg AD. A leptin missense mutation associated with hypogonadism and morbid obesity. Nat Genet. 1998;18(3):213-5.

Stryer L. Biochemie. 1. Korrigierter Nachdruck der 4. Auflage 1996. Spektrum Akademischer Verlag GmbH, Heidelberg – Berlin, 1999.

Stunkard AJ, Sørensen TI, Hanis C, Teasdale TW, Chakraborty R, Schull WJ, Schulsinger F. An adoption study of human obesity. N Engl J Med. 1986;314(4):193-8.

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(21):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(15):1837-44.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-

based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545-50.

Sundblom E, Petzold M, Rasmussen F, Callmer E, Lissner L. Childhood overweight and obesity prevalences levelling off in Stockholm but socioeconomic differences persist. Int J Obes (Lond). 2008;32(10):1525-30.

Suomalainen A, Elo JM, Pietiläinen KH, Hakonen AH, Sevastianova K, Korpela M, Isohanni P, Marjavaara SK, Tyni T, Kiuru-Enari S, Pihko H, Darin N, Õunap K, Kluijtmans LA, Paetau A, Buzkova J, Bindoff LA, Annunen-Rasila J, Uusimaa J, Rissanen A, Yki-Järvinen H, Hirano M, Tulinius M, Smeitink J, Tyynismaa H. FGF-21 as a biomarker for muscle-manifesting mitochondrial respiratory chain deficiencies: a diagnostic study. Lancet Neurol. 2011;10(9):806-18.

Suzuki H, Hosokawa Y, Nishikimi M, Ozawa T. Existence of common homologous elements in the transcriptional regulatory regions of human nuclear genes and mitochondrial gene for the oxidative phosphorylation system. J Biol Chem. 1991;266(4):2333-8.

Szibor R, Plate I, Heinrich M, Michael M, Schöning R, Wittig H, Lutz-Bonengel S. Mitochondrial D-loop (CA)n repeat length heteroplasmy: frequency in a German population sample and inheritance studies in two pedigrees. Int J Legal Med. 2007;121(3):207-13.

Taylor RW, Turnbull DM. Mitochondrial DNA mutations in human disease. Nat Rev Genet. 2005;6(5):389-402. Review.

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(1):18-24.

Torkamani A, Topol EJ, Schork NJ. Pathway analysis of seven common diseases assessed by genome-wide association. Genomics. 2008;92(5):265-72.

Torroni A, Lott MT, Cabell MF, Chen YS, Lavergne L, Wallace DC. mtDNA and the origin of Caucasians: identification of ancient Caucasian-specific haplogroups, one of which is prone to a recurrent somatic duplication in the D-loop region. Am J Hum Genet. 1994;55(4):760-76.

Troiano RP, Flegal KM. Overweight children and adolescents: description, epidemiology, and demographics. Pediatrics. 1998;101(3 Pt 2):497-504. Review.

United Nations, Department of Economic and Social Affairs, Population Division (2011). World Population Prospects: The 2010 Revision, Volume I: Comprehensive Tables. ST/ESA/SER.A/313.

van Oven M, Kayser M. 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum Mutat 30(2):E386-E394. http://www.phylotree.org .

Veyrieras JB, Kudaravalli S, Kim SY, Dermitzakis ET, Gilad Y, Stephens M, Pritchard JK. High-resolution mapping of expression-QTLs yields insight into human gene regulation. PLoS Genet. 2008;4(10):e1000214.

Vimaleswaran KS, Tachmazidou I, Zhao JH, Hirschhorn JN, Dudbridge F, Loos RJ. Candidate genes for obesity-susceptibility show enriched association within a large genome-wide association study for BMI. Hum Mol Genet. 2012;21(20):4537-42.

Völzke H, Wallaschofski H, Wolff B, Berger K, John U, Dörr M. Thyroid function and serum ferritin levels: the study of health in Pomerania. Thyroid. 2006;16(7):681-6.

Wan J, Jiang L, Lü Q, Ke L, Li X, Tong N. Activation of PPARdelta up-regulates fatty acid oxidation and energy uncoupling genes of mitochondria and reduces palmitate-induced apoptosis in pancreatic beta-cells. Biochem Biophys Res Commun. 2010;391(3):1567-72.

Wang K, Li M, Bucan M. Pathway-based approaches for analysis of genomewide association studies. Am J Hum Genet. 2007;81(6):1278-83.

Wanrooij S, Fusté JM, Farge G, Shi Y, Gustafsson CM, Falkenberg M. Human mitochondrial RNA polymerase primes lagging-strand DNA synthesis in vitro. Proc Natl Acad Sci U S A. 2008;105(32):11122-7.

Watson PE, Watson ID, Batt RD. Obesity indices. Am J Clin Nutr. 1979;32(4):736-7.

Wellek S, Ziegler A. Cochran-Armitage test versus logistic regression in the analysis of genetic association studies. Hum Hered. 2012;73(1):14-7.

White R, Leonardsson G, Rosewell I, Ann Jacobs M, Milligan S, Parker M. The nuclear receptor co-repressor nrip1 (RIP140) is essential for female fertility. Nat Med. 2000;6(12):1368-74.

WHO Technical Report Series 894, Obesity: Preventing and managing the global epidemic, World Health Organization, Geneva, 2000.

WHO Fact sheet N° 311, updated March 2013.

Wiesner RJ, Rüegg JC, Morano I. Counting target molecules by exponential polymerase chain reaction: copy number of mitochondrial DNA in rat tissues. Biochem Biophys Res Commun. 1992;183(2):553-9.

Wigginton JE, Cutler DJ and Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. Am. J Hum Genet 2005;76: 887-893.

Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010;26(17):2190-1.

Wilson MR, Stoneking M, Holland MM, DiZinno JA, Budowle B. Guidelines for the use of mitochondrial DNA sequencing in forensic science. Crime Lab Digest 1993;20 : 68–77

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. 2011a;43(6):519-25.

Yang TL, Guo Y, Shen H, Lei SF, Liu YJ, Li J, Liu YZ, Yu N, Chen J, Xu T, Cheng Y, Tian Q, Yu P, Papasian CJ, Deng HW. Genetic association study of common mitochondrial variants on body fat mass. PLoS One. 2011b;6(6):e21595.

Yoza BK, Bogenhagen DF. Identification and in vitro capping of a primary transcript of human mitochondrial DNA. J Biol Chem. 1984;259(6):3909-15.

Yu J, Auwerx J (2009) The role of sirtuins in the control of metabolic homeostasis. Ann N Y Acad Sci 1173: E10–19. Review.

Zollo O, Tiranti V, Sondheimer N. Transcriptional requirements of the distal heavy-strand promoter of mtDNA. Proc Natl Acad Sci U S A. 2012;109(17):6508-12.

Zonta LA, Jayakar SD, Bosisio M, Galante A, Pennetti V. Genetic analysis of human obesity in an Italian sample. Hum Hered. 1987;37(3):129-39.

Zuk O, Hechter E, Sunyaev SR, Lander ES. The mystery of missing heritability: Genetic interactions create phantom heritability. Proc Natl Acad Sci U S A. 2012;109(4):1193-8.

8 Appendix

15961 [a]	**1** [b]	gaaaaagtct	ttaactccac	cattagcacc	caaagctaag	attctaattt	aaactattct
16021	**61**	ctgttctttc	atggggaagc	agatttgggt	accacccaag	tattgactca	cccatcaaca
16081	**121**	accgctatgt	atttcgtaca	ttactgccag	ccaccatgaa	tattgtacgg	taccataaat
16141	**181**	acttgaccac	ctgtagtaca	taaaaaccca	atccacatca	aaacccccctc	cccatgctta
16201	**241**	caagcaagta	cagcaatcaa	ccctcaacta	tcacacatca	actgcaactc	caaagccacc
16261	**301**	cctcacccac	taggatacca	acaaacctac	ccacccttaa	cagtacatag	tacataaagc
16321	**361**	catttaccgt	acatagcaca	ttacagtcaa	atcccttctc	gtccccatgg	atgacccccc
16381	**421**	tcagataggg	gtcccttgac	caccatcctc	cgtgaaatca	atatcccgca	caagagtgct
16441	**481**	actctcctcg	ctccgggccc	ataacacttg	ggggtagcta	aagtgaactg	tatccgacat
16501	**541**	ctggttccta	cttcagggtc	ataaagccta	aatagcccac	acgttcccct	taaataagac
16561	**601**	atcacgatg					
1	**610**	gatcacaggt	ctatcaccct	attaaccact	cacgggagct	ctccatgcat	ttggtatttt
61	**670**	cgtctggggg	gtatgcacgc	gatagcattg	cgagacgctg	gagccggagc	accctatgtc
121	**730**	gcagtatctg	tctttgattc	ctgcctcatc	ctattattta	tcgcacctac	gttcaatatt
181	**790**	acaggcgaac	atacttacta	aagtgtgtta	attaattaat	gcttgtagga	cataataata
241	**850**	acaattgaat	gtctgcacag	ccactttcca	cacagacatc	ataacaaaaa	atttccacca
301	**910**	aacccccccct	cccccgcttc	tggccacagc	acttaaacac	atctctgcca	aaccccaaaa
361	**970**	acaaagaacc	ctaacaccag	cctaaccaga	tttcaaattt	tatcttttgg	cggtatgcac
421	**1030**	ttttaacagt	caccccccaa	ctaacacatt	attttcccct	cccactccca	tactactaat
481	**1090**	ctcatcaata	caaccccgc	ccatcctacc	cagcacacac	acaccgctgc	taaccccata
541	**1150**	ccccgaacca	accaaacccc	aaagacaccc	cccacagttt	atgtagctta	cctcctcaaa
601	**1210**	gcaatacact	gaaaatgttt	agacgggctc	acatcacccc	ataaacaaat	aggtttggtc
661	**1270**	ctagcctttc	tattagctct	tagtaagatt	acacatgcaa	gcatccccgt	tccagtgagt

Figure 8.1 Evaluation sheet of re-sequencing of mitochondrial D-loop

D-loop of rCRS from http://www.ncbi.nlm.nih.gov/nuccore/251831106. Each cell comprises 10 and each line 60 nucleotides. Deviations from rCRS were noted on a separate sheet for each individual. For complete re-sequencing of mtDNA the procedure was performed accordingly.

[a] numbering as found in rCRS

[b] consecutive numbering as found in alignment of reference and the four re-sequenced sequences

Table 8.1 Enrichment of association signals above 95th percentile in 705 family-based trios with initial genotype calling in eight batches

Gene set	total number of genes	effective number of genes	number of SNPs involved	% of all autosomal SNPs involved [a]	$P^{GSEA,WMW}$ Wilcoxon-Mann-Whitney test	$P^{GSEA,KS}$ Kolmogorov-Smirnov-test	$P^{GSEA,t}$ t-test	$P^{GSEA,95}$ 95th percentile cut-off test [b]	$P^{GSEA,75}$ 75th percentile cut-off test [c]	$P^{GSEA,50}$ 50th percentile cut-off test [d]
Gene set 1	16	16	1,039	0.1	**0.0210**	**0.0224**	**0.0415**	**0.0419**	0.0826	0.1085
All auto-somal genes [e]	17,680	9,517	470,499	65.1	-	-	-	-	-	-

[a] 723,288 autosomal SNPs were included

[b] cut-off = 0.0095

[c] cut-off = 0.1283

[d] cut-off = 0.3788

[e] reference gene set

GSEA p-values below 0.05 are highlighted in bold

Danksagung

Allen voran gilt mein besonderer Dank Frau Prof. Anke Hinney für die Überlassung dieses spannenden Themas. Ich danke ihr für die exzellente fachliche Betreuung, Diskussionsbreitschaft und Unterstützung meiner Arbeit, sowie den Einblick in die Genetik der Adipositas, den sie und die Bearbeitung dieses Themas mir ermöglichten.

Auch möchte ich mich bei Prof. Johannes Hebebrand für die Begleitung meiner Doktorarbeit und die wertvollen Hinweise und Diskussionen bedanken. So danke ich auch allen Mitarbeitern der Kinder- und Jugendpsychiatrie für die Zusammenarbeit sowie die interessanten freitäglichen Vorträge und Diskussionen.

Mein Dank gilt auch Dr. Susann Scherag nicht nur für ihre fachliche Unterstützung, sondern auch für jede private Hilfe, die mein Ankommen und Leben hier in Essen erleichtert haben. Anna-Lena Volckmar danke ich für die 3 gemeinsamen Jahre unserer Doktorandenzeit, in der sie mir immer mit Rat und Tat zur Seite stand und stets ein offenes Ohr für mich hatte. Dr. Sandra Majno danke ich für unsere gemeinsame Zeit im Büro.

Jitka Andrä, Sieglinde Düerkop und Beate Kirschbaum danke ich für die Hilfe im Labor und für die „Versüßung“ meiner Doktorandenzeit durch die köstliche Versorgung mit Kaffee und Kuchen. Jitka Andrä und Sieglinde Düerkop möchte ich auch noch einmal besonders für ihre unermüdliche Unterstützung bei der Auswertung der D-loop Elektropherogramme danken. Nur so ließ sich dieses Projekt „relativ schnell“ abschließen.

Mein besonderer Dank gilt PD Dr. André Scherag und Ivonne Jarick für die statistische Beratung und die Berechnungen. Ohne deren Einsatz und Geduld wäre diese Doktorarbeit nicht möglich gewesen.

Dem BMBF danke ich für die finanzielle Unterstützung.

Außerdem danke ich allen Verantwortlichen und Mitwirkenden der populationsbasierten Kollektive KORA, POPGEN und SHIP, für die Bereitstellung der Affymetrix-Chip Daten und Unterstützung jeglicher Art bei der Verwendung dieser Daten, sowie allen Kooperationspartnern des NGFN-plus Netzwerks „Molekulare Mechanismen der Adipositas“.

Mein Dank gilt auch allen Studienteilnehmern.

Raha Vatanparast danke ich vielmals fürs Korrekturlesen meiner Arbeit.

Meiner Mutter und meinem Bruder danke ich, dass sie immer für mich da sind und fürs Rückenstärken.

Nicht zuletzt danke ich Stephan für unsere gemeinsame Zeit, für all seine Geduld, sein Verständnis, jede Aufmunterung und Zuversicht vor allem während des letzten halben Jahres. Ich freue mich auf die kommende Zeit.

Curriculum vitae

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

Ehrenwörtliche Erklärung

Erklärung:

Hiermit erkläre ich, gem. § 6 Abs. 2, f der Promotionsordnung der Math.-Nat. Fakultäten zur Erlangung der Dr. rer. nat., dass ich das Arbeitsgebiet, dem das Thema „Variation in Mitochondrial Genes in Obesity“ zuzuordnen ist, in Forschung und Lehre vertrete und den Antrag von Nadja Knoll befürworte.

Essen, den _____ _____ _____

Name des wissenschaftl. Betreuers/Mitglieds der Universität Duisburg-Essen	Unterschrift d. wissenschaftl. Betreuers/ Mitglieds der Universität Duisburg-Essen

Erklärung:

Hiermit erkläre ich, gem. § 7 Abs. 2, c und e der Promotionsordnung der Math.-Nat. Fakultäten zur Erlangung des Dr. rer. nat., dass ich die vorliegende Dissertation selbständig verfasst und mich keiner anderen als der angegebenen Hilfsmittel bedient habe und alle wörtlich oder inhaltlich übernommenen Stellen als solche gekennzeichnet habe.

Essen, den _____ _____

Unterschrift des/r Doktoranden/in

Erklärung:

Hiermit erkläre ich, gem. § 7 Abs. 2, d und f der Promotionsordnung der Math.-Nat. Fakultäten zur Erlangung des Dr. rer. nat., dass ich keine anderen Promotionen bzw. Promotionsversuche in der Vergangenheit durchgeführt habe, dass diese Arbeit von keiner anderen Fakultät abgelehnt worden ist, und dass ich die Dissertation nur in diesem Verfahren einreiche.

Essen, den _____ _____

Unterschrift des/r Doktoranden/in