

# Genetics of dietary habits and obesity

- a twin study

*Ann Louise Hasselbalch*

This review has been accepted as a thesis together with six previously published papers by University of Copenhagen 5th of November 2009 and defended on 11th of January 2010.

Tutors: Thorkild IA Sørensen, Kirsten Ohm Kyvik and Karri Silventoinen.

Official opponents: Simon Francis Thomsen, Finn Rasmussen and Kim Overvad.

Correspondence: Institute of Preventive Medicine, Copenhagen Capital Region, Copenhagen University Hospitals, Øster Søgade 18, 1, 1357 Copenhagen, Denmark.

E-mail: awj@ipm.regionh.dk

Dan Med Bull 2010;57: B4182

## THE SEKS ORIGINAL PAPERS ARE:

- 1: Hasselbalch AL, Benyamin B, Visscher PM, Heitmann BL, Kyvik KO, Sørensen TIA. Common genetic components of obesity traits and serum leptin. *Obesity* 2008; 16(12): 2723-2729.
- 2: Hasselbalch AL, Heitmann BL, Kyvik KO, Sørensen TIA. Studies of Twins Indicate That Genetics Influence Dietary Intake. *Journal of Nutrition* 2008; 138(12): 2406-12.
- 3: Hasselbalch AL, Silventoinen K, Kesitalo K, Pietiläinen KH, Rissanen A, Heitmann BL, Kyvik KO, Sørensen TIA, Kaprio J. Twin Study of Heritability of Eating Bread in Danish and Finnish Men and Women. *Twin Research and Human Genetics* 2010; 13(2): 163-167.
- 4: Hasselbalch AL, Ängquist L, Christiansen L, Heitmann BL, Kyvik KO, Sørensen TIA. A Variant in the Fat Mass and Obesity-Associated Gene (FTO) and Variants near the Melanocortin-4 Receptor Gene (MC4R) Do Not Influence Dietary Intake. *Journal of Nutrition* 2010; 140: 831-834.
- 5: Hasselbalch AL, Heitmann BL, Kyvik KO, Sørensen TIA. Associations between dietary intake and body fat independent of genetic and familial environmental background. *International Journal of Obesity* 2010; 34:892-898.
- 6: Silventoinen K, Hasselbalch AL, Lallukka T, Bogl L, Pietiläinen KH, Heitmann BL, Schousboe K, Rissanen A, Kyvik KO, Sørensen TIA, Kaprio J. Modification effects of physical activity and protein intake on heritability of body size and composition. *Am J Clin Nutr* 2009; 90:1096-103.

## INTRODUCTION

### Trends of obesity

The prevalence of obesity (body mass index, BMI  $\geq 30$  kg/m<sup>2</sup>) has over the past decades increased dramatically world-wide in both sexes, all age groups, all ethnicities and all educational levels (1). The International Obesity Task Force estimated that over one billion people globally were overweight (BMI  $\geq 25$  kg/m<sup>2</sup>) in 2003 (IOTF analysis of data gathered for the WHO Global Burden of Disease ([www.iotf.org](http://www.iotf.org))).

In the Danish population there has been a pronounced, distinctive increase in the prevalence of obesity during the last 25-30 years (2). The latest estimation of the prevalence of obesity in the adult population based on objective measures of height and weight obtained in 1999-2001 demonstrates a prevalence of 17.7% in men and 17.8% in women (3). Overweight and obesity has increased not only in the adult population but also among school age children. In 2003 every fifth girl and more than every seventh boy in Copenhagen were overweight (2).

Obesity appears to be a complex and composite phenotype in which different compartments of adipose tissue exhibits different biology and hence may have different genetic and environmental determinants of their sizes. The pathogenesis of obesity is complex and results from combined effects of genes, environment, lifestyle, and their interactions. The obvious differences between men and women in average body fat distribution reflect the biological complexity. Particularly, the differences between the trunk, especially the intra-abdominal fat mass, and the peripheral fat mass is important because of its implications for risk of cardiovascular disease and death (4). Moreover, the waist circumference for given levels of BMI is positively correlated with mortality, even within the normal range of BMI (5).

Given the association with several chronic diseases, obesity constitutes a major public health problem. Obesity has been found to increase the risk of type-2 diabetes, cardiovascular diseases, hypertension, some types of cancer and other health consequences including reduced mental and social health (6). These health consequences result in decreased quality of life, stigmatization, reduced working ability and early death (7). In addition, the increased prevalence of obesity results in increased expenses for prevention and treatment of the condition and its co-morbidities both at the individual and the societal level (8). The Danish health care expenses related to obesity and co-morbidities constitute approximately 3% of the total annual health care expenses (9). Also indirect costs related to reduced function and withdrawal from the labour market increase with the increased prevalence (8).

### Genetic epidemiology of obesity

The phenotype commonly used in population-based studies of obesity is BMI. BMI does, however, not specify whether excess body mass is due to excess fat mass and how the body fat is distributed. It is rather easy to assess BMI by self-report, since most people know their approximate height and weight. Using self-reported information on height and weight does, however, introduce bias, since both men and women tend to over-report their height and especially women tend to underreport their weight with increasing level of overweight (10). When calculating BMI this misreporting leads to an underestimation of BMI.

Waist circumference is becoming more and more common in population-based studies because of the increasing evidence of the association between waist circumference and especially waist for given BMI and increased risk of co-morbidity and mortality (5). Waist circumference can be assessed by self-report, but the best data are obtained by objective measurement. Other anthropometric phenotypes such as skin fold thicknesses, fat body mass and lean body mass are not assessable by self-report and the techniques (DXA, MRI) for assessing body fat mass are timely and costly.

Large family-based studies in different populations have consistently demonstrated a familial correlation in adult body mass index BMI, at about 0.2 between parents and offspring and at about 0.3 between siblings (11). The fact that both genes and environment influence human anthropometry has long been established (11). The amount of genetic influence on body fatness and body fat distribution has previously been established to be ~70% for BMI based on studies of twins (11,12). Some studies have found that the extent to which anthropometry is under genetic influence differs between men and women (13) and there seem also to be differences in the amount of genetic influence in different age groups (14). In addition, the extent to which the common environment influences anthropometry differs between age groups. The common environment seem to play a role in childhood, but during adolescence the influences by the common environment is reduced and by early adulthood the effect disappears (14). Fewer studies have addressed the genetic and environmental influences on body shape, assessed by body circumferences and skinfold measurements (15-18), and on body composition of the fat and lean mass (19-23). There is, however, a considerable variation in the results across the studies and they may be biased for several reasons.

The development within the field of molecular genetics has made genotyping more accessible and affordable. It is now possible to perform genome-wide-scans of even more than 300,000 single nucleotide polymorphisms (SNPs) on large numbers of subjects. This has led to identification of ~20 common SNPs associated with BMI (24,25). The SNP with the largest effect size is the rs99395609 in the fat mass and obesity associated (*FTO*) gene (26). Also SNPs (rs17782313, rs17700633 and rs12970134) near the melanocortin-4 receptor (*MC4R*) gene have been found to be associated with increased BMI (27). The function of the SNPs identified in the genome-wide-association studies (GWAS) is currently unknown and need to be further investigated.

The *FTO* gene is located on chromosome 16. The SNP for which the strongest association with obesity has been demonstrated is the rs99395609. The findings of Frayling et al. (26) have been replicated in several independent cohorts. The verification of the association between the *FTO* gene and body mass lends strong support to the suggestion that this gene has common variants (the AA and AT genotypes) that predispose to obesity, relative to the wild (TT) genotype. Expression studies indicate that *FTO* is

widely expressed in many tissues, but has its highest expression in the brain, particularly the arcuate nucleus of the hypothalamus (26).

The *MC4R* is located on chromosome 18 and is highly expressed in the hypothalamus, where the central control of feeding and energy balance is located (28). Genome-wide-association studies of obesity have identified loci located near the *MC4R* gene (27;29). The loci near the *MC4R* gene, that have been found to be associated with fat mass, are the rs17782313, rs17700633 (27) and rs12970134 (29).

### Genetic epidemiology of dietary habits

In contrast to the evidence of the heritability of obesity and body composition, the literature on the heritability of dietary intake is much more scarce. A reason for this could be, that human eating behaviour is a more complicated phenotype to study than body fat (30). The estimates of the heritability of dietary intake in the existing literature differ to a great extent, which is primarily due to the methods by which the food intake was measured and the relative sample sizes (30). The existing studies show that the heritability of the preference for macronutrients is larger than that for the individual food items (31). No sex difference in the heritability of overall dietary intake has been shown, but sex differences in the heritability of intake of food items has been demonstrated (32).

As mentioned in the previous section, consistent associations between body weight and variation in the *FTO* gene and near-*MC4R* gene have been found (26,27). The discovery of these common variants has turned attention to the genes' functional effects. A few studies have focused on associations between variation in *FTO* and dietary intake (33-38) and found associations between *FTO* and total energy intake (33,34,38,39), weight of food intake (36), intake of energy from dietary fat (34) and reduced satiety responsiveness (35). With regard to associations between the near *MC4R* SNPs and dietary intake, one study (40) have focused on the SNPs rs17700633 and rs17782313 and found positive associations for rs17782313 and total energy intake, intake of protein and intake of fat.

### Epidemiology of the relation between dietary habits and obesity

Traditionally, the development of overweight and obesity has been thought to be a consequence of positive energy balance, where persons consume more energy from food and beverages than they spend through their basal metabolism and physical activity resulting in an accumulation of energy stored as fat in the adipose tissue. Cross-sectional studies have demonstrated significant positive correlations between energy-adjusted fat intake and various measures of obesity (41). However, the results from prospective studies of dietary fat and weight changes have been inconsistent (42). To date, sugar-sweetened soft drinks are the only dietary component for which consistent evidence of an association with weight gain has been well established (43). When studying the relation between habitual dietary intake and anthropometry it is, however, important to recognize that the genetic background influences both diet and body fat mass. In an overfeeding study of 12 pairs of healthy, male monozygotic twins the variance in weight gain was three times higher between pairs than within pairs, indicating that weight gain originate from shared genetic and environmental influences (44).

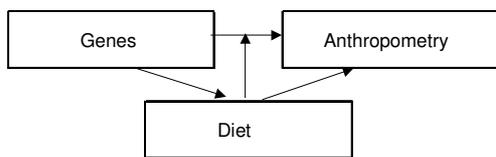
To get a better understanding of the complex interplay between genes, diet and obesity we need to take into account the genetic background of the subjects in the population-based studies of associations between habitual dietary intake and obesity. This can

be done by using data on genetically related individuals, such as twins, or by actual genotyping of unrelated subjects. In the present thesis the focus has been on analyses of twin data.

### Aims and research questions

The objective of the PhD-study is to investigate in a twin population the quantitative genetic, shared and unique environmental influences on dietary intake and on development of obesity, and thereby assess the extent to which habitual dietary intake can account for the environmental influences on the development of obesity. This includes investigation of common, correlated and interacting genetic and environmental effects on the dietary habits and the development of obesity.

Figure 1:



As illustrated in the previous sections there remain a number of unanswered questions:

- Are different obesity measures under the control of a common genetic component and is this the same across the two sexes (paper 1)?
- Is the habitual dietary intake heritable? And is the heritability of these factors the same across the two sexes (paper 2 and 3)?
- Does genes demonstrated to be associated with obesity measures have an impact on dietary intake (paper 4)?
- Is it possible to demonstrate associations between dietary intake and obesity measures if control for genetic background is included in the analysis (paper 5)?
- Does dietary intake modify the genetic influence on obesity (paper 6)?

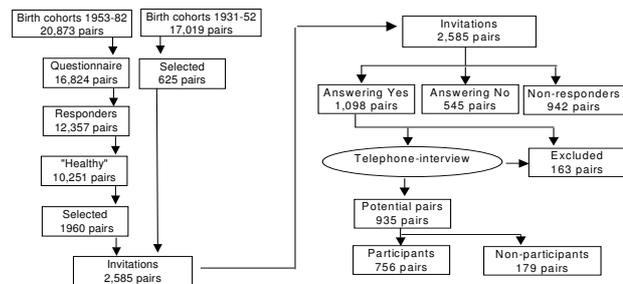
## MATERIALS

### The GEMINAKAR study

This PhD-study is primarily based on the Danish twin study GEMINAKAR. GEMINAKAR consists of 756 complete twin pairs recruited from the national, population based twin registry (Figure 2). GEMINAKAR's primary inclusion criteria were that at least one of the twins in a pair should be living no more than 100 km from Odense or Copenhagen, where the examinations took place, and that the twins were not participating in other ongoing studies. From the birth cohorts 1931-52 a total of 625 pairs were randomly selected. The birth cohorts 1953-1982 had recently received a nationwide questionnaire survey about physical health and health-related behavior (Omnibus1). Out of the 12,357 complete pair, who responded to the questionnaire, 10,251 were regarded as "healthy", as none of the twins had reported diabetes, rheumatic disease, epilepsy, asthma, chronic bowel disease, or anorexia nervosa. 1960 pairs were selected from the "healthy" group received invitations for the GEMINAKAR study. Of the 2,585 invited twin pairs, 942 pairs were non-responders in that one or both of the twins did not reply (Figure 2). Among 545 twin pairs at least one of the twins were not willing to participate in the study, leaving 1098 complete pairs (42%), who were willing to participate. These twins were contacted by telephone, given

oral information about the study and furthermore asked a few screening questions about their present health status. Either based on written replies or the telephone interview, 163 twin pairs were excluded due to diabetes mellitus, cardiovascular disease, pregnancy, breastfeeding, alcohol/drug abuse, or conditions making it difficult to participate in the clinical examination, which also included a bicycle exercise test.

Figure 2: Flow diagram for GEMINAKAR\*



\*Dropout can be the result of at least one twin in a pair being unwilling or unfit to participate in the study. Incomplete twin pairs were excluded from the sample.

Remaining 935 complete twin pairs were willing and able to participate. In order to obtain an equal distribution of twin pairs along the age span in the different zygosity groups, the sampling from this group was stratified according to age and sex. A total of 756 twin pairs (151 monozygotic (MZ) male pairs, 148 dizygotic (DZ) male pairs, 158 MZ female pairs, 168 DZ female pairs and 131 DZ opposite sex pairs (DZOS)) underwent examination. The participants were examined in either Copenhagen or Odense in the period of August 1997 to November 2000. Twins in a pair were examined on the same date.

The examination day started in the morning with the subjects fasting since having a standard meal at 9 p.m. the evening before. Blood samples were drawn for measuring insulin and glucose levels, lipid profile, growth hormone as well as blood serum and plasma samples for later analyses of e.g. DNA and serum leptin level. The blood samples were analyzed for a number of other compounds such as blood lipids, insulin, glucose etc. These data are not used in the present thesis. DNA-based microsatellite markers with the PE Applied Biosystems AmpFISTR Profiler Plus kit were used to determine the zygosity of the twins. Serum leptin concentration was derived from the blood samples using the BIOMOL leptin (human) ELISA AK-153 kit (BIOMOL, Plymouth, PA).

During the day the subjects underwent oral glucose tolerance test, measurements of height, weight, waist and hip circumference, skin fold measures at four measurement-points, bio-impedance measurement, ECG recording, measurement of blood pressure and a bicycle test. In addition the participants filled out questionnaires concerning family history, general health, housing, work, lifestyle exposures and socio-economic status (45).

### Anthropometric measurements

The participants' height to the nearest cm and their weight to the nearest 0.1 kg were measured, and BMI was calculated as weight (kg) / height<sup>2</sup> (m<sup>2</sup>). Skin fold thickness to the nearest 0.1 mm was measured on four different sites (biceps, triceps, sub-scapular and supra-iliac) using a Harpenden caliper. The skin fold measures were repeated three times at each measurement-point and the

average were used in the analyses to reduce measurement error. The sum of truncal skin fold thicknesses was calculated by summing the sub-scapular and supra-iliac skin fold thicknesses. The sum of extremity skin folds was calculated by summing the biceps and triceps skin fold thicknesses. Waist and hip circumference as well as pulse and blood pressure were also measured. The bio-electrical impedance was measured using a BIA-103 RJL-system analyzer (RJL-systems Detroit) with a 50 kHz, 800 micro Ampere device, following the instructions given by the manufacturer. Body fat assessed by bioelectrical impedance was estimated for women and men using the following equation, where R50 is the electrical impedance with a 50 kHz, 800 micro Ampere device, and where sex is 0 for women and 1 for men: Fat mass =  $0.819 * \text{weight} - 0.064 * \text{sex} * \text{weight} - 0.279 * (\text{height}^2 / \text{R50}) - 0.231 * \text{height} + 0.077 * \text{age} + 14.941$  (46). Owing to technological device problems, measurements of bioelectrical impedance were carried out in 517 twin pairs. Fat-free mass was calculated by subtracting the fat mass from the total body weight. Fat free mass index (FFMI) and fat mass index (FMI) were calculated as fat mass (kg) / height<sup>2</sup> (m<sup>2</sup>) and fat-free mass (kg) / height<sup>2</sup> (m<sup>2</sup>), respectively.

#### Information about habitual dietary intake

Information about the participants dietary intake was assessed through a very extensive food frequency questionnaire (FFQ), that was initially designed for the Danish "European Prospective Investigation into Cancer and Nutrition" (EPIC) study (47) and validated against two 7-d weighed diet records in the EPIC study (48). The questionnaire included 247 foods and recipes for which the respondents indicated frequency of consumption as the number of intakes per day, week, or month. Dietary calculations were made using the FOODCalc programme (49), which is based on values from the Danish National Food Tables 1996 (50). Replies on the food frequency questionnaire were obtained from 1212 subjects (600 complete twin pairs and 12 incomplete twin pairs). The FFQ was based on 1-month recall and included detailed information regarding: Beverage intake, breakfast products, bread and fat on bread, products to put on the bread, hot meals, accessories, vegetables, desserts, fruits, cakes, snacks and sweets, type of fat for bread and cooking, cooking style, fast food, meal structure, dietary habits, organic food habits and dietary supplements.

Based on the collected FFQs a diet-database was constructed. The diet database includes: Total energy with or without alcohol, intake of macronutrients (in grams, energy (kJ) and in energy%), fiber intake, dietary energy density (DED) calculated as total energy intake divided by total weight of intake (51), glycemic index (GI) and glycemic load (GL) and energy from 20 food groups (kJ). The criteria for the groupings are based on assumptions of communality of various nutritional aspects between the individual food items assigned to each group (table 1) (52).

#### Genotyping

Genotyping of the four SNPs (rs9939609, rs17782313, rs17700633 and rs12970134) were conducted by allelic discrimination using pre-designed Taqman<sup>®</sup> SNP genotyping assays (Applied Biosystems) applying the conditions described by the manufacturer. PCR was performed in the ABI Prism 7700 and analyzed using the Sequence Detection System software (Applied Biosystems). The GEMINAKAR sample has a lot of advantages in relation to the aims of this PhD-study. The sample is relatively large (756 twin pairs), it includes males and females, as well as MZ, DZ and DZOS

twin pairs. The subjects have been thoroughly examined, and dietary information of high quality has been collected. The sample includes lean, overweight and obese subjects in the age range of 18-67 years.

**Table 1:** Description of food groups

Food groups	Food items contributing to food groups
Whole grain	Whole grain bread, rye bread, whole grain flour, oatmeal, corncobs, muesli, whole grain crisp bread
Refined grain	White (wheat) bread, wheat flour, rice, pudding rice, potato flour, corn flour/starch, pasta breadcrumbs, crisp bread (wheat)
Potatoes	All kinds of potatoes
Vegetables	Fruity vegetables, leafy vegetables, root vegetables, cabbages, mushrooms, stalk vegetables, onions, legumes, soy sauce, herbs
Fruits	Citrus and other fruits (fresh and canned), nuts
Juices	Vegetable and fruit juice
Red meat	All kinds of beef, pork, lamb, veal and offal
Processed meat	Bacon, smoked ham, sausages, liver pate
Poultry	Chicken, turkey and game
Fish	Fresh and processed lean and fatty fish
Eggs	Eggs
Lean milk products	Skimmed (<0.5g) and semi-skimmed milk (1.5g), buttermilk, low-fat yogurts (1.5 g), cottage cheese (5 g), firm cheese (16), milk ice (3 g), processed cheese (17 g), unripened smoked cheese (0.5 g)
Fat milk products	Whole milk (3.5 g), whole milk yogurt, cream, firm cheese (25 g), brie (28 g), camembert (22 g), Roquefort/Danablu (30 g), cream cheese (37 g), ice cream (10 g), unripened smoked cheese (10 g)
Butter	Butter and butter products (mix of butter (75%) and vegetable oils (25%)), lard
Vegetable oil	All kinds of vegetable oils, mayonnaise
Margarine	All kinds of margarines (diet and full fat)
Jam, syrups, sugar	Jam, sugar, honey, fruit syrups
Soft drinks	Sugar-sweetened soft drinks
Candy, chocolate, snacks	Chocolates, sweets, liquorices, fruit gum, toffees etc.
Rest/Other	Coffee, tea, not specified

#### The Omnibus studies

In 1994 (Omnibus1) and 2002 (Omnibus2) twins in the Danish Twin Registry received a mailed questionnaire. In the questionnaires the participants were asked a number of questions i.a. about their height and weight, and in Omnibus2 also their waist circumference. For the subjects in GEMINAKAR who also participated in the previous and/or the following questionnaire-study, it is possible to analyze the weight development before and after as well as changes in waist circumference after the dietary information was collected.

#### FinnTwin16

The PhD-study is part of the EU-funded integrated research project Diogenes where data from the Finnish twin study FinnTwin16 (University of Helsinki, Department of Public Health) is included. The FinnTwin16 study is a population-based, longitudinal study of five consecutive Finnish twin cohorts born in 1975-1979. Baseline data collection was conducted within 60 days of their 16<sup>th</sup> birthday. The fourth follow-up questionnaire, which was carried out when the twins were 23-27 years of age, included a food frequency questionnaire based on one-year recall as well as questions about height and weight. The study consists of 2010 MZ, DZSS and DZOS twin pairs with information on zygosity, dietary intake, and height and weight. Data from the FinnTwin16 study are included in two sub-studies of the present PhD-thesis (papers 3 and 6).

#### METHODS

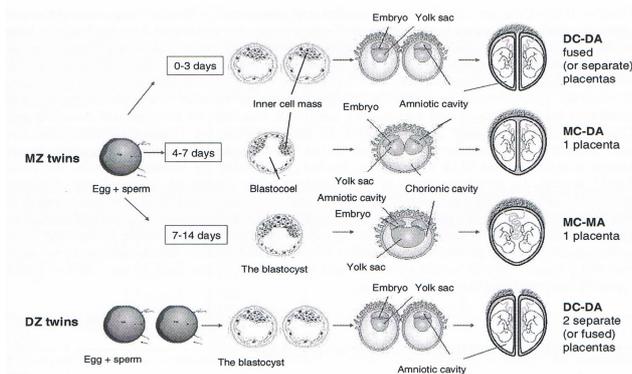
To answer the research questions of the PhD-study, twin studies were used as a method to quantify the influence of genes, dietary

intake and the interactions on the risk of development of common forms of obesity.

### Monozygotic and dizygotic twins

Twins are two offspring resulting from the same pregnancy born in close succession. They can be either monozygotic (MZ) or dizygotic (DZ). The rate of MZ twinning is relatively stable, occurring in approximately four pregnancies out of 1000 across countries (53). DZ twinning rates, by contrast, vary by geographical region; in Asia about 6 in 1000, in Europe and USA about 10-20 in 1000 and in Africa about 40 in 1000 pregnancies are DZ twin pregnancies (54). The tendency to give birth to DZ twins is inherited and increases with maternal age and use of fertility drugs or in vitro fertilization procedures (53). About one-third of twins born are MZ, one-third are DZ same-sex (DZSS), and the remaining one-third are DZ opposite-sex (DZOS) (53). DZ twins almost invariably have their own placentas and are dichorionic and diamniotic (DC-DA) in placental membrane structure (54) (Figure 3).

**Figure 3:** Development of monozygotic (MZ) and dizygotic (DZ) twins. Placental membranes are dichorionic and diamniotic (DC-DA), monochorionic and diamniotic (MC-DA) or monochorionic and monoamniotic (MC-MA).



MZ twins, on the other hand, may be DC-DA, monochorionic and diamniotic (MC-DA) or monochorionic and monoamniotic (MC-MA). The timing of ovum division determines whether MZ twins have DC-DA, MC-DA or MC-MA membrane structure (54) (Figure 3). Approximately 30% of MZ co-twins separate within the first three days of fertilization, yielding DC-DA placentation, 70% divide between 4 and 7 days resulting in MC-DA and 1-2% of cases divide later (7-14 days), yielding twins with only one set of membranes (MC-MA) (54).

Monochorionic twins have higher perinatal mortality than dichorionic twins (53), and in survivors, the health of the twins may be compromised due to vascular anatomoses and drainage of blood from one twin to another (Twin-Twin-Transfusion syndrome) (53,55)

Because of nutritional imbalances between the co-twins in utero, MC MZ twins typically have larger intrapair differences in birth weight than DC MZ pairs (56). Further, MC MZ twins tend to weigh approximately 100 g less at birth than DC MZ pairs, who in turn are almost 100 g lighter than DZ twins (53,57). Twins are on average born three weeks pre-term and are 1000 g lighter than singletons (58). Thus, for size at birth MZ twin and DZ twins differ and in general newborn twins are not representative of newborn singletons. When studying body fat mass in adulthood these differences, however, seem to have disappeared. In the GEMINA-

KAR twins no significant differences between the zygosity groups in adults BMI, fat mass and other anthropometric measures (For more details see table 1 in paper 1) could be found. The mortality in childhood is higher among twins than singletons, but this difference disappears with age (59).

### Twin methodology

Quantitative genetic analyses examine the nature of individual differences as well as similarities between family members and other relatives. In order to address the question regarding genetic and environmental influences on dissimilarities, the variance of a trait is studied. To do so it is necessary to perform studies on subjects with different degrees of genetic and environmental relationships. The twin-design is, therefore, of great use for genetic studies as twins are sampled from the same gene pool and they share same genes, although to a different degree. MZ twins share in principal all of their genes, and DZ twins share, on average, half of their segregating genes such as ordinary full siblings, but unlike ordinary full siblings, twins are matched on age.

Twin studies aim to explain the inter-individual variation in a trait. Studying twins offers a unique opportunity to study genetic and environmental effects on multifactorial polygenic traits such as body weight and habitual dietary intake. Genetic effects may arise from cumulative effects of multiple genes (additive genetic effects), or because of interaction between the alleles of these genes (dominance genetic effects). Environmental differences may arise from the environment unique to the individual (non-shared environment) making the twins unlike or from the environment common to co-twins (shared environment) making the twins alike. The greater similarity of MZ than DZ twin pairs is regarded to result from genetic effects. This is the basic principle of twin methodology.

The classic twin study compares phenotypic resemblances of MZ and DZ twins. Comparing the resemblance of MZ twins for a trait with the resemblance of DZ twins offers the first estimate of the extent to which genetic variation determines phenotypic variation of the trait. If MZ twins resemble each other more than do DZ twins, then the narrow heritability ( $h^2$ ) of the phenotype can be estimated from twice the difference between MZ and DZ correlations. The proportion of the variance that is due to a shared environment is the difference between the total twin correlation and the part that is explained by heritability (60).

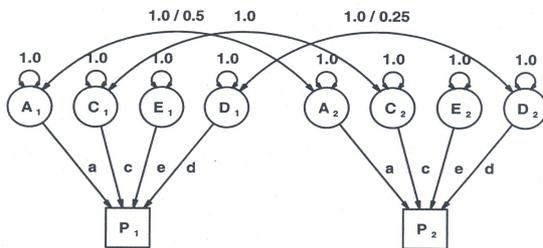
### Structural equation models

Structural equation models are widely used to estimate the genetic and environmental components in twin studies. The aim is to estimate the genetic and environmental components of variance in a given trait among twins. The advantages of applying structural equation modelling as compared to only comparing twin correlations is that the variance of a trait can be partitioned into several components, the fit of different models can be tested against each other to find the model which best explains the data, and confidence intervals for the estimates can be computed. Further, this method offers an opportunity to analyse more complex processes such as common genetic background of several correlated traits or developmental processes (61). Structural equation modelling is based on a path-analysis approach, which summarises the variance and the intrapair covariance data of MZ and DZ twin pairs. The expected variances and co-variances generated by the model are fit to the observed data and tested against the null hypothesis that the data and the model do not differ significantly.

Parameters contributing to the inter-individual variance of a phenotype (P) are the additive genetic effects (A), consisting of cumulative effects of several single genes, the dominant genetic effects (D), resulting from interactions between alleles in the same loci, the common environmental effect (C), shared by the twins and the specific environmental effect (E) not shared by the twins. Measurement errors are modelled as part of E. Epistatic effects, that is interactions between alleles in different loci, are modelled as part of dominance genetic effects if the loci is not linked. In the case of close linkage between loci, they segregate together and subsequently any epistatic effect between them is modelled as part of additive genetic factors.

A univariate path-analysis model for additive genetic, shared and non-shared environmental and dominance genetic effects are presented in figure 4.

**Figure 4:** Univariate model



P<sub>1</sub>: Phenotype for twin 1  
P<sub>2</sub>: Phenotype for twin 2

If only data on twin pairs reared together are available, it is not possible to model dominance genetic and shared environmental effects simultaneously. This is because the common environmental effects increase and the dominance genetic effects decrease the relative size of the DZ correlation as compared to the MZ correlation when compared to a situation where the resemblance within twin pairs is totally due to additive genetic effects. Although the method does not allow modeling of C and D simultaneously, the models does not rule out the possibility of that influence of both C and D are present for the phenotype under study.

The three models based on combinations of parameters used in this study are:

- AE-model: Additive genetic and non-shared environment effects on phenotypic variance
- ACE-model: Additive genetic, shared and non-shared environmental effects on phenotypic variance
- ADE-model: Additive and dominant genetic, and non-shared environmental effects on the phenotypic variance

The model parameters are estimated by using maximum likelihood method and the fit of the model is tested using Chi-squared-statistics ( $\chi^2$ ). The degrees of freedom are based on the number of unique statistics (variances and co-variances) and the number of estimated parameters. The best-fitting model is the simplest model with the largest statistical probability and similarity to the data under study. The hypothesis that the data fit to the tested model is rejected if the Chi-squared test is significant at a chosen level of statistical significance. The Mx-programme (62) is an

example of a statistical program especially suited for the structural equation approach for twin and family studies. The estimates of heritability obtained by structural equation models are generally more reliable than estimates obtained by the classic twin study method (61). The use of structural equation models requires that the general twin study assumptions are valid in the study sample.

#### Assumptions of twin studies

The twin study approach is based on basic assumptions that apply to both the classic twin study and the structural equation approach. The validity of these assumptions is dependent on the trait under study, and may vary in different study populations at different times.

##### *Monozygotic twins are genetically identical*

Implicit in the twin study method is the assumption that MZ twins are genetically identical. During cell division, however, genetic mutations can occur. In addition, even though MZ twins are genetically identical, epigenetic differences may occur during their lifetime resulting in differences in gene expression (63).

##### *Equal environments*

###### Non-shared environment

The assumption that the effect of the non-shared environment on the phenotype is similar in MZ and DZ twins are fundamental for the design.

A violation of this assumption could be the result of either genetic pre-disposition, or of a conscious choice by one of the groups. The genetic pre-disposition to differential non-shared environment would occur if, for example, MZ twins were more inclined to make choices that resemble each other. This could either make the non-shared environment more similar or less similar leading to a biased estimation of the shared environment.

###### Shared environment

The assumption states that the effect of the shared environment on the phenotype is similar in MZ and DZ twins. Circulation differences are known to be present between MZ and DZ twins during foetal life. All DZ twins have separate placentas, whereas two thirds of MZ twins have only one placenta and one chorion (64). Such monozygotic twins also have a lower birth weight than dizygotic MZ and DZ pairs (65). Furthermore, the mean difference in birth weight is larger among MZ than DZ twins due to the twin-twin transfusion syndrome, i.e., the common circulation among monozygotic twins (66). However, the difference in bodyweight between monozygotic and dizygotic twins has been observed to diminish during childhood (67). Moreover, it is possible that MZ twins are treated more similarly than DZ twins in childhood.

Potential violation of the assumption of equal environment, where MZ environments are more similar than DZ environments, leads to overestimation of genetic effects, because the greater similarity of MZ pairs does not arise from genetic effects only.

##### *No gene-environment interaction*

The assumption of no gene-environment interaction assumes that an individual's response to the environment is not dependent on the genotype, or that the expression of genotype is not dependent on the environment (GxE effects). Gene-environment interactions are modelled as part of the genetic effects if the environmental exposure is shared by the co-twins and as part of the specific environment if it is not shared (68).

### Random mating

The assumption of random mating assumes that mating between individuals is random for the trait under study, e.g. lean men randomly marry lean and overweight women. Assortative mating exaggerates the similarity of the DZ pairs since they share more than 50% of their segregating genes, and leads to underestimation of the proportion of variance attributable to heritability and overestimation of the shared environmental effect.

### Representativeness

#### Representativeness of the sample of twins

The twin sample under study should be representative of the general population of twins. There exists a risk of violation of this assumption in twin registries established on voluntary basis. The Danish Twin Registry is sampled from the Danish parish books, the Civil Registration System and the Medical Birth Registry including on average 70% of twins born before 1968 when the Civil Registration System was established and all twins born after 1968. The participation in the various research projects of the Danish Twin Registry is voluntary, while registration into the registry is not.

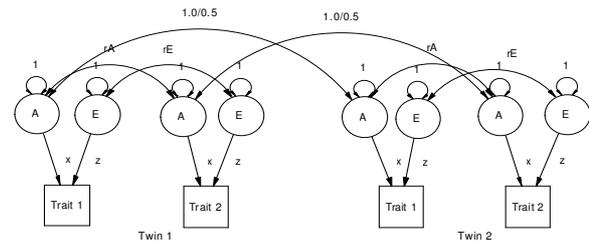
#### Representative of the general population

The individuals in the sample of twins should be representative of the general population with regard to the phenotype under study. This assumption can be tested by comparison of whether means and variances of continuous data and incidence and prevalence of disease are the same as in the general population.

### Bivariate analyses

In the study of common genetic and environmental influence on a variety of anthropometric measures (paper 1), the bivariate correlation model was used. The bivariate correlation model addresses to what extent the variance components A, C, D and E affect two different phenotypes and to what extent there is an overlap between the phenotypes (figure 5). In other words the bivariate model estimates to what extent the observed covariance between the traits can be accounted for by correlations between the genetic effects and/or correlations between the environmental effects on the traits. As the anthropometric measures are correlated with age (69), the age of the twins was controlled for in the model (62,70). Effects of age were incorporated in the mean model by linear regression of the phenotypes on age. The bivariate analyses were conducted separately for men and women and the outcome of the analyses were not compared statistically between the two sexes for several reasons. First, there are obvious differences between men and women in average body fat distribution, which have a genetically determined biological and physiological basis. Second, including the two sexes in the same analyses would require that sex-specific effects were taken into account by analysing also the opposite-sex dizygotic twin pairs, and this would rest on the untestable assumption that differences in intra-class correlations between the same-sex and opposite-sex twin pairs were due to either genetic or environmental sources of variance. Third, the inclusion of sex-specific effects in the bivariate models would make them too complex and lead to too unreliable estimates with the present sample size.

Figure 5: Bivariate correlation model



### Regression analyses

In the study of associations between the obesity related variations in a SNP (rs9939609) in the *FTO* gene and three SNP (rs17782313, rs17700633 and rs12970134) near the *MC4R* gene (paper 4) the analytical strategy was somewhat different from the traditional twin methods.

In this study departure from Hardy Weinberg Equilibrium was tested for all SNPs and the LD-pattern between the three SNPs near *MC4R* was tested.

Associations between the *FTO* SNP and the near *MC4R* SNPs, including suitable adjustment for age and sex, were assessed by regression analyses. Some of the dietary variables were transformed by natural logarithm to achieve a more normal distribution. Non-transformed variables and variables transformed by natural logarithm were analyzed by linear regression analyses. Some of the food groups were dichotomized, based on the 80% percentile, because their distributions were skewed. In other words, in each case, the first four quintiles (low values) give an observation of 0 (zero) and the fifth quintiles (high values) give an observation of 1. The variables that were dichotomized were analyzed by logistic regression analysis. The analytical approach assumed an additive genetic model and the analyses were run with and without adjustment for BMI and with and without adjustment for total energy intake in the analyses of macronutrient energy intake and energy intake from the 20 food groups. Since the subjects in the study sample are members of twin pairs, we adjusted for the mutual dependency of the twin individuals in the twin pairs by using a twin cluster method in the analyses.

### Correlation analyses of MZ intrapair differences

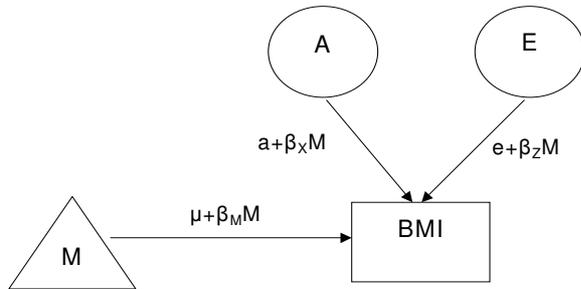
In the study of associations between intrapair differences in dietary intake and intrapair differences in anthropometric measures (paper 5) partial correlation analyses were performed. By computing the partial correlations between the MZ-intrapair differences in dietary intake and intrapair differences in anthropometric phenotypes we eliminate any potential influence of genetic effects and environmental effects shared by the twins, since any intrapair difference by definition (the MZ twins in a pair are genetically identical) only can be due to unshared environmental effects. The partial correlations were adjusted for age. As an additional analysis we excluded pairs that were discordant for obesity, smoking and physical activity to examine whether such differences would confound the results.

### Modification analyses

In the study of the modification effect of protein intake and physical activity on the genetic variation of BMI, waist circumference and body fat mass (paper 6), we used the gene-environment interaction model as presented in figure 6 (71). The moderator factor, e.g. protein intake, is denoted as M. This factor can affect the mean trait value ( $\beta_M$ ) but also modify the effects of genetic

( $\beta_x$ ) and unique environment ( $\beta_z$ ) on the trait. In practice this means that protein intake can affect both the mean BMI and the variances. The genetic models were carried out using Mx statistical package (62). Only linear modification effects were fitted in the models since our data are not large enough to test non-linear effects.

**Figure 6:** Gene-environment interaction model



**Multiple testing**

In some of the studies of this thesis multiple testing not driven by particular hypotheses derived from the literature are performed. The results are therefore evaluated using the Bonferroni correction for mass-significance.

**RESULTS**

The main results of papers 1-6 will be summarized in this section, whereas a more detailed description is given in the papers.

**Genetic and environmental influence on anthropometry (paper 1)**

*Specific aims*

The aim of this study was to estimate to what extent the genetic and environmental factors underlying different anthropometric traits are the same. This was estimated by performing bivariate variance decomposition on a number of anthropometric phenotypes. Included in the study were (i) the overall fat mass (BMI and FMI), (ii) truncal fat mass (waist circumference and truncal skin fold thicknesses), peripheral fat mass (hip circumference and extremity skin fold thicknesses), (iv) FFMI and (v) serum leptin level. The analyses were performed separately for men and women and all analyses were adjusted for age.

*Results*

The univariate analyses demonstrated genetic and environmental influences on all anthropometric measures in both men and women. For all anthropometric traits the majority of the variation could be explained by additive genetic variation (table 2). The common environmental influence on the traits was close to zero. The only trait for which the common environmental influence was significant was hip circumference. For all other traits the common environmental effect was insignificant. Exclusion of the common environmental component for all other trait than hip circumference did not result in a significantly worse model fit.

**Table 2:** Estimates of standardised variance components with 95% confidence intervals in parentheses from the univariate genetic analyses

Phenotype	$h^2$	$c^2$	$e^2$
<b>BMI (kg/m<sup>2</sup>)</b>			
Men	0.71 (0.71-0.77)	-	0.29 (0.23-0.29)
Women	0.76 (0.70-0.81)	-	0.24 (0.19-0.30)
<b>FMI (kg/m<sup>2</sup>)</b>			
Men	0.66 (0.54-0.75)	-	0.34 (0.25-0.46)
Women	0.74 (0.66-0.80)	-	0.26 (0.20-0.34)
<b>Hip circumference (cm)</b>			
Men	0.55 (0.33-0.80)	0.25 (0.01-0.45)	0.20 (0.16-0.26)
Women	0.50 (0.28-0.75)	0.24 (0.01-0.43)	0.26 (0.21-0.33)
<b>Sum of extremity skin folds(cm)</b>			
Men	0.70 (0.62-0.77)	-	0.30 (0.23-0.38)
Women	0.65 (0.55-0.73)	-	0.35 (0.28-0.45)
<b>Waist circumference (cm)</b>			
Men	0.60 (0.50-0.69)	-	0.40 (0.31-0.50)
Women	0.67 (0.59-0.74)	-	0.33 (0.26-0.41)
<b>Sum of truncal skin folds (cm)</b>			
Men	0.70 (0.61-0.77)	-	0.30 (0.23-0.39)
Women	0.72 (0.64-0.78)	-	0.28 (0.22-0.36)
<b>FFMI (kg/m<sup>2</sup>)</b>			
Men	0.92 (0.90-0.94)	-	0.08 (0.06-0.10)
Women	0.88 (0.86-0.89)	-	0.12 (0.10-0.15)
<b>Leptin (mmol/l)</b>			
Men	0.38 (0.24-0.51)	-	0.62 (0.49-0.76)
Women	0.59 (0.48-0.67)	-	0.41 (0.33-0.52)

All anthropometric traits correlate on the phenotypic level with phenotypic correlations  $r_p$  been moderate to high (for details see table 4 in paper 1). Bivariate correlation analyses were performed in order to estimate the amount of common genetic and environmental effects on the traits under study. Since the univariate analyses demonstrated that the best fitting model was the AE-model this model was used in the bivariate analyses.

The additive genetic correlations between the measures of total fatness (BMI and FMI, kg/m<sup>2</sup>) were very high, 0.94 in men and 0.98 in women (table 3). The genetic correlations between the measures of total fatness and waist circumference were very high in men ( $r_A$  (BMI and waist) = 0.83 and  $r_A$  (FMI and waist) = 0.82). In women the measures of total fatness had high genetic correlations with both waist and hip circumference ( $r_A$  (BMI and waist) = 0.87,  $r_A$  (BMI and hip) = 0.81 and  $r_A$  (FMI and waist) = 0.84 and  $r_A$  (FMI and hip) = 0.84). In women, also the genetic correlation between waist and hip circumference was very high ( $r_A$  (waist and hip) = 0.82).

Even though, the genetic correlations between some of the anthropometric traits were very high, all correlations were significantly < 1.00 as judged from the confidence intervals.

The results indicate that the phenotypes of BMI, FMI, and waist circumference are closely genetically related in men. In women also hip circumference are closely genetically related to BMI, FMI, and waist circumference. Generally the results of the study show that the anthropometric traits are moderately to highly correlated, and the results thereby suggest that there exists strong common genetic factors for all obesity related traits, but also that there is a residual genetic influence on the various fatness measures that cannot be explained exclusively by the genetic influence on overall fatness. The results of the study underline the importance of taking the phenotype under study into careful consideration when searching for genes associated with excess body fatness.

**Table 3:** Matrix of additive genetic correlations of obesity related traits with 95% confidence intervals in parentheses; results for men above and for women below the diagonal.

	BMI (kg/m <sup>2</sup> )	FMI (kg/m <sup>2</sup> )	Hip (cm)	Sum of extremity skin folds (cm)	Waist (cm)	Sum of truncal skin folds (cm)	FFMI (kg/m <sup>2</sup> )	Leptin (mmol/l)
BMI (kg/m <sup>2</sup> )	-	0.94 (0.92-0.96)	0.60 (0.51-0.67)	0.49 (0.37-0.59)	0.83 (0.78-0.87)	0.61 (0.51-0.69)	0.60 (0.50-0.68)	0.60 (0.45-0.74)
FMI (kg/m <sup>2</sup> )	0.98 (0.97-0.98)	-	0.73 (0.64-0.80)	0.61 (0.49-0.71)	0.82 (0.75-0.87)	0.73 (0.63-0.80)	0.52 (0.40-0.63)	0.54 (0.36-0.69)
Hip (cm)	0.81 (0.77-0.85)	0.84 (0.79-0.88)	-	0.47 (0.35-0.57)	0.68 (0.60-0.74)	0.61 (0.51-0.69)	0.22 (0.09-0.34)	0.69 (0.60-0.77)
Sum of extremity Skin folds (cm)	0.70 (0.62-0.76)	0.74 (0.66-0.80)	0.70 (0.61-0.77)	-	0.49 (0.36-0.60)	0.72 (0.64-0.79)	0.31 (0.18-0.44)	0.67 (0.56-0.76)
Waist (cm)	0.87 (0.84-0.90)	0.84 (0.79-0.88)	0.82 (0.77-0.86)	0.61 (0.51-0.70)	-	0.57 (0.44-0.66)	0.31 (0.19-0.43)	0.65 (0.55-0.74)
Sum of truncal Skin folds (cm)	0.74 (0.67-0.79)	0.79 (0.73-0.84)	0.76 (0.69-0.81)	0.70 (0.61-0.77)	0.71 (0.62-0.77)	-	0.38 (0.25-0.49)	0.74 (0.65-0.81)
FFMI (kg/m <sup>2</sup> )	0.60 (0.52-0.68)	0.59 (0.49-0.67)	-0.01 (-0.16-0.14)	0.21 (0.07-0.34)	0.22 (0.06-0.37)	0.31 (0.18-0.44)	-	0.32 (0.17-0.46)
Leptin (mmol/l)	0.71 (0.63-0.78)	0.57 (0.45-0.67)	0.48 (0.32-0.62)	0.75 (0.59-0.92)	0.64 (0.48-0.77)	0.68 (0.52-0.84)	0.28 (0.14-0.40)	-

FFMI, fat-free mass index; FMI, fat mass index

## Genetic and environmental influence on dietary intake (papers 2, 3 and 4)

### Specific aims

The aim of the studies related to genetic and environmental influence on dietary intake was to estimate the genetic and environmental contribution to habitual dietary intake. The studies included confirmation of the results in men and women and in two countries. In addition, associations between habitual dietary intake and variations in one obesity-associated SNP in the *FTO* gene and three obesity-associated SNPs near the *MC4R* gene were studied.

### Results

#### Nutritional factors and food groups

Based on information about habitual diet from the FFQ the genetic influence on total energy intake, macronutrient energy intake, Glycemic Index (GI), Glycemic Load (GL) and Dietary energy density (DED) as well as intake of energy from 20 food groups was calculated.

**Table 4:** Estimates of additive genetic effects ( $a^2$ ), non-additive genetic effects ( $d^2$ ), shared environmental effects ( $c^2$ ) and un-shared environmental effects ( $e^2$ ) on nutritional factors in all twin pairs for men and women.

	$a^2$	$d^2$	$c^2$	$e^2$
<b>Men</b>				
Energy Intake	0.38 (0.24-0.51)	-	-	0.62 (0.49-0.76)
Protein E%	0.28 (0.12-0.43)	-	-	0.72 (0.57-0.88)
Carbohydrate E%	0.36 (0.22-0.49)	-	-	0.64 (0.51-0.78)
Fat E%	0.01 (0.00-0.45)	0.36 (0.00-0.50)	-	0.64 (0.50-0.79)
Alcohol E%	0.46 (0.32-0.58)	-	-	0.54 (0.42-0.69)
GI	0.30 (0.14-0.44)	-	-	0.70 (0.56-0.86)
GL	0.25 (0.11-0.38)	-	-	0.75 (0.62-0.89)
Dietary Fiber	0.41 (0.27-0.52)	-	-	0.59 (0.48-0.73)
DED	0.17 (0.00-0.49)	-	0.24 (0.00-0.45)	0.59 (0.47-0.72)
<b>Women</b>				
Energy Intake	0.32 (0.12-0.48)	-	-	0.68 (0.52-0.88)
Protein E%	0.01 (0.00-0.39)	0.55 (0.15-0.66)	-	0.44 (0.34-0.58)
Carbohydrate E%	0.49 (0.35-0.61)	-	-	0.51 (0.39-0.65)
Fat E%	0.01 (0.00-0.57)	0.53 (0.00-0.64)	-	0.46 (0.36-0.59)
Alcohol E%	0.06 (0.00-0.52)	0.61 (0.13-0.75)	-	0.33 (0.25-0.44)
GI	0.36 (0.21-0.48)	-	-	0.64 (0.52-0.79)
GL	0.33 (0.16-0.48)	-	-	0.67 (0.52-0.84)
Dietary Fiber	0.49 (0.34-0.60)	-	-	0.51 (0.40-0.66)
DED	0.12 (0.00-0.39)	-	0.13 (0.00-0.31)	0.75 (0.59-0.89)

The univariate analysis revealed sex differences in genetic and environmental influences on habitual dietary intake. The proportion of variation in dietary intake explained by variation in genes differed between the dietary traits under study but for the majority of dietary variables the genetic influence was 20-50% (table 4, 5 and 6).

In table 4 the results of the univariate analyses of total energy intake, macronutrient energy intake, GI, GL, fiber and DED are shown. The analyses demonstrated that the best-fitting model for

some of the macronutrients were the ADE-model, indicating dominant genetic effect of the intake. For DED the best-fitting model for both sexes was the ACE-model indicating an effect of the common environment.

The univariate analyses of the 20 food groups showed that for some food groups (fruit, poultry, eggs and margarine) no significant genetic influence could be found for men (table 5). In women only the food group consisting of fruit was not genetically influenced (table 6). In both men and women significant effect of the common environment could be demonstrated on a number of food groups.

**Table 5:** Estimates of additive genetic effects ( $a^2$ ), non-additive genetic effects ( $d^2$ ), shared environmental effects ( $c^2$ ) and un-shared environmental effects ( $e^2$ ) on food group intake in men

	$a^2$	$d^2$	$c^2$	$e^2$
Food groups <sup>1</sup>				
Whole grain	0.24 (0.09-0.37)	-	-	0.76 (0.63-0.91)
Refined grain	0.19 (0.03-0.35)	-	-	0.81 (0.65-0.97)
Potatoes	0.68 (0.59-0.74)	-	-	0.32 (0.26-0.41)
Vegetables	0.24 (0.01-0.53)	-	0.40 (0.13-0.59)	0.37 (0.29-0.46)
Fruit	-	-	0.44 (0.35-0.52)	0.56 (0.48-0.65)
Juices	0.36 (0.23-0.48)	-	-	0.64 (0.53-0.77)
Red meat	0.34 (0.19-0.48)	-	-	0.66 (0.52-0.81)
Processed meat	0.47 (0.31-0.59)	-	-	0.53 (0.41-0.69)
Poultry	-	-	0.39 (0.30-0.48)	0.61 (0.52-0.70)
Fish	0.17 (0.01-0.42)	-	0.41 (0.18-0.57)	0.42 (0.34-0.51)
Eggs	-	-	-	1.00 (1.00-1.00)
Low fat milk	0.39 (0.22-0.52)	-	-	0.61 (0.48-0.78)
High fat milk	0.37 (0.23-0.49)	-	-	0.63 (0.51-0.77)
Butter	0.35 (0.22-0.47)	-	-	0.65 (0.53-0.78)
Vegetable oil	0.48 (0.37-0.57)	-	-	0.52 (0.43-0.63)
Margarine	-	-	0.23 (0.11-0.34)	0.77 (0.66-0.89)
Sugar, jam, syrups	0.45 (0.31-0.57)	-	-	0.55 (0.43-0.69)
Soft drinks	0.26 (0.09-0.41)	-	-	0.74 (0.59-0.91)
Candy, chocolate	0.22 (0.00-0.48)	-	0.17 (0.00-0.42)	0.61 (0.50-0.73)
Other	0.63 (0.45-0.75)	-	0.10 (0.01-0.25)	0.27 (0.21-0.35)

<sup>1</sup> Food groups are adjusted for total energy intake

**Table 6:** Estimates of additive genetic effects ( $a^2$ ), non-additive genetic effects ( $d^2$ ), shared environmental effects ( $c^2$ ) and un-shared environmental effects ( $e^2$ ) on food group intake in women

	$a^2$	$d^2$	$c^2$	$e^2$
Food groups <sup>1</sup>				
Whole grain	0.12 (0.00-0.27)	-	-	0.88 (0.73-0.99)
Refined grain	0.20 (0.03-0.37)	-	-	0.80 (0.63-0.97)
Potatoes	0.28 (0.04-0.57)	-	0.34 (0.08-0.54)	0.38 (0.30-0.48)
Vegetables	0.14 (0.00-0.41)	-	0.46 (0.22-0.61)	0.40 (0.32-0.49)
Fruit	-	-	0.41 (0.32-0.50)	0.59 (0.50-0.68)
Juices	0.00 (0.00-0.20)	0.61 (0.37-0.70)	-	0.39 (0.30-0.53)
Red meat	0.33 (0.17-0.47)	-	-	0.67 (0.53-0.83)
Processed meat	0.29 (0.14-0.43)	-	-	0.71 (0.57-0.86)
Poultry	0.38 (0.15-0.58)	-	0.23 (0.07-0.41)	0.39 (0.31-0.49)
Fish	0.61 (0.52-0.69)	-	-	0.39 (0.31-0.48)
Eggs	0.00 (0.00-0.42)	0.56 (0.10-0.66)	-	0.44 (0.34-0.58)
Low fat milk	0.39 (0.23-0.53)	-	-	0.61 (0.47-0.77)
High fat milk	0.32 (0.20-0.44)	-	-	0.68 (0.56-0.80)
Butter	0.42 (0.28-0.55)	-	-	0.58 (0.45-0.72)
Vegetable oil	0.09 (0.00-0.40)	-	0.46 (0.18-0.60)	0.45 (0.36-0.55)
Margarine	0.53 (0.40-0.63)	-	-	0.47 (0.37-0.60)
Sugar, jam, syrups	0.23 (0.07-0.38)	-	-	0.77 (0.62-0.93)
Soft drinks	0.30 (0.18-0.40)	-	-	0.70 (0.60-0.82)
Candy, chocolate	0.27 (0.00-0.55)	-	0.28 (0.05-0.49)	0.45 (0.36-0.57)
Other	-	-	0.37 (0.28-0.46)	0.63 (0.54-0.72)

<sup>1</sup> Food groups are controlled for total energy intake

These results support that dietary intake of macronutrients as well as food groups are influenced by genetic differences between individuals. In addition the study found significant contribution of the shared environment, which is rarely found in studies of adult twins. As in our study, previous studies, that have found effects of the shared environment, have primarily found these effects for fruits and vegetables (72) and sweets (73). In children shared environmental effects have been demonstrated for desserts, fruits and vegetables (74). Our results are therefore in agreement with the few previous studies that suggest that dietary

habits established in childhood are actually maintained to some extent in adulthood.

Sex differences were demonstrated in the magnitude of genetic influence and in the genes underlying food consumption. For some food-groups (juices, poultry, eggs and margarine) genetic effects were found for women, but not for men. These findings are in agreement with previous studies demonstrating sex differences in the heritability of intake of food items (32;73). Fruit intake was not genetically influenced in either men or women.

#### Bread intake and choice

The analyses of bread consumption and bread type as a specific food item among Danish and Finnish men and women demonstrated moderate heritability of total intake of bread as slices per day (table 7).

The results of the study demonstrate that bread intake as well as choice of type of bread are partially under genetic control, whereas the environment shared by the twins seemed not to influence bread intake frequency or bread choice.

**Table 7:** Standardised variance components from the univariate genetic analyses of all traits.

$h^2$  refers to the heritability, i.e. the proportion of variance attributable to genetic effects, while  $e^2$  is the proportion of variance due to environmental effects.

		Men		Women	
		$h^2$	$e^2$	$h^2$	$e^2$
Total bread intake	DK	0.26 (0.10-0.40)	0.74 (0.60-0.90)	0.23 (0.07-0.38)	0.77 (0.62-0.93)
n slices per day	FI	0.37 (0.28-0.45)	0.63 (0.55-0.72)	0.40 (0.33-0.47)	0.60 (0.53-0.67)
White bread	DK	0.24 (0.09-0.38)	0.77 (0.62-0.91)	0.28 (0.12-0.42)	0.72 (0.58-0.88)
n slices per day	FI	0.31 (0.21-0.41)	0.69 (0.59-0.79)	0.30 (0.22-0.38)	0.70 (0.62-0.78)
Rye bread	DK	0.45 (0.30-0.57)	0.56 (0.43-0.70)	0.24 (0.08-0.39)	0.76 (0.61-0.92)
n slices per day	FI	0.41 (0.31-0.49)	0.60 (0.51-0.69)	0.33 (0.25-0.41)	0.67 (0.59-0.75)
Preference for rye bread	FI	0.48 (0.33-0.60)	0.52 (0.40-0.67)	0.29 (0.14-0.43)	0.71 (0.57-0.86)
rye bread / total bread	DK	0.37 (0.27-0.46)	0.63 (0.54-0.73)	0.27 (0.18-0.34)	0.74 (0.66-0.82)

DK: Denmark

FI: Finland

#### Variations in *FTO* and near-*MC4R*

With regards to the analyses of the SNP in the *FTO* gene and the three SNPs near the *MC4R* gene the test for deviation from Hardy Weinberg Equilibrium (HWE) showed that there was no evidence for deviance in the data. The p-values for the HWE-tests were the following:  $p = 0.9$  for *FTO* rs9939609,  $p = 0.34$  for rs12970134,  $p = 0.08$  for rs17700633 and  $p = 0.55$  for rs17782313. The test of linkage pattern between the 3 SNPs near *MC4R* showed that the LD between the loci rs12970134 and rs17782313 was high ( $D' = 0.96$  and  $R$ -square = 0.79).

The only SNP under study for which it was possible to find a significant association with BMI and waist circumference was the near-*MC4R* SNP rs12970134.

The analyses of associations between the polymorphisms and habitual dietary intake demonstrated that the SNPs have some influence on diet. In the analysis of the *FTO* rs9939609 SNP and the nutritional factors and food groups it was not possible to detect significant associations. In addition, no significant effect of sex was demonstrated.

With regard to the associations between the three near-*MC4R* SNPs and the nutritional factors, there were however some indication of a negative association between whole grain intake and the near-*MC4R* SNPs rs12970134 and rs17700633.

Overall no sex differences were demonstrated.

Since multiple testing of the food groups not driven by particular hypotheses derived from the literature is performed, we evaluate the results using Bonferroni correction for mass-significance. For the associations between the SNPs and the food groups to reach

a significant level the cut-point is ( $\alpha$  / number of tests = 0.05 / 20)  $2.5 * 10^{-3}$ . None of the associations reached this significance level.

The results of the study indicate that the obesity-associated SNPs in *FTO* and near-*MC4R* do not influence habitual dietary intake.

The results suggest that individuals with an allele that increases their risk for being overweight avoid food items with high levels of dietary fat and sugar or may underreport intake of certain food items. We did, however, not find an association between the genetic variants and total energy intake. Possible underreporting of specific food items leading to spurious associations with the genetic variants would therefore require over-reporting of other food items.

The results of the models with adjustment for only BMI or total energy intake, or neither BMI or total energy intake showed essentially the same overall results.

#### Genetic influence on the relation between habitual dietary intake and obesity (papers 5 and 6)

##### Specific aims

The last two papers of this thesis focus on the interplay between genes and habitual dietary intake and the relation to obesity. This was initially assessed by analyzing whether intrapair differences in habitual dietary intake was associated with intrapair differences in anthropometry in the MZ twins only in order to control for potential confounding by genetic influence and influence by environmental factors shared by the twins. Secondly, a gene-environment interaction model including protein intake and physical activity was constructed to estimate the possible moderating effect of protein intake and physical activity on mean BMI, waist circumference and body fat mass as well as on the genetic and residual environmental influence on BMI, waist circumference and body fat mass.

##### Results

In order to control for potential confounding by genetic variation and shared environment on the association between habitual diet and body fat, the MZ twin pairs were selected and the association between the intrapair differences in habitual dietary intake and intrapair differences in anthropometry was studied. For the majority of dietary traits no or weak associations were found. The study showed, however, consistent positive associations between intake of sugar-sweetened soft drink and BMI, FMI and waist circumference in men (for more detail see figure 2 in paper 5). Intake of sugar-sweetened soft drinks was, however, not statistically significantly associated with subsequent changes in BMI and waist circumference in men as it has been found in other studies (43).

Exclusion of pairs discordant on obesity, smoking and physical activity did not change the results.

The gene-environment interaction models showed that physical activity was inversely associated with means of BMI, waist circumference and percentage body fat indicating slightly lower mean values in physically active persons. However, the effect was statistically significant only for waist circumference in Danish men and Finnish women and percentage body fat in Danish women. Physical activity also reduced genetic and environmental variances for most of the traits; the exceptions were genetic variance of waist circumference in Danish men and environmental variance of percentage body fat in Danish men and waist circumference in Danish women. Because in most of the cases physical activity decreased genetic variance of BMI and waist circumference but this effect was not, taken individually, always significant, we decided to conduct a meta-analysis of the results from the

two countries to avoid type II error. The overall estimate for modification of genetic variance by physical activity including all four estimates (males and females in Denmark and Finland) was -0.18 (95% CI -0.31, -0.05) for BMI and -0.14 (95% CI -0.22, -0.05) for waist circumference (for more details see table 3 in paper 6). High proportion of protein in the habitual diet was systematically associated with higher mean values of the obesity traits, although most of these effects on the means were not significant (for more details see table 4 in paper 6). In Danish men, we found that high proportion of protein in diet reduced genetic and environmental variances of BMI and waist circumference. In women high proportion of protein in diet increased strongly environmental variance of BMI; however because of high heritability of BMI, the absolute size of the environmental variance is low, which may explain this high point estimate. In Finland the size of the modification effects were lower than in Denmark and generally not significant.

## DISCUSSION

### General findings

#### *Genetic and environmental influence on anthropometry*

The smaller studies examining the common genetic influences on measures of overall fatness and leptin concentration, although using measured phenotypic traits, were usually based on selected groups of subjects with an unclear relation to the background population (75,76). Our estimates of the heritability of fasting serum leptin concentration and its relation to other obesity-related traits were in agreement with the result of the study by Kaprio et al (77), but in contrast to the findings in a previous much smaller study, based on only 33 pairs (75).

FMI and FFMI proved to be fairly highly correlated, both on the phenotypic level and with regard to the underlying genetic and environmental effects. The increased FFMI may be a consequence of the various endocrine and metabolic disturbances that follows development of obesity, and it seems likely that underlying pleiotropic mechanisms may operate at some steps in the pathways from the genes through to the phenotypes.

Previous studies have suggested sex-differences in genetic effects on BMI (78,79). Our results may suggest sex-differences also on other obesity-related traits as well as on the underlying common genetic and environmental influences on these traits. However, as mentioned, the present study is not suitable for a more detailed assessment of sex-differences of the genetic architecture behind the obesity-related traits.

The results of this study indicate that there are significant genetic correlations between different measures of fatness. However, for all the obesity traits included in this study, the genetic correlation was significantly below 1.00, indicating significant residual genetic influence specific for each trait. This stresses the importance of taking the specific obesity-phenotype into account when searching for obesity-related genes. A similar conclusion can be drawn from the analyses of the unique environmental influences.

#### *Genetic and environmental influences on dietary intake*

##### Nutritional factors and food groups

Contrary to expectations a common conclusion in the existing literature is that the shared environment does not contribute to adult eating behavior (73,80-83), but a number of limitations can be identified in the existing studies. In some of the studies the recruitment of study subjects have taken place through word of mouth or newspaper ads (80,83) and not through a population-based twin registry. One study has a rather small sample size (200

MZ and DZ twin pairs) (81). A number of studies focuses on a limited age-range, by either focusing on young children (74), young adults (73), middle-aged or older population (83). One study include only women (72) and one study use a Food Frequency Questionnaire with a limited number of food items (73). Habitual dietary intake appears, as demonstrated, to be controlled by both genetic and environmental factors. Genes influence the chemical senses (taste, smell, and chemical irritation) and large individual differences exist in human sensitivity towards specific compounds (84). However, information on only few specific genes controlling the chemical senses can be found in the existing literature. One prominent example is the inherited ability to taste the bitter synthetic compounds phenylthiocarbamide (PTC) and 6-n-propylthiouracil (PROP) (84). *TAS2R38*, the gene that control PTC taste sensitivity, has three single-nucleotide polymorphisms (A49P, A262V and V296I) that give rise to the major haplotypes PAV (the taster variant) and AVI (the non-taster variant). Environmental factors suggested to have an impact on taste and food preferences in children are cultural differences in weaning practice as well as exposure to tastes through breastfeeding and weaning (85,86).

##### Bread intake and choice

Sex-differences in the heritability of intake of rye bread were found in both Denmark and Finland, indicating that the magnitude of the genetic influences on intake of rye bread were higher in men than in women. The greater heritability estimate for use of rye bread versus use of white bread, especially in men, may be due to genetically determined liability to perceive compounds, such as bitter compounds (87), specific to rye bread. These compounds remain unknown, and should be further pursued using genetically informative study samples. Rye is typically prepared as dry rye crisp bread or as sourdough bread. We did not ask about the specific types of rye bread that were consumed. In Finland and Denmark, rye breads are generally not sweetened unlike in Sweden. Thus, the genetic effects on rye bread choice are not confounded by genetic effects on sweet choice (88).

The study showed genetic effects on overall bread use that were very similar in men and women from Finland and Denmark, whereas the common environment, e.g. acquired bread preferences during childhood, did not seem to play a role. Among men, genetic effects on rye bread consumption were further greater than for other bread variables. This could be due to inherited differences in taste preferences.

##### Variations in *FTO* and near-*MC4R*

Previous studies of the association between the *FTO* SNP and diet have primarily focused on total energy intake (33,34,37,38) or total weight of intake (36) and found positive associations. In the present study we were not able to detect an effect of the *FTO* SNP on total energy intake. A few studies have analyzed the effect of the SNP on dietary energy density (DED) and found no effect (33,37), similar to the findings of the present study. The association to macronutrient intake has also been studied previously, where some studies found no association with macronutrient intake (37,38), and one study found a positive association for intake of dietary fat but not for protein and carbohydrate intake (34). Most of the studies of associations between the *FTO* rs9939609 and dietary intake conducted so far have primarily been conducted in children (33-37). One study, focusing on the *FTO* SNP in a sample of adults, showed a positive association between the *FTO* SNP and total energy intake (38). With regard to associations between the near *MC4R* SNPs and dietary intake,

one study (40) have focused on the SNPs rs17700633 and rs17782313 and found positive associations for rs17782313 and total energy intake, intake of protein and intake of fat. This study did, however, only include women.

In the existing literature the focus have primarily been on total energy intake, dietary energy density and macronutrient intake. In our study we have had the opportunity to analyse a large variety of additional dietary factors including GI, GL, fiber intake and intake of energy from food groups. Including food groups in the analyses gives us an opportunity to study preferences for specific food items, which has not been possible in the existing literature.

The outcome of the study shows that the genetic variants underlying common forms of obesity are not associated with habitual dietary intake.

#### *Genetic influence on the relation between habitual dietary intake and obesity*

When studying intrapair correlations between MZ twins, we eliminated the possible effect of genes and common environment. This study design is a unique opportunity to identify components of dietary intake that are potentially modifiable on the individual level, since any differences between the twins per definition are due to environmental factors that are not shared. The main results of our study are that the association between intake of sugar-sweetened soft drinks is operating within the range of the individually modifiable environmental influences given the possible genetic and preceding family-shared environmental influences. The limited number of associations identified in this study may be because the differences in habitual dietary intake leading to overweight and obesity are so small that they are not possible to detect with the methods of assessment used in the present study (89).

With the well-known methodological reservations in mind, our study suggests that the previously demonstrated association between sugar-sweetened soft drinks and obesity is driven not only by genetic and preceding shared environmental influences but by individually modifiable environment, which enhances the prospect of addressing this behavior as part of a preventive program. We did not find any other similar associations, which of course does not exclude that there may be some, but also in this regard our study is in accordance with the previous literature. It has been suggested that high protein diet may protect against obesity (90,91). However using the gene-environment interaction model, we found that high proportion of protein in diet was associated with higher rather than lower mean BMI, waist circumference and percentage body fat. One possible explanation for this discrepancy is reporting bias, where the heavier twin misreports the habitual dietary intake. It is also noteworthy that while many previous studies have compared very high and low protein diets, our study analyzes protein content in habitual diet.

We found that high proportion of protein in the diet reduced genetic and environmental variance in Danish men, but no evidence for this was found in Danish women or in Finnish men and women. It is, however, noteworthy that protein data in Denmark were based on much more detailed assessment than in Finland, which may explain the more modest effect sizes in the Finnish cohort.

In the study we also found that high physical activity was inversely associated with BMI, waist circumference and percentage body fat. These results are in accordance with previous studies, which have also suggested that physical activity is inversely associated with obesity (92). However, although there is evidence that

physical inactive lifestyle predicts the development of obesity (93), there are also studies suggesting that physical activity does not predict weight change (94) and that obesity may rather lead to low physical activity (95).

The reduction of the specific environmental part of the variation is expected, and it shows that since mean values of the obesity indicators are lower at the higher level of physical activity there is also less variation. Similar decrease in specific environmental variance of BMI and waist circumference was also found in another cohort of Finnish twins in females but not in males (96). However, reduction of genetic part of the variation shows that the reduction in variance has been more similar within MZ than within DZ pairs. This gives evidence that also the expression of genes differs at the different levels of physical activity. A recent study found that the rs9939609 polymorphism of the *FTO* gene increases BMI in sedentary persons but not in physically active persons (97).

The results have certain implications for further efforts to find candidate genes behind obesity. Even though BMI shows high heritability (1), the findings of candidate genes have shown to be difficult to consistently replicate. Genes identified in large genome-wide association studies of large numbers of individuals, including *FTO* variants and the variants near *MC4R*, account for less than 1 % of the variation of BMI in Caucasian populations (24,25). One possibility for these difficulties is that the function of genes varies as a function of environmental exposures and thus can be found only in the presence of certain environments. The results gave some evidence of this since genetic variation was higher in sedentary than physically active persons. This suggests that selecting sedentary persons for further studies of genetics of obesity as well as more careful assessment of physical activity would increase power to find candidate genes.

It is noteworthy that even when we have treated physical activity and proportion of protein in diet as environmental moderators in the statistical models, they are not pure environmental factors. Previous studies have shown that there is a moderate genetic component both behind dietary energy intake (73) and physical exercise (98). The moderation effect of physical activity on genetic variation of phenotypes is, however, more likely to be because of phenotypic moderation than gene-gene interactions, even when our model cannot exclude this possibility.

#### *Design*

The studies in the present thesis are primarily based on a relatively large population-based twin sample (GEMINAKAR) identified through registers, with a wide age range (18-67 years), with detailed assessment of habitual dietary intake and anthropometry, and with DNA-based zygosity determination. The precision of the direct measurements is phenotype dependent, which will be discussed in detail in the following sections. In attempt to remove the confounding of disease status and/or medical treatment on the intra-class correlations, we excluded subjects having diagnosed diabetes or cardiovascular disease.

The anthropometric measures collected at the same time as the dietary assessment were all objectively measured. The good quality of the assessment of the anthropometric measures is a major strength of the study since it has been demonstrated that using self-reported data on height and weight introduces bias due to underreporting (10,99). In the existing literature, studies have shown lower repeatability of skin fold measurements than for other anthropometric measurements, which may lead to overestimation of the nonshared environment, and conversely, underes-

timation of the genetic influence on skin folds (100,101). In the GEMINAKAR it has been attempted to reduce the measurement error by repeating the measure at each skin fold site three times and subsequently used the average of three measures in the analyses. In general, however, biases in the various estimates due to errors of measurements, which may also be correlated, cannot be excluded although we expect them to be of minimal influence. The data from the Omnibus-studies that were used for calculation of change in weight and waist circumference before and after dietary assessments were self-reported. This might explain why the change in BMI after dietary assessment was lower than expected.

The dietary information collected for the study is based on an extensive food-frequency questionnaire (FFQ). The information from the FFQ is however based on self-report. This is the major limitation of the study, since previous studies have demonstrated that subjects tend to underreport their food intake in general as well as differential with regard to which foods are reported on and that the degree of underreporting increases with the degree of overweight (102,103). This introduces a systematic bias, resulting in the paradoxical observation that obese individuals appear to eat less than lean people. This might also be the case for the women in the present study, since negative associations between anthropometry and total energy intake as well as intake of candy, chocolate and snacks were found in women.

While recognizing that the food groupings are essentially based on narrative interpretation of the existing literature on food groupings and current knowledge about the individual food items assigned to each group, an empirical statistical justification of the groupings may be considered in future studies. Thus, a pattern analyses (using factor analysis, principal component analysis or cluster analysis) (104) coping with the mutual correlations between the various aspects of the food intake may be considered, and can possibly lead to indications of aspects of the food intake that might share genetic and environmental determinants. However, the main problem is that the results of such analyses cannot be translated to identifiable food intake at the individual level. We have therefore not included this type of analysis in the present thesis, but consider it a future option in relation to the perspectives regarding the mechanisms behind the findings.

## CONCLUSIONS

Based on the results of this Ph.D. thesis it can be concluded that:

- There are significant genetic correlations between different measures of fatness. However, for all the obesity traits included in this study, the genetic correlation was significantly below 1.00, indicating significant residual genetic influence specific for each traits. This stresses the importance of taking the specific obesity-phenotype into account when searching for obesity-related genes.
- Almost half of the variation in most components of dietary intake in adults appears to be heritable. Environmental factors shared by siblings, e.g. childhood family environment and common friends, also influence intake of certain food groups. While allowing for the inclusion of measurement errors, it appears that for all components of dietary intake the non-shared environment has a considerable influence. In a public health perspective this indicates that there exists – despite the genetic in-

fluence – a potential for establishing healthy eating habits both at individual and family levels.

- The genetic effects on overall bread use were moderate in men and women from Finland and Denmark, whereas acquired bread preferences during childhood did not seem to play a role. Among men, genetic effects on rye bread consumption were further greater than for other bread types.
- Polymorphisms in *FTO* and near-*MC4R* do not have a role in the regulation of habitual dietary intake.
- Only few associations between habitual dietary intake and anthropometry could be found, even when controlling for genetic background. Our study does, however, suggest that the previously demonstrated association between sugar-sweetened soft drinks and obesity is driven not only by genetic and preceding shared environmental influences but by individually modifiable environment, which enhances the prospect of addressing this behavior as part of a preventive program
- While high physical activity is associated with a down-regulation of the effect of genes predisposing to obesity, such effects were not found for protein intake. Our results highlight the importance of investigating the role of regular physical activity in the maintenance of healthy weight, especially in persons with genetic susceptibility to obesity.

The findings presented in this thesis may contribute to a better understanding of the genetic and environmental contribution to dietary habits, anthropometry, and the interplays between genes, diet and obesity.

## PERSPECTIVES AND FUTURE RESEARCH TOPICS

Many questions remain for which this study cannot provide answers due to lack of suitable data or other limitations. First, the longitudinal data on changes in anthropometry after the dietary assessment used in paper 5 is based on self-reported data on weight and waist circumference. A thorough follow-up of the GEMINAKAR study will be initiated in the autumn/winter 2009. This follow-up will include objective measures of weight, waist circumference, hip circumference and body fat assessment by bio-impedance. Obtaining these high-quality data will provide excellent opportunities to analyze the interplay between habitual dietary intake and genes in relation to longitudinal changes in weight and waist circumference.

There is a need for population-based studies of gene-diet interaction in relation to obesity, including genotyping for known obesity associated genes. This type of research is currently ongoing in the large EU-funded integrated research project Diogenes (Diet, Obesity and Genes) research line 3 (Contract no. FP6-513946). In this project EPIC-cohorts (European Prospective Investigation into Cancer and Nutrition) from five European countries participate in studies of gene-diet interaction analyses of weight change among adults. The results of these studies will be published in the coming years.

Generally, the knowledge about the possible mechanisms behind the genetic and environmental influences on the habitual food intake is far from coherent and comprehensive. Studies to identify compounds in food items making individuals genetically susceptible to eat certain food items are needed and our findings also encourage investment in nutrigenetics and nutrigenomics to

enhance our understanding how genes and environment may determine the food intake profile.

With the increasing knowledge of the genetic influences on body weight and body composition, the ongoing challenge will be to identify the genes responsible for the genetic variance and map their functions. Currently, a great challenge lies in finding the function of the SNPs identified through genome-wide-association studies of obesity.

#### ABBREVIATIONS

A	additive genetic effects, cumulative effects of several genes ( $A = a_2$ )
$a_2$	proportion of variance explained by additive genetic effects
ACE model	model with additive genetic effects and shared and unique environmental effects
ADE model	model with additive and dominant genetic effects and unique environmental effects
AE model	model with additive genetic effects and unique environmental effects
BMI	body mass index
C	shared environmental effects ( $C = c_2$ )
$c_2$	proportion of variance explained by shared environmental effects
CE model	model with shared and unique environmental effects
$\chi^2$	chi-squared
CI	confidence interval
cova	covariance explained by additive genetic effects
covc	covariance explained by shared environmental effects
covd	covariance explained by dominant genetic effects
cove	covariance explained by unique environmental effects
D	dominant genetic effects, allelic interaction within a loci ( $D = d_2$ )
$d_2$	proportion of variance explained by dominant genetic effects
DED	dietary energy density
Diogenes	EU 6th framework programme (Diet, Obesity and Genes)
DA	diamniotic
DC	dichorionic
DZ	dizygotic
DZOS:	dizygotic opposite sex
E	unique environmental effects ( $E = e_2$ )
$e_2$	proportion of variance explained by unique environmental effects
E%	energy percentage
FFQ	food frequency questionnaire
FFMI	fat free mass index
FMI	fat mass index
GI	glycemic index
GL	glycemic load
GWAS	genome-wide association study
GxE	gene-environment interaction
$h^2$	heritability (proportion of variance explained by additive genetic effects, $h^2 = a_2$ )
MA	monoamniotic
MC	monochorionic
MZ	monozygotic
r	correlation coefficient
$r_A$	correlation between additive genetic effects
$r_C$	correlation between shared environmental effects
$r_E$	correlation between unique environmental effects
$r_P$	phenotypic correlation
SNP	single nucleotide polymorphisms
WHO	World Health Organisation

#### SUMMARY

Obesity has become a major health concern due to the increased risk of co-morbidities, resulting in decreased quality of life, stigmatization, reduced working ability and early death. This causes a great challenge for the health care systems and results in increased direct costs related to treatment of obesity and co-morbidities, as well as increased indirect costs related to reduced function and withdrawal from the labour market.

Both between and within societies, large variation in the prevalence of overweight and obesity exists. This variation is caused by differences in environmental exposures as well as genetic differences between individuals, resulting in differentiated susceptibility to environmental exposures. The evidence for genetic influence on anthropometry has previously been established and has been estimated to be 60-70% based on twin studies. These inter-individual differences can, however, not explain the increase in obesity prevalence during the past 70 years. Environmental factors must therefore play an important role in the obesity epidemic. Habitual diet is one of many environmental factors that potentially contribute to the inter-individual differences in body fat mass, but only limited evidence for associations between habitual dietary intake and anthropometry exists.

Differences in habitual dietary intake are also partly determined by differences in genes influencing smell and taste preferences.

But, so far, only few studies have investigated genetic influences on dietary intake in adults and the interplay between diet, genes and obesity.

The focus of the thesis was to investigate the genetic and environmental influence on habitual diet and obesity as well as the association between habitual diet and anthropometry. The thesis is based on structural equation modelling of twin data from the Danish Twin Registry with special focus on the GEMINAKAR twin study that was performed in 1997-2000. In this study, anthropometric traits of the twin pairs were measured and habitual dietary intake was assessed through a food frequency questionnaire (FFQ).

When studying body fat mass in population-based studies, the phenotype used is often the body mass index (BMI). This measure does, however, not specify whether excess body mass is due to excess fat mass and how the body fat is distributed. Studying the genetic and environmental correlations between the anthropometry measures in the GEMINAKAR sample showed that the genetic correlations between BMI, fat mass index (FMI) and waist circumference were high in men and that the genetic correlations between BMI, FMI, waist and hip circumference were high in women. For all anthropometric phenotypes, significant residual genetic influence existed.

Based on information about habitual diet from the FFQ the genetic influence on total energy intake, macronutrient intake, as well as intake of energy from 20 food groups, was estimated. The proportion of variation in dietary intake explained by variation in genes differed between the dietary traits under study but for the majority of dietary variables the genetic influence was 20-50%. Accordingly, both diet and anthropometry is influenced by genetic variation. In order to control for potential confounding by genetic variation and shared environment on the association between habitual diet and body fat, the monozygotic twin pairs were selected and the associations between intrapair differences in dietary intake and intrapair differences in anthropometry were studied. For the majority of dietary traits, no associations or only weak associations were found. The study showed, however, consistent positive associations between intake of sugar-sweetened soft drink and BMI, FMI and waist circumference in men.

Gene-environment interaction models showed that while high physical activity is associated with a down-regulation of genes predisposing to obesity, such effects were not found for protein intake.

In conclusion, the studies included in this thesis contribute to the relatively limited existing literature, with insight into genetic determinants of habitual dietary intake, pleiotropic influences on anthropometry, and the interplay between diet, genes and obesity.

#### REFERENCES

1. Mokdad AH, Ford ES, Bowman BA, Dietz WH, Vinicor F, Bales VS, et al. Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. *JAMA* 2003 Jan 1;289(1):76-9.
2. Due P, Heitmann BL, Sorensen TI. Prevalence of obesity in Denmark. *Obes Rev* 2007 May;8(3):187-9.
3. Jorgensen ME, Glumer C, Bjerregaard P, Gyntelberg F, Jorgensen T, Borch-Johnsen K. Obesity and central fat pattern among Greenland Inuit and a general population of Denmark (Inter99): relationship to metabolic risk

- factors. *Int J Obes Relat Metab Disord* 2003 Dec;27(12):1507-15.
4. Larsson B, Svarsdudd K, Welin L, Wilhelmsen L, Bjorn-torp P, Tibblin G. Abdominal adipose tissue distribution, obesity, and risk of cardiovascular disease and death: 13 year follow up of participants in the study of men born in 1913. *Br Med J (Clin Res Ed)* 1984 May 12;288(6428):1401-4.
  5. Bigaard J, Tjonneland A, Thomsen BL, Overvad K, Heitmann BL, Sorensen TI. Waist circumference, BMI, smoking, and mortality in middle-aged men and women. *Obes Res* 2003 Jul;11(7):895-903.
  6. Obesity: preventing and managing the global epidemic. Report of a WHO consultation. *World Health Organ Tech Rep Ser* 2000;894:i-253.
  7. Kolotkin RL, Meter K, Williams GR. Quality of life and obesity. *Obes Rev* 2001 Nov;2(4):219-29.
  8. Pedersen KM, Worre-Jensen AL. [Adiposity and Health Economics.]. *Ugeskr Laeger* 2006 Jan 9;168(2):196-201.
  9. Worre-Jensen AL, Jensen NB, Heitmann BL, Sorensen TI. [The cost of obesity on the Danish health care system]. *Ugeskr Laeger* 2007 Aug 13;169(33):2634-7.
  10. Bostrom G, Diderichsen F. Socioeconomic differentials in misclassification of height, weight and body mass index based on questionnaire data. *Int J Epidemiol* 1997 Aug;26(4):860-6.
  11. Maes HH, Neale MC, Eaves LJ. Genetic and environmental factors in relative body weight and human adiposity. *Behav Genet* 1997 Jul;27(4):325-51.
  12. Schousboe K, Visscher PM, Erbas B, Kyvik KO, Hopper JL, Henriksen JE, et al. Twin study of genetic and environmental influences on adult body size, shape, and composition. *Int J Obes Relat Metab Disord* 2004 Jan;28(1):39-48.
  13. Souren NY, Paulussen AD, Loos RJ, Gielen M, Beunen G, Fagard R, et al. Anthropometry, carbohydrate and lipid metabolism in the East Flanders Prospective Twin Survey: heritabilities. *Diabetologia* 2007 Oct;50(10):2107-16.
  14. Lajunen HR, Kaprio J, Keski-Rahkonen A, Rose RJ, Pulkkinen L, Rissanen A, et al. Genetic and environmental effects on body mass index during adolescence: a prospective study among Finnish twins. *Int J Obes (Lond)* 2009 May;33(5):559-67.
  15. Poulsen P, Kyvik KO, Vaag A, Beck-Nielsen H. Heritability of type II (non-insulin-dependent) diabetes mellitus and abnormal glucose tolerance--a population-based twin study. *Diabetologia* 1999 Feb;42(2):139-45.
  16. Rose KM, Newman B, Mayer-Davis EJ, Selby JV. Genetic and behavioral determinants of waist-hip ratio and waist circumference in women twins. *Obes Res* 1998 Nov;6(6):383-92.
  17. Selby JV, Newman B, Quesenberry CP, Jr., Fabsitz RR, King MC, Meaney FJ. Evidence of genetic influence on central body fat in middle-aged twins. *Hum Biol* 1989 Apr;61(2):179-94.
  18. Selby JV, Newman B, Quesenberry CP, Jr., Fabsitz RR, Carmelli D, Meaney FJ, et al. Genetic and behavioral influences on body fat distribution. *Int J Obes* 1990 Jul;14(7):593-602.
  19. Arden NK, Spector TD. Genetic influences on muscle strength, lean body mass, and bone mineral density: a twin study. *J Bone Miner Res* 1997 Dec;12(12):2076-81.
  20. Bouchard C, Perusse L, Leblanc C, Tremblay A, Theriault G. Inheritance of the amount and distribution of human body fat. *Int J Obes* 1988;12(3):205-15.
  21. Carey DG, Nguyen TV, Campbell LV, Chisholm DJ, Kelly P. Genetic influences on central abdominal fat: a twin study. *Int J Obes Relat Metab Disord* 1996 Aug;20(8):722-6.
  22. Perusse L, Despres JP, Lemieux S, Rice T, Rao DC, Bouchard C. Familial aggregation of abdominal visceral fat level: results from the Quebec family study. *Metabolism* 1996 Mar;45(3):378-82.
  23. Rice T, Daw EW, Gagnon J, Bouchard C, Leon AS, Skinner JS, et al. Familial resemblance for body composition measures: the HERITAGE Family Study. *Obes Res* 1997 Nov;5(6):557-62.
  24. Thorleifsson G, Walters GB, Gudbjartsson DF, Steinthorsdottir V, Sulem P, Helgadóttir A, et al. Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. *Nat Genet* 2009 Jan;41(1):18-24.
  25. Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, Heid IM, et al. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat Genet* 2009 Jan;41(1):25-34.
  26. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 2007 May 11;316(5826):889-94.
  27. Loos RJ, Lindgren CM, Li S, Wheeler E, Zhao JH, Prokopenko I, et al. Common variants near MC4R are associated with fat mass, weight and risk of obesity. *Nat Genet* 2008 Jun;40(6):768-75.
  28. Bell CG, Walley AJ, Froguel P. The genetics of human obesity. *Nat Rev Genet* 2005 Feb 10.
  29. Chambers JC, Elliott P, Zabaneh D, Zhang W, Li Y, Froguel P, et al. Common genetic variation near MC4R is associated with waist circumference and insulin resistance. *Nat Genet* 2008 Jun;40(6):716-8.
  30. Keller KL, Pietrobelli A, Faith MS. Genetics of food intake and body composition: lessons from twin studies. *Acta Diabetol* 2003 Oct;40 Suppl 1:S95-100.
  31. Reed DR, Bachmanov AA, Beauchamp GK, Tordoff MG, Price RA. Heritable variation in food preferences and their contribution to obesity. *Behav Genet* 1997 Jul;27(4):373-87.
  32. de Castro JM. Genes and environment have gender-independent influences on the eating and drinking of free-living humans. *Physiol Behav* 1998 Feb 1;63(3):385-95.
  33. Johnson L, van Jaarsveld CH, Emmett PM, Rogers IS, Ness AR, Hattersley AT, et al. Dietary energy density affects fat mass in early adolescence and is not modified by FTO variants. *PLoS ONE* 2009;4(3):e4594.
  34. Timpson NJ, Emmett PM, Frayling TM, Rogers I, Hattersley AT, McCarthy MI, et al. The fat mass- and obesity-associated locus and dietary intake in children. *Am J Clin Nutr* 2008 Oct;88(4):971-8.
  35. Wardle J, Carnell S, Haworth CM, Farooqi IS, O'Rahilly S, Plomin R. Obesity associated genetic variation in FTO is associated with diminished satiety. *J Clin Endocrinol Metab* 2008 Sep;93(9):3640-3.

36. Wardle J, Llewellyn C, Sanderson S, Plomin R. The FTO gene and measured food intake in children. *Int J Obes (Lond)* 2009 Jan;33(1):42-5.
37. Cecil JE, Tavendale R, Watt P, Hetherington MM, Palmer CN. An obesity-associated FTO gene variant and increased energy intake in children. *N Engl J Med* 2008 Dec 11;359(24):2558-66.
38. Speakman JR, Rance KA, Johnstone AM. Polymorphisms of the FTO gene are associated with variation in energy intake, but not energy expenditure. *Obesity (Silver Spring)* 2008 Aug;16(8):1961-5.
39. Cecil JE, Tavendale R, Watt P, Hetherington MM, Palmer CN. An obesity-associated FTO gene variant and increased energy intake in children. *N Engl J Med* 2008 Dec 11;359(24):2558-66.
40. Qi L, Kraft P, Hunter DJ, Hu FB. The common obesity variant near MC4R gene is associated with higher intakes of total energy and dietary fat, weight change and diabetes risk in women. *Hum Mol Genet* 2008 Nov 15;17(22):3502-8.
41. Bray GA, Paeratakul S, Popkin BM. Dietary fat and obesity: a review of animal, clinical and epidemiological studies. *Physiol Behav* 2004 Dec 30;83(4):549-55.
42. Togo P. Food intake patterns and development of obesity University of Copenhagen; 2003.
43. Malik VS, Schulze MB, Hu FB. Intake of sugar-sweetened beverages and weight gain: a systematic review. *Am J Clin Nutr* 2006 Aug;84(2):274-88.
44. Perusse L, Bouchard C. Gene-diet interactions in obesity. *Am J Clin Nutr* 2000 Nov;72(5 Suppl):1285S-90S.
45. Schousboe K. Body Composition and Glucose Homeostasis: Genes and Environment Faculty of Health Sciences, University of Southern Denmark; 2003.
46. Heitmann BL. Evaluation of body fat estimated from body mass index, skinfolds and impedance. A comparative study. *Eur J Clin Nutr* 1990 Nov;44(11):831-7.
47. Overvad K, Tjonnelland A, Haraldsdottir J, Ewertz M, Jensen OM. Development of a semiquantitative food frequency questionnaire to assess food, energy and nutrient intake in Denmark. *Int J Epidemiol* 1991 Dec;20(4):900-5.
48. Tjonnelland A, Overvad K, Haraldsdottir J, Bang S, Ewertz M, Jensen OM. Validation of a semiquantitative food frequency questionnaire developed in Denmark. *Int J Epidemiol* 1991 Dec;20(4):906-12.
49. Lauritsen J. FoodCalc version 1.3. Available from: [www.ibt.dk/jesper/foocalc](http://www.ibt.dk/jesper/foocalc) 2007 Dec 1.
50. Møller A, Saxholt E. Levnedsmiddeltabeller (Food composition tables). 4 ed. Copenhagen: Levnedsmiddelstyrelsen (National Food Agency); 1996.
51. Iqbal SI, Helge JW, Heitmann BL. Do energy density and dietary fiber influence subsequent 5-year weight changes in adult men and women? *Obesity (Silver Spring)* 2006 Jan;14(1):106-14.
52. Halkjaer J, Tjonnelland A, Overvad K, Sorensen TI. Dietary predictors of 5-year changes in waist circumference. *J Am Diet Assoc* 2009 Aug;109(8):1356-66.
53. Blickstein I, Keith LG. Multiple Pregnancy: Epidemiology, Gestation, and Perinatal Outcome, Second Edition. Informa HealthCare; 2005.
54. Hall JG. Twinning. *Lancet* 2003 Aug 30;362(9385):735-43.
55. Benirschke K. The biology of the twinning process: how placentation influences outcome. *Semin Perinatol* 1995 Oct;19(5):342-50.
56. Loos RJ, Beunen G, Fagard R, Derom C, Vlietinck R. The influence of zygosity and chorion type on fat distribution in young adult twins consequences for twin studies. *Twin Res* 2001 Oct;4(5):356-64.
57. Ramos-Arroyo MA, Ulbright TM, Yu PL, Christian JC. Twin study: relationship between birth weight, zygosity, placentation, and pathologic placental changes. *Acta Genet Med Gemellol (Roma)* 1988;37(3-4):229-38.
58. Kyvik KO. Generalisability and assumptions of twin studies. In: Spector TD, Snieder H, Macgregor AJ, editors. *Advances in twin and sib-pair analysis*. London: Greenwich Medical Media Ltd; 2000.
59. Christensen K, Vaupel JW, Holm NV, Yashin AI. Mortality among twins after age 6: fetal origins hypothesis versus twin method. *BMJ* 1995 Feb 18;310(6977):432-6.
60. Boomsma D, Busjahn A, Peltonen L. Classical twin studies and beyond. *Nat Rev Genet* 2002 Nov;3(11):872-82.
61. Plomin R, DeFries JC, McClearn GE, Rutter M. *Behavior Genetics*. New York: Freeman WH and Co.; 1997.
62. Neale MC, Boker SM, Xie G, Maes H. *Mx: Statistical Modeling*. 6 ed. Richmond: 2004.
63. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, et al. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A* 2005 Jul 26;102(30):10604-9.
64. Derom R, Derom C, Vlietinck R. Placentation. In: Keith L, Papiernik E, Keith D, Luke B, editors. *Multiple Pregnancy: Epidemiology, Gestation & Perinatal Outcome*. London: The Parthenon Publishing Group; 1995.
65. Sebire NJ, Snijders RJ, Hughes K, Sepulveda W, Nicolaides KH. The hidden mortality of monochorionic twin pregnancies. *Br J Obstet Gynaecol* 1997 Oct;104(10):1203-7.
66. Burn J, Corney G. Zygosity determination and the types of twins. In: MacGillivray I, Campbell DM, Thompson B, editors. *Twinning and Twins*. Chichester: John Wiley & Sons; 1988.
67. Falkner F, Matheny A. The long-term development of twins: Anthropometric factors and cognition. In: Keith L, Papiernik E, Keith D, Luke B, editors. *Multiple Pregnancy: Epidemiology, Gestation & Perinatal Outcome*. London: The Parthenon Publishing Group; 1995.
68. Jaffee SR, Price TS. Gene-environment correlations: a review of the evidence and implications for prevention of mental illness. *Mol Psychiatry* 2007 May;12(5):432-42.
69. Korkeila M, Kaprio J, Rissanen A, Koskenvuo M. Effects of gender and age on the heritability of body mass index. *Int J Obes* 1991 Oct;15(10):647-54.
70. Neale MC, Cardon LR. *Methodology for genetic studies of twins and families*. 1992. Dordrecht, Kluwer. Ref Type: Generic
71. Purcell S. Variance components models for gene-environment interaction in twin analysis. *Twin Res* 2002 Dec;5(6):554-71.
72. Teucher B, Skinner J, Skidmore PM, Cassidy A, Fairweather-Tait SJ, Hooper L, et al. Dietary Patterns and Heritability of Food Choice in a UK Female Twin Cohort. *Twin Res Hum Genet* 2007 Oct;10(5):734-48.

73. Keskitalo K, Silventoinen K, Tuorila H, Perola M, Pietilainen KH, Rissanen A, et al. Genetic and environmental contributions to food use patterns of young adult twins. *Physiol Behav* 2008 Jan 28;93(1-2):235-42.
74. Breen FM, Plomin R, Wardle J. Heritability of food preferences in young children. *Physiol Behav* 2006 Jul 30;88(4-5):443-7.
75. Narkiewicz K, Szczech R, Winnicki M, Chrostowska M, Pawlowski R, Lysiak-Szydłowska W, et al. Heritability of plasma leptin levels: a twin study. *J Hypertens* 1999 Jan;17(1):27-31.
76. Jordan J, Brabant G, Brinsuk M, Tank J, Horn R, Luft FC, et al. Heritability of free and receptor-bound leptin in normal twins. *Am J Physiol Regul Integr Comp Physiol* 2005 May;288(5):R1411-R1416.
77. Kaprio J, Eriksson J, Lehtovirta M, Koskenvuo M, Tuomilehto J. Heritability of leptin levels and the shared genetic effects on body mass index and leptin in adult Finnish twins. *Int J Obes Relat Metab Disord* 2001 Jan;25(1):132-7.
78. Korkeila M, Kaprio J, Rissanen A, Koskenvuo M. Effects of gender and age on the heritability of body mass index. *Int J Obes* 1991 Oct;15(10):647-54.
79. Schousboe K, Willemsen G, Kyvik KO, Mortensen J, Boomsma DI, Cornes BK, et al. Sex differences in heritability of BMI: a comparative study of results from twin studies in eight countries. *Twin Res* 2003 Oct;6(5):409-21.
80. de Castro JM. Genetic influences on daily intake and meal patterns of humans. *Physiol Behav* 1993 Apr;53(4):777-82.
81. Heller RF, O'Connell DL, Roberts DC, Allen JR, Knapp JC, Steele PL, et al. Lifestyle factors in monozygotic and dizygotic twins. *Genet Epidemiol* 1988;5(5):311-21.
82. Hur YM, Bouchard TJ, Jr., Eckert E. Genetic and environmental influences on self-reported diet: a reared-apart twin study. *Physiol Behav* 1998 Jul;64(5):629-36.
83. van den Bree MB, Eaves LJ, Dwyer JT. Genetic and environmental influences on eating patterns of twins aged  $\geq 50$  y. *Am J Clin Nutr* 1999 Oct;70(4):456-65.
84. Tepper BJ. Nutritional Implications of Genetic Taste Variation: The Role of PROP Sensitivity and Other Taste Phenotypes. *Annu Rev Nutr* 2008 Apr 11.
85. Harris G. Development of taste and food preferences in children. *Curr Opin Clin Nutr Metab Care* 2008 May;11(3):315-9.
86. Wardle J, Cooke L. Genetic and environmental determinants of children's food preferences. *Br J Nutr* 2008 Feb;99 Suppl 1:S15-S21.
87. Heinio R-L, Liukkonen K-H, Myllymaki O, Pihlava J-M, Adlercreutz H, Heinonen L, et al. Quantities of phenolic compounds and their impacts on the perceived flavour attributes of rye grain. *Journal of Cereal Science* 2008;47(3):566-75.
88. Keskitalo K, Knaapila A, Kallela M, Palotie A, Wessman M, Sammalisto S, et al. Sweet taste preferences are partly genetically determined: identification of a trait locus on chromosome 16. *Am J Clin Nutr* 2007 Jul;86(1):55-63.
89. Sorensen TI. Conference on "Multidisciplinary approaches to nutritional problems". Symposium on "Diabetes and health". Challenges in the study of causation of obesity. *Proc Nutr Soc* 2009 Feb;68(1):43-54.
90. Krieger JW, Sitren HS, Daniels MJ, Langkamp-Henken B. Effects of variation in protein and carbohydrate intake on body mass and composition during energy restriction: a meta-regression 1. *Am J Clin Nutr* 2006 Feb;83(2):260-74.
91. Merchant AT, Anand SS, Vuksan V, Jacobs R, Davis B, Teo K, et al. Protein intake is inversely associated with abdominal obesity in a multi-ethnic population. *J Nutr* 2005 May;135(5):1196-201.
92. Jebb SA, Moore MS. Contribution of a sedentary lifestyle and inactivity to the etiology of overweight and obesity: current evidence and research issues. *Med Sci Sports Exerc* 1999 Nov;31(11 Suppl):S534-S541.
93. Pietilainen KH, Kaprio J, Borg P, Plasqui G, Yki-Jarvinen H, Kujala UM, et al. Physical inactivity and obesity: a vicious circle. *Obesity (Silver Spring)* 2008 Feb;16(2):409-14.
94. Luke A, Dugas LR, Ebersole K, Durazo-Arvizu RA, Cao G, Schoeller DA, et al. Energy expenditure does not predict weight change in either Nigerian or African American women. *Am J Clin Nutr* 2009 Jan;89(1):169-76.
95. Mortensen LH, Siegler IC, Barefoot JC, Gronbaek M, Sorensen TI. Prospective associations between sedentary lifestyle and BMI in midlife. *Obesity (Silver Spring)* 2006 Aug;14(8):1462-71.
96. Mustelin L, Silventoinen K, Pietilainen K, Rissanen A, Kaprio J. Physical activity reduces the influence of genetic effects on BMI and waist circumference: a study in young adult twins. *Int J Obes (Lond)* 2009 Jan;33(1):29-36.
97. Andreasen CH, Stender-Petersen KL, Mogensen MS, Torekov SS, Wegner L, Andersen G, et al. Low physical activity accentuates the effect of the FTO rs9939609 polymorphism on body fat accumulation. *Diabetes* 2008 Jan;57(1):95-101.
98. Stubbe JH, Boomsma DI, Vink JM, Cornes BK, Martin NG, Skytthe A, et al. Genetic influences on exercise participation in 37,051 twin pairs from seven countries. *PLoS ONE* 2006;1:e22.
99. Gorber SC, Tremblay M, Moher D, Gorber B. A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. *Obes Rev* 2007 Jul;8(4):307-26.
100. Hansen S, Cold S, Petersen PH, Rose C. Estimates of the sources of variation (variance components) of bioelectric impedance and anthropometric measurements in an epidemiological case-control study of breast cancer. *Eur J Clin Nutr* 1997 Nov;51(11):764-70.
101. Nordhamn K, Sodergren E, Olsson E, Karlstrom B, Vessby B, Berglund L. Reliability of anthropometric measurements in overweight and lean subjects: consequences for correlations between anthropometric and other variables. *Int J Obes Relat Metab Disord* 2000 May;24(5):652-7.
102. Heitmann BL, Lissner L. Dietary underreporting by obese individuals--is it specific or non-specific? *BMJ* 1995 Oct 14;311(7011):986-9.
103. Westterterp KR, Goris AH. Validity of the assessment of dietary intake: problems of misreporting. *Curr Opin Clin Nutr Metab Care* 2002 Sep;5(5):489-93.
104. Togo P, Osler M, Sorensen TI, Heitmann BL. Food intake patterns and body mass index in observational studies. *Int J Obes Relat Metab Disord* 2001 Dec;25(12):1741-51.

