universität wien

DIPLOMARBEIT

EPIDEMIOLOGIC APPRAISAL OF SINGLE NUCLEOTIDE POLYMORPHISMS & THE RELEVANCE OF EPIGENETIC REGULATION IN SELECTED EMPHASIS OF THE METABOLIC SYNDROME

angestrebter akademischer Grad:

Magistra der Naturwissenschaften (Mag. rer.nat.)

Verfasserin:	Stefanie Doris Englert
Matrikel-Nummer:	a0301430
Institut:	Institut für Ernährungswissenschaften
Betreuer:	Univ.-Doz. Dr. Alexander Haslberger

Wien, Oktober 2010

Danksagung

Mit der Unterstützung und Mithilfe von Freunden, Familie und Kollegen konnte ich in der Vergangenheit viele Dinge erreichen. Sie haben mir Türen geöffnet, mich unterstützt und an mich geglaubt.

Dank gebührt Herrn Prof. Dr. Alexander Haselberger dafür dass er mich als Diplomandin unter seine Fittiche genommen hat und mir damit den Weg in ein spannendes Gebiet geebnet hat.

Meinen Kollegen, Mag. Eva Aumüller und Mag. Berit Hippe danke ich, dass sie mich an ihrer fachlichen Kompetenz teilhaben ließen und für ihren Beistand in organisatorischen Problemen.

Dank auch an Matt Parrish der seine kostbare Zeit geopfert hat um die eine oder andere Korrekturlesung auf sich zu nehmen.

Katharina Gössl, Gregor Denkmayr und David Fisslthaler haben mir oft den Rücken gestärkt. Durch ihre Unterstützung wurde mir einiges ermöglicht und Vieles erleichtet.

Dank gebührt auch meiner Familie. Vor allem meine Mutter, Brigitte Englert hat immer an mich geglaubt und mich aufopfernd unterstützt. Sie war mir in vielen Belangen eine große Hilfe.

Jeder von Euch war auf seine eigene Art und Weise in den letzten Jahren für mich da. Ihr habt mir geholfen, meine Ziele zu verfolgen und nicht aus den Augen zu verlieren.

Ohne Euch hätte ich das nicht geschafft. Vielen Dank!

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Tables results

ABBREVATIONS

ADRB3	beta-3 adrenergic receptor
ATP	adenosine triphosphate
BAT	brown adipose tissue
BMI	body mass index
bw	body weight
CEBPα	CCAAT/enhancer-binding protein alpha
CI	confidence interval
cM	centiMorgan
DNA	Deoxyribonucleic acid
EBM	evidence based medicine
FABP4	fatty acid binding protein 4
FFM	fat free mass
FM	fat mass
GWA	genome wide association study
H3	histone 3
H3K14	histone 3 lysine 14
H3K27	histone 3 lysine 27
H3K4	histone 3 lysine 4
H3K9	histone 3 lysine 9
H4	histone 4
H4K16	histone 4 lysine 16
H4K20	histone 4 lysine 20
HAT	histone acetylase
HC	hip circumference
HDAC	histone deacetylases
LD	linkage disequilibrium
LPL	lipoprotein lipase
n.a.	not applicable
ob	obese
Ob	obesity
OR	odds ratio
ov	overweight
P	P-value
PGC1α	peroxisome proliferated activator receptor gamma coactivator 1 α
PPARγ	peroxisome proliferated activator receptor gamma
PPRE	PPAR response element
RNA	Ribonucleic acid
RR	relative risk
RXRα	retinoid acid X receptor alpha
SNP	single nucleotide polymorphism
UCP	uncoupling protein
WC	waist circumference
WHR	waist-to-hip ratio

INTRODUCTION

Human health and an individual's susceptibility to a disease are determined by genetic variation, as well as by variation in the influencing environment. Epidemiologic research deals with *"the distribution and determinants of health-related states or events in specified populations, and the application of this study to control of health problems"* (Last, Spasoff et al. 2001).

In developed countries, the majority of morbidity and mortality consists of common yet complex diseases such as cardiovascular disease, cancer, and autoimmune disorders as well as metabolic conditions such as obesity and diabetes, not to mention psychiatric and neurological disorders (Zhang, De et al.). Such complex traits are characterized by the interplay of multiple genes and multiple environmental factors.

During the last decades, epidemiologic research has more and more focused on unraveling the genetic background of diseases – with limited success for complex traits. After more than 50 years of research still, a big part of heritability is missed or not understood. Problems with study designs, methods of data analysis and interpretation result in consistent errors in studies and a lack of replication across big data sets (Ioannidis 2005).

The close-up on strategies and approaches in general and human genome epidemiology in this work will help to understand challenges of previous, present, and future research. The brief introduction to molecular epidemiology focuses on epigenetics, a young field of research that is dealing with mechanisms that modify gene expression due to environmental changes (among other roles).

The ''early origins hypothesis'' postulates that nutrition and a range of other environmental factors during prenatal and early postnatal development program cellular plasticity through epigenetic changes (Dolinoy and Jirtle 2008). Moreover, the epigenome is influenced by environmental factors during the entire life. This alters susceptibilities to adult cardiovascular disease, type 2

diabetes, obesity, and other chronic diseases independently of the genetic predisposition. Previous findings indicate that genetic polymorphisms are more likely to induce a *shift* in between the working capacities of all networks that are involved in the etiology of a complex disease rather than de-/increasing the overall risk for the disease (Stoger 2008).

A case study on the UCP1 network in adaptive thermogenesis in the metabolic syndrome is an excellent example to discuss uncertainties of association studies and to compare results with evidence on epigenetic regulation in the same network. Epigenetic markers might be able to increase the quality of human genome epidemiology on complex diseases. Furthermore, validated epigenetic biomarkers might increase the predictive value of genetic tests and in matters of public health, they might help to unravel problems of complexity.

.

ENVIRONMENT, GENES & THE PHENOTYPE

For a human being, "*environment"* consists of all external surroundings shaping the course of life. The functionality of the living organism depends on the interaction of (more or less) essential physical, chemical, and biological agents, but also social, cultural, and political factors (Davis 1989). "*Environmental health* is the totality of influences on human health and disease that are determined by environmental factors (1947).

The international Epidemiological Association defines *environment* as:

> *"All that which is external to the human host".* […] *"The environment provides the food people eat, the water they drink, the air they breathe, the energy they command, the plagues and pests they combat and the mountains, seas, lakes, streams, plants and animals that they enjoy and depend upon"* (Last, Spasoff et al. 2001).

According to this definition, *"environment"* includes anything that is not genetic. The genetic information of an organism for development and functioning is coded on a nucleotide sequence, called DNA. The structural flexibility of the DNA allows its packing with structural proteins, e.g. histone proteins to higher order structures called chromatin (Lander, Linton et al. 2001).

The human genome consists of 3 billion base pairs and encodes for 20,000 to 25,000 genes. Of each gene a person owns two copies (alleles) – one inherited from each parent. The sequence varies for 0.2 % and such variations manly occur in non-coding regions. If the frequency of a specific genetic variant in a population reaches 1 % or more, it is referred to as a 'polymorphism'. About 90 % of all DNA polymorphisms occur as single base pair substitutions, called single nucleotide polymorphisms (SNPs) (Brookes 1999). The rest is due to restriction fragment polymorphisms, short tandem repeats, insertion-deletion variants, and other structural variants (Frazer, Murray et al. 2009). Polymorphisms can be functional (change amino acid code, alter expression or stability), or in linkage disequilibrium (LD) with functional variations. The LD

describes the non-random allocation of alleles at two or more loci to individual chromosomes as a result of recent mutation, genetic drift or selection. It is the difference between observed and expected allelic frequencies and measured as the variation of one SNP explained by the other (Hedrick and Kumar 2001).

The "*genotype*" is the genetic constitution of each cell of an individual with reference to a specific trait under consideration. The "*phenotype*" is any observable trait of an individual. It is the visible and quantifiable effect of gene expression, environmental influence and the interaction between those two. Many different genotypes can result in one phenotype, or vice versa, many different phenotypes can be the result of one genotype. For example, obesity[1] phenotypes are anthropometric measurements like body mass index (BMI), waist circumference (WC), waist-to-hip ration (WHR), body fat content, skin fold thickness, or features like "higher weight gain" or "lower weight loss" (Rankinen, Zuberi et al. 2006). All this phenotypes are multifactorial traits each characterized by multiple genes and multiple environmental factors. Such a polygenetic background may involve major genes and a subset of genes with lower penetrance. The trait is caused by different genetic loci acting together, or a gene may be manifest only in the presence of a particular environmental exposure (Khoury, Little et al. 2004).

Complex diseases are characterized by a heterogeneous etiology with different genetic backgrounds and environmental influences resulting in the same outcome. A profile describing a combination of risk factors is very unlikely to be found in more than one person.

[1] Obesity is a complex metabolic condition characterized by accumulation of lipids in body cells, markedly in adipocytes.

GENERAL EPIDEMIOLOGY

The word *epidemiology* is based on the ancient Greek words “epi” meaning “on top of”, “demos” a root meaning “the people”, and “logos”, a suffix meaning “the study of”. Taken literally, the word refers to the study of what befalls people. One of the most commonly used definitions for epidemiology is “the study of the distribution and determinants of health-related states or events in populations, and the application of this study to control health problems” (Last, Spasoff et al. 2001).

History

The Greek physician Hippocrates of Cos lived around 2,500 years ago and became the father of medicine and the first epidemiologist known. In his texts, he examined associations between environmental influences and the occurrence of disease. This idea of environmental contribution to the etiology of diseases may have been clouded by explanations with supernatural causes until the rediscovery in the 19th century. In the meanwhile several important key players helped shaping the discipline of epidemiology, as it is known today. Besides Hippocrates, some other pioneers in this area are worth mentioning. John Graunt (1620-1674) used systematic methods to describe diseases and calculate life expectancies, and Thomas Sydenham (1624-1689) was the first who used observational study approaches. In the 18th century, group comparisons and population thinking emerged when James Lind (1716-1794) observed the effect of time, place, weather, and diet on spreading diseases in the British army. Since then, epidemiologists have developed and refined concepts of research elucidating questions related to human health (Morabia 2004).

Flow of research

An epidemiologic theory gives opportunities to generate, test or screen for a hypothesis on a causal association between an exposure and a disease. Several approaches may be achieved such as, a descriptive (observational) or an analytical investigation. Furthermore, studies can be distinguished by the time of conduction (prospective/retrospective) (Rothman 2008). First steps when planning to conduct an epidemiologic study are definitions of aims and the nature of the outcome variable, and decisions on study design and methods of analysis. The probability of a study finding a true outcome depends on its prior probability (before doing the study) of being true, and its statistical power and significance (Pigeot 2007).

Study designs

For the purpose of this work, observational study designs will be the only designs discussed. Previous knowledge on causal relationships between environmental exposures and human diseases comes primarily from the two epidemiologic study designs: cohort and case-control study (Rothman 2008).

The *cohort study design* often begins with a single group of persons heterogeneous to each other with respect to exposure history (*source population*). All participants are equal on developing the disease. They are followed over time and according to exposure status, divided into subgroups to ascertain disease occurrence (prospective). Cohorts can be defined from membership of administrative units, such as doctors or nurses, or defined from employment records, such as a cohort from mine workers.

In a *case-control study* only a part of the population at disease risk is observed. Diseased cases arise from a source population. A control group with the same exposure distribution is chosen and the two groups are compared. The approach can be prospective as well as retrospective – depending on time of exposure recording the occurrence of disease. Cases and controls selected from a cohort make the study "*nested"* (Pigeot 2007; Rothman 2008).

If case-control and cohort designs are population based, cases are representative of all cases in the population and controls are representative of

the source population of cases. The benefit of this approach it the interpretation of the outcomes for a whole population (Rothman 2008).

Another commonly used observational study design on population level is the *cross-sectional design*. Without regard to exposure or disease status, a representative sample of the population is ascertained at a time. The different exposure subpopulations are compared with respect to their disease prevalence (Rothman 2008).

Sample size strongly determines study power. The problem in planning study sizes is finding a balance between greater precision in results and greater costs. When planning a study, computational sample-size statistics serve as rough guidelines. These approaches relate study size to the design, study population, and to the desired power of precision, but do not account for anything that is not included as a variable in the formula (Rothman 2008).

Data analysis

After editing and summarizing the data, it is usually modeled in a 2x2 table (table 1) and analyzed for the estimation of the measurements of interest (discussed below).

Table 1 The epidemiologic 2x2 table. Building subgroups stratifies the number in cells. Measurements are calculated as follows, absolute risk a/(a+b), RR a/(a+b) / c/(c+d), OR a/b / c/d.

	Disease	No Disease
Exposed	**a**	**b**
Not exposed	**c**	**d**

Statistical *testing of hypothesis* focuses on the null hypothesis postulated as “no association between two variables”. In the alpha level of analysis the significance is approved usually based on a cutoff point (α, usually 5 %). The P-value is the continuous measure of whether or not the hypothesis and data are compatible. If $P > \alpha$, the hypothesis has to be rejected (Rothman 2008). The α-error or type I error defines the probability of claiming a relationship when none truly exists. In the next level, the probability that a negative assumption is false

is tested (type II or β error). The probability of a study finding a true relationship reflects 1 – β. If the null hypothesis is rejected, the (more often very unspecific) alternative hypothesis is favored.

The statistical *estimation of measurements* is another method to interpret data. A confidence interval (CI) is used to indicate the reliability of an estimate as 100(1-α) %. If the model is correct and there is no bias, a CI will contain the true parameter with a frequency no less than its level (Rothman 2008).

Modern epidemiology often relies on measurement estimations from *regression analysis*. Here, the focus is on the relationship between a dependent and one or more independent variables. Especially, when the data under analysis is limited, a *Bayesian approach* may be achieved. Such an analysis requires a prior probability distribution to combine with a likelihood function allowing the production of a new set of certainties for the parameter.

Measures of occurrence, effect, and association

The objectives of an epidemiologic study are measurements for valid and precise estimation of a disease frequency or the effect of an exposure on disease outcome.

Ratios, risks, rates and odds that describe the occurrence of an event in the source population are the measurements computed from a 2x2 table (table 1). This measurements differ by the nature of their numerator and denominator and describe disease risk (absolute risk), association between exposure and disease risk (relative risk), or the impact of exposure-disease association (attributable risk) (Morabia 2004). The absolute risk is the probability to get diseased if exposed. The relative risk (RR) is the gold standard, eminently interpretable on the individual level as a given-fold increase in disease risk. What cannot be directly calculated from RR is the probability of being not diseased. An approximation of the RR is the odds ratio (OR). Only for very rare diseases the OR is the same as the RR. In contrast to the RR, the OR is symmetric so that the OR for the disease is the inverse of the OR for no disease. Furthermore, an OR based on exposure probabilities equals the OR based on disease probabilities. Impact measurements estimate the contribution

of an exposure to the incident rate[1] at population level. The attributable risk is the most commonly used. In public health it is interpreted as a measure of the disease burden attributable, or at least related to one or several exposures (Pigeot 2007)

Technically, all of the discussed measurements may be calculated from a cohort study. Case-control studies do only provide enough data for estimating RRs. Only if data is complemented by follow-up or population data (as for example in nested or population-based designs), an estimation of exposure-specific rates is possible (Rothman 2008).

Limitations

The accuracy and reliability of estimates depends on unpredictable (*random error*) and predictable errors (*systematic error*) that may occur (Rothman 2008). In standard practice, each epidemiologic study is treated for the unpredictable *sampling error* that results from the assumption that each subject included in the study is a sample of people who could have been included in the study. Random error might also be due to fluctuations in the readings of a measurement apparatus (Porta, Greenland et al. 2008). An estimate with little random error is referred to as precise. Improving precision of a study will improve reproducibility of collected data and increase reliability. Random errors can be decreased by calculating means such as standard deviations, P-values, and CIs. The testing for hypothesis reduces the probability for random errors. Also, a common way to increase precision is to enlarge the sample size of the study (Rothman 2008).

A *systematical error* (also referred to as bias) is typically constant, or proportional to the true value. An estimate with a little systematic error can be described as valid. Validity may refer to the source population (internal validity), or to the people outside that population (external validity or generalizability) (Porta, Greenland et al. 2008).

[1] *Incidence rate* is the number of new cases among subjects with a risk of disease development in the source population in a defined period of time.

Confounders are external factors associated with the disease and exposure of interest and are not intermediate steps in the exposure's pathway to disease. Regardless of whether or not the exposure has an effect, confounding occurs if external factors become mixed with it. They can be eliminated or reduced later in data analysis (Pigeon 2007) (Rothman 2008).

Selection bias occurs if control subjects were not representative of the study base from which the cases arose, or if not all the cases from the study base were identified. The distribution of environmental factors of interest among case patients and control subjects is not comparable due to the fact that the relation between exposure and disease is different (Pigeon 2007) (Rothman 2008).

Information bias: errors in measuring information can occur throughout each phase of a study and may result in the substantial increase in sample size requirement for adequate statistical power. (Garcia-Closas, Rothman et al. 1999) Measurements errors are for instance, questionnaire design and administration, biologic sample collection, processing and storage, sample labeling, through data coding, data entry, and analysis. For categorical factors, e.g. sex or body weight, an error in measurement is usually referred to as *misclassification*. Interaction of two or more factors can result in over or underestimation of the parameter. If a study relies on subject memory, time may affect recall and reporting, called '*recall bias*' (Pigeon 2007) (Rothman 2008).

An observational study is characterized by the examination of a population under a defined effect. The absence of randomization makes a causal inference largely speculative. At least, an epidemiologic study can only go to prove that an exposure could have caused (but not that it did cause) an effect in any particular case. Despite the main hypothesis, additional thoughts on sources of errors and bias are helpful and will be required.

Addressing bias quantitatively in *sensitivity analysis* is an important process. Systematic errors might have larger effects than random errors, especially, in larger size studies. A lot of methods have been developed to adjust for confounders by standardization of rates and risks,or by stratification. The difficulty in bias analysis is the involvement of the analyst as they require many

choices and judgments from the investigators. Despite analytical tools, the Bradford Hill criteria (9 fundamental principles) are widely excepted and a helpful tool to provide adequate evidence of causality as well (Bradford Hill 1965).

Applications for findings

Epidemiologic studies increase insights into biological processes. Data obtained is used for further investigations, and clinically relevant information is used in primary, secondary, and tertiary prevention and in clinical practice.

Evidence based medicine (EBM) refers to clinical decision making based on the best available evidence about health care outcomes. Scales for evaluating the quality of evidence have been developed: Best evidence for evaluating medical interventions comes from well-designed randomized clinical trials. Best evidence for prevalence and risk assessment comes from large population-based samples. And finally, best evidence for diagnostic accuracy of a test comes from studies comparing the test in representative affected and unaffected populations (Woolf, DiGuiseppi et al. 1996). Whether or not a biomarker achieved from research is valid for clinical use, has to be assessed individually. The *ACCE model process* (figure 1) for the evaluation of data was developed in 2000 by the Secretary's Advisory committee on genetic testing but also applies to general epidemiology. The ACCE acronym represents the four main criteria: **a**nalytical validity, **c**linical validity, **c**linical utility, and **e**thical/legal/social issues.

Analytical validity of a biomarker refers to its accuracy to measure what it is supposed to measure (analytical sensitivity[1] and specificity[2]). The clinical validity includes its sensitivity and specificity in measuring a clinical or subclinical endpoint. The clinical utility refers to the risk for developing the disease with and without the use of the biomarker and accompanying interventions.

[1] Sensitivity is the probability that a test finds all positives.
[2] Specificity measures the proportion of negatives which are correctly identified.

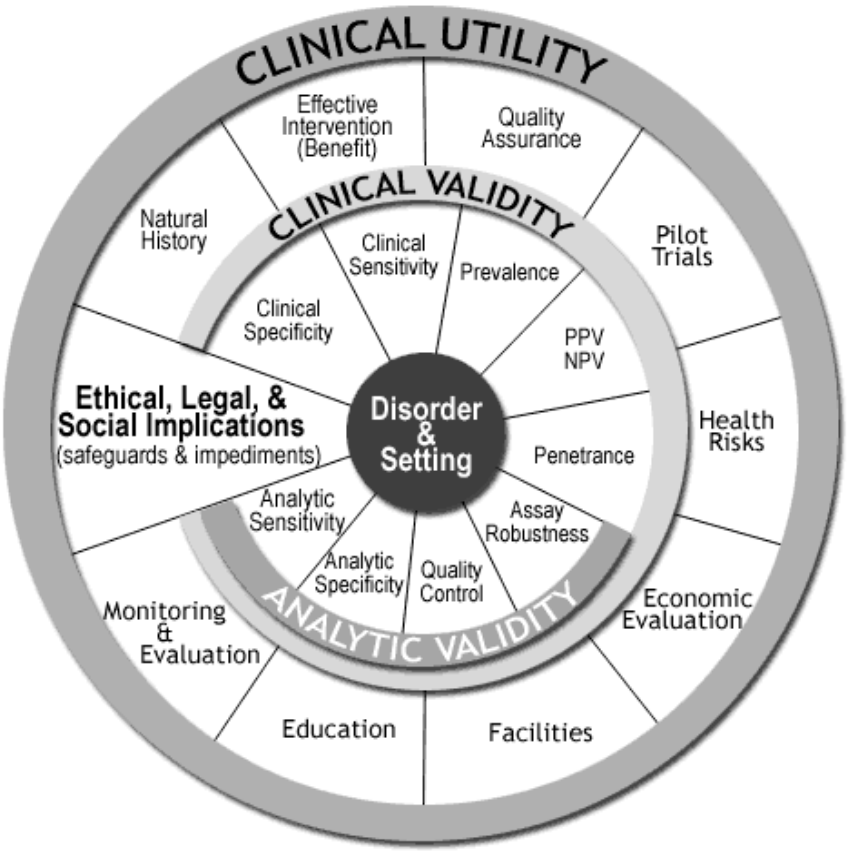

Figure 1 Graphical representation of the ACCE model system. PPV, positive predictive value; NPV, negative predictive value (http://www.cdc.gov, 31st August 2010).

Translating information from epidemiologic studies from individual to population level is the core subject of the "surveillance system", also referred to as *public health surveillance*. *Public health* deals with preventive aspects of health on population level, rather than to treat a disease on individual level. It is 'the science and art of preventing disease, prolonging life and promoting health through the organized efforts and informed choices of society, organizations, public and private, communities and individuals' (Winslow 1920). Surveillance systems provide information about where health problems occur and who is affected. The aim is the monitoring of incidence or prevalence of specified health problems, and the characterization of affected people and those at risk (Rothman 2008). For Public health specialists measurements like risk factors, incidence, and prevalence give a feedback on how health systems respond to current population health issues. Population-based screening is an important component in this scenario.

Areas of concern

Already in 1995 epidemiology has been occasionally viewed as having reached its limits (Taubes 1995).

Common observational *study designs* are often incapable of answering the research questions which they are chosen for. Each approach has its advantages and problems. Many common diseases develop over long time periods making a follow-up expensive (Pigeot 2007; Porta, Greenland et al. 2008). Furthermore, these study designs often have difficulties in locating people or their records many years after their enrollment into the study (Khoury, Little et al. 2004). For the diseases, the cohort design hardly gives sufficient power. In cross-sectional studies, cases with long duration are often overrepresented while those with short duration of illness are underrepresented. A big disadvantage of the case-control design is an overestimation of relative risks for common diseases (Pearson and Manolio 2008). Which design fits best depends on the research question, and the opportunities researcher have.

The great flexibility in study designs, definitions, outcomes, and analytical modes makes them less likely to find true relationships (Ioannidis 2005). Furthermore, a lot of mistakes can be made with respect to study conduct, analysis and dissemination (Pigeot 2007).

The *statistical methods* used in observational data analyzes have been developed for randomized trials. In randomized experiments, systematic errors are of little concern. In contrast, uncontrolled systematic variation in observational studies is the source of confounding factors and other sources of systematic errors. Most uncontrolled biases cannot be analyzed unless additional data for validation is available (which is usually not available or very limited). On the other hand, such validation studies may be subject to systematic errors themselves (Rothman 2008)

The lack in replicating study results is a well recognized problem in the epidemiologic research community and can be caused by several factors. A research finding is less likely to be true the greater the financial and other

interests and prejudices. Replication rates drop further, the more research teams are involved in the same field (Ioannidis 2005).

An obvious tendency of results being published when statistically positive or “interesting” entails a sort of “publication bias” that leads stakeholders to overemphasize results of individual studies (Rothman 2008).

HUMAN GENOME EPIDEMIOLOGY: GENETIC EPIDEMIOLOGY

Flow of research

Genetic epidemiology asks for the genetic contribution to a phenotype. Study designs are pedigree based. The flow of research starts with evidence for familial aggregation of a phenotype in *nuclear families* (index case and parents). If the pattern of correlation is consistent with a possible effect of genes, the mode of segregation is analyzed with e.g. *affected relative pairs* (sibs, cousins, or any two members of the family). Modes of inheritance for single gene disorders can be non-Mendelian (mitochondrial, imprinted) or Mendelian (autosomal recessive/dominat, X-linked) (Khoury, Little et al. 2004).

If a clear mode of Mendelian inheritance can be described, the next step is a coarse mapping (> 1cM[1]) of the genome to narrow the region on the genome where the susceptibility locus is. Such a *linkage study* in related individuals (siblings or extended pedigrees) uses polymorphic markers for identifying biological mechanism for transmitting a trait or the approximate location of the involved gene. A *linkage disequilibrium mapping* is more precise about the location on the genome. The linkage approach has been proven to be very successful in identifying genes of large effect (e.g. cystic fibrosis). In common diseases with a complex background it is of limited success. The large size of the chromosomal regions shared among family members makes it difficult to narrow the linkage signal sufficiently in order to identify a causative gene (Pearson and Manolio 2008).

A linkage analysis is a property of loci and requires family pedigrees; an *association study* deals with fine mapping (< 1cM) of alleles. Aim is the discovery of an association between an allelic variant and a trait, or to find LD between a susceptibility allele and a marker. Two approaches can be achieved. With an a-priori-knowledge, the candidate gene approach examines the role of

[1] Centimorgan, a unit of recombinant frequency for measuring genetic linkage.

a specific gene by associating it with a phenotype. For the last 15 years, candidate gene and genome-wide linkage studies have been the two main genetic epidemiological approaches to identify genetic loci for common traits. Since technical improvements allow for high throughput screenings with lower financial input, more and more articles report findings from *genome wide association studies* (GWAs) (Khoury, Bertram et al. 2009). Performing a whole genome screen in a Caucasian population requires 500,000 loci. As a consequence of LD, it is necessary to genotype only this subset of markers in order to capture common variation in the genome (Spencer, Su et al. 2009). This hypothesis-screening approach seeks for associations between genetic variants and phenotypes and does not require a-priori-knowledge (Manolio). GWAs detect common variants with low penetrance. Furthermore, GWAs are able to detect a gene-gene interaction, or a modification of an association by another variant (Pearson and Manolio 2008).

The signal of a positive association is narrowed to potential causative variants in the haplotype block represented by the tag SNP (fine mapping). This determines whether one of the SNPs has a stronger association (than the tag SNP), or an established functional effect (Manolio). In the next step, a more valid study has to replicate the SNP to identify false positive findings. In such a multi-stage design, positive findings from a GWA (discovery set) are genotyped in an independent population sample (replication set 1). This yields to a smaller subset of significantly associated SNPs which are then tested in a third tier (replication set 2). The second replication is attempted in a study as similar as possible to the initial report. Later it is extended to related phenotypes, different populations, or different study designs (Chanock, Manolio et al. 2007). To detect variants with small effects, large consortia are often needed. Positive associations may also be investigated for their functional implications experimentally (Pearson and Manolio 2008). These processes help to minimize false positive associations.

For the purpose of this work, an association study will be referred to as a study of any polymorphism across the entire human genome, designed to identify one or more associations with an observable phenotype.

Depending on whether the study population is family, or population based, several different designs, screening, and data analyzing methods may be achieved.

Study designs in association studies

Designs such as population-based case-control (most frequently used) and cohort designs do not differ in principle from general epidemiologic studies. Family-based designs are for example, sibling-controls in which affected individuals are cases and unaffected siblings are controls. Studying unrelated controls is more powerful to than studies using unaffected siblings as controls (Teng and Risch 1999; Li, Gail et al. 2002). A general problem of family-based association studies is the difficulty to apply to disease of middle or old age. Different reasons for availability of a sibling can be related to a great variety of factors such as social relationships, or socioeconomic and occupational status related to migration (Wacholder, Rothman et al. 2002).

Unlike other markers of exposure, genetic variants can be seen as stable indicators of host susceptibility making the case-control approach particularly well suited. On the other hand, due to changes in environmental exposures over time, the cohort study design may be preferred. The nested case-control design merges both advantages, but is less cost intensive compared to a cohort study (Rothman 2008). Another opportunity to save costs is the case-cohort design (using a cohort as control group) (Khoury, Little et al. 2004).

The cross-sectional design is used to estimate allele and genotype frequencies in populations. Further, relationships between frequencies and exposure levels or frequencies and phenotypes can be assessed. Although, cross-sectional studies may guide research and health policies, they are unable to identify incidence or causal inference. (Rothman 2008)

Migrant, adoption, and twin study designs have been proven well in identifying any environmental contribution by identifying rapid changes that cannot be attributed to genetic factors (Khoury, Little et al. 2004). Case-control and cohort studies are the most often used designs to evaluate gene-environment interaction (Langholz, Rothman et al. 1999).

Statistical power in association studies

A major criterion when detecting a causative variant is statistical power. It is determined by the effect size of the polymorphis, and the sample size of the study. Required sample sizes are normally calculated based on haplotypes in HapMap (http://hapmap.ncbi.nlm.nih.gov/). In the simulation modeling one has to think of a particular disease model and effect size, genotype frequencies in cases and the power of commercially available chips (Spencer, Su et al. 2009). Generally, the sample size tends to be very high according to allele frequency in the population, strength of association and further, to guarantee a low P-value. To detect interactions between common genetic polymorphisms and environmental exposures, required sample size is even larger (Manolio).

Data modeling & analysis in association studies

The *2x4 table* is the tool of choice to evaluate if included factors are independent of each other (table 2).The role of each factor is assessed in terms of both, individual association and potential attributable fraction. ORs can be calculated for each possible combination (Rothman 2008)

In the null hypothesis it is assumed that the vast majority of SNPs are not associated with the phenotype. Rejecting the null hypothesis of a locus implies that the SNP is associated. The analysis is based on the principles of LD, which means that the OR does not necessarily reflect the risk associated with the true mutation (it may be in LD with a causal SNP)

Table 2 Layout for a 2x4 table in a case-control association study to assess the effect of a genotype (G) and an environmental factor (E).

G	*E*	*Cases*	*Control*	*OR*		*Contrast*	*Main Information*
+	+	a	b	ah/bg	A	A vs. D	Joint genotype and environmental factor vs. none
+	-	c	d	ch/dg	B	B vs. D	genotype alone vs. none
-	+	e	f	eh/fg	C	C vs. D	Environmental factor alone vs. none
-	-	g	h	1	D		Common reference

Other measures	*OR*	*Main information*
Case-only OR	ag/ce	Departure from multiplicative model of interaction
Control-only OR	bh/df	Independence of factors in population
Multiplicative interaction	A/(B*C)	Deviation from multiplicative model of interaction
Additive interaction	A – (B + C – 1)	Deviation from additive model of interaction
Stratified 1-a	ad/bc	Association with environmental factor among people *with* genotype
Stratified 1-b	eh/fg	Association with environmental factor among people *without* genotype
Stratified 2-a	af/be	Association with genotype among people *exposed* to environmental factor
Stratified 2-b	ch/dg	Association with genotype among people *not exposed* to environmental factor

adapted from (Khoury, Little et al. 2004)

Similar to general epidemiology, *regression analyses* is a core method in genome epidemiology. Data is modeled in additive or multiplicative ways to overcome numerical limitations of simpler methods. An extension of traditional regression analysis is the *partitioning method* (joint analysis of multiple genes for quantitative traits). It relies on genotypic partitioning based on multiple loci (Nelson, Kardia et al. 2001)

Limitations in association studies

A successful genome screen is determined by the effect of the marker allele and the density of markers used (Pearson and Manolio 2008) (Chanock, Manolio et al. 2007). Additionally, adopted approaches and methods have their limitations as well.

Study designs

Pros and cons of common observational study designs have been discussed in previous chapters. Despite difficulties, case-control and cohort studies are best suited to evaluate common variants with small effects on complex traits.

Sample size & study power

To guarantee sufficient power, association studies require high sample sizes to be able to detect the modest effects of SNPs in a complex condition like (Manolio). Sample sizes are often limited by the number of well characterized clinical samples to include into the study population as examples of (Fogelholm, Valve et al. 1998), (Ukkola, Tremblay et al. 2001) and (Deeb and Brunzell 2009) demonstrate. Aside from that, sample size is often calculated based on HapMap chromosomes which results in oversampling those risk alleles (Spencer, Su et al. 2009).

The 2x4 table is the epidemiologic tool most researchers used to work with. One can estimate the departure of the joint effect from specific models of interaction (e.g., additive or multiplicative) only by dichotomizing factors. With three factors, the exposure combinations become 8 (2^3). With increasing complexity the number of possible combinations grows (2^n for n factors) and corresponding tables rapidly become cumbrous (Khoury, Little et al. 2004). For example, the simultaneous testing of 5 genetic variants, with 3 genotypes each, gives 3^5 or 243 potential genotype combinations. Testing of 9 variants with 3 genotypes each gives 19,683 combinations (Janssens 2008). Sample size requirements are an inescapable challenge (figure 2).

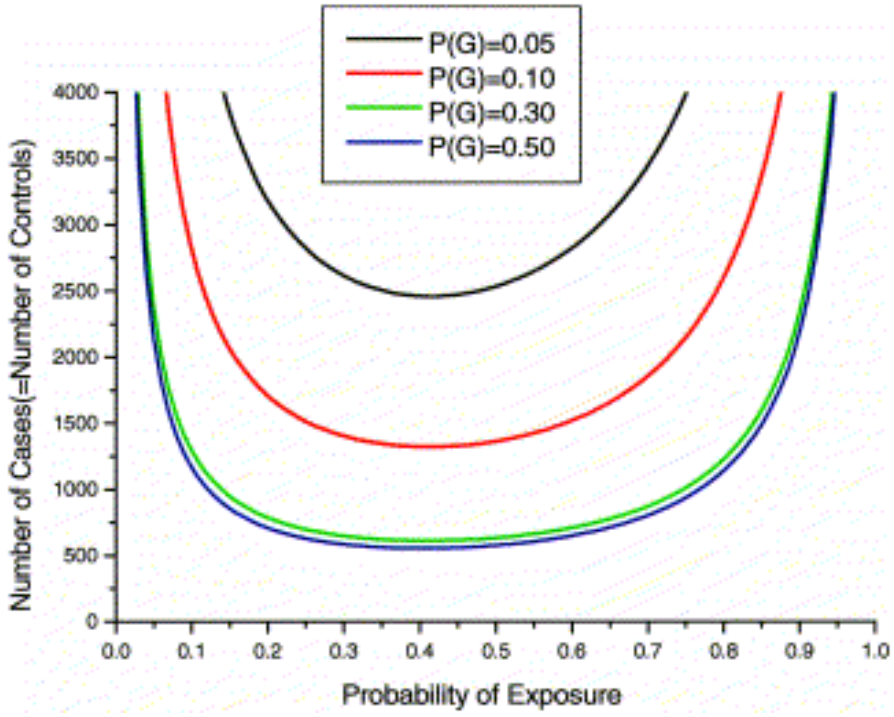

Figure 2 displays required case and control numbers to detect a 1.5-fold multiplicative interaction for varying prevalence values (P) of the "at risk" genotype (assuming a two-sided test with 80 % power and an overall α-level of 5 %). The ORs are 1.0 for the at-risk genetic variant in the absence of exposure, 2.0 for the exposure without the at-risk allele, and 3.0 for having both factors. Adapted from (Rothman, Wacholder et al. 2001)

An increasing number of factors under study quickly reduce the per-stratum size within a fixed total number of subjects, and decreases study power. The more factors are involved, the less likely it is that interactions are described by simple multiplicative or additive models (Khoury, Little et al. 2004). Additionally, relations can be other than dichotomous, for example graded or continuous (dose-response) enhancing complexity.

Modeling data

As a matter of fact, the complexity of pathways often makes it difficult to examine the effects of exposures or genes at a time, without allowing for the influence of other factors also biasing the data (Khoury, Little et al. 2004). In the presence of unaccounted etiologic factors it will also be difficult to unravel interaction partners and joint effects (Pearson 2008).

Errors & bias

Studies of human genome epidemiology hold new sources of errors and bias, additional to the problems that have been discussed in the general epidemiology chapter.

Population stratification is a confounding bias by ethnicity and the most cited reason for inconsistency in results. Genotype distribution and disease vary across ethnic groups and correlate with each other. Estimations of allele effects on disease risk fail if heterogeneity in the study population is missed. But, even in the broad categories of ethnicity (e.g., Caucasians, Asians, etc) strong variations in allele frequencies may occur. If subgroups are identifiable (by genotyping or self-reported information) a standard technique such as statistical stratification deals with the problem (Rothman 2008; Khoury, Little et al. 2004). Another *confounder* is an unknown allele that is associated with the disease and in LD with the allele under study (Khoury, Little et al. 2004).

Measurement error of genotype determinations is usually referred to as *misclassification*. This information bias can be reduces by including duplicate samples for quality control when performing genotype assays. By testing if the control population is in Hardy-Weinberg equilibrium[1] can also help to identify such problems (Khoury, Little et al. 2004). Misclassification of case participants can markedly reduce study power and bias study results towards no association, particularly when large numbers of unaffected individuals are misclassified as affected (Pearson and Manolio 2008). Misclassification can also be due to LD, especially when the genotype is measured at DNA level (Rothman 2008).

Genotyping errors result in spurious associations. They strongly decrease the ability to detect true findings.

[1] It is expected that in a population, genotype and allele proportions remain stable over generations.

Applications for findings, genetic testing

Findings from GWAs enlarge the body of knowledge of biologic mechanisms in normal health and development and of disease causation (Pearson and Manolio 2008). Their potential applications to improve health and prevent diseases should not be underestimated. They encourage the development of better drugs and tailor drug use to maximize benefits and to minimize harms. In response to several levels of exposure, genetic markers are stable indicators when measuring inter-individual variability. They are predictors of an individual's risk for developing a disease.

A *genetic test* determines whether a person is a carrier of a mutation or not. In general, testing can be distinguished as followed: diagnostic testing (to confirm a diagnosis), predictive testing (presymptomatic testing for an individual's susceptibility to a disease), carrier testing (test for presence of mutation), and identity testing (forensics). Furthermore, tests may predict optimal therapeutic regimens (the likelihood of response or toxicity). With respect to time, genetic tests are carried out prenatal, postnatal, and through the entire life.

Today, for more than 2,160 diseases, genetic tests are available, 1,898 of them for clinical use (http://www.ncbi.nlm.nih.gov/sites/GeneTests/, 31th Aug 2010). A growing number of tests are looking at multiple genes that increase or decrease a person's risk to develop complex diseases, such as several different types of cancers or diabetes. Altogether they have the potential to improve the health of individuals and entire populations. During this new era the ACCE model[1] and 44 questions (directed at each of the preceding categories of the model) have been developed to evaluate data on emerging genetic tests.

In the public health sector, the enhanced identification of subgroups based on genotypes will help to develop more effective interventions to prevent disease onset or progression (Caporaso and Goldstein 1995).

[1] Introduced on pages 18 and 19.

Areas of concern

In the community of human genome epidemiologists, the lack of replicating study results is a well recognized problem. A meta-analysis included 370 studies addressing 36 genetic associations for various diseases. It demonstrated that the first study often had suggested a stronger genetic effect than was found by subsequent studies (Ioannidis, Ntzani et al. 2001). A literature review of over 600 gene-disease associations by Hirschhorn et al revealed that most reported associations were not robust. Of the 166 associations studied at least three times, only 6 were consistently replicated (Hirschhorn, Lohmueller et al. 2002). Additionally, a research finding is less likely to be true the more research teams are involved which might be due to the competition and publishing philosophies of editors (Ioannidis 2005). This lack of reproducibility has been varyingly attributed to some serious problems:

In human genome association studies no stringent consensus exists on the study design fitting best a particular research question, the way to model and analyze data, and on what information to include in the report. Such flexibilities tend to reduce the probability of a research finding to be true (Ioannidis 2005).

The methodological shortcomings of *traditional study designs* raise skepticism of the ability to identify and characterize genetic determinants of disease risk. The most widely used contemporary approach in the field is the population based case-control association study. It carries the most assumptions. If they are not met, substantial biases and spurious associations occur (Pearson and Manolio 2008). Another disadvantage is the difficulty of control group selection (Rothman 2008). In population based studies of this design, relative risks are overestimated quite often (Pearson and Manolio 2008). Cohort studies spanning over many years present logistic and financial problems adversely affecting validity.

SNP associations identified in one *population* are often not transferable to other populations. If an environmental or genetic factor varies with ethnicity and with risk of disease, it acts as surrogate. In different populations diverse mechanisms can contribute to the same phenotypes. Furthermore, variations in

allele frequency defined by ethnicity or geographic origin falsely identify subgroup-associated genes related to disease (Pearson and Manolio 2008).

In *data analysis*, correcting for multiple testing assumes independent associations of each SNP which is simply false because individual SNPs correlate to some degree due to LD (Pearson and Manolio 2008).

The large number of possible allele combinations and non-genetic risk factors result in high rates of false positives (Khoury, Little et al. 2004) (Yang, Kelly et al. 2007). Association studies often carry limited and biased information on environmental exposures or other non-genetic risk factors (for example, inquiring ethnicity or food frequencies[1]). Stratifying data on such information results in the problem of sparse data.

Emerging technologies allow the simultaneous measuring of hundreds and thousands of SNPs, gene expression profiles, and protein patterns. A large amount of data on each individual is the rule rather than the exception. The emerging *complexity in data analyses*, especially in complex networks, raises concerns. 2x4 tables, stratified analyses, and logistic regression analysis are already facing limitations (Rothman 2008). For instance, regression modeling is limited in the number of variables that it examines simultaneously. The results of Nelson et al suggest that traditional methods of building multiloci models simply fail to identify combinations that best predict a trait's variability (Nelson, Kardia et al. 2001).

Although, estimation of effects should be favored, researchers use probabilities of data as measures of compatibility between data and hypotheses, or as measures of the relative support that data provide hypotheses (Rothman 2008).

Statistical methods for analyzing data from observational studies were developed for randomized trials. In randomized experiments, systematic errors

[1] Self-reported ethnicity is a better surrogate for environmental risk factors than genome-based methods (people's behavior is likely to be closer to the ethnicity they identified themselves with, than to the ethnic composition of their genome) (Wacholder Rothman, et al 2002).

are of little concern. In observational studies, the uncontrolled systematic variation is the source of confounding factors and other sources of systematic errors. Most uncontrolled biases cannot be analyzed unless additional data for validation is available (which is usually not available or very limited). Furthermore, such validation studies may be subject to systematic errors themselves (Rothman 2008).

Association studies generally rely on the LD. After several generations, genetic recombination leads to complete independence between a marker allele and a disease allele in the same region. This fact is accelerated by population admixture, genetic drift, rate of mutation, non-random mating, selection, and genetic linkage (Vogel 1996). The use of marker alleles in case-control studies leads to misclassification and a dilution of the OR (Rothman 2008).

Many features in design and analysis of GWAs aim the minimization of type I error while maintaining the power to identify true-positive associations. Such efforts could be demonstrated to result in overlooking a true association. In a multi-stage design, the most robust findings often are not the most statistically significant associations in the initial scan (Pearson and Manolio 2008). The diligent search and correction for biases is important as they are able to reduce the magnitude of an observed association towards null.

Genotyping a population of European ancestors with a commercial platform of 500,000 markers (recommended marker density) fails to capture structural and rare variants, and relatively young mutations (Pearson and Manolio 2008). If a true causative SNP is not captured, several markers on the chip correlating with it may give a signal of significant association and hence allow detection of the locus. The complexity of a LD does not allow analytically capturing of this effect so that it has to be accessed via simulation studies (Spencer, Su et al. 2009). These problems can lead to a missing bulk of the heritable fraction of in complex traits.

High rates of false positive results are also due to multiple statistical testing on the one side and a large number of possible allele combinations in a complex trait on the other side (Khoury, Little et al. 2004) (Yang, Kelly et al. 2007). The

large-scale data mining with overestimation of statistical significance testing leads to numerous positive associations that are not replicated (Rothman 2008).

Genome association studies provide little information on the *functional basis* and the link between genetic polymorphism and a complex trait (Pearson and Manolio 2008).

Genetic testing

An increasing development of commercial platforms selling genome-wide association-based screenings directly to the consumer can be recognized. In the beginning, most of the tests (were designed and thus) applied only for families with a strong history of a specific disease. However, in recent years, the growing list of susceptible genes made tests available for everyone interested. Today, many companies offer clinically available tests for a range of disorders and diseases, and/or extend to genetic tests for nutrition, behavior and aging. However, resulting reports often include information based on inadequate evidence. Predictions are based on average risks which are difficult to apply to an individual person. Nevertheless companies already give advice on how the risk for these diseases can be reduced using medical, behavioral, and environmental interventions (Khoury, Little et al. 2004) (Manolio). A study of the Government Accountability Office for example, investigated results of direct-to-consumer test results. 14 hypothetical consumer profiles, generated from only two sources, were sent to four different web sites offering genetic tests. The report concluded that the information passed on to consumers was "misleading" with respect to the terminology and the inconsistency of disease prediction for identical samples. Furthermore, the recommended, “personalized”, nutritional supplements offered by some of the companies were in fact not customized (http://www.genome.gov/19518344).

The little information on analytical validity of many tests makes it difficult to decide whether different studies of the same genetic trait used comparable laboratory methods (Burke, Atkins et al. 2002). For the accuracy of the prediction, quality of samples and medical history collected from family

members are as important as the technical aspects of the testing process in the laboratory.

SNPs uncovered by GWA analysis are noisy measurements of the genotype at the position of the causal variant and the risk associated with its occurrence is unclear. Nevertheless, evidence from GWAs is often used to predict diseases. Other genetic tests are developed from studies observing variations associated with disease susceptibility in multiple affected family members. Such high-risk families represent the severe end of a spectrum of risks associated with mutations and tempt to overestimate the risk by not integrating other factors that may modify risk (Burke, Atkins et al. 2002).

Another problem of available tests is that they are merely not ‘yes or no’ decisions and rather predict relative risks (giving a probability for developing the disease) (Khoury, Little et al. 2004) In other words: people carrying a polymorphism associated with a complex disease not necessarily develop the disease.

Currently, a huge amount of causal variants has not been identified and it is impossible to assay all known human genome variations. The consumer may still carry a susceptibility variant and develop the disease even if the test tells him he is not. Commercially available genotyping chips further limit the variations tested. Furthermore, the risk of disease will only marginally differ between carriers and non-carriers of risk variants of one single susceptibility gene (Janssens and van Duijn, 2008). The complex interactions also influence and result in differences in clinical validity.

Even if a risk factor is well established, the clinical utility might still be low. For many of the diseases which can be tested, no effective preventive or therapeutic treatment exists yet. Genetic variance alone probably cannot account for disease susceptibility without the addition of pre- and postnatal environmental and/or behavioral factors. Moreover, an interaction between genetic and environmental factors has to be accounted for, e.g., in genomewide association studies. Advanced and new tools are needed for unraveling complex diseases.

MOLECULAR EPIDEMIOLOGY

Molecular epidemiology deals with genetic and environmental factors contributing to diseases on the molecular level. It entails the inclusion of molecular measurements in epidemiologic research. Molecular techniques are used to measure exposure, biologic responses, or host characteristics influencing susceptibility to adverse health outcomes. The aim is not to formulate new etiologic hypotheses but to reveal mechanisms and events that occur along the theoretical continuum between exposure and the associated disease. It offers higher resolution answers in relation to disease causation (McMichael 1994) (Vineis and Perera 2007).

Biomarkers

Molecular biomarkers have first been applied in cancer epidemiologic research in the 1980s (Perera and Weinstein 1982). They can be categorized as markers of internal dose, markers of biologically effective dose, markers of early response/effect, markers of susceptibility, and markers of altered structure and function (1987).

Epigenetics investigates heritable changes in gene expression that alter chromatin without changing the DNA sequence. The Epigenome is the overall epigenetic state of a cell including all mechanisms and marks such as non-coding RNAs, the histone code, and DNA methylation (Holliday 2006). The epigenotype is tissue and cell-type specific and varies over time as a function of time, environmental factors, and random processes (Foley, Craig et al. 2009). In the germline, during embryogenesis, and the first years of life the epigenome undergoes reorganization. Cued by surrounding environments, it results in the sensitization for later environmental factors – a so called “programming of the genotype” (Hales and Barker 2001). A number of studies demonstrated that nutritional components and lifestyle factors (Misík 2010; Rust 2010; Stefanska 2010; Thaler 2010), toxins (Bursch 2010) and endocrine disruptors (Sato, Fukata et al. 2006), radiation (Schofield 1998), air pollution (Bursch 2010), and

maternal behavior (Szyf, Weaver et al. 2007) influence the epigenetic state of cells dynamically during the entire life. With age, a general decrease in tissue/gene specific DNA methylation[1] has been reported and genome-wide and specific methylation patterns could be identified (Foley, Craig et al. 2009; Komninou 2010). This developmental plasticity makes it possible to develop a range of phenotypes from a single genotype. Epigenetic modifications can be stably passed over numerous cycles of cell division and some alterations can even be inherited from one generation to the next (Kaati 2010). They have been suggested to form an alternative, "soft inheritance" system (Jablonka and Raz 2009).

Disruption of the epigenetic profile raises the probability of dysfunctioning of gene activity control and nuclear architecture with severe consequences for health (Esteller 2008) (Espada and Esteller). Epigenetic research had its breakthrough in cancer research where the difference in tumor suppressor gene methylation between cancerous and noncancerous tissue are sometimes close to 100%. Specific methylation profiles have also been associated with predicting factors of prognosis (Esteller 2008; Bursch 2010; Komninou 2010).

The epigenome is already a field of research in several other complex diseases as for example, neurodegenerative disorders (Migliore 2010), asthma (Miller 2010), viral infections (Huber 2010) and the metabolic condition obesity (Campion, Milagro et al. 2009).

In this work the following types of epigenetic markers will be discussed:

- *DNA methylation:* is a dynamic marker of the transcriptionally repressed gene (Bird 2002). Methylation is measured at specific sites, or as the overall level of methylation pattern in a tissue (Foley, Craig et al. 2009).
- *Histone modifications*: Depending on type, amount, site of modification and its reversibility the modification results in different transcriptional outcomes (table 3) (He and Lehming 2003).

[1] the covalent addition of a methyl group to a cytosine that precedes a guanosine in the DNA sequence (the CpG dinucleotide)

Table 3 Examples of histone modification and their transcriptional outcome.

	monomethylation	dimethylation	trimethylation	acetylation
H3K4	**Activation** (Bernstein, Humphrey et al. 2002)		**Activation** (Koch, Andrews et al. 2007)	
H3K9	**Activation** (Lachner, O'Carroll et al. 2001)	**Repression** (Rosenfeld, Wang et al. 2009)	**Repression** (Barski, Cuddapah et al. 2007)	**Activation** (Kurdistani, Tavazoie et al. 2004)
H3K27	**Activation** (Lachner, O'Carroll et al. 2001)	**Repression** (Rosenfeld, Wang et al. 2009)	**Repression** (Barski, Cuddapah et al. 2007)	
H4K20	**Activation** (Margueron, Trojer et al. 2005)			

Validation of biomarkers

Biomarkers are validated for their ability to predict for preventive or clinical purposes and, in intervention studies, reflect the modification of the natural course of disease (Vineis and Perera 2007). *Technical validation* deals with intrinsic measurement error and analytic sensitivity, for example, the ability to detect exposures at levels that exist in real human populations (Albertini, Nicklas et al. 1996). *Epidemiologic validation* is about the marker's behavior in the population, depending on variability within the population, for example, variability between subjects (inter-individual) or within subjects (intra-individual). In order to identify intergroup differences (e.g., exposed vs. unexposed, healthy vs. diseased), the total intragroup variation needs to be minimized. (Vineis and Perera 2007).

The validation process is three major steps starting from the development of the biomarker, to its characterization, and to the final step of longitudinal[1] studies. The gold standard of study design is cohort studies, but also nested case-control and case-cohort designs are common (Bonassi and Au 2002).

Techniques

Molecular epidemiologic research is equipped with the standard epidemiologic tools and methods that have been discussed in previous chapters. Biomarkers measure factors, outcomes, confounders or effect modifiers. To be able to adjust for the intragroup variation several strategies are pursued. The collection of repeat samples, documentation on information on subject characteristics influencing intersubject variation, documentation on information on conditions under which samples have been collected and laboratory analyses are common methods (Vineis and Perera 2007).

Limitations

Studies are subject to similar problems and pitfalls as general epidemiologic studies. With the addition of biomarkers, pertinent issues that have to be considered are for example, analytic batch effects, storage effects, and freeze-thaw cycles. They can create random noise or bias with the same consequences as misclassification of exposure in epidemiologic studies and have an impact on study validity (Rundle, Vineis et al. 2005).

Areas of concern

Already discussed in the chapter about general epidemiology, concerns regard the limitations of study designs, required sample sizes, statistical analysis, and the reporting and interpretation of results as discussed in previous chapters.

The availability of validated biomarkers is still limited. Although in use, most available biomarkers require further validation. Main concerns are, the relevance to the exposure of interest, its specificity and sensitivity (Boffetta).

[1] A longitudinal study involves repeated observations of the same trait over long periods of time, mostly decades.

An enormous amount of studies even fails to use validated biomarkers. The complexity of a study increases enormously when integrating multiple markers, each reflecting a different stage or mechanism in the etiology (Perera and Weinstein 2000).

New markers introduce increased potential for bias and confounding (Vineis and Perera 2007). For instance, the quality of the biomarker data might not be comparable in the case series and the reference group (Rundle, Vineis et al. 2005). Further, timing issues such as when the laboratory begins the analytic runs of the samples, or the life course of the underlying cohort influence batching decisions, differences in storage duration between subgroups, and the likelihood that sample aliquots will have endured freeze-thaw cycles (Rundle, Vineis et al. 2005). Most biomarker-based studies of both, prospective and retrospective design rely on a single biological sample (Boffetta).

Applications for biomarkers

In cross-sectional, retrospective, prospective and nested case-control epidemiologic studies biomarkers offer new opportunities to overcome limitations of general epidemiologic methods as demonstrated in cancer research (Hemminki K. 1997; Hagmar 2005).

In epidemiologic studies, once biomarkers are validated, they can be used to assess exposure more validly and precisely than exposure information is gathered via questionnaire and historical information such as employment records and monitoring data. Markers can show biological plausibility for an exposure proposed as etiologically relevant (Vineis and Perera 2007). Newer biomarkers can be combined with the more traditional biomarkers in hypothesis-testing studies (Bonassi and Au 2002).

Once biomarkers are validated properly, they are also used for clinical purposes like early diagnosis, prognosis, and follow-up and to predict potential risks for a group or population more precisely and to introduce better interventions (Perera and Weinstein 1982) (Perera and Weinstein 2000) (Vineis and Perera 2007).

CASE STUDY: ADAPTIVE THERMOGENESIS IN OBESITY

The potential of results coming from genome association studies to predict the risk of complex diseases has been anticipated since the first publication, but this application is problematic. Currently known variants explain only little of the heritability of diseases. The resulting lack in replicating results leads to skepticism all over the research community. Unraveling the complex multi-gene-environment interactions responsible for most diseases of public health importance seems to be more challenging than expected (Manolio).

Obesity is a common complex condition. It is a major risk factor for common chronic diseases such as diabetes, diseases of the cardiovascular and musculoskeletal system, and some cancers (Haslam and James 2005). The World Health Organization (WHO) defines *overweight* as body mass index (BMI) 25 – 30 kg/m^2 and *obesity* ≥ 30 kg/m^2. Once considered a problem only in high-income countries, overweight and obesity are now dramatically on the rise in low- and middle-income countries, particularly in urban settings. WHO projects that by 2015, approximately 2.3 billion adults will be overweight and more than 700 million will be obese.

Obesity is a highly heritable trait with estimates ranging from 50 – 70 % for BMI (Allison 1996) and from 75 – 80% for total body fat (Stunkard 1986). In 2006 the last update of the Obesity Gene Map Database summarized positive associations for DNA sequence variation with obesity phenotypes in 127 candidate genes. Among them are (those showing replications in at least 10 studies): PPARG (30 studies), ADRB3 (29), ADRB2 (20), LEPR (16), GNB3 (14), UCP3 (12), ADIPOQ (11), LEP (11), UCP2 (11), HTR2C (10), NR3C1 (10), and UCP1 (10) (Rankinen, Zuberi et al. 2006). Since then, many more loci have been identified, but they still account for little of the heritability, often less than 1 %. Yet, the scientific community is a long way from reaching a comprehensive picture of the heritable components of obesity and advancing from plain statistical significance into a biological understanding where the true contribution to a trait is recognized.

Body fat content, energy expenditure & obesity

The human body harvests energy as fat in adipocytes as a reserve in undersupplied times. The regulation of body fat content results from an integrative network influenced by the central nervous system, the gastrointestinal tract, the adipocytes itself, skeletal muscles, taste receptors and by multiple nutrient, drug, toxin, sensory, and hormonal inputs. Furthermore, circadian rhythm and physical, gut and social environments influence the network.

Adipocytes can be found in tissues mixed with other cells and in stereotypic depots all throughout the body and can be defined by function, brown or white adipocytes. Brown adipose tissue (BAT) serves as thermogenic organ by dissipating energy while white adipose tissue (WAT) has the contrary physiological function: energy storage (Rosen and Spiegelman 2006). Brown adipocytes contain multiple small droplets of triglycerides and a high number of mitochondria giving them the ability to oxidize high amounts of fatty acids (Silva 2006). Mature white adipocytes exclusively transcribe the full machinery of enzymes and regulatory proteins needed to carry out lipolysis and de novo lipogenesis (Kopelman 2000). In mice, the occurrence of cold induced brown adipocytes in WAT is well established (Cousin 1992; Guerra 1998). Recently it could be demonstrated to be due to a transdifferentiation from adult white to brown adipocytes with similar morphology (Barbatelli, Murano et al.).

The mechanisms underlying differentiation from a multipotent mesenchymal stem cell to a preadipocyte and a mature adipocyte are not fully understood yet, Although, the orphan nuclear receptor peroxisome proliferator activator receptor gamma (PPARγ) and the transcription factor CCAAT/enhancer-binding protein alpha (CEBPα) have been demonstrated to be essential (reviewed in Rosen and Spiegelman 2006).

Accumulation of WAT results in excessive weight gain and is a major risk factor for chronic diseases such as: cardiovascular disease (mainly heart disease and stroke), diabetes, musculoskeletal system disorders and some cancers

(endometrial, breast, and colon) (Haslam and James 2005). Energy expenditure is the crucial factor to main energy balance and a healthy weight. The skeletal muscle and its aerobic oxidation of glucose and fatty acids is the dominant mechanism for dissipating energy. Adaptive thermogenesis in BAT protects the body from hypothermia and is responsible for 1 – 2 % of the overall energy expenditure, preventing a weight gain of 1 – 2 kg/year (Gesta 2007).

Core elements in adaptive thermogenesis in BAT thermogenesis

In BAT mitochondria, uncoupling of oxidative phosphorylation from ATP synthesis by the uncoupling protein 1 (UCP1) enhances heat production (Silva 2006). Participating factors and interaction partner in the network of the UCP1 transcription activation are displayed in figure 2. The UCP1 gene holds (among other regulatory elements) a response element for the nuclear receptor family of PPARs. PPARγ is a member of this transcription factor family and acts as sensor for a variety of lipids and lipid like compounds (e.g. ω-3 fatty acids). Different promoters in the gene and alternative splicing result in several isoforms. For instance, PPARγ2 is expressed exclusively in adipose tissue while PPARγ1 can also be found in liver and other tissues (McClelland, Shrivastava et al. 2009). Ligand binding to the receptor initiates the gene transcription process through conformational changes, a heterodimerization with the retinoid acid X receptor alpha (RXRα) and the assembly with coactivators to a multiprotein complex (Kliewer, Umesono et al. 1992). PGC1α acts as a ligand-independent coactivator for PPARs and is expressed preferentially in BAT. Histone acetylase (HAT) activity of coactivators leads to a modification of the chromatin structure of the proximate DNA allowing the complex to bind to a respond element and to initiate transcription. In the absence of ligands, a PPAR interacts with corepressor proteins and builds a complex with histone deacetylases (HDACs) which results in a locally more compact chromatin packaging (Li, Pascual et al. 2000). The binding of the PPARγ2/RXRα/PGC1α complex to the respond element of UCP1, followed by its transcriptional activation, have been demonstrated to be essential in brown adipogenesis (Barak, Nelson et al. 1999). Other PPARγ target genes are, for example, the extracellular lipoprotein lipase (*LPL*) which mediates the uptake of

fatty acids into the cell (Wang and Eckel 2009), and the fatty acid binding protein 4 (*FABP4*) which is expressed exclusively in adipocytes and macrophages. Although the role of the fatty acid carrier protein FABP4 is not well understood, its function could be correlated to lipid metabolism (Shin, Li et al. 2009). *LPL* is also a target gene of CEBPα who is also a transcription factor for the beta-3 adrenergic receptor gene *(ADRB3)* (Rosen and Spiegelman 2006). For rapid adjustment the ADRB3 in BAT is targeted by catecholamines in response to cold (Silva, 2006). The activation of the receptor signaling results in an increased expression of the *PGC1α* gene (Cao, Daniel et al. 2004).

The capacity of the described network strongly depends on environmental influences. Fatty acids are the only substrate for oxidative phosphorylation in BAT and also act as ligands of the nuclear receptor PPARγ. RXRα is a high affinity nuclear receptor for 9-cis retinoic acid, a metabolite of retinoic acid. Increased uncoupling of the respiratory chain from ATP production depends on the activation of the ADRB3 in response to cold (Silva 2006). Generally, the thermogenic activity of BAT depends on the degree of stimulation by the sympathetic nervous system, the amount of UCP1, and available substrate. (Trayhurn, Jones et al. 1982)

Less activity of UCP1 results in a more efficient energy coupling and lower metabolic rates, both predictors of weight gain (Cypess and Kahn) UCP1 deficient mice are cold intolerant, but surprisingly not ob. When fed with a high-fat diet, they gain less weight than wild type (WT) controls. This Ob resistance is ambient temperature dependent because when mice reared at room temperature deficient mice gain weight faster than WT (Silva 2006). In obese humans a reduced expression of UCP1 mRNA down to 50 % was demonstrated (Oberkofler 1997). It is speculated that mutations reducing the activity of the protein diminish energy expenditure (Dalgaard 2001).

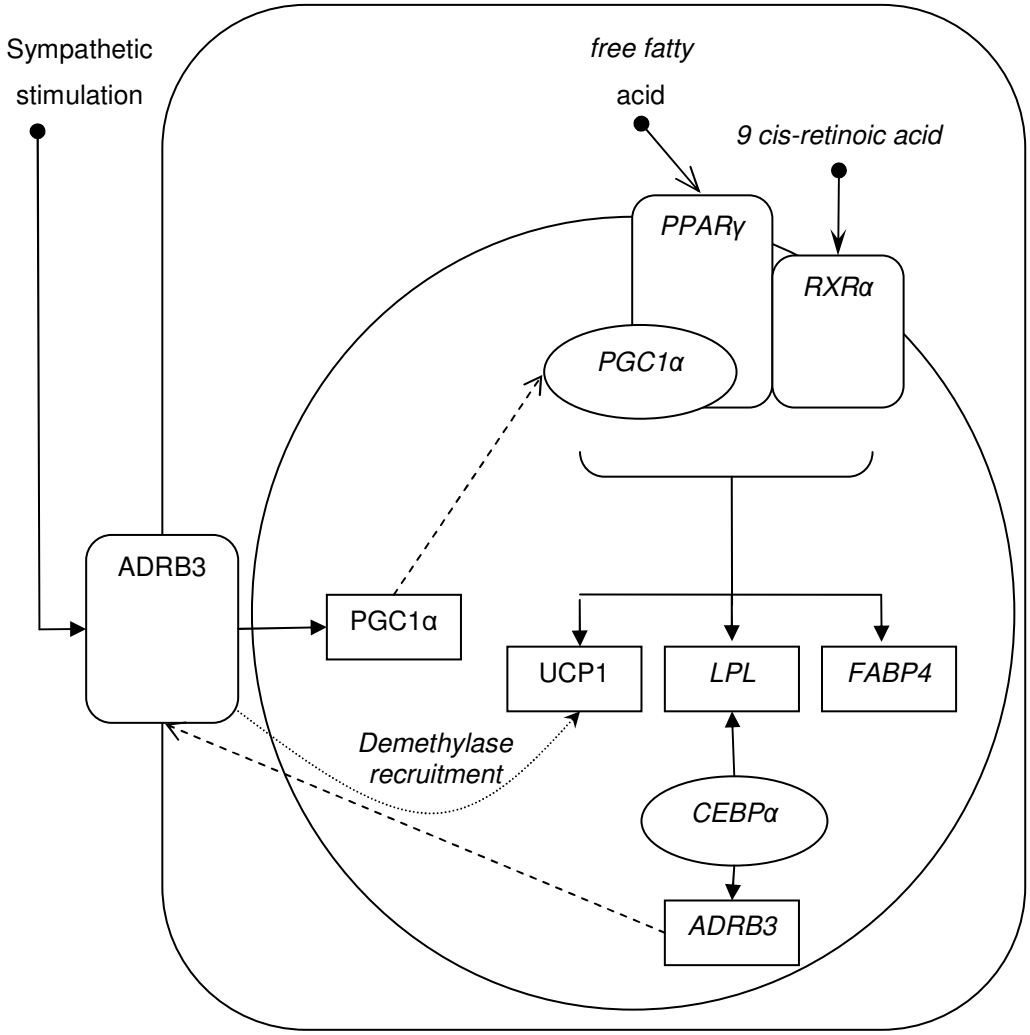

Figure 3 Participating factors and interaction partner in the network of the UCP1 transcription activation.

Nowadays, consumers can already test their risk for developing obesity by companies like *23andMe*[1]. They for example, predict by testing for the SNP rs3751812 in the FTO gene. Markers in the area of this gene appear to be associated with BMI and carriers of the variant weight about 3 kg more. Although, it is not known yet how the gene affects BMI. Furthermore, it has only been replicated in Caucasians (Frayling 2007). *23andMe* tests two other, even more speculative, variants (rs7566605 and rs6232). Each of these SNPs have been related to obesity of Caucasians in one study only. The first study could be replicated once, while four studies disputed the result. The other SNP (rs6232) has even only associated once.

[1] www.23andme.com

OBJECTIVES

Issue of this work is the integration of epigenetic markers in human genome epidemiologic studies.

In a case study of an important network in adaptive thermogenesis and obesity the synergisms of genetic and epigenetic regulation will be demonstrated. This will highlight the need for a systematic concept that integrates different types of biomarkers into epidemiologic research.

The benefits for public health specialists and clinicians with special focus on genetic testing shall be discussed.

METHODS

A systematic literature focused on most frequently observed SNPs and evidence for epigenetic regulation in genes of adaptive thermogenesis (table 3). The literature search revealed observational studies associating the polymorphisms with anthropometric measurements and other related ob phenotypes (English language). Studies were included that reported positive and negative associations in Caucasians and mixed ethnicities, by not distinguishing between self-reported and genome based methods to evaluate ethnicity. If no ethnicity was reported, a study was included if conducted in central Europe, and if it could be assumed to include Caucasians. Evidence for epigenetic regulation of the same genes was searched in *PubMed* by using the terms “gene” and/or “gene name” and “epigenetic” and/or “methylation” in title and/or abstract. Additionally, reference lists from articles were searched for further suitable studies. English publications giving sufficient information on DNA methylation and histone modifications in the genes were summarized. No other variations (e.g. frame shift mutations or copy number variations) than selected SNPs and no other data on posttranscriptional modification than epigenetic modifications are included.

Table 4 List of genes under observation.

Gene name	*Gene*	*Gene map locus*	*SNP*
Uncoupling protein 1	UCP1	4q31	A-3826G
Retinoid acid X receptor alpha	RXRα	9q34	-
Peroxisome proliferated activator receptor gamma coactivator 1 alpha	PGC1α	4p15	rs8192678
Peroxisome proliferated activator receptor gamma	PPARγ	3p25	rs1801282
Lipoprotein lipase	LPL	8p22	rs285
			rs320
			rs328
Fatty acid binding protein 4	FATBP4	8q21	rs1054135
CCAAT/enhancer-binding protein alpha	CEBPα	19q13	rs12691
			rs41490344
			rs41504144
			rs41344151
			rs16967952
			rs34508287
			rs41428545
			rs34529039
			rs41367646
			rs1049969
			rs707656
Beta-3 adrenergic receptor	ADRB3	8p12-p11	rs4994

RESULTS

Results are summarized in tables R1 – R18 on page 54 – 67.

PPARγ

Genetic polymorphisms: The Pro12Ala (rs1801282) is one of the most prominent SNPs in the *PPARγ* gene. It has been observed for associations with phenotypes of the metabolic syndrome for many years. The literature research revealed a meta-analysis that included over 19,600 subjects and positively associated the Ala variant with increased BMI (table R1).

Epigenetic modifications: In the past, six studies observed DNA methylation, and eleven studies observed histone modifications of *PPARγ.* In adipocytes, the correct epigenetic modification of the promoter region seems to be essential during cell differentiation and for the maintenance of the adult state of the cell. In tables R2 and R3, studies observing epigenetic regulations are summarized. In cells from a human umbilical cord where PPARγ2 is not expressed, its promoter region is hypermethylated. In human gastric mucosa cells of ulcer and cancer patients, hypomethylation of the promoter region was observed which concurs with findings from cancer research where cancerous cells become globally hypomethylated and tumorsuppressor genes hypermethylated (Esteller 2008). In human adipose stem cells, the promoter region is hypermethylated and gradually demethylated upon induction of differentiation in 3T3-L1 cultures. Furthermore, promoter methylation in lean mice was increased compared to ob mice. Studies on histone modifications unravel a similar picture. Several independent research groups replicated the enzymatically coordinated deacetylation and methylation events at specific sites of H3 and H4 facilitating the transcription of the gene during adipogenesis (R3).

RXRα

The search for RXRα, the heterodimerization partner of PPARγ, has not revealed much literature.

Genetic polymorphisms: When screening the gene for DNA variants, Hegele et al (Hegele and Cao 2001) identified 3 common variations in Caucasians: -25 A>G, 69 G>A, and 1470 C>T. Though, the receptor has an important role in development and metabolism, no study associated any genetic variation with obesity traits.

Epigenetic modifications: The search for evidence on epigenetic regulation in the gene revealed one study examining the effects of green tea on demethylation of the promoter region in cancerous mice (Volate, Muga et al. 2009).

C/EBPα

Genetic polymorphisms: The gene encoding for C/EBPα holds several polymorphisms that have been associated with different types of cancers. Two groups have studied their contribution to obesity (table R4).

Epigenetic modifications: Eleven studies reported on evidence for DNA methylation and six studies on histone modifications (tables R5 and R6). Heterogeneous responses to inhibitors of histone and DNA modifying enzymes demonstrate both mechanisms beeing involved in the transcriptional regulation of the gene. C/EBPα is incapable of a direct interaction with DNA-Mtase[1] (Hervouet, Vallette et al. 2009) but several studies demonstrated that a C/EBPα-HDAC complex is able to silence several protein coding genes due to aging (Jin, Wang et al. 2009).

ADRB3

Genetic polymorphisms: Studies observing associations of the SNP Trp64Arg (rs4994) with obesity are summarized in tables R7 and R8. The last meta-analysis, published in 2008, included 53 studies in Caucasians and had an overall size of about 23,000 individuals. No positive association/s could be detected. Since then, ten more studies had been published associating the SNP with obesity traits (table R8). Seven studies with an overall all size of 8,323 subjects did not detect an effect, while 4 studies (1,267 subjects) each did detect positive associations.

[1] DNA methyltransferases are enzymes transferring a methylgroup to CpGs.

Epigenetic modifications: Searching for epigenetic evidence of the ADRB3 gene itself did not reveal any studies. However, the stimulation of the receptor induces the expression and recruitment of a demethylase to the promoter of the UCP1 gene and facilitates the ability of PPARγ, RXRα and their co-activators to bind to the response element on the DNA.

UCP1

Genetic polymorphisms: Groups observing genetic variation in the *UCP1* gene merely focused on the A-3826G substitution which was associated in fifteen different studies (table R9). A gender specific effect observed by two groups (total 701 women) could not be replicated by two other groups (total 77 women). Furthermore, a study including 722 men did not detect any effect. Eleven studies with an overall size of 4,466 subjects could not positively associate the polymorphism with obesity, even though 3 studies (all together 557 subjects) did.

Epigenetic modifications: Two observations of the promoter region on DNA methylation and three on histone modifications (table R10) in several different tissues revealed tissue specificity. In 3T3-L1 cultures gene, silencing by DNA methylation and repressive histone modifications have been demonstrated by different groups.

PGC1α

Genetic polymorphisms: The coactivator PGC1α holds the SNP rs8192678 which changes the amino acid sequence at position 482 from glycin to serin. It is the most examined polymorphism in the gene in relation to obesity. The thirteen studies revealed by literature are summarized in table R11. No effect of the variant was detected by studies (all together 16,000 individuals) that included patients with attributes of the metabolic syndrome (one meta-analysis included). When associating healthy obese carriers, no association was observed in over 800 persons, while positive associations were found in over 2,200 persons. Also, gender specific effects had been detected. In 467 women the variant had been positively associated with a higher BMI and related phenotypes. On the other hand, higher weight gain could be detected in male

but not in female carriers of the polymorphism. Literature search revealed no study observing histone modifications of *PGC1α*.

Epigenetic modifications: Four studies reported on evidence for DNA methylation (table R12), while the literature search revealed no study on histone modification. DNA methylation of the promoter region in skeletal muscle and pancreatic cells from diabetics could be demonstrated to be increased compared to cells from healthy controls. A high fat diet and free fatty acids increased methylation even more.

LPL

Genetic polymorphisms: The *LPL* gene holds several polymorphisms. Three studies observed them with obesity phenotypes. Two different studies detected gender specific effects, each on a different SNP (rs320 and rs328) (summarized in table R13).

Epigenetic modifications: Epigenetic evidence for the regulation of *LPL* revealed 3 studies on DNA methylation and 3 studies on histone modifications (table R14). In adipose stem cells, DNA methylation in the promoter regions could be demonstrated. Furthermore, during adipogenesis, histone modification of the gene changes. Treatment of 3T3-L1 cultures with diallyl sulfide (a compound in garlic) results in decreased activity of HDAC which correlates with gene expressions.

FABP4

Genetic polymorphisms: One study observed the rs1054135 SNP in the gene of FABP4 in 309 subjects of mixed ethnicity and found a positive association with BMI (table R15). Epigenetic evidence for the regulation of *LPL* and *FABP4* are summarized in tables R14 and R16.

Epigenetic modifications: Epigenetic evidence for the regulation of *FABP4* revealed 3 studies on DNA methylation and 3 studies on histone modifications (table R16). In adipose stem cells, DNA methylation in the promoter regions could be demonstrated. Furthermore, during adipogenesis, histone modifications of the gene changes. Treatment of 3T3-L1 cultures with diallyl

sulfide (a compound in garlic) results in decreased activity of HDAC which correlates with gene expressions.

Joint effects of observed polymorphisms

Genetic association studies observing joint effects of polymorphisms are listed in tables R17 and R18. Carrier of the Pro12Ala variant in the *PPARγ* and the Trp64Arg variant in the ADRB3 gene had a higher BMI. Gender specific effects were detected for polymorphism rs320 in the *LPL* gene (independent of the simultaneous occurrence of the *ADRB3* SNP (rs4994)). Most, but not all, reports suggest a synergistic effect of the variant in the UCP1gene and the *ADRB3* variant.

Table R1 Meta-analyses observing associations between SNP rs1801282 in the PPARγ gene and BMI.

Year	*Subjects*	*Phenotype*	*+/-*	*P-value*	*Additional information*	*References*
2003	19,136	BMI	+	0.019	Mixed ethnicities diabetics and non-diabetics	(Masud and Ye 2003)
2006	19,699	BMI	+	0.015		(Tonjes, Scholz et al. 2006)

Table R2 Evidence on DNA methylation of the *PPARγ* gene.

Tissue	*Evidence*	*References*
Rat hepatocytes	due to dietary protein restriction of female pregnant rats PPARy1 promoter methylation in offspring did not differ	(Lillycrop, Phillips et al. 2005)
Human adipose stem cells	Overall CpG hypomethylation of PPARy2 due to senescence	(Noer, Boquest et al. 2007)
wt- mice Lep+/Lep+ mice Diet-induced ob mice	Increased promoter methylation in wt-mice	(Fujiki, Kano et al. 2009)
3T3-L1	PPARy2 promoter hypermethylation. Progressive demethylation upon induction of differentiation	
Human umbilical cord	Hypermethylation of selected CpGs in promoter region	(Gemma, Sookoian et al. 2009)
Adipose stem cells	Largely unmethylated in PPARy2 promoter region	(Sorensen, Timoskainen et al. 2009)
Human gastric mucosa of ulcer patients, cancer patients, and controlls	Hypomethylation of promoter region in ulcer patients and cancer patients compared to methylation status in controls	(Hong, Oh et al.)

Table R3 Evidence on histone modifications at the PPARγ gene.

Modification	*Tissue*	*Evidence*	*References*
H3K4 acetylation	Mammary rat tumors	Cancerogenesis: in cancer-rats increased acetylation and phosphorylation associated with less expression	(Tikoo, Kumar et al. 2009)
	3T3-L1	Adipogenesis: hyperacetylation in PPARy2 gene during ongoing differentiation and hyperacetylation in PPARy1 gene before induction of differentiation.	(Okamura, Kudo et al. 2009)
H3K4 methylation	embryonic fibroblast	Adipogenesis: association with 2 methyltransferases	(Cho, Hong et al. 2009)
	Long-term cultured undifferentiated adipose stem cells	Inactive PPARy2 promoter	(Noer, Lindeman et al. 2009)
	Differentiated adipose stem cells	No change of trimethylation in PPARy2 promoter region when differentiated	
H3K4 phosphorylation	Mammary rat tumors	Cancerogenesis: in cancer rats increased acetylation and phosphorylation associated with less expression	(Tikoo, Kumar et al. 2009)
H3K9 acetylation	Retrovirus infected fibroblasts 3T3-L1	Adipogenesis: Hyperacetylation during differentiation (PPARy2 promoter activation)	(Salma, 2007)
	NIH 3T3 3T3-L1	Adipogenesis: acetylation upon differentiation (PPARy2 promoter)	(Badri, Zhou et al. 2008)
	Long-term cultured undifferentiated adipose stem cells	Inactive PPARy2 promoter	(Noer, Lindeman et al. 2009)
	Differentiated adipose stem cells	Hyperacetylation in promoter region	
H3K9 methylation	Mesenchymal stem cells	Osteoblastogensis/adipogenesis: inactivation of PPARy through methylation	(Takada, Suzawa et al. 2007)

Table R3 continued

Modification	*Tissue*	*Evidence*	*References*
H3K14 acetylation	Retrovirus infected fibroblasts 3T3-L1	Adipogenesis: Hyperacetylation during differentiation (PPARy2 promoter activation)	(Salma, Xiao et al. 2004)
	NIH 3T3 3T3-L1	Adipogenesis: acetylation upon differentiation (PPARy2 promoter)	(Badri, Zhou et al. 2008)
H3K27 methylation	embryonic fibroblasts	Adipogenesis: association with H3K27 specific demethylase	(Cho, Hong et al. 2009)
	Long-term cultured undifferentiated adipose stem cells	Inactive PPARy2 promoter	(Noer, Lindeman et al. 2009)
	Differentiated adipose stem cells	Demethylation in PPARy2 promoter region	
	Hepatic stellate cells	Cell differentiation: repression mediates anti-adipogenic transdifferentiation	(Zhu, Wang et al.)
	Myofibroblasts	Cell differentiation: repression mediates anti-adipogenic transdifferentiation	(Mann, 2007) (Mann, 2010)
H4 acetylation	Retrovirus infected fibroblasts 3T3-L1	Adipogenesis: tetra-acetylated before onset of differentiation (PPARγ2 promoter activation)	(Salma, Xiao et al. 2004)
H4K16 acetylation	NIH 3T3 3T3-L1	Adipogenesis: acetylation upon differentiation (PPARy2 promoter)	(Badri, Zhou et al. 2008)
H4K20 methylation	3T3-L1	Adipogenesis: PPARγ self upregulation	(Wakabayashi, Okamura et al. 2009)

Table R4 Studies observing associations between SNPs in the C/EBPα gene and ob phenotypes.

SNP	*Phenotype*	*+/-*	*P*	*Subjects*	*Additional information*	*References*
rs12691	BMI	-	-.	147 male 381 female	SOS study	(Olofsson, Orho-Melander et al. 2008)
		-	-	2248 male 2618 female	Botnia study 30 % diabetes mellitus 2 patients	
rs41490344	WHR	+	0.02	537	Leeds family study	(Bennett, Nsengimana et al.)
rs41504144		-	-			
rs41344151						
rs16967952						
rs34508287						
rs41428545						
rs34529039						
rs41367646						
rs1049969						
rs707656						

Table R5 Evidence on DNA methylation in the C/EBPα gene.

Tissue	*Evidence*	*References*
Granulocytes from human acute myeloid leukemia patients	Aberrant gene methylation in infrequent	(Chim, Wong et al. 2002)
Endometrial cancer cell line	DNA methyltransferases inhibitor increased gene expression	(Takai, Kawamata et al. 2005)
Human hepatoma cell line	Gene expression was not induced after treatment with DNA methyltransferase inhibitor but together with HDAC inhibitor.	(Dannenberg and Edenberg 2006)
Lung cancer cell lines and human lung primary tumor	Methylation in upstream promoter region is associated with low or absent gene expression Treatment with DNA methyltransferases inhibitor increased gene expression	(Tada, Brena et al. 2006)
Bone marrow samples from acute myeloid leukemia	No correlation with upstream promoter methylation and gene expression level	(Hackanson, Bennett et al. 2008)
Placenta and Cord blood		(Katari, Turan et al. 2009)
Bone marrow or spleen from leukemic, transgenic mice	Downregulation of expression is due to methylation, but not to HDAC	(Guibal, Alberich-Jorda et al. 2009)
Human head and neck squamous cell carcinoma	Downregulation of expression is due to methylation in upstream promoter region.	(Bennett, Romigh et al. 2009)
Pancreatic cancer cell line	Upregulated gene expression after treatment with DNA methyltransferases inhibitor due to upstream promoter demethylation. Effect was greater in combination when additionally treated with HDAC inhibitor.	(Kumagai, Akagi et al. 2009)
Hematopoietic tumor cell line	Hypermethylation of proximal promoter is association with transcriptional silencing	(Jost, do et al. 2009)
Human hepatocellular carcinomas	Expression regulation due to methylation of upstream promoter region	(Lu, Leung et al.)

Table R6 Evidence on histone modifications of the C/EBPα gene.

Modification	*Tissue*	*Evidence*	*References*
H3 acetylation	Lung cancer cell lines and human lung primary tumor	Histone acetylation was directly associated with DNA methylation of the upstream promoter and inversely with gene expression.	(Tada, Brena et al. 2006)
	Pancreatic cancer cell line	HDAC inhibitors increase gene expression	(Kumagai, Wakimoto et al. 2007)
	Pancreatic cancer cell line	Upregulated gene expression after treatment with HDAC inhibitor. Effect was enhanced in combination with DNA methyltransferases inhibitor.	(Kumagai, Akagi et al. 2009)
	Human hepatocellular carcinomas	Upregulation of expression depends on deacetylation of H3.	(Lu, Leung et al.)
-	Human hepatoma cell line	HDAC inhibitor increases gene expression	(Dannenberg and Edenberg 2006)
H3K4me3	Mouse embryonic fibroblasts and preadipocytes	Enrichment of trimethylation of H3K4 in promoter region ends up in defects in adipogenesis	(Cho, Hong et al. 2009)
H4 acetylation	Lung cancer cell lines and human lung primary tumor	Histone acetylation was directly associated with DNA methylation of the upstream promoter and inversely with gene expression.	(Tada, Brena et al. 2006)

Table R7 Meta-analysis observing associations between SNP rs4994 in the ADRB3 gene and BMI in Caucasians.

Year	*Subjects*	*Phenotype*	*+/-*	*P-value*	*Additional information*	*References*
1998	3,866	BMI	-	-		(Allison, Heo et al. 1998)
2008	23,198	BMI	-	-		(Kurokawa, Young et al. 2008)

Table R8 Studies associating obesity phenotypes and the SNPs rs4994 since the publication of (Kurokawa, 2008).

Phenotype	*+/-*	*P*	*Subjects*	*Additional information*	*References*
BMI, WC, WHR	-	-	2,903 male 2,919 female		(Gjesing, Andersen et al. 2007)
BMI	+	< 0.001	100 male (+25 controls) 165 female (+53 controls)	Severly ob. Increased BMI was related to age in ob.	(Bracale, Pasanisi et al. 2007)
BMI	+	< 0.01	514	387 patients with obstructive sleep apnoea syndrome	(Pierola, Barcelo et al. 2007)
BMI, FM, WC Android/gynoid fat deposit	-	-	200 female	Postmenopausal.	(Lwow, Dunajska et al. 2007)
BMI Weight WHR	-	-	1,519	EPIC study	(Zafarmand, van der Schouw et al. 2008)
BMI, FM, WC, WHR, Weight	+	< 0.05	63 male 154 female	Ob.	(de Luis, Aller et al. 2008)
FFM	-	-	63 male 154 female	Ob.	
BMI	-	-	60 ob 33 lean	Children	(Zawodniak-Szalapska, Stawerska et al. 2008)
BMI WC	-	-	284 female	Postmenopausal.	(Dunajska, Lwow et al. 2008)
Weight loss	+	< 0.05	48 male 145 female	Ob. Less response to 2 month hypocaloric diets	(de Luis, Gonzalez Sagrado et al. 2009)
BMI Weight, FFM, FM, WC, WHR	-	-	93 male 131 female	Ob.	(de Luis, Ballesteros et al.)

Table R9 Studies observing associations between the SNP rs in the UCP1 gene and ob phenotypes.

Phenotype	*+/-*	*P*	*Subjects*	*Additional information*	*References*
BMI	+	0.02	526 female	Ov and ob	(Heilbronn, Kind et al. 2000)
	+	0.008	172 female	Lean and ob	(Ramis, Gonzalez-Sanchez et al. 2004)
	+	n.a.	295	With and without family history of type 2 diabetes	(Sramkova, Krejbichova et al. 2007)
	-	-	562 male	Juvenile-onset ob or lean and healthy	(Urhammer, Fridberg et al. 1997)
			985	Lean and ob	(Gagnon, Lago et al. 1998)
			1,020	Diabetomobile study	(Schaffler, Palitzsch et al. 1999)
			193	Non-diabetics and type 2 diabetes	(Sivenius, Valve et al. 2000)
			379	Lean and ob	(Urhammer, Hansen et al. 2000)
			118	Ov and ob relatives	(Kiec-Wilk, Wybranska et al. 2002)
			308	Lean and ob	(Nieters, Becker et al. 2002)
			162	Lean and ob	(Herrmann, Wang et al. 2003)
BMI	-	-	122	40 ob families	(Malczewska-Malec, Wybranska et al. 2004)
			160 male		(Sramkova, Krejbichova et al. 2007)
Others	-	-	24	12 pairs of identical twins	(Valve, Heikkinen et al. 1998)
Higher weight gain	+	0.02	238	ob	(Clement, Ruiz et al. 1996)
	-	-	77 female	Ob	(Fogelholm, Valve et al. 1998)
	-	-	985	Lean and ob	(Schaffler, Palitzsch et al. 1999)
Lower weight loss	+	n.a.	24	12 pairs of identical twins	(Ukkola, Tremblay et al. 2001)
	-	-	170	Ob	(Valve, Heikkinen et al. 1998)

Table R10 Evidence for epigenetic regulations of the UCP1 gene.

Modification	*Tissue*	*Evidence*	*References*
CpG methylation	3T3-L1	72-80 % enhancer methylation 40 % promoter methylation	(Kiskinis, Hallberg et al. 2007)
	3T3-L1 HIB-1B Ex vivo mouse tissue (BAT, WAT, liver)	Tissue specific DNA methylation pattern in enhancer and promoter	(Shore, Karamitri et al.)
Histone modification	3T3-L1	Repressive histone modifications in enhancer and promoter region.	(Kiskinis, Hallberg et al. 2007)
Histone methylation	3T3-L1 HIB-1B Ex vivo mouse tissue (BAT, WAT, liver)	Tissue specific methylation of histones in enhancer and promoter region	(Shore, Karamitri et al.)
H3K9 methylation	ex vivo, mouse tissue	Dynamical methylation state of H3K9 regulates UCP1 activation	(Tateishi, Okada et al. 2009)

Table R11 Studies observing associations between the SNP rs8192678 in the PGC1α gene and ob phenotypes.

Phenotype	*+/-*	*P*	*Subjects*	*Additional information*	*References*
BMI WC HC FM	+	<0,005	467 female	Association only in female heterozygote for the allele	(Esterbauer, Oberkofler et al. 2002)
	-	-	591 male	Association only in female heterozygote for the allele	
BMI	-	-	2,349	participants suffering of Metabolic syndrome	(Ambye, Rasmussen et al. 2005)

Table R11 continued.

Phenotype	*+/-*	*P*	*Subjects*	*Additional information*	*References*
BMI WC	-	-	79 male 197 female	Participants were morbidly ob	(Vohl, Houde et al. 2005)
BMI	+	0,033	156		(Pihlajamaki, Kinnunen et al. 2005)
BMI	-	-	3,718 diabetics 4,818 controls	Meta-analysis for type 2 diabetes association	(Barroso, Luan et al. 2006)
BMI	-	-	200 women	100 gestational diabetes 100 controls	(Leipold, Knoefler et al. 2006)
FM FFM	-	-	691		(Franks, Ekelund et al. 2007)
Change in bw and FM	-	-	63 male 73 female	9 month intervention (diet & physical activity). BMI >27 kg/m^2 and a family history of type 2 diabetes.	(Stefan, Thamer et al. 2007)
mRNA expression	+	< 0,001	34 non-diabetics	Allele carrier had lower expression	
Mitochondrial DNA content	-	-	157		(Gianotti, Sookoian et al. 2008)
BMI	-	-	13,949	Meta-analysis on genotype blood pressure association. Mixed ethnicities	(Vimaleswaran, Luan et al. 2008)
BMI WC FM FFM Change in bw	-	-	93 male 87 female	8 week low-caloric diet. Participants were ov and ob.	(Goyenechea, Crujeiras et al. 2008)
Higher weight gain	+ -	0.0045 -	302 male 280 female	Mean weight gain: 17 kg over 6.2 yrs intensive therapy. Type 1 diabetics.	(Deeb and Brunzell 2009)
WC	+	0.02	2,101	Boys and girls from the European Youth Heart Study	(Brito, Vimaleswaran et al. 2009)

Table R12 Evidence for DNA methylation of the PGC1α gene.

Tissue	*Evidence*	*References*
Human pancreatic islets from non-diabetic and type 2 diabetic subjects	Twofold promoter methylation in diabetic compared to non diabetic subjects.	(Ling, Del Guerra et al. 2008)
Human skeletal muscle tissue from non-diabetic and type 2 diabetic subjects	Promoter hypermethylation in diabetic subjects.	(Barres, Osler et al. 2009)
Myocytes from healthy male humans	Free fatty acids induce cytosine methylation at non-CpG sites. DNMT3b silencing prevented methylation.	
Human umbilical cord	No association among characteristics of newborns and status of promoter methylation. Positive correlation between maternal BMI and promoter methylation.	(Gemma, Sookoian et al. 2009)
Skeletal muscle tissue of male humans	After 5 days high fat overfeeding subjects born with low birth weight had higher promoter methylation.	(Brons, Jacobsen et al.)

Table R13 Studies observing associations between SNPs in the LPL gene and ob phenotypes.

SNP	*Phenotype*	*+/-*	*P*	*Subjects*	*Additional information*	*References*
rs285	Abdominal ob	+	0.013	58 female 47 male	No information about ethnicity	(Sertic, Juricic et al. 2009)
rs320	BMI Lower risk ov	+	<0.02	587 female	Association only in female	(Corella, Guillen et al. 2001)
		-	-	476 male		
rs328	BMI FM	+	0.01	249 female	20 weeks physical activity Association only in female	(Garenc, Perusse et al. 2001)
		-	-	231 male		

Table R14 Evidence for epigenetic regulation of the LPL gene.

Modification	*Tissue*	*Evidence*	*References*
CpG methylation	Human adipose stem cells	Promoter hypomethylation in clonally cultured adipose stem cells	(Noer, Sorensen et al. 2006)
		Promoter hypermethylation in senescent undifferentiated adipocytes	(Noer, Boquest et al. 2007)
	Human prostate cancer cells	Promoter methylation observed in prostate cancer cells vs. no methylation in controls	(Kim, Cheng et al. 2009)
H3 acetylation	3T3-L1 treated with diallyl disulfide	Decreased HDAC activity upon treatment correlates with gene expression	(Lee, Kim et al. 2007)
	Murine embryonic fibroblasts	Recruitement of HDAC1, HDAC3 and histone methyltransferase upon induction of adipogenesis	(Fu, Rao et al. 2005)
H3K4 methylation	Differentiated adipose stem cells	No change of trimethylation in promoter region when differentiated	(Noer, Lindeman et al. 2009)
	Undifferentiated adipose stem cells	Enrichment in promoter	
H3K9 acetylation	Murine embryonic fibroblasts	Recruitement of HDAC1, HDAC3 and histone methyltransferase upon induction of adipogenesis	(Fu, Rao et al. 2005)
	Undifferentiated adipose stem cells	No modification in promoter region	(Noer, Lindeman et al. 2009)
	Differentiated adipose stem cells	Hyperacetylation of H3K9 and demethylation of H3K27 in promoter region	
H3K9 methylation	Undifferentiated adipose stem cells	No modification in promoter region	
H3K27 methylation	Differentiated adipose stem cells	Hyperacetylation of H3K9 and demethylation of H3K27 in promoter region	
	Undifferentiated adipose stem cells	Enrichment in promoter	
H4 acetylation	3T3-L1 treated with diallyl disulfide	Decreased HDAC activity upon treatment correlates with gene expression	(Lee, Kim et al. 2007)

Table R15 A study observing the association between the SNP rs1054135 in the FABP4 gene and BMI.

Phenotype	*+/-*	*P*	*Subjects*	*Additional information*	*References*
BMI	+	<0.001	182 ob 127 non-ob	5-7 year old children of mixed ethnicities	(Khalyfa, Bhushan et al.)

Table R16 Evidence for epigenetic regulation of the FABP4 gene.

Modification	*Tissue*	*Evidence*	*References*
CpG methylation	adipose stem cells	Promoter hypermethylation remains stable upon differentiation	(Noer, Sorensen et al. 2006)
	Senescent undiff. adipose stem cells	Variation in promoter methylation	(Noer, Boquest et al. 2007)
	Adipose stem cells	Promoter region is largely unmethylated	(Sorensen, Timoskainen et al. 2009)
H3 acetylation	3T3-L1	Decreased HDAC activity upon treatment with diallyl disulfide correlates with gene expression	(Lee, Kim et al. 2007)
H3K4 methylation	Mouse embryonic fibroblasts	Methylation is essential for PPARγ-dependent adipogenesis	(Lee, Saha et al. 2008)
	Undiff. adipose stem cells	Enrichment in promoter	(Noer, Lindeman et al. 2009)
	Differentiated adipose stem cells	No change of trimethylation upon differentiation	
H3K9 acetylation	Differentiated adipose stem cells	Hyperacetylation of H3K9 in promoter region	
	Undiff. adipose stem cells	No modification in promoter region	
H3K9 methylation	Undiff. adipose stem cells	No modification in promoter region	
H3K27 methylation	Undiff. adipose stem cells	Enrichment in promoter	
	Differentiated adipose stem cells	Demethylation of H3K27 in promoter region	
H4 acetylation	3T3-L1	Decreased HDAC activity upon treatment with diallyl disulfide correlates with gene expression	(Lee, Kim et al. 2007)

Table R17 Studies observing the association of joint effects in SNPs and ob phenotypes.

SNP1	*SNP2*	*Phenotype*	*+/-*	*P*	*Subjects*	*Additional information*	*References*
rs4994 (ADRB3)	rs1801282 (PPARγ)	BMI	+	0.04	453	10 Mexican-American families	(Hsueh, Cole et al. 2001)
			+	0.039	185 ob 185 controls	Ob children (5-18 y) with sex- and age-matched controls	(Ochoa, Marti et al. 2004)
		FM WC	-	-	453	10 Mexican-American families	(Ek, Andersen et al. 2001)
rs1801282 (PPARγ)	rs8192678 (PGC1α)	BMI	-	-	430 male 254 female 239 male controls 270 female controls	Type 2 diabetics and controls	(Ek, Andersen et al. 2001)
rs4994 (ADRB3)	rs320 (LPL)	BMI	+	0.026	587 female	Gender-specific effect	(Corella, Guillen et al. 2001)

Table R18 Studies observing additive effects of the SNP rs in the UCP1 gene and rs4994 in the ADRB3 gene and associations with ob phenotypes.

Phenotype	*+/-*	*P*	*Subjects*	*Additional information*	*References*
BMI	+	0.060	379	Lean and ob.	(Urhammer, Hansen et al. 2000)
	-	-	118	Ov and ob relatives	(Kiec-Wilk, Wybranska et al. 2002)
High weight gain	+	0.05	238	Ob.	(Clement, Ruiz et al. 1996)
		0.051	77 female	Ob.	(Fogelholm, Valve et al. 1998)
		0.005	985	Lean and ob.	(Schaffler, Palitzsch et al. 1999)
		0.036	379	Lean and ob.	(Urhammer, Hansen et al. 2000)
Low basal metabolic rate	+	0.002	170	Ob.	(Valve, Heikkinen et al. 1998)
Leptin levels	-	-	172 female	Lean and ob.	(Ramis, Gonzalez-Sanchez et al. 2004)

DISCUSSION

Aim of this work was to highlight the need for going beyond the scope in human genome epidemiology of complex traits and diseases. Therefore, in a case study of the metabolic syndrome, a pathway of adaptive thermogenesis had been chosen to demonstrate the great role of the epigenome on gene expression compared to SNP influence.

Evidence for obesity associated SNPs & epigenetic modification

During adipogenesis, enzymatic coordination of changes in histone modification and DNA demethylation in the promoter regions of *PPARγ* and *C/EBPα* are essential. If polymorphisms in these genes have effects on the activity of the transcription factors will have to be warranted in further studies. Previous studies have not been properly sized to detect effects. Although, a meta-analysis is a powerful tool to confirm evidence, 19,600 subjects might still not be enough to draw a clear conclusion on the variants ability to increase the risk for developing Ob. Similar uncertainties remain with Trp64Arg (rs4994) in *ADRB3* and rs8192678 in the *PGC1α*. Meta-analysis reported big in-between study heterogeneity or did not distinguish between ethnic subgroups. The literature search did not reveal any study observing epigenetic modification in the *ADRB3.* I suggest future research focusing on epigenetic regulation of the *ADRB3.* More studies are also needed on effects of variations in the genes *FABP4*, *LPL,* and *RXRα*. The two PPARγ target genes, *LPL* and *FABP4,* are important factors in pathways ensuring steady substrate availability in mitochondria. Polymorphisms in the *LPL* gene probably influence the enzyme's capacity to facilitate free fatty acids. Epigenetic regulation of *PGC1α, FABP4*, and *LPL* was demonstrated several times.

Stimulation of the ADRB3 changes the epigenetic state of the *UCP1* gene and facilitates it for transcription factors inducing expression. The A-3826G substitution of the *UCP1* gene had been associated with reduced UCP1 mRNA

expression (Esterbauer, Oberkofler et al. 1998). Regardless, association studies represent no clear evidence on its contribution to Ob. These inconsistencies will have to be elucidated further.

Current techniques allow simultaneous testing of many SNPs. A synergistic effect of the UCP1 and the ADRB3 polymorphisms could be detected if they reduced functions or expressions of the genes. Anyhow, by now, no study had appropriate sample sizes to detect differences.

Characteristically shortcomings in human genome epidemiology

General concerns in human genome epidemiology have already been discussed in previous chapters. Additionally to them, some interesting yet not discussed problems occurred and have to be mentioned.

The re-evaluation of existing evidence has turned out to be an efficient and suitable approach to extract information from a great amount of small studies. Literature search revealed all together five meta-analysis in the genes PPARγ, ADRB3, and PGC1α. Each was reported to have big in-between study heterogeneity. Such heterogeneities make it difficult to synthesize information across studies in an appropriate way.

The inconsistency in results of studies observing joint effects let suggest that traditional methods of building multiloci models simply fail to identify combinations.

None of the studies reported any validation study. As previously discussed, they are a helpful tool to analyze uncontrolled bias.

This review limited the search to Caucasians because SNP associations identified in one population are often not transferable to another population. Surprisingly, the comparison of results from studies on mixed ethnicities did not reveal any differences. Is this because there are no differences? Obesity is assumed to develop from a heterogeneous background, many different ways and environments involved. A profile describing a combination of risk factors is extremely unlikely to be found in more than one person.

Much more effort and money had been put into association studies of genetic variation than in observations of epigenetic variation. Though, the majority of observed polymorphisms could not be demonstrated to alter the network qualitatively.

Characteristics in studies of epigenetic epidemiology

Although, in epigenetic epidemiology high-throughput technologies already allow for the massive investigations of the epigenome, currently no consensus to the most appropriate way to model and present data exists. Approaches tend to be discovery-oriented rather than oriented to testing specific hypotheses. In studies revealed by the literature search, DNA methylation was merely measured at specific sites and hardly as the overall level of methylation pattern in a tissue. If DNA methylation was documented in detail, it was mostly referred to as % methylation of site. Histone modifications have been specified by histone number, position and modification. Some of the articles reported in- or decreased concentrations of specific histone modifying enzymes.

Epigenetic modifications are able to regulate gene transcription in adaptive thermogenesis. How much environmental factors influence this part of the epigenome is still unclear.

CONCLUSION (I) *CASE STUDY*

The results of the case study raise the question whether epigenetic mechanisms will explain the inconsistencies in human genome epidemiology. BAT content is highest in older infants and young children compared to preterm infants, stillborns, or human adults in whom WAT dominates (Lean, 1986). The less prominent role of thermogenesis in BAT in adult humans might be due to transcriptional downregulation of *UCP1*. A lesser activity of UCP1 is followed by more efficient energy coupling and lower metabolic rates – both predictors of weight gain (Cypess and Kahn). UCP1 deficient mice that are reared at room temperature gain weight faster than WT (Silva 2006) and obese humans have reduced expression of UCP1 mRNA down to 50 %. It is speculated that in adult humans, as little as 50 g of BAT could utilize up to 20% of basal caloric needs if maximally stimulated (Rothwell, Stock et al. 1983). If transdifferentiation from WAT to BAT was regulated by epigenetic mechanisms due to environmental influences, epigenetics could explain at least some of the inconsistencies between studies of human genome epidemiology on obesity.

CONCLUSION (II) *FUTURE EPIDEMIOLOGIC STRATEGIES*

Network approaches

An alternative to the forward genetics approach is the construction of molecular networks. In a system of multiple tissues, genetic, and environmental factors are acting together, the partitioning into component networks is able to increase the chance for elucidating variations in susceptibility to the phenotype. In recent studies the reconstruction of networks revealed how genetic variation, gene expression, and clinical traits may be combined and are able to explain causal associations between expression and disease traits (Schadt, Sachs et al. 2005; Chen, Zhu et al. 2008). These integrative approaches demonstrate the illumination of complex systems by incorporating multiple sources of information instead of considering each source independently. Integration of epigenetic data in discovery orientated approaches can reveal information that is not captured by previous observations. Figueroa et al combined epigenetic data with expression arrays in acute leukemia of distinct cell lineages. The analysis revealed hundreds of additional differentiated expressed genes that were missed by single-platform microarray studies. Furthermore, it enhanced the detection and statistical significance of biological pathways that were deregulated conditions (Figueroa, Reimers et al. 2008).

Validating epigenetic markers

The Human Epigenome project (www.epigenome.org/) aims to identify, catalogue and interpret tissue specific genome wide DNA methylation patterns. Markers are being classified by the type of variation (e.g. heritable) and by intra- and inter individual variants. Information o relationships (e.g. epigenotype – haplotype association) on the markers are of interest (Bjornsson, Fallin et al. 2004) will have to be catalogued as well. Further, influences of time and other environmental factors (for example, nutritional compounds) will have to be analyzed systematically. In unraveling the histone code and the landscape of RNAs, such projects are still missing.

Animal models combine advances of genetic purity and environmental flexibility. Lately, a new approach was introduced to analyze large gene groups across species. It allows the identification of higher order systems of disease and phenotypes (Zhang, De et al.). Though, genetic or epigenetic variation has not been integrated yet, this method is a promising starting point to apportion findings from animal models if human tissue is unavailable.

The annotation of validated epigenetic marker for gene maps will make it easier to assess the importance of including measured epigenotypes in gene discovery, or for evaluating the impact of time and environment on a genetic predisposition.

Integrating epigenetic data to human genome epidemiology

In a complex condition the knowledge of all genetic and environmental factors and of their interactions is required to explain the full etiology and identify subgroups. In 2008, Khoury et al (Rothman 2008) predicted a new level of complexity in epidemiology by simultaneously studying biomarkers of exposure, susceptibility, and outcomes. The integration of epigenetic markers into large-scale population based studies that test for hypothesis in human genome epidemiology could be a solution for the problems of incomplete penetrance and variable expressivity.

The common disease genetic and epigenetic (CDGE) model (figure 4) overlies the ‘common genetic variant, common disease’ hypothesis with an epigenetic component interacting. It will be helpful when designing such a study in the future (Bjornsson, Fallin et al. 2004).

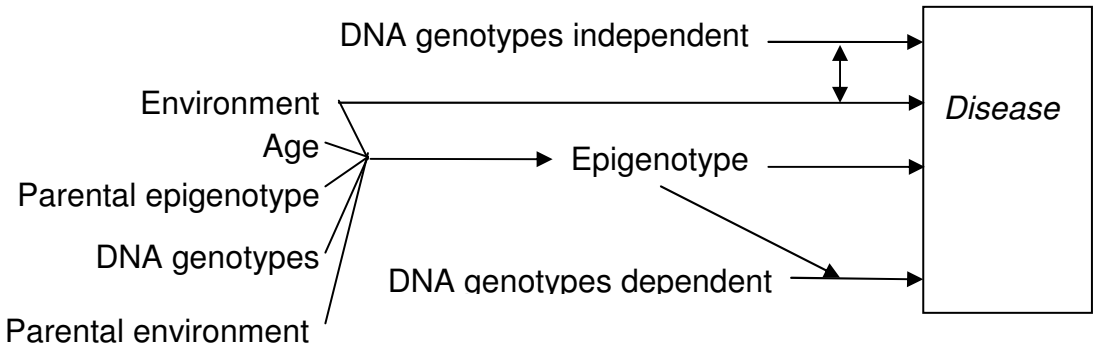

Figure 4 The CDGE model demonstrates factors (genetic, environmental, stochastic) contributing to epigenetic variation, epigenome and genome interacting indirectly and the environments influence on the disease. In some cases the epigenome will not influence the disease (DNA genotypes independent). It provides a starting framework for including epigenetic marker in epidemiologic studies.

Study designs

Unlike genetic variants, epigenetic and other markers cannot be seen as stable indicators of host susceptibility. On the one hand, diseased populations are a rich source for molecular phenotype data, while on the other hand, cohorts have been obtained and stored biologic specimens since decades and are interesting for genetic association studies (Langholz, Rothman et al. 1999). A “meet in the middle approach” emphasizes the complementarity between different types of markers and study designs and has the potential of simultaneous testing of markers of multiple levels (Vineis and Perera 2007). If population based, designs have the important advantage to provide estimates of absolute risk and population attributable risk in addition to estimates of relative risk (Khoury, Little et al. 2004).

Data analysis

With information on genotypes, molecular profiles and clinical information collected from large-scale populations will generate many data points for each sample. The complexity in data analyzes poses novel statistical challenges, including an increased likelihood to produce false-positive results. (Boffetta) Approaches of pharmacological dose-response studies will be useful (Conti, Cortessis et al. 2003). Also, stochastically models in carcinogenesis describing

the relationship of time, exposure and response are a good starting point. Furthermore, clustering approaches, or networks for analyzing complex molecular and genetic datasets will become increasingly interesting (Zhang, De et al.).

CONCLUSION (III) *GENETIC TESTING & PUBLIC HEALTH*

Consumer orientated genetic testing

The 'common disease – common variant hypothesis' postulates that the genetic influence on the susceptibility to common diseases is attributable to a limited number of variants present in more than 1% to 5% of the population. For instance, one of the first GWAs on age-related macular degeneration identified five major variants. Each variant is associated with a two to three times the risk for a person to develop the disease (Manolio). These positive associations could be replicated easily. Most common conditions have not been as amenable to investigate and results cannot be effectively translated into risk assessments and preventive health policies at present. Patients who acquire genomewide association testing should be advised that as of yet, the majority of results have no value in predicting risks and have no clinical direction.

Increasing the predictive value

A large number of biomarkers for exposure, absorption, metabolism, distribution, critical target interaction (i.e., DNA damage and repair), genetic changes, and disease have been revealed by epidemiologic research. The combined use of genetic and molecular profiles with clinical data provides more convincing results. Epigenetic key drivers with disease-specific occurrence could easily be combined with environmental risk factors and genetic susceptibility variants. Other epigenetic markers are unbiased information when assessing environmental influences. Integrating such biomarkers in genetic testing would increase their predictive value (Burke, Atkins et al. 2002) (Schadt, Sachs et al. 2005).

Challenges for public health

Today, without regards to inter-individual variations, results from previous epidemiologic investigations are applied to entire populations. In almost every population studied, a degree of disease heterogeneity is present indicating the mixture of different subtypes of the disease. With respect to a disease

phenotype, individuals may appear similar but underlying molecular processes vary widely between individuals. This dispels the “one-drug-fits-all” view and explains the lack of success of many intervention programs.

Human genome epidemiology has got the potential to unravel the heterogeneous etiologies of common complex diseases. Increasing sensitivity and specificity of epidemiological studies will help to detect different levels of diseases and to identify subgroups. Guidelines and recommendations for the evaluation and integration of data from human genome epidemiology have already been published (Burke, Atkins et al. 2002). Rothman et al further suggested an ethical duty of granting agencies and editors to require a thorough quantitative assessment of relevant literature and their systematic errors (Rothman 2008).

By now, genotype prevalence data is not concrete enough to work with on a population-level. A view from inter-individual levels offers opportunities for understanding the relationship between cumulative exposures and population. Public health will have to deal with the meaning of altered levels of biomarkers at an individual level. Such powerful tools for risk assessment provide the ability for stronger prevention strategies in the future.

The importance of following good epidemiologic practice and implementing rigorous quality assurance and control procedures cannot be overstated and helps to guarantee the successful integration of genetic tests and services into the health care system. Accurate reporting will allow the evaluation of new tests with the assurance that decisions are based on the available evidence. For instance, the “Screening Brief” featured in the *Journal of Medical Screening* could be applied to genetic tests, or the Human Genome Epidemiology reviews which are a regular feature in the *American Journal of Epidemiology* (Burke, Atkins et al. 2002).

In the United States, the National Genomic Applications in Practice and Prevention (EGAPP) Initiative was established to support the development and implementation of an evidence-based process for evaluating genetic tests and other genomic applications for clinical and public health practice. (Teutsch,

Bradley et al. 2009). Such initiatives will also need to develop guidelines (e.g., clinical guidelines) with respect to differences in health care priorities and resources in different health care settings (Burke, Atkins et al. 2002) for a scientifically tenable integration of genetic testing into the health care system.

SUMMARY

A complex trait consists of the interplay of several networks, tissues, and environments. The Epigenome includes all mechanisms and marks that alter chromatin without changing the DNA sequence. It is tissue and cell-type specific and varies as a function of time, environmental factors, and random processes.

Human genome epidemiology offers frameworks for understanding joint impacts of genes and environments regarding disease risk. Genome-wide association studies have identified many genetic variations and provided insights into the genetic architecture of diseases. However, modest effect sizes of common mutations and inadequate power to overcome the heterogeneity of genetic effects could not be revealed by stringent genome wide significance thresholds.

This work approached epigenetic influence on transcriptional regulation in a case study on adaptive thermogenesis and obesity on a network level. Results from a systematic literature search were compared to research findings from genome association studies on SNPs common in Caucasians. Whether these polymorphisms diminish energy expenditure remains unclear and has to be warranted in further studies. Altogether, the results of the analysis let suggest that hereditary genetic- and environment responsive regulation is involved. Integration of epigenetic markers in epidemiologic research could thus help to unravel multi-gene-environment interactions in the development of obesity.

Future research strategies that simultaneously study biomarkers of exposure, susceptibility, outcomes in appropriately planned, designed, and conducted studies will extend beyond current approaches. Nevertheless, sample size requirements remain an inescapable challenge. The integration of epigenetic markers in genetic testing should increase their predictive value for complex diseases with public health importance. Effective organization is vital for disseminating genetic tests to the health care system.

ZUSAMMENFASSUNG

Komplexe Krankheiten zeichnen sich durch das Zusammenwirken mehrerer genetischer und umweltbedingter Faktoren aus. In der Epigenetik werden Veränderungen des Chromatins untersucht, welche die DNA Sequenz nicht verändern jedoch auf Umweltfaktoren und Alterungsprozesse reagieren und somit die Genexpression beeinflussen.

Moderne epidemiologische Forschung versucht Möglichkeiten zu entwickeln um komplexe Gen-Umwelt Interaktionen zu verstehen. In der Vergangenheit konnten in Genom-weiten Studien einige bahnbrechende Erfolge erzielt werden. Trotzdem gab es bisher noch keine zufriedenstellenden Ansätze um ständig wiederkehrende Probleme zu lösen. Zu den Schwierigkeiten zählen die oft kleinen Effekte einzelner Polymorphismen, die heterogene Ätiologie einer komplexen Krankheit und zu kleine Studiengrößen um Effekte zu detektieren. Eine vielversprechende Herangehensweise ist die gleichzeitige Untersuchung epigenetischer und genetischer Marker.

In dieser Arbeit wurde ein Netzwerk der adaptiven Thermogenese gewählt in dem das Zusammenspiel von Umwelt und Genetik klar ersichtlich ist. Die Ergebnisse einer systematischen Literatursuche sowohl nach Assoziationsstudien als auch nach Studien epigenetischer Modifikationen ergaben, dass Epigenetik eine nicht zu vernachlässigende Rolle spielt, wogegen die Studien an Polymorphismen keine klare Evidenz für deren Beteiligung an der Entstehung von Übergewicht lieferten.

Zukünftige Forschungsansätze in gewissenhaft geplanten Studien werden verschiedene Biomarker einsetzen müssen. Jedoch stellen auch hier die Studiengrößen weiterhin ein Problem dar. Den Erfordernissen eines modernen Public Health Sektors dies dienlich sein.

Des Weiteren haben validierte epigenetische Marker großes Potential zur Steigerung der Prognosesicherheit genetischer Tests.

BIBLIOGRAPHY

(1947). "CONSTITUTION of the World Health Organization." Chron World Health Organ **1**(1-2): 29-43.

(1987). "Biological markers in environmental health research. Committee on Biological Markers of the National Research Council." Environ Health Perspect **74**: 3-9.

Albertini, R. J., J. A. Nicklas, et al. (1996). "Future research directions for evaluating human genetic and cancer risk from environmental exposures." Environ Health Perspect **104 Suppl 3**: 503-10.

Allison, D. B., M. Heo, et al. (1998). "Meta-analysis of the association of the Trp64Arg polymorphism in the beta3 adrenergic receptor with body mass index." Int J Obes Relat Metab Disord **22**(6): 559-66.

Allison, D. B., J. Kaprio, M. Korkeila, M. Koskenvuo, M. C. Neale, K. Hayakawa (1996). "The heritability of body mass index among an international saple of monozygotic twins reared apart." Metab Disord **20**: 501-6.

Ambye, L., S. Rasmussen, et al. (2005). "Studies of the Gly482Ser polymorphism of the peroxisome proliferator-activated receptor gamma coactivator 1alpha (PGC-1alpha) gene in Danish subjects with the metabolic syndrome." Diabetes Res Clin Pract **67**(2): 175-9.

Badri, K. R., Y. Zhou, et al. (2008). "Effects of the SANT domain of tension-induced/inhibited proteins (TIPs), novel partners of the histone acetyltransferase p300, on p300 activity and TIP-6-induced adipogenesis." Mol Cell Biol **28**(20): 6358-72.

Barak, Y., M. C. Nelson, et al. (1999). "PPAR gamma is required for placental, cardiac, and adipose tissue development." Mol Cell **4**(4): 585-95.

Barbatelli, G., I. Murano, et al. "The emergence of cold-induced brown adipocytes in mouse white fat depots is determined predominantly by white to brown adipocyte transdifferentiation." Am J Physiol Endocrinol Metab **298**(6): E1244-53.

Barres, R., M. E. Osler, et al. (2009). "Non-CpG methylation of the PGC-1alpha promoter through DNMT3B controls mitochondrial density." Cell Metab **10**(3): 189-98.

Barroso, I., J. Luan, et al. (2006). "Meta-analysis of the Gly482Ser variant in PPARGC1A in type 2 diabetes and related phenotypes." Diabetologia **49**(3): 501-5.

Barski, A., S. Cuddapah, et al. (2007). "High-resolution profiling of histone methylations in the human genome." Cell **129**(4): 823-37.

Bennett, C. E., J. Nsengimana, et al. "CCAAT/enhancer binding protein alpha, beta and delta gene variants: associations with obesity related phenotypes in the Leeds Family Study." Diab Vasc Dis Res **7**(3): 195-203.

Bennett, K. L., T. Romigh, et al. (2009). "Activator protein 2 alpha (AP2alpha) suppresses 42 kDa C/CAAT enhancer binding protein alpha (p42(C/EBPalpha)) in head and neck squamous cell carcinoma." Int J Cancer **124**(6): 1285-92.

Bernstein, B. E., E. L. Humphrey, et al. (2002). "Methylation of histone H3 Lys 4 in coding regions of active genes." Proc Natl Acad Sci U S A **99**(13): 8695-700.

Bird, A. (2002). "DNA methylation patterns and epigenetic memory." Genes Dev **16**(1): 6-21.

Bjornsson, H. T., M. D. Fallin, et al. (2004). "An integrated epigenetic and genetic approach to common human disease." Trends Genet **20**(8): 350-8.

Boffetta, P. "Biomarkers in cancer epidemiology: an integrative approach." Carcinogenesis **31**(1): 121-6.

Bonassi, S. and W. W. Au (2002). "Biomarkers in molecular epidemiology studies for health risk prediction." Mutat Res **511**(1): 73-86.

Bracale, R., F. Pasanisi, et al. (2007). "Metabolic syndrome and ADRB3 gene polymorphism in severely obese patients from South Italy." Eur J Clin Nutr **61**(10): 1213-9.

Bradford Hill, A. (1965). "The Environment and Disease: Association or Causation?" Proceedings of the Royal Society of Medicine **58**: 295-300.

Brito, E. C., K. S. Vimaleswaran, et al. (2009). "PPARGC1A sequence variation and cardiovascular risk-factor levels: a study of the main genetic effects and gene x environment interactions in children from the European Youth Heart Study." Diabetologia **52**(4): 609-13.

Brons, C., S. Jacobsen, et al. "Deoxyribonucleic acid methylation and gene expression of PPARGC1A in human muscle is influenced by high-fat overfeeding in a birth-weight-dependent manner." J Clin Endocrinol Metab **95**(6): 3048-56.

Brookes, A. J. (1999). "The essence of SNPs." Gene **234**(2): 177-86.

Burke, W., D. Atkins, et al. (2002). "Genetic test evaluation: information needs of clinicians, policy makers, and the public." Am J Epidemiol **156**(4): 311-8.

Bursch, W. (2010). Genotoxic, Non-Genotoxic and Epigenetic Mechanisms in Chemical Hepatocarcinogenesis: Implications for Safety Evaluation. Epigenetics and Human Health. A. G. Haslberger, S. Gressler. Weinheim, WILEY-VCH Verlag.

Campion, J., F. I. Milagro, et al. (2009). "Individuality and epigenetics in obesity." Obes Rev **10**(4): 383-92.

Cao, W., K. W. Daniel, et al. (2004). "p38 mitogen-activated protein kinase is the central regulator of cyclic AMP-dependent transcription of the brown fat uncoupling protein 1 gene." Mol Cell Biol **24**(7): 3057-67.

Caporaso, N. and A. Goldstein (1995). "Cancer genes: single and susceptibility: exposing the difference." Pharmacogenetics **5**(2): 59-63.

Chanock, S. J., T. Manolio, et al. (2007). "Replicating genotype-phenotype associations." Nature **447**(7145): 655-60.

Chen, Y., J. Zhu, et al. (2008). "Variations in DNA elucidate molecular networks that cause disease." Nature **452**(7186): 429-35.

Chim, C. S., A. S. Wong, et al. (2002). "Infrequent hypermethylation of CEBPA promotor in acute myeloid leukaemia." Br J Haematol **119**(4): 988-90.

Cho, Y. W., S. Hong, et al. (2009). "Histone methylation regulator PTIP is required for PPARgamma and C/EBPalpha expression and adipogenesis." Cell Metab **10**(1): 27-39.

Clement, K., J. Ruiz, et al. (1996). "Additive effect of A-->G (-3826) variant of the uncoupling protein gene and the Trp64Arg mutation of the beta 3-adrenergic receptor gene on weight gain in morbid obesity." Int J Obes Relat Metab Disord **20**(12): 1062-6.

Conti, D. V., V. Cortessis, et al. (2003). "Bayesian modeling of complex metabolic pathways." Hum Hered **56**(1-3): 83-93.

Corella, D., M. Guillen, et al. (2001). "Gender specific associations of the Trp64Arg mutation in the beta3-adrenergic receptor gene with obesity-related phenotypes in a Mediterranean population: interaction with a common lipoprotein lipase gene variation." J Intern Med **250**(4): 348-60.

Cousin, B., S. Cinti, M. Morroni, S. Raimbault, D. Ricquier, L. Penicaud, L. Casteilla (1992). "Occurrence of brown adipocytes in rat white adipose tissue: molecular and morphological characterization." J Cell Sci **103**(4): 931-42.

Cypess, A. M. and C. R. Kahn "The role and importance of brown adipose tissue in energy homeostasis." Curr Opin Pediatr **22**(4): 478-84.

Dalgaard, L. T. a. O. P. (2001). "Uncoupling proteins: functional characteristics and role in the pathogenesis of obesity and Type II diabetes." Diabetologia **44**(8): 946-65.

Dannenberg, L. O. and H. J. Edenberg (2006). "Epigenetics of gene expression in human hepatoma cells: expression profiling the response to inhibition of DNA methylation and histone deacetylation." BMC Genomics **7**: 181.

Davis, F. (1989). Tabler's cyclopedic medical dictionary. Philadelphia, FA Davis Company.

de Luis, D. A., R. Aller, et al. (2008). "Relation of Trp64Arg polymorphism of beta 3-adrenergic receptor gene to adipocytokines and fat distribution in obese patients." Ann Nutr Metab **52**(4): 267-71.

de Luis, D. A., M. Ballesteros, et al. "Polymorphism Trp64Arg of beta 3 adrenoreceptor gene: allelic frequencies and influence on insulin resistance in a multicenter study of Castilla-Leon." Nutr Hosp **25**(2): 299-303.

de Luis, D. A., M. Gonzalez Sagrado, et al. (2009). "Influence of Trp64Arg polymorphism of beta 3-adrenoreceptor gene on insulin resistance, adipocytokines and weight loss secondary to two hypocaloric diets." Ann Nutr Metab **54**(2): 104-10.

Deeb, S. S. and J. D. Brunzell (2009). "The Role of the PGC1alpha Gly482Ser Polymorphism in Weight Gain due to Intensive Diabetes Therapy." PPAR Res **2009**: 649286.

Dolinoy, D. C. and R. L. Jirtle (2008). "Environmental epigenomics in human health and disease." Environ Mol Mutagen **49**(1): 4-8.

Dunajska, K., F. Lwow, et al. (2008). "beta(3)-adrenergic receptor polymorphism and metabolic syndrome in postmenopausal women." Gynecol Endocrinol **24**(3): 133-8.

Ek, J., G. Andersen, et al. (2001). "Mutation analysis of peroxisome proliferator-activated receptor-gamma coactivator-1 (PGC-1) and relationships of

identified amino acid polymorphisms to Type II diabetes mellitus." Diabetologia **44**(12): 2220-6.

Espada, J. and M. Esteller "DNA methylation and the functional organization of the nuclear compartment." Semin Cell Dev Biol **21**(2): 238-46.

Esteller, M. (2008). "Epigenetics in cancer." N Engl J Med **358**(11): 1148-59.

Esterbauer, H., H. Oberkofler, et al. (2002). "Peroxisome proliferator-activated receptor-gamma coactivator-1 gene locus: associations with obesity indices in middle-aged women." Diabetes **51**(4): 1281-6.

Esterbauer, H., H. Oberkofler, et al. (1998). "Uncoupling protein-1 mRNA expression in obese human subjects: the role of sequence variations at the uncoupling protein-1 gene locus." J Lipid Res **39**(4): 834-44.

Figueroa, M. E., M. Reimers, et al. (2008). "An integrative genomic and epigenomic approach for the study of transcriptional regulation." PLoS One **3**(3): e1882.

Fogelholm, M., R. Valve, et al. (1998). "Additive effects of the mutations in the beta3-adrenergic receptor and uncoupling protein-1 genes on weight loss and weight maintenance in Finnish women." J Clin Endocrinol Metab **83**(12): 4246-50.

Foley, D. L., J. M. Craig, et al. (2009). "Prospects for epigenetic epidemiology." Am J Epidemiol **169**(4): 389-400.

Franks, P. W., U. Ekelund, et al. (2007). "PPARGC1A coding variation may initiate impaired NEFA clearance during glucose challenge." Diabetologia **50**(3): 569-73.

Frayling, T. M., N.J. Timpson, M. N. Weedon, E. Zeggini, R. M. Freathy, C. M. Lindgren, J. R. Perry, K. S. Elliott, H. Lango, N. W. Rayner, B. Shields (2007). "A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity." Science **316**(5826): 889-94.

Frazer, K. A., S. S. Murray, et al. (2009). "Human genetic variation and its contribution to complex traits." Nat Rev Genet **10**(4): 241-51.

Fu, M., M. Rao, et al. (2005). "Cyclin D1 inhibits peroxisome proliferator-activated receptor gamma-mediated adipogenesis through histone deacetylase recruitment." J Biol Chem **280**(17): 16934-41.

Fujiki, K., F. Kano, et al. (2009). "Expression of the peroxisome proliferator activated receptor gamma gene is repressed by DNA methylation in visceral adipose tissue of mouse models of diabetes." BMC Biol **7**: 38.

Gagnon, J., F. Lago, et al. (1998). "DNA polymorphism in the uncoupling protein 1 (UCP1) gene has no effect on obesity related phenotypes in the Swedish Obese Subjects cohorts." Int J Obes Relat Metab Disord **22**(6): 500-5.

Garcia-Closas, M., N. Rothman, et al. (1999). "Misclassification in case-control studies of gene-environment interactions: assessment of bias and sample size." Cancer Epidemiol Biomarkers Prev **8**(12): 1043-50.

Garenc, C., L. Perusse, et al. (2001). "Evidence of LPL gene-exercise interaction for body fat and LPL activity: the HERITAGE Family Study." J Appl Physiol **91**(3): 1334-40.

Gemma, C., S. Sookoian, et al. (2009). "Maternal pregestational BMI is associated with methylation of the PPARGC1A promoter in newborns." Obesity (Silver Spring) **17**(5): 1032-9.

Gesta, S., Y. H. Tseng, C. R. Kahn (2007). "Develpmental origin of fat: tracking obesity to its source." Cell **135**(2): 366.

Gianotti, T. F., S. Sookoian, et al. (2008). "A decreased mitochondrial DNA content is related to insulin resistance in adolescents." Obesity (Silver Spring) **16**(7): 1591-5.

Gjesing, A. P., G. Andersen, et al. (2007). "Studies of associations between the Arg389Gly polymorphism of the beta1-adrenergic receptor gene (ADRB1) and hypertension and obesity in 7677 Danish white subjects." Diabet Med **24**(4): 392-7.

Goyenechea, E., A. B. Crujeiras, et al. (2008). "Enhanced short-term improvement of insulin response to a low-caloric diet in obese carriers the Gly482Ser variant of the PGC-1alpha gene." Diabetes Res Clin Pract **82**(2): 190-6.

Guerra, C., R. A. Koza, K. Walsh, D. M. Kurtz, P. A. Wood, L. P. Kozak (1998). "Abnormal nonshivering thermogenesis in mice with inherited defects of fatty acid oxidation." J Clin Invest **102**(9): 1742-31.

Guibal, F. C., M. Alberich-Jorda, et al. (2009). "Identification of a myeloid committed progenitor as the cancer-initiating cell in acute promyelocytic leukemia." Blood **114**(27): 5415-25.

Hackanson, B., K. L. Bennett, et al. (2008). "Epigenetic modification of CCAAT/enhancer binding protein alpha expression in acute myeloid leukemia." Cancer Res **68**(9): 3142-51.

Hagmar, L. E. W., B. Paulsson, M. Tornqvist (2005). "Differences in hemoglobin adduct levels of acrylamide in the general population with respect to dietary intake, smoking habits and gender." Mutat Res **580**: 157-65.

Hales, C. N. and D. J. Barker (2001). "The thrifty phenotype hypothesis." Br Med Bull **60**: 5-20.

Haslam, D. W. and W. P. James (2005). "Obesity." Lancet **366**(9492): 1197-209.

He, H. and N. Lehming (2003). "Global effects of histone modifications." Brief Funct Genomic Proteomic **2**(3): 234-43.

Hedrick, P. and S. Kumar (2001). "Mutation and linkage disequilibrium in human mtDNA." Eur J Hum Genet **9**(12): 969-72.

Hegele, R. A. and H. Cao (2001). "Single nucleotide polymorphisms of RXRA encoding retinoid X receptor alpha." J Hum Genet **46**(7): 423-5.

Heilbronn, L. K., K. L. Kind, et al. (2000). "Association of -3826 G variant in uncoupling protein-1 with increased BMI in overweight Australian women." Diabetologia **43**(2): 242-4.

Hemminki K., D. C., Karlsson S., et al. (1997). "Aromatic-DNA adducts in foundry workers in relation to exposure, lifestyle and CYP1A1 and glutathione transferase M1 genotype." Carcinogenesis **18**: 345-50.

Herrmann, S. M., J. G. Wang, et al. (2003). "Uncoupling protein 1 and 3 polymorphisms are associated with waist-to-hip ratio." J Mol Med **81**(5): 327-32.

Hervouet, E., F. M. Vallette, et al. (2009). "Dnmt3/transcription factor interactions as crucial players in targeted DNA methylation." Epigenetics **4**(7): 487-99.

Hirschhorn, J. N., K. Lohmueller, et al. (2002). "A comprehensive review of genetic association studies." Genet Med **4**(2): 45-61.

Holliday, R. (2006). "Epigenetics: a historical overview." Epigenetics **1**(2): 76-80.

Hong, S. J., J. H. Oh, et al. "DNA methylation patterns of ulcer-healing genes associated with the normal gastric mucosa of gastric cancers." J Korean Med Sci **25**(3): 405-17.

Hsueh, W. C., S. A. Cole, et al. (2001). "Interactions between variants in the beta3-adrenergic receptor and peroxisome proliferator-activated receptor-gamma2 genes and obesity." Diabetes Care **24**(4): 672-7.

Huber, K. R. (2010). Viral Infections and Epigenetic Control Mechanisms. Epigenetics and Human Health. A. G. Haslberger, S. Gressler. Weinheim, WILEY-VCH Verlag.

Ioannidis, J. P. (2005). "Why most published research findings are false." PLoS Med **2**(8): e124.

Ioannidis, J. P., E. E. Ntzani, et al. (2001). "Replication validity of genetic association studies." Nat Genet **29**(3): 306-9.

Jablonka, E. and G. Raz (2009). "Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution." Q Rev Biol **84**(2): 131-76.

Janssens, A. C., M. Gwinn, L. A. Bradley, B. A. Oostra, C. M. van Duijn, M. J. Khoury (2008). "A critical appraisal of the scientific basis of commercial genomic profiles used to assess health risks and personalize health interventions." Am J Hum Genet **82**: 593-599.

Jin, J., G. L. Wang, et al. (2009). "GSK3beta-cyclin D3-CUGBP1-eIF2 pathway in aging and in myotonic dystrophy." Cell Cycle **8**(15): 2356-9.

Jost, E., O. N. do, et al. (2009). "Aberrant DNA methylation of the transcription factor C/EBPalpha in acute myelogenous leukemia." Leuk Res **33**(3): 443-9.

Kaati, G. (2010). Case Studies on Epigenetic Inheritance. Epigenetics and Human Health. A. G. Haslberger, S. Gressler. Weinheim, WILEY-VCH Verlag.

Katari, S., N. Turan, et al. (2009). "DNA methylation and gene expression differences in children conceived in vitro or in vivo." Hum Mol Genet **18**(20): 3769-78.

Khalyfa, A., B. Bhushan, et al. "Fatty-acid binding protein 4 gene variants and childhood obesity: potential implications for insulin sensitivity and CRP levels." Lipids Health Dis **9**: 18.

Khoury, M. J., L. Bertram, et al. (2009). "Genome-wide association studies, field synopses, and the development of the knowledge base on genetic variation and human diseases." Am J Epidemiol **170**(3): 269-79.

Khoury, M. J., J. Little, et al. (2004). Human genome epidemiology : a scientific foundation for using genetic information to improve health and prevent disease. New York ; Oxford, Oxford University Press.

Kiec-Wilk, B., I. Wybranska, et al. (2002). "Correlation of the -3826A >G polymorphism in the promoter of the uncoupling protein 1 gene with obesity and metabolic disorders in obese families from southern Poland." J Physiol Pharmacol **53**(3): 477-90.

Kim, J. W., Y. Cheng, et al. (2009). "Genetic and epigenetic inactivation of LPL gene in human prostate cancer." Int J Cancer **124**(3): 734-8.

Kiskinis, E., M. Hallberg, et al. (2007). "RIP140 directs histone and DNA methylation to silence Ucp1 expression in white adipocytes." EMBO J **26**(23): 4831-40.

Kliewer, S. A., K. Umesono, et al. (1992). "Convergence of 9-cis retinoic acid and peroxisome proliferator signalling pathways through heterodimer formation of their receptors." Nature **358**(6389): 771-4.

Koch, C. M., R. M. Andrews, et al. (2007). "The landscape of histone modifications across 1% of the human genome in five human cell lines." Genome Res **17**(6): 691-707.

Komninou, D., J.P. Richie (2010). Epigenetic Dysregulation in Aging and Cancer. Epigenetics and Human Health. A. G. Haslberger, S. Gressler. Weinheim, WILEY-VCH Verlag.

Kopelman, P. G. (2000). "Obesity as a medical problem." Nature **404**(6778): 635-43.

Kumagai, T., T. Akagi, et al. (2009). "Epigenetic regulation and molecular characterization of C/EBPalpha in pancreatic cancer cells." Int J Cancer **124**(4): 827-33.

Kumagai, T., N. Wakimoto, et al. (2007). "Histone deacetylase inhibitor, suberoylanilide hydroxamic acid (Vorinostat, SAHA) profoundly inhibits the growth of human pancreatic cancer cells." Int J Cancer **121**(3): 656-65.

Kurdistani, S. K., S. Tavazoie, et al. (2004). "Mapping global histone acetylation patterns to gene expression." Cell **117**(6): 721-33.

Kurokawa, N., E. H. Young, et al. (2008). "The ADRB3 Trp64Arg variant and BMI: a meta-analysis of 44 833 individuals." Int J Obes (Lond) **32**(8): 1240-9.

Lachner, M., D. O'Carroll, et al. (2001). "Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins." Nature **410**(6824): 116-20.

Lander, E. S., L. M. Linton, et al. (2001). "Initial sequencing and analysis of the human genome." Nature **409**(6822): 860-921.

Langholz, B., N. Rothman, et al. (1999). "Cohort studies for characterizing measured genes." J Natl Cancer Inst Monogr(26): 39-42.

Last, J. M., R. A. Spasoff, et al. (2001). A dictionary of epidemiology. Oxford, Oxford University Press.

Lee, J., P. K. Saha, et al. (2008). "Targeted inactivation of MLL3 histone H3-Lys-4 methyltransferase activity in the mouse reveals vital roles for MLL3 in adipogenesis." Proc Natl Acad Sci U S A **105**(49): 19229-34.

Lee, J. H., K. A. Kim, et al. (2007). "Diallyl disulfide accelerates adipogenesis in 3T3-L1 cells." Int J Mol Med **20**(1): 59-64.

Leipold, H., M. Knoefler, et al. (2006). "Peroxisome proliferator-activated receptor gamma coactivator-1alpha gene variations are not associated with gestational diabetes mellitus." J Soc Gynecol Investig **13**(2): 104-7.

Li, M., G. Pascual, et al. (2000). "Peroxisome proliferator-activated receptor gamma-dependent repression of the inducible nitric oxide synthase gene." Mol Cell Biol **20**(13): 4699-707.

Li, Z., M. H. Gail, et al. (2002). "Statistical properties of Teng and Risch's sibship type tests for detecting an association between disease and a candidate allele." Hum Hered **53**(3): 114-29.

Lillycrop, K. A., E. S. Phillips, et al. (2005). "Dietary protein restriction of pregnant rats induces and folic acid supplementation prevents epigenetic modification of hepatic gene expression in the offspring." J Nutr **135**(6): 1382-6.

Ling, C., S. Del Guerra, et al. (2008). "Epigenetic regulation of PPARGC1A in human type 2 diabetic islets and effect on insulin secretion." Diabetologia **51**(4): 615-22.

Lu, G. D., C. H. Leung, et al. "C/EBPalpha is up-regulated in a subset of hepatocellular carcinomas and plays a role in cell growth and proliferation." Gastroenterology **139**(2): 632-43, 643 e1-4.

Lwow, F., K. Dunajska, et al. (2007). "Post-exercise oxidative stress and obesity in postmenopausal women: the role of beta3-adrenergic receptor polymorphism." Gynecol Endocrinol **23**(10): 597-603.

Malczewska-Malec, M., I. Wybranska, et al. (2004). "Analysis of candidate genes in Polish families with obesity." Clin Chem Lab Med **42**(5): 487-93.

Manolio, T. A. "Genomewide association studies and assessment of the risk of disease." N Engl J Med **363**(2): 166-76.

Margueron, R., P. Trojer, et al. (2005). "The key to development: interpreting the histone code?" Curr Opin Genet Dev **15**(2): 163-76.

Masud, S. and S. Ye (2003). "Effect of the peroxisome proliferator activated receptor-gamma gene Pro12Ala variant on body mass index: a meta-analysis." J Med Genet **40**(10): 773-80.

McClelland, S., R. Shrivastava, et al. (2009). "Regulation of Translational Efficiency by Disparate 5' UTRs of PPARgamma Splice Variants." PPAR Res **2009**: 193413.

McMichael, A. J. (1994). "Invited commentary--"molecular epidemiology": new pathway or new travelling companion?" Am J Epidemiol **140**(1): 1-11.

Migliore, L. a. F. C. (2010). The Impact of Genetic and Environmental Factors in Neurodegeration: Emerging Role of Epigenetics. Epigenetics and Human Health. A. G. Haslberger, S. Gressler. Weinheim, WILEY-VCH Verlag.

Miller, R. L. a. J. H. (2010). Epigenetic Mechanisms in Asthma. Epigenetics and Human Health. A. G. Haslberger, S. Gressler. Weinheim, WILEY-VCH Verlag.

Misík, M., A. Nersesyan, W. Parzefall, S. Knasmüller (2010). Carcinogens in Foods: Occurrence, Modes of Action and Modulation of Human Risks by Genetic Factors and Dietary Constituents. Epigenetics and Human Health. A. G. Haslberger, S. Gressler. Weinheim, WILEY-VCH Verlag.

Morabia, A. (2004). A history of epidemiologic methods and concepts. Basel, Birkhäuser Verlag.

Nelson, M. R., S. L. Kardia, et al. (2001). "A combinatorial partitioning method to identify multilocus genotypic partitions that predict quantitative trait variation." Genome Res **11**(3): 458-70.

Nieters, A., N. Becker, et al. (2002). "Polymorphisms in candidate obesity genes and their interaction with dietary intake of n-6 polyunsaturated fatty acids affect obesity risk in a sub-sample of the EPIC-Heidelberg cohort." Eur J Nutr **41**(5): 210-21.

Noer, A., A. C. Boquest, et al. (2007). "Dynamics of adipogenic promoter DNA methylation during clonal culture of human adipose stem cells to senescence." BMC Cell Biol **8**: 18.

Noer, A., L. C. Lindeman, et al. (2009). "Histone H3 modifications associated with differentiation and long-term culture of mesenchymal adipose stem cells." Stem Cells Dev **18**(5): 725-36.

Noer, A., A. L. Sorensen, et al. (2006). "Stable CpG hypomethylation of adipogenic promoters in freshly isolated, cultured, and differentiated mesenchymal stem cells from adipose tissue." Mol Biol Cell **17**(8): 3543-56.

Oberkofler, H., A. Beer, D. Breban, E. Hell, F. Krempler, W. Patsch (1997). "Human obese gene expression: alternative splicing of mRNA and relation to adipose tissue localization." Obes Surg **7**(5): 390-6.

Ochoa, M. C., A. Marti, et al. (2004). "Gene-gene interaction between PPAR gamma 2 and ADR beta 3 increases obesity risk in children and adolescents." Int J Obes Relat Metab Disord **28 Suppl 3**: S37-41.

Okamura, M., H. Kudo, et al. (2009). "COUP-TFII acts downstream of Wnt/beta-catenin signal to silence PPARgamma gene expression and repress adipogenesis." Proc Natl Acad Sci U S A **106**(14): 5819-24.

Olofsson, L. E., M. Orho-Melander, et al. (2008). "CCAAT/enhancer binding protein alpha (C/EBPalpha) in adipose tissue regulates genes in lipid and glucose metabolism and a genetic variation in C/EBPalpha is associated with serum levels of triglycerides." J Clin Endocrinol Metab **93**(12): 4880-6.

Pearson, T. A. and T. A. Manolio (2008). "How to interpret a genome-wide association study." JAMA **299**(11): 1335-44.

Perera, F. P. and I. B. Weinstein (1982). "Molecular epidemiology and carcinogen-DNA adduct detection: new approaches to studies of human cancer causation." J Chronic Dis **35**(7): 581-600.

Perera, F. P. and I. B. Weinstein (2000). "Molecular epidemiology: recent advances and future directions." Carcinogenesis **21**(3): 517-24.

Pierola, J., A. Barcelo, et al. (2007). "beta3-Adrenergic receptor Trp64Arg polymorphism and increased body mass index in sleep apnoea." Eur Respir J **30**(4): 743-7.

Pigeot, A. W. a. I. (2007). Handbook of epidemiology, Springer-Verlag Berlin Heidelberg.

Pihlajamaki, J., M. Kinnunen, et al. (2005). "Haplotypes of PPARGC1A are associated with glucose tolerance, body mass index and insulin sensitivity in offspring of patients with type 2 diabetes." Diabetologia **48**(7): 1331-4.

Porta, M. S., S. Greenland, et al. (2008). A dictionary of epidemiology. Oxford, Oxford University Press.

Ramis, J. M., J. L. Gonzalez-Sanchez, et al. (2004). "The Arg64 allele of the beta 3-adrenoceptor gene but not the -3826G allele of the uncoupling

protein 1 gene is associated with increased leptin levels in the Spanish population." Metabolism **53**(11): 1411-6.
Rankinen, T., A. Zuberi, et al. (2006). "The human obesity gene map: the 2005 update." Obesity (Silver Spring) **14**(4): 529-644.
Rosen, E. D. and B. M. Spiegelman (2006). "Adipocytes as regulators of energy balance and glucose homeostasis." Nature **444**(7121): 847-53.
Rosenfeld, J. A., Z. Wang, et al. (2009). "Determination of enriched histone modifications in non-genic portions of the human genome." BMC Genomics **10**: 143.
Rothman, K. J., S. Greenland and T. L. Lash (2008). Modern Epidemiology. Philadelphia, Lippincott Williams & Wilkins.
Rothman, N., S. Wacholder, et al. (2001). "The use of common genetic polymorphisms to enhance the epidemiologic study of environmental carcinogens." Biochim Biophys Acta **1471**(2): C1-10.
Rothwell, N. J., M. J. Stock, et al. (1983). "Decerebration activates thermogenesis in the rat." J Physiol **342**: 15-22.
Rundle, A. G., P. Vineis, et al. (2005). "Design options for molecular epidemiology research within cohort studies." Cancer Epidemiol Biomarkers Prev **14**(8): 1899-907.
Rust, P. (2010). Health Determinants Throughout the Life Cycle. Epigenetics in Human Health. A. G. Haslberger, S. Gressler. Weinheim, WILEY-VCH Verlag.
Salma, N., H. Xiao, et al. (2004). "Temporal recruitment of transcription factors and SWI/SNF chromatin-remodeling enzymes during adipogenic induction of the peroxisome proliferator-activated receptor gamma nuclear hormone receptor." Mol Cell Biol **24**(11): 4651-63.
Sato, K., H. Fukata, et al. (2006). "Neonatal exposure to diethylstilbestrol alters the expression of DNA methyltransferases and methylation of genomic DNA in the epididymis of mice." Endocr J **53**(3): 331-7.
Schadt, E. E., A. Sachs, et al. (2005). "Embracing complexity, inching closer to reality." Sci STKE **2005**(295): pe40.
Schaffler, A., K. D. Palitzsch, et al. (1999). "Frequency and significance of the A-->G (-3826) polymorphism in the promoter of the gene for uncoupling protein-1 with regard to metabolic parameters and adipocyte transcription factor binding in a large population-based Caucasian cohort." Eur J Clin Invest **29**(9): 770-9.
Schofield, P. N. (1998). "Impact of genomic imprinting on genomic instability and radiation-induced mutation." Int J Radiat Biol **74**(6): 705-10.
Sertic, J., L. Juricic, et al. (2009). "Variants of ESR1, APOE, LPL and IL-6 loci in young healthy subjects: association with lipid status and obesity." BMC Res Notes **2**: 203.
Shin, J., B. Li, et al. (2009). "Comparative analysis of fatty acid-binding protein 4 promoters: conservation of peroxisome proliferator-activated receptor binding sites." J Anim Sci **87**(12): 3923-34.
Shore, A., A. Karamitri, et al. "Role of Ucp1 enhancer methylation and chromatin remodelling in the control of Ucp1 expression in murine adipose tissue." Diabetologia **53**(6): 1164-73.

Silva, J. E. (2006). "Thermogenic mechanisms and their hormonal regulation." Physiol Rev **86**(2): 435-64.

Sivenius, K., R. Valve, et al. (2000). "Synergistic effect of polymorphisms in uncoupling protein 1 and beta3-adrenergic receptor genes on long-term body weight change in Finnish type 2 diabetic and non-diabetic control subjects." Int J Obes Relat Metab Disord **24**(4): 514-9.

Sorensen, A. L., S. Timoskainen, et al. (2009). "Lineage-Specific Promoter DNA Methylation Patterns Segregate Adult Progenitor Cell Types." Stem Cells Dev.

Spencer, C. C., Z. Su, et al. (2009). "Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip." PLoS Genet **5**(5): e1000477.

Sramkova, D., S. Krejbichova, et al. (2007). "The UCP1 gene polymorphism A-3826G in relation to DM2 and body composition in Czech population." Exp Clin Endocrinol Diabetes **115**(5): 303-7.

Stefan, N., C. Thamer, et al. (2007). "Genetic variations in PPARD and PPARGC1A determine mitochondrial function and change in aerobic physical fitness and insulin sensitivity during lifestyle intervention." J Clin Endocrinol Metab **92**(5): 1827-33.

Stefanska, B. M., K. Fabianowska-Majewska (2010). Effects of Dietary Natural Compounds on DNA Methylation Related to Cancer Chemoprevention and Anticancer Epigenetic Therapy. Epigenetics and Human Health. A. G. Haslberger, S. Gressler. Weinheim, WILEY-VCH Verlag.

Stoger, R. (2008). "The thrifty epigenotype: an acquired and heritable predisposition for obesity and diabetes?" Bioessays **30**(2): 156-66.

Stunkard, A. J., T. T. Foch, Z. Hrubec (1986). "A twin study of human obesity." JAMA **256**: 51-4.

Szyf, M., I. Weaver, et al. (2007). "Maternal care, the epigenome and phenotypic differences in behavior." Reprod Toxicol **24**(1): 9-19.

Tada, Y., R. M. Brena, et al. (2006). "Epigenetic modulation of tumor suppressor CCAAT/enhancer binding protein alpha activity in lung cancer." J Natl Cancer Inst **98**(6): 396-406.

Takada, I., M. Suzawa, et al. (2007). "Suppression of PPAR transactivation switches cell fate of bone marrow stem cells from adipocytes into osteoblasts." Ann N Y Acad Sci **1116**: 182-95.

Takai, N., N. Kawamata, et al. (2005). "Discovery of epigenetically masked tumor suppressor genes in endometrial cancer." Mol Cancer Res **3**(5): 261-9.

Tateishi, K., Y. Okada, et al. (2009). "Role of Jhdm2a in regulating metabolic gene expression and obesity resistance." Nature **458**(7239): 757-61.

Taubes, G. (1995). "Epidemiology faces its limits." Science **269**(5221): 164-9.

Teng, J. and N. Risch (1999). "The relative power of family-based and case-control designs for linkage disequilibrium studies of complex human diseases. II. Individual genotyping." Genome Res **9**(3): 234-41.

Teutsch, S. M., L. A. Bradley, et al. (2009). "The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: methods of the EGAPP Working Group." Genet Med **11**(1): 3-14.

Thaler, R., E. Aumüller, C. Berner, A.G. Haslberger (2010). Interaction of Hereditary and Epigenetic Mechanisms in the Regulation of Gene Expression. Epigenetics and Human Health. A. G. Haslberger, S. Gressler. Weinheim, Wiley-VCH Verlag.

Tikoo, K., P. Kumar, et al. (2009). "Rosiglitazone synergizes anticancer activity of cisplatin and reduces its nephrotoxicity in 7, 12-dimethyl benz{a}anthracene (DMBA) induced breast cancer rats." BMC Cancer **9**: 107.

Tonjes, A., M. Scholz, et al. (2006). "Association of Pro12Ala polymorphism in peroxisome proliferator-activated receptor gamma with Pre-diabetic phenotypes: meta-analysis of 57 studies on nondiabetic individuals." Diabetes Care **29**(11): 2489-97.

Trayhurn, P., P. M. Jones, et al. (1982). "Effects of overfeeding on energy balance and brown fat thermogenesis in obese (ob/ob) mice." Nature **295**(5847): 323-5.

Ukkola, O., A. Tremblay, et al. (2001). "Genetic variation at the uncoupling protein 1, 2 and 3 loci and the response to long-term overfeeding." Eur J Clin Nutr **55**(11): 1008-15.

Urhammer, S. A., M. Fridberg, et al. (1997). "Studies of genetic variability of the uncoupling protein 1 gene in Caucasian subjects with juvenile-onset obesity." J Clin Endocrinol Metab **82**(12): 4069-74.

Urhammer, S. A., T. Hansen, et al. (2000). "Studies of the synergistic effect of the Trp/Arg64 polymorphism of the beta3-adrenergic receptor gene and the -3826 A-->G variant of the uncoupling protein-1 gene on features of obesity and insulin resistance in a population-based sample of 379 young Danish subjects." J Clin Endocrinol Metab **85**(9): 3151-4.

Valve, R., S. Heikkinen, et al. (1998). "Synergistic effect of polymorphisms in uncoupling protein 1 and beta3-adrenergic receptor genes on basal metabolic rate in obese Finns." Diabetologia **41**(3): 357-61.

Vimaleswaran, K. S., J. Luan, et al. (2008). "The Gly482Ser genotype at the PPARGC1A gene and elevated blood pressure: a meta-analysis involving 13,949 individuals." J Appl Physiol **105**(4): 1352-8.

Vineis, P. and F. Perera (2007). "Molecular epidemiology and biomarkers in etiologic cancer research: the new in light of the old." Cancer Epidemiol Biomarkers Prev **16**(10): 1954-65.

Vogel, F. a. A. G. M. (1996). Human Genetics. Problems and Approaches, Springer-Verlag Berlin Heidelberg New York.

Vohl, M. C., A. Houde, et al. (2005). "Effects of the peroxisome proliferator-activated receptor-gamma co-activator-1 Gly482Ser variant on features of the metabolic syndrome." Mol Genet Metab **86**(1-2): 300-6.

Volate, S. R., S. J. Muga, et al. (2009). "Epigenetic modulation of the retinoid X receptor alpha by green tea in the azoxymethane-Apc Min/+ mouse model of intestinal cancer." Mol Carcinog **48**(10): 920-33.

Wacholder, S., N. Rothman, et al. (2002). "Counterpoint: bias from population stratification is not a major threat to the validity of conclusions from epidemiological studies of common polymorphisms and cancer." Cancer Epidemiol Biomarkers Prev **11**(6): 513-20.

Wakabayashi, K., M. Okamura, et al. (2009). "The peroxisome proliferator-activated receptor gamma/retinoid X receptor alpha heterodimer targets the histone modification enzyme PR-Set7/Setd8 gene and regulates adipogenesis through a positive feedback loop." Mol Cell Biol **29**(13): 3544-55.

Wang, H. and R. H. Eckel (2009). "Lipoprotein lipase: from gene to obesity." Am J Physiol Endocrinol Metab **297**(2): E271-88.

Winslow, C. E. (1920). "The Untilled Fields of Public Health." Science **51**(1306): 23-33.

Woolf, S. H., C. G. DiGuiseppi, et al. (1996). "Developing evidence-based clinical practice guidelines: lessons learned by the US Preventive Services Task Force." Annu Rev Public Health **17**: 511-38.

Yang, W., T. Kelly, et al. (2007). "Genetic epidemiology of obesity." Epidemiol Rev **29**: 49-61.

Zafarmand, M. H., Y. T. van der Schouw, et al. (2008). "T64A polymorphism in beta3-adrenergic receptor gene (ADRB3) and coronary heart disease: a case-cohort study and meta-analysis." J Intern Med **263**(1): 79-89.

Zawodniak-Szalapska, M., R. Stawerska, et al. (2008). "Association of Trp64Arg polymorphism of beta3-adrenergic receptor with insulin resistance in Polish children with obesity." J Pediatr Endocrinol Metab **21**(2): 147-54.

Zhang, Y., S. De, et al. "Systematic analysis, comparison, and integration of disease based human genetic association data and mouse genetic phenotypic information." BMC Med Genomics **3**: 1.

Zhu, N. L., J. Wang, et al. "Necdin-Wnt pathway causes epigenetic PPAR{gamma} repression in hepatic stellate cells." J Biol Chem.

ATTACHED MANUSCRIPT: *SYNERGISM OF GENETIC AND EPIGENETIC VARIATION IN THERMOGENESIS*

ABSTRACT

Background: Complex traits are characterized by several different pathways where multiple genetic and environmental factors interact. Epidemiologic research has been facing severe limitations in unraveling complex conditions when focusing either on genetic or environmental factors. As a result of many genome wide association studies a missing heritability has been discussed and epigenetic factors are considered to explain gene-environment interactions.

Objective: The example of adaptive thermogenesis in obesity was analyzed to demonstrate synergisms of genetic and epigenetic regulation in obesity.

Method: A systematic literature search was done for genetic and epigenetic variation of genes involved in adaptive thermogenesis: gens included beta-3 adrenergic receptor (*ADRB3*), uncoupling protein 1 (*UCP1*), lipoprotein lipase (*LPL*) and fatty acid binding protein 4 (*FABP4*) and the transcription factors peroxisome proliferated activator receptor gamma (*PPARγ*), peroxisome proliferated activator receptor gamma coactivator 1 alpha (*PGC1α*), CCAAT/enhancer-binding protein alpha (*C/EBPα*), and retinoid acid X receptor alpha (*RXRα*).

Results: SNPs as well as epigenetic mechanisms seem to regulate adaptive thermogenesis. Genetic association studies had serious problems of inconsistencies with results. Studies of epigenetic evidence were discovery orientated rather than hypothesis testing but demonstrate existing epigenetic modulation in obesity and adaptive thermogenesis.

Discussion: Results from this analysis suggest that hereditary genetic- and environment responsive regulation is involved in adaptive thermogenesis. The integration of epigenetic markers in epidemiologic research could thus help to unravel multi-gene-environment interactions. They also should have the potential to increase the predictability of genetic testing for complex diseases with public health importance.

INTRODUCTION

Epidemiologic research of complex interactions

In developed countries complex diseases make up the majority of morbidity and mortality (Zhang, De et al.). They are characterized by the joint effects of multiple genes and multiple environmental factors in different pathways and networks. Genetic variations identified so far, only explain a small proportion of the heritability of complex conditions. Furthermore, the replication of results has turned out to be more challenging than expected (Hirschhorn, Lohmueller et al. 2002) (Ioannidis 2005). Genetic association studies often carry only limited information about environmental exposures or other non-genetic risk factors. In the presence of such unknown etiologic factors it is difficult to unravel interactive partners and joint effects (Pearson and Manolio 2008). In a complex disease, the knowledge of all genetic and environmental factors and their complex interactions might be required to explain the full etiology. But even if the exact combination of factors that contributed to a disease phenotype has been identified in one person, it is unlikely that this etiologic profile fits another person. Such heterogeneous etiologies make it even more difficult to identify interactions between the environment and genes or environmental modifications of genotype associations.

Epigenetics

Epigenetics studies heritable changes in gene expression that are due to altered chromatin without changing the DNA sequence and it offers promising tools to unravel inconsistencies in human genome epidemiology (Bjornsson, Fallin et al. 2004) (Dolinoy and Jirtle 2008) (Figueroa, Reimers et al. 2008) (Foley, Craig et al. 2009). Epigenetic modifications vary with several environmental influences and can be stably passed over numerous cycles of cell division (Jablonka and Raz 2009). Some alterations can even be inherited from one generation to the next. They have been suggested to form an alternative, “soft inheritance” system (Jablonka and Raz 2009). The ‘‘fetal basis of adult disease hypothesis’’ postulates that environmental influences during prenatal and early postnatal development model cellular plasticity, thereby

altering susceptibility to adult diseases like cardiovascular disease, type 2 diabetes, and (the metabolic condition) obesity (Ob) (Dolinoy and Jirtle 2008). The most prominent example is that of the agouti mouse model where maternal dietary genistein supplementation during gestation results in hypermethylation of an upstream promoter region of the Agouti gene in the offspring. This epigenetic marker persists into adulthood and protects offspring from obesity (Dolinoy and Jirtle 2008).

The Epigenome is the overall epigenetic state of a cell including all mechanisms and marks such as non-coding RNAs, the histone code, and DNA methylation. The latter is the covalent addition of a methyl group to a cytosine that precedes a guanosine in the DNA sequence (the CpG dinucleotide) and is a dynamic marker of the transcriptionally repressed gene (Bird 2002). Histone modifications result in different transcriptional outcomes depending on type, amount, site of modification and its reversibility (reviewed in (He and Lehming 2003).

Genetics & Epigenetics in Obesity

In humans, Ob is defined as body mass index (BMI) >30 kg/m^2 and works as a major risk factor for common chronic diseases such as diabetes, diseases of the cardiovascular and musculoskeletal system, and some cancers (Haslam and James 2005). It is characterized by an excessive accumulation of lipids in body cells, markedly in adipocytes and mostly results form an imbalance between energy intake and expenditure (obesogenic environment). The regulation of body fat content results from an integrative network influenced by the central nervous system, the gastrointestinal tract, the adipose tissue itself and skeletal muscles in responses to multiple nutrient, drug, toxin, sensory and hormonal inputs. Further it is influenced by circadian rhythm and physical, gut and social environments. For instance, pathways in energy homeostasis, adipogenesis and hormonal signaling, have been suggested to contribute to the development of Ob. In 2006 the last update of the Obesity Gene Map Database summarized positive associations for DNA sequence variation with Ob phenotypes in 127 candidate genes. Among them (those showing replications in

at least 10 studies) have been: PPARG (30 studies), ADRB3 (29), ADRB2 (20), LEPR (16), GNB3 (14), UCP3 (12), ADIPOQ (11), LEP (11), UCP2 (11), HTR2C (10), NR3C1 (10), and UCP1 (10) (Rankinen, Zuberi et al. 2006). Since then, many more loci have been identified, but they still account for little of the heritability, often less than 1 % (Speliotes, Willer et al 2010) (Heid, Jackson 2010).

Ob is expected to develop from a heterogeneous background. That means, with respect to all cases in an obesogenic environment, in any individual with the obese phenotype, only one or a subset of these pathways might be involved. So in other words, in the same obesogenic environment individuals with different gene combinations might be equally well adapted (Manolio). Additionally, associations between polymorphisms and phenotypes tend to have modest effects with effects less than OR = 2. Such variations rather tend to induce a shift in between the working capacities of the networks influencing the development of Ob. On the other hand, epigenetic programming adjusts the phenotype without integrating the genetic determination of the trait. Extreme environmental conditions, postnatal and during early life, lead to an establishment of the trait further away from the genetically determined (Stoger 2008). During the entire life diet, lifestyle factors, maternal behavior, air pollution, toxins, and other environmental factors may modify the epigenome and contribute to the development of Ob (Haslberger 2010). This developmental plasticity allows the development of a range of phenotypes from a single genotype.

The Network of adaptive thermogenesis in BAT

Energy balance is crucial for constant body weight, strongly depends on the work of the skeletal muscle, and is a main element in the development of Ob. Also, adaptive thermogenesis in brown adipose tissue (BAT) dissipates energy in the respiratory chain due to the oxidation of fatty acids (Silva 2006). BAT's thermogenic activity depends on the degree of stimulation by the sympathetic nervous system and the amount of uncoupling protein 1 (UCP1) in the inner mitochondrial membrane which uncouples respiration from ATP synthesis to

produce heat (Silva 2006). Less activity of UCP1 results in a more efficient energy coupling and lower metabolic rates, both predictors of weight gain. (Cypess and Kahn) UCP1 deficient mice are cold intolerant, but surprisingly not ob. When fed with a high-fat diet, they gain less weight than wild type (WT) controls. This Ob resistance is ambient temperature dependent because when reared at room temperature deficient mice gain weight faster than WT (Silva 2006). Downregulation of UCP1 expression in adult humans might result in its less prominent role of thermogenesis in BAT (Silva 2006). In obese persons a reduced expression of UCP1 mRNA down to 50 % was demonstrated. During the last few years, studies have revealed a sort of a transdifferentiation from adult white to brown adipocytes with similar morphology and that might be due to epigenetic regulation (Barbatelli, Murano et al.).

The UCP1 gene holds (among other regulatory elements) a response element for the orphan nuclear receptor family of peroxisome proliferated activator receptors (PPARs). PPARγ is a member of these transcription factors and acts as sensor for a variety of lipids and lipid like compounds (e.g. ω-3 fatty acids). Different promoters in the gene and alternative splicing result in several isoforms. For instance, PPARγ2 is expressed exclusively in adipose tissue while PPARγ1 can also be found in liver and other tissues (McClelland, Shrivastava et al. 2009). Ligand binding to the receptor initiates the gene transcription process through conformational changes, a heterodimerization with the retinoid acid X receptor alpha (RXRα) and the assembly with coactivators to a multiprotein complex (Kliewer, Umesono et al. 1992). PGC1α acts as a ligand-independent coactivator for PPARs and is expressed preferentially in BAT. Histone acetylase (HAT) activity of coactivators leads to a modification of the chromatin structure of the proximate DNA which allows the large complex to bind to a respond element and to initiate transcription. In the absence of ligands, a PPAR interacts with corepressor proteins and builds a complex with histone deacetylases (HDACs) which results in a locally more compact chromatin packaging (Li, Pascual et al. 2000). The binding of the PPARγ2/RXRα/PGC1α complex to the respond element of UCP1, followed by its transcriptional activation, have been demonstrated to be essential in brown

adipogenesis (Barak, Nelson et al. 1999). Other PPARγ target genes are, for example, the extracellular lipoprotein lipase (*LPL*) which mediates the uptake of fatty acids into the cell (Wang and Eckel 2009), and the fatty acid binding protein 4 (*FABP4*) which is expressed exclusively in adipocytes and macrophages. Although the role of this fatty acid carrier protein is not yet well understood, its function could be correlated to lipid metabolism (Shin, Li et al. 2009).

Another essential transcription factor in adipogenesis is CCAAT/enhancer-binding protein alpha (CEBPα) whose target genes are *LPL* and *ADRB3*, among others (Rosen and Spiegelman 2006). For rapid adjustment the ADRB3 in BAT is targeted by catecholamines in response to cold (Silva 2006). The activation of the receptor signaling results in an increased expression of the *PGC1α* gene (Cao, Daniel et al. 2004).

The capacity of this network strongly depends on environmental influences. Fatty acids are the only substrate for oxidative phosphorylation in BAT and act as ligands of the nuclear receptor PPARγ. RXRα is a high affinity nuclear receptor for 9-cis retinoic acid, a metabolite of retinoic acid. Increased uncoupling of the respiratory chain from ATP production depends on the activation of the ADRB3 in response to cold (Silva 2006).

METHOD

In a systematic literature review information was screened for in most frequently observed SNPs in genes involved in adaptive thermogenesis (table 1) in Caucasians. The literature search revealed observational studies associating the polymorphisms with anthropometric measurements and related ob phenotypes. We included studies reporting observations on Caucasians and mixed ethnicities. If no ethnicity was reported, a study was included if it was conducted in central Europe and it could be assumed to include Caucasians. Evidence for epigenetic regulation of the same genes was searched in the same way by using the terms “gene” and/or “gene name” and “epigenetic” and/or “methylation” in title and/or abstract. Additionally, reference lists from articles were searched for further suitable studies. English publications that

gave sufficient information on DNA methylation and histone modifications in the genes were summarized. No other variations (e.g. frame shift mutations or copy number variations) than selected SNPs and no other data on posttranscriptional modification than epigenetic modifications have been included.

Table1 genes under observation

Gene name	Gene	Gene map locus	SNP
Beta-3 adrenergic receptor	ADRB3	8p12-p11	rs4994
CCAAT/enhancer-bindingproteinalpha	CEBPα	19q13	rs12691 rs41490344 rs41504144 rs41344151 rs16967952 rs34508287 rs41428545 rs34529039 rs41367646 rs1049969 rs707656
Fatty acid binding protein 4	FATBP4	8q21	rs1054135
Lipoprotein lipase	LPL	8p22	rs285 rs320 rs328
Peroxisome proliferated activator receptor gamma coactivator 1 alpha	PGC1α	4p15	rs8192678
Peroxisome proliferated activator receptor gamma	PPARγ	3p25	rs1801282
Retinoid acid X receptor alpha	RXRα	9q34	-
Uncoupling protein 1	UCP1	4q31	A-3826G

RESULTS

The Pro12Ala (rs1801282) is one of the most prominent SNPs in the *PPARγ* gene. It has been observed for associations with phenotypes of the metabolic syndrome for many years. The literature research revealed a meta-analysis that included over 19,600 subjects and positively associated the Ala variant with increased BMI (table R1).

In adipocytes the correct epigenetic modification of the promoter region in the *PPARγ* gene seems to be essential during cell differentiation and for the maintenance of the adult state of the cell. In tables R2 and R3, studies observing epigenetic regulations of the gene are summarized. In cells from a human umbilical cord where PPARγ2 is not expressed, its promoter region is hypermethylated. In human gastric mucosa cells of ulcer and cancer patients, hypomethylation of the promoter region was observed which concurs with findings from cancer research where cancerous cells become globally hypomethylated and tumorsuppressor genes hypermethylated (Esteller 2008). In human adipose stem cells the promoter region is hypomethylated and gradually demethylated upon induction of differentiation in 3T3-L1 cultures. Furthermore, in lean mice, promoter methylation was increased when compared to that of ob mice. The results from studies observing histone modifications give a similar picture: during adipogenesis enzymatically coordinated deacetylation and methylation events at sites on H3 and H4 facilitate the transcription of the gene and have been replicated by several groups (R3).

The search for RXRα, the heterodimerization partner of PPARγ, has not revealed sufficient literature to draw a clear conclusion for its susceptibility to contribute to the development of Ob. When screening the gene for DNA variants, Hegele et al identified 3 common variations in Caucasians: -25 A>G, 69 G>A, and 1470 C>T (Hegele and Cao 2001). Though, this receptor plays a key role in development and metabolism, no study associated any genetic variation with Ob traits. The search for epigenetic evidence in the gene revealed only one study examining the effects of green tea on demethylation of the promoter region in cancerous mice (Volate, Muga et al. 2009).

The C/EBPα gene holds several polymorphisms that have been associated with different types of cancers. Their contribution to Ob have been under study by only two groups (table R4). Evidence for epigenetic regulation of the gene exists as well. Heterogeneous responses to inhibitors of histone and DNA modifying enzymes demonstrate that both mechanisms are involved in the transcriptional regulation of the gene (tables R5 and R6). C/EBPα is incapable of a direct interaction with DNAmethyltransferases (Hervouet, Vallette et al. 2009) but several studies demonstrated that a C/EBPα-HDAC complex is able to silence several protein coding genes due to aging (Jin, Wang et al. 2009). One of C/EBPα's target genes is *ADRB3.* Studies observing the SNP Trp64Arg (rs4994) and its association with Ob are summarized in tables R7 and R8. The last meta-analysis, published in 2008, summarized 53 studies in Caucasians and had an overall size of about 23,000 individuals. No positive association/s could be detected. Since then, further studies have been published associating the SNP with Ob (table R8). Seven studies with an overall all size of 8,323 subjects did not detect an effect, while 4 studies (1,267 subjects) did find positive associations. Searching for epigenetic evidence did not reveal any studies of the ADRB3 gene itself. However, the stimulation of the receptor induces the expression and recruitment of a demethylase to the promoter of the UCP1 gene and facilitates the ability of PPARy, RXRα and their co-activators to bind to the response element on the DNA.

Groups observing genetic variation in the *UCP1* gene merely focused on the A-3826G substitution (table R9). The gender specific effect observed by two groups (total 701 women) could not be replicated by two other groups (total 77 women). As well, in a study including 722 men no effect had been detected. Furthermore, eleven studies with an overall size of 4,466 subjects could not positively associate the polymorphism with Ob, even though 3 studies (all together 557 subjects) did.

Observations of DNA methylation and histone modifications (table 10) of the promoter region in several different tissues revealed tissue specificity. This reflects previous observations of the epigenome (Esteller 2008). In 3T3-L1

cultures gene silencing by DNA methylation and repressive histone modifications were demonstrated by different groups.

The coactivator PGC1α holds the SNP rs8192678 that changes the amino acid sequence at position 482 from glycin to serin. It is the most examined polymorphism in the gene in relation to Ob. The studies revealed by literature are summarized in table R11. No effect of the variant was detected by studies (all together 16,000 individuals) that included patients with attributes of the metabolic syndrome (one meta-analysis included). When associating healthy obese carriers, no association was observed in over 800 persons, while positive associations were found in over 2,200 persons. Also, gender specific effects had been detected. In 467 women the variant had been positively associated with a higher BMI and related phenotypes. On the other hand, higher weight gain could be detected in male but not in female carriers of the polymorphism. An observation that focused on histone modifications of the *PGC1α* was not detected. DNA methylation of the promoter region in skeletal muscle and pancreatic cells from diabetics could be demonstrated to be increased compared to cells from healthy controls. A high fat diet and free fatty acids increased methylation even more. In table R12 findings from the literature search are listed.

The *LPL* gene holds several polymorphisms. Two different studies observed gender specific effects, each on a different SNP (rs320 and rs328) (summarized in table R13). One study observed the rs1054135 SNP in the gene of FABP4 in 309 subjects of mixed ethnicity and found a positive association with BMI (table R15). Epigenetic evidence for the regulation of *LPL* and *FABP4* are summarized in tables R14 and R16. In adipose stem cells DNA methylation in both promoter regions could be demonstrated. Further, histone modifications of each gene change during adipogenesis. Treatment of 3T3-L1 cultures with diallyl sulfide (a compound in garlic) results in decreased activity of HDAC correlating with gene expression of both, *LPL* and *FABP4*.

Observed joint effects of SNPs

Genetic association studies observing joint effects of polymorphisms are listed in tables R17 and R18. No matter of ethnicity, carrier of both, the Pro12Ala variant in the *PPARγ* and the Trp64Arg variant in the ADRB3 gene had a higher BMI. Gender specific effects were detected for the polymorphism rs320 in the *LPL* gene (independent of the simultaneous occurrence of the *ADRB3* SNP (rs4994)). Most, but not all, reports suggest a synergistic effect of the variant in the UCP1gene and the *ADRB3* variant.

DISCUSSION

Discussion network: SNPs vs. epigenetic evidence

In adipogenesis, the enzymatic coordination of changes in histone modification and DNA demethylation in the promoter regions of *PPARγ* and *C/EBPα* are essential. If genetic variations have effects on their activities as transcription factors will have to be warranted in further studies. Previous studies have not been properly sized to detect little effects. Although, a meta-analysis is a powerful tool to confirm evidence, 19,600 subjects might still be not enough to draw a clear conclusion. Similar uncertainties remain with theTrp64Arg (rs4994) in *ADRB3* and the SNP (rs8192678) in the *PGC1α*. Meta-analysis reported big in-between study heterogeneity or did not distinguish between ethnic subgroups. The literature search did not reveal any study observing epigenetic modification in the *ADRB3.* We believe that future epigenetic research will have to focus more on the *ADRB3* gene as it is an interesting candidate in the field. Further observations on genetic variations are also needed for *FABP4*, *LPL,* and *RXRα*. The two PPARγ target genes, *LPL* and *FABP4,* are important factors in pathways ensuring the steady substrate availability for mitochondria. Polymorphisms in the *LPL* gene probably influence the enzyme's capacity to facilitate free fatty acids. Epigenetic regulation of *PGC1α, FABP4*, *LPL,* and *RXRα* could be demonstrated several times. Stimulation of the ADRB3 changes the epigenetic state of the *UCP1* gene and facilitates it for transcription factors to induce expression. The A-3826G substitution of the *UCP1* gene had also been associated with reduced UCP1 mRNA expression (Esterbauer, Oberkofler

et al. 1998). Although, it is a promising candidate for an Ob susceptibility variant, findings do not show clear evidence for associations with Ob phenotypes. These inconsistencies will have to be elucidated in further research.

Current techniques allow the simultaneous testing of many SNPs. A synergistic effect of the UCP1 and the ADRB3 polymorphisms could occur if the functions or expressions of the genes were reduced. But by now, no study had the appropriate sample size for reliable findings.

Limitations of summarized epidemiology

Much more effort and money had been put into association studies of genetic variation than in observations of epigenetic variation. Though, evidence for epigenetic mechanisms potentially shaping the network capacity is strong. The majority of observed genetic polymorphisms have not shown qualitative alterations in the network yet. This is a well-recognized problem in the community of human genome epidemiologists. Often, the excitement about a strong first finding is followed by inconsistencies (Ioannidis 2005). The general lack of reproducibility has been attributed to some serious problems like the heterogeneity of disease etiology, limitations in study designs, massive amount of errors, bias and confounder decreasing statistical power, and many problems in data analysis (Khoury 2004) (Pearson and Manolio 2008) (Manolio). A big challenge, especially when observing complex traits, are interactions between multiple genetic and environmental factors which raise problems of cost effective sample size requirements (Khoury 2004). Till now, the approaches in epigenetic epidemiologic studies tend to be discovery-oriented rather than oriented to testing specific hypotheses. And although high-throughput epigenetic technologies already allow massive investigations of the epigenome, there is still no consensus to an appropriate way to model and present data. This problem has already been discussed by (Foley, Craig et al. 2009) and has not changed till now. The epigenome probably accounts for some of the heritability not explained by previous and current human genome epidemiologic approaches and holds measurable and unbiased markers for investigations. In

the Human Epigenome project (www.epigenome.org), tissue specific genome wide DNA methylation patterns are identified, catalogued and interpreted. Marks are already being classified by the type of variation (e.g. heritable) and by intra- and inter individual variations. In a next step, information of relationships (e.g. epigenotype – haplotype association) will have to be examined (Bjornsson, Fallin et al. 2004). Similar efforts are being made for the histone code and miRNAs and will allow for assessing the importance to include epigenotypes in gene discovery, or for evaluating the impact of time and environment on a genetic predisposition. Integration of epigenetic data in discovery orientated approaches can reveal information that is not captured by previous observations as demonstrated by Figueroa et al in acute leukemias of distinct cell lineages. By combining epigenetic data with expression arrays, the analysis revealed hundreds of additional differentiated expressed genes that were missed by single-platform microarray studies and it also enhanced the detection and statistical significance of biological pathways that were deregulated in the conditions (Figueroa, Reimers et al. 2008). An integrative approach has promising solutions in cases of incomplete penetrance and variable expressivity of complex diseases.

CONCLUSION

We selected a molecular network in BAT which is under strong environmental influences. Current knowledge of SNPs in the network does not reveal conclusive information, while epigenetic regulation clearly influences the network's thermogenic capacity. Evidence for the contribution to Ob will have to be warranted in both epigenetic and genetic epidemiology.

In such a complex condition, genetic and environmental factors will have to be observed simultaneously using unbiased methods. Therefore, tools of epigenetic epidemiology might be helpful. The integration of validated epigenetic biomarkers into databases will make it easier to assess the importance of including measured epigenotypes in gene discovery or for evaluating the impact of time and environment on a genetic predisposition.

In 2008, Khoury et al (Rothman 2008) predicted a new level of complexity in epidemiology by simultaneously studying biomarkers of exposure, susceptibility, and outcomes. High-throughput methods and innovative approaches for analyzing big amounts of data are making this easier (Figueroa, Reimers et al. 2008) (Zhang, De et al.). Epigenetics is a promising feature to expand the scientific frontiers on etiology, disease risk prediction and prevention of health threatening conditions.

CURRICULUM VITAE

geb. Salzburg, Österreich am 18. September 1983

Ausbildung

Sep 1997 – Juni 2003	Bundeshandelsakademie I, Salzburg
Okt 2003 – Okt 2004	Internationale Entwicklung, Universität Wien
Okt 2004 – Okt 2010	Diplomstudium Ernährungswissenschaften, Universität Wien

Besondere Auszeichnungen & Auslandsaufenthalte zu Studienzwecken

Nov 2008 Stipendium für kurze wissenschaftliche Aufenthalte der Universität Wien
Universita Cattolica del Sacro Cuore, Faculty of Medicine A.Gemelli, Institute of Hygiene. Rome, Italy
International courses in Special Topics in Epidemiology and Biostatistics: Evidence-based genome research in medicine: from genetic epidemiology research to public health application.

Aug 2009 Netherlands Institute of Health Sciences/ESP fellowship 2009
Erasmus University Rotterdam
3 Weeks in Quantitative Medical Research for Health Professionals: genetic epidemiology

Okt 2009 Teilnahme und Posterpräsenation am *3rd Congress of the International Society of Nutrigentics/Nutrigenomics*, Bethesda, Washington CD, USA.

Wissenschaftliche Tätigkeiten

- Wissenschaftliche Konferenzen mit eigenem Beitrag:
  19th International Congress of Nutrition, 4th-9th Oct 2009 Bangkok, Thailand
  Poster: *Combined hereditary/genetic & epigenetic aspects in genetic testing of complex diseases and nutrigenetics*
  3rd Congress of the International Society of Nutrigenetics/Nutrigenomics, 21st-23st Oct 2009, Bethesda/Maryland, USA
  Poster: *Combined hereditary/genetic & epigenetic aspects in genetic testing of complex diseases and nutrigenetics: PPARgamma*
- seit Nov 2008: wissenschaftliche Assistenz und Laborassistenz in der Arbeitsgruppe von Prof. Dr. Alexander Haslberger am Institut für Ernährungswissenschaften der Universität Wien, Althanstraße, 1090 Wien
- Publizierte Texte: *Linking hereditary, environmental and nutritional aspects (Nutrigenetics)*
  Carolin Berner, Stefanie Englert, Alexander G. Haslberger. Journal für Ernährungsmedizin, 12. Jahrgang Nummer 1 /2010. Plus Zeitung, p.b.b, Verlagsort 1010 Wien; GZ 10z038394P