Genetic and environmental effects on body mass index from infancy to the onset of adulthood: an individual-based pooled analysis of 45 twin cohorts participating in the COLaborative project of Development of Anthropometrical measures in Twins (CODATwins) study 1–3

Environmental factors shared by co-twins affect BMI variation remained roughly similar across different regions. East Asia, but the relative proportion of genetic variation to total BMI was highest in North America and Australia and lowest in developing countries (2). Previous twin and family studies have found shared obesogenic environments.

INTRODUCTION

Childhood obesity is a major public health problem throughout the world. In the United States, >30% of children and adolescents were classified as overweight or obese in 2011–2012 (1), and childhood obesity is also a growing problem in many developing countries (2). Previous twin and family studies have shown that both genetic and environmental factors contribute to obesity. As early as 1923, the tendency toward obesity was found to vary between families, suggesting a role of genetic factors (3), and a meta-analysis of 31 twin studies showed that for adults, the heritability estimates of BMI (in kg/m²), i.e., total BMI variation explained by genetic variation, ranged from 47% to 80% (4). However, much less is known about the variation of the genetic architecture of BMI during childhood and adolescence. A meta-analysis of 9 twin studies found that the environmental factors shared by co-twins contributed to BMI in infancy and early childhood but were not evident after mid-childhood, when genetic factors become more important (5). An individual-based analysis of 4 twin cohorts found shared environmental contributions to BMI from 3 to 8 y of age, which disappeared at 9–19 y of age (6). Somewhat different results were found in a Finnish longitudinal study, which showed that shared environmental affect BMI at 11–12 and 14 y of age but was no longer evident at 17 y of age (7). Thus, previous twin studies suggest that the effect of shared environmental factors

ABSTRACT

Background: Both genetic and environmental factors are known to affect body mass index (BMI), but detailed understanding of how their effects differ during childhood and adolescence is lacking. Objectives: We analyzed the genetic and environmental contributions to BMI variation from infancy to early adulthood and the ways they differ by sex and geographic regions representing high (North America and Australia), moderate (Europe), and low levels (East Asia) of obesogenic environments. Design: Data were available for 87,782 complete twin pairs from 0.5 to 19.5 y of age from 45 cohorts. Analyses were based on 383,092 BMI measurements. Variation in BMI was decomposed into genetic and environmental components through genetic structural equation modeling. Results: The variance of BMI increased from 5 y of age along with increasing mean BMI. The proportion of BMI variation explained by additive genetic factors was lowest at 4 y of age in boys ($a^2 = 0.42$) and girls ($a^2 = 0.41$) and then generally increased to 0.75 in both sexes at 19 y of age. This was because of a stronger influence of environmental factors shared by co-twins in midchildhood. After 15 y of age, the effect of shared environment was not observed. The sex-specific expression of genetic factors was seen in infancy, but was most prominent at 13 y of age and older. The variance of BMI was highest in North America and Australia and lowest in East Asia, but the relative proportion of genetic variation to total variation remained roughly similar across different regions. Conclusions: Environmental factors shared by co-twins affect BMI in childhood, but little evidence for their contribution was found in late adolescence. Our results suggest that genetic factors play a major role in the variation of BMI in adolescence among populations of different ethnicities exposed to different environmental factors related to obesity. Am J Clin Nutr doi: 10.3945/ajcn.116.130252.

Keywords: BMI, children, genetics, international comparisons, twins
influencing BMI disappears in late adolescence when genetic factors explain ~80% of the variation of BMI.

However, little is known about the universality of these results considering that the 2 previous multinational analyses were primarily based on Western populations, with the exception of one Korean twin cohort. A multinational study pooling 8 cohorts of adolescent twins found that the heritability estimates of BMI were approximately similar in Western and East Asian populations even when the mean BMI and total variation of BMI were higher in Western populations (8). However, it is still unknown whether the genetic architecture is similar at earlier ages. Furthermore, because of a lack of data in the previous multinational analyses (5, 6), it is still unclear how genetic influences on BMI differ between boys and girls in infancy and childhood.

To answer these questions on differences in the genetic architecture of BMI during childhood and adolescence, we conducted an individual-based analysis pooling twin cohorts from different countries. Our very large sample size allowed us to estimate the proportions of BMI variation explained by genetic and environmental factors by using 1-y age groups in boys and girls separately. We aimed 1) to estimate how the genetic architecture of BMI changes from infancy to the onset of adulthood, 2) to study age- and sex-differences in the contributions of genetic and environmental factors, and 3) to analyze whether these estimates are similar in different geographic–cultural regions representing different levels of obesogenic environment.

METHODS

The data were derived from the CODATwins (COllaborative project of Development of Anthropometrical measures in Twins) database described elsewhere (9). Briefly, the CODATwins project was intended to collect height and weight measurements from all twin cohorts in the world having information both on monozygotic and dizygotic twins. For the present analysis, we selected 45 twin cohorts from 20 countries with height and weight measurements available from 0.5 to 19.5 y of age for ≥50 twin individuals. We divided these cohorts into 3 geographic–cultural regions: Europe, North America and Australia, and East Asia. The prevalence of obesity and overweight is lowest in East Asia, thus representing a lesser obesogenic environment, and highest in North America and Australia, thus representing a more obesogenic environment (10). We had 20 cohorts from Europe, 15 cohorts from North America and Australia, and 8 cohorts from East Asia. Furthermore, we had 1 cohort from Africa and 2 from the Middle East. However, during the course of the study, we found that in a large Chinese National Twin Cohort Study, the heritability estimates of BMI were substantially lower than in other East Asian cohorts as reported previously (11). Given this heterogeneity, we presented the East Asian results both without (main results) and with (supplemental results) this cohort.

The names of the cohorts included in the main analyses are given in the footnotes of Supplemental Table 1, and more information on these cohorts is available elsewhere (9). The construction of the study cohort is presented as a flow diagram (Supplemental Figure 1). We eliminated impossible values and outliers in each age and sex group based on visual inspection allowing the BMI distribution to be positively skewed. We removed 1151 measurements as outliers representing 0.3% of the measurements. Further, we selected only one observation per twin individual for each 1-y age group. For 87,782 twin pairs, we had information for both co-twins (36% monozygotic twins, 37% same-sex dizygotic, and 27% opposite-sex dizygotic twins), and for 4826 twin pairs, information for only 1 twin. These incomplete twin pairs were removed from all genetic analyses. In the final analyses, we had 383,092 BMI measurements for 175,564 twin individuals (46% female). Thus, on average, we had 2 BMI measurements/individual, but the

**HERITABILITY OF BMI IN CHILDHOOD AND ADOLESCENCE**

3 of 9
number of longitudinal measures varied between the cohorts as described elsewhere (9). To test the effect of having multiple measurements for the same individual, we also repeated the genetic analyses after randomly selecting only one observation for each twin pair.

The number of complete twin pairs by age, zygosity, and region is presented in Supplemental Table 1. The number of BMI measurements varied from 6174 at 6 y of age to 31,708 at 1 y of age. The largest number of measurements was available from Europe ($n = 278,479$), followed by North America and Australia ($n = 66,204$), and finally East Asia ($n = 36,528$). In the additional analyses including the Chinese National Twin Cohort Study, the number of BMI measurements in East Asia was 55,756. From all BMI measurements, 57% were done in the year 2000 or later and 88% in the year 1990 or later. The majority of the BMI measurements were based on self report (66%) or parental report (20%), and only a minority were clinically measured (14%). Because the collaborators were asked to send height and weight measures, no missing cases existed for BMI.

The data were analyzed by using classic genetic twin modeling based on linear structural equations (12). Genetic twin modeling is based on the fact that monozygotic twins share virtually the same DNA sequence, whereas dizygotic twins share, on average, 50% of their genes identical-by-descent. Dizygotic within-pair correlations of BMI were more than half of the monozygotic correlations, suggesting the presence of environmental effects shared by co-twins (Supplemental Table 1). Thus, we decomposed the trait variation into 1) an additive genetic component (A), which is the sum of the effects of all alleles affecting the trait, 2) a shared environmental component (C), including all environmental factors shared by co-twins, and 3) a unique environmental component (E), reflecting the effects of all environmental factors that make co-twins dissimilar including measurement error. The additive genetic correlation is 1 between monozygotic co-twins and 0.5 between dizygotic co-twins, whereas the correlation between the shared environmental factors is 1, and that between unique environmental factors is 0 both in monozygotic and dizygotic co-twins. All genetic models were fitted with the OpenMx package, version 2.0.1, which is part of the R statistical platform (13). All parameter estimates and corresponding 95% CIs were estimated by raw-data maximum likelihood method allowing nonsymmetric 95% CIs. Heritability is defined as the proportion of total variation accounted for by additive genetic variation.

BMI showed an increasing skew to the right from 1 to 18 y of age, and thus we used a log-transformation to normalize the BMI distribution at all ages when calculating the relative proportions of genetic and environmental variation. Further, we adjusted BMI for age and study cohort differences within each 1-y age and sex group by calculating regression residuals. Cohort differences, i.e., differences in mean BMI between cohorts, were adjusted for by including a group of dummy variables in the regression models. We tested the technical assumptions of twin modeling by comparing the ACE model to the saturated model, which specifies an unconstrained model for trait means, variances, and covariances between co-twins. The fit of nested models was compared by calculating differences in $-2\log$-likelihood values, which follows the chi-square distribution with a difference in df that corresponds to the difference in the number of free parameters estimated. As reported previously, dizygotic twins had slightly higher mean BMIs and higher SDs than monozygotic twins at some ages over childhood and adolescence (14). We therefore allowed different means for monozygotic and dizygotic twins, but in the genetic models we allowed constrained variance components to be the same in all zygosity groups within sex.

The model fit results are presented in Supplemental Table 2. At most ages, the fit of the full ACE model was significantly poorer than the fit of the saturated model because of the higher SDs of BMIs in dizygotic twins. Even when the differences were small, they were statistically significant because of our very large sample size. Moreover, we tested possible sex differences by constraining the A, C, and E parameter estimates to be equal in boys and girls. We found that at most ages, the fit of this model was poor, suggesting that these variance components differed between sexes. We also tested whether this difference was because of different variances of the natural logarithm of BMI (logBMI) in boys and girls by fitting a scaled model allowing different sizes of variance components but fixing the relative size of these components to be equal. This model also showed significant differences compared with the full ACE model. Accordingly, we presented results separately for boys and girls. Finally, we tested whether a partly different set of genes affects BMI in boys and girls by fitting a sex-limitation model. This model tests whether the genetic correlation of opposite-sex dizygotic twins is $<0.5$. We found evidence of a sex-specific genetic effect at some ages seen also as lower opposite-sex dizygotic correlations (Supplemental Tables 1 and 2). Therefore, sex-specific genetic effects were allowed at all ages.

The pooled analysis was approved by the ethical board of the Department of Public Health, University of Helsinki. The data collection procedures of participating twin cohorts were approved by local ethical boards following the regulations in each country. Only anonymized data were delivered to the data management center at University of Helsinki (9).

RESULTS

Mean BMI decreased from infancy, reaching a nadir at 5 y of age in boys and girls before increasing until 19 y of age in the pooled data (Table 1). Along with the increasing mean BMI, the variance of BMI also started to increase after 5 y of age. The increase in mean BMI started in Europe after 5 y of age but slightly later in East Asia (6 y) and in North America and Australia (7 y). Boys had higher BMI than girls from 1 to 4 y of age and again from 17 to 19 y of age, but at other ages sex differences were small. In Europe and North America and Australia, BMI variances were higher in girls than in boys, especially in adolescence and early adulthood. North American and Australian boys and girls had the highest mean BMI at all ages, and this difference increased after 7 y of age. European boys and girls had also slightly higher BMI than East Asian boys and girls at most ages. Similar differences were also seen in the BMI variation, and at all ages variances were highest in North America and Australia.

Figure 1 presents the relative proportions of logBMI variation explained by additive genetic, shared environmental, and unique environmental factors in the pooled data (the estimates with 95% CIs are presented in Supplemental Table 3). The heritability estimate of logBMI was lowest at 4 y of age in boys ($a^2$: 0.42; 95% CI: 0.37, 0.47) and girls ($a^2$: 0.41; 95% CI: 0.35, 0.46). They started to increase after 8 y of age, and at 19 y of age...
In the heritability estimates between ages somewhat increased. However, the heritability estimates in these analyses were statistically significantly lower in midchildhood than in late adolescence and onset of adulthood (Supplemental Table 4).

Genetic correlations within opposite-sex dizygotic pairs were generally <0.5, suggesting sex-specific genetic effects, especially in adolescence (Figure 2). Wide upper 95% CIs were seen at ages 10 y, 12 y, and 14 y. This was because of a difference in shared environmental variation between boys and girls at these ages; if this difference increases, it can compensate for the effect of increasing additive genetic correlation for opposite-sex pairs in the statistical model. However, higher additive genetic correlation in opposite-sex than in same-sex dizygotic pairs is not biologically plausible.

We then fitted similar models for logBMI by region. Only the estimates of additive genetic factors are presented in Figure 3,
but all estimates with 95% CIs are available in Supplemental Table 3. In Europe and in North America and Australia, the age-related differences in the heritability estimates were largely similar to those in the pooled data. At 4 y of age, the heritability estimates in Europe were 0.41 (95% CI: 0.35, 0.47) in boys and 0.42 (95% CI: 0.35, 0.49) in girls, whereas in North America and Australia they were 0.38 (95% CI: 0.30, 0.47) and 0.27 (95% CI: 0.18, 0.37), respectively. After childhood, the heritability estimates generally increased and at 19 y of age were 0.78 (95% CI: 0.73, 0.80) in males and 0.75 (95% CI: 0.66, 0.82) in females in Europe and 0.65 (95% CI: 0.55, 0.77) and 0.82 (95% CI: 0.67, 0.85), respectively, in North America and Australia. The heritability estimates were even higher at some other ages in adolescence, but they did not differ statistically significantly from the estimates at 19 y of age. In East Asia, the pattern was not as clear because of the smaller sample size, but the heritability estimates showed some increase. However, especially after 12 y of age, the number of twin pairs was small in this region, leading to wide 95% CIs, and the results were not generally statistically significant. Despite the roughly similar age patterns, the proportions of logBMI variation explained by genetic and environmental factors were significantly different between the regions at all ages (Supplemental Table 2). When the Chinese National Twin Cohort Study was included in the East-Asia region, the proportion of genetic factors decreased and shared environmental factors increased; the change was from 0.1 to 0.4 units depending on the age group (Supplemental Table 5).

**DISCUSSION**

In this very large study of nearly 400,000 BMI measurements in nearly 88,000 complete twin pairs from 20 countries, we demonstrated that heritability of BMI was lower in midchildhood than in late adolescence and onset of adulthood, which has been suggested previously also by 2 international studies (5, 6). The increasing role of genetic factors is consistent with previous molecular genetic studies that have found that the variants of the FTO gene, which account for the largest fraction of variance in BMI among the known candidate genes for BMI (15), and other obesity-related candidate genes have increasing effects on BMI after 6 y of age (16–19). Evidence of increasing heritability of BMI from 4 to 10 y of age has also been reported in genome-wide complex trait analysis (20).

However, this increasing role of genetic factors in BMI with age does not negate the importance of health behavior associated with childhood obesity, because genetic factors can affect BMI by modifying food intake and other behavioral...
factors. For example, the variants of the *FTO* gene, which act on the actual functional gene *IRX3* (21), were found to be associated with food-intake self-regulation and eating styles in childhood, which are further associated with weight gain (22). Although not yet conclusive, there is evidence that common genetic risk variants of BMI are active in the hypothalamus, pituitary gland, hippocampus, and limbic system, i.e., areas of the brain that have an important role in appetite regulation, learning, cognition, emotion, and memory (23). It has also been found that shared environmental factors have effects on nutritional intake in childhood (24), but they disappear in adulthood when genetic factors become more important (25, 26). The increasing proportion of genetic variation of BMI may thus reflect the increasing independence of children from their parents in eating and other behavioral factors associated with the variation of BMI. However, the associations between energy intake and obesity are complex and still an object of debate (27). Differences in DNA methylation have also been found between lean and obese co-twins (28), and epigenetic processes by themselves are in part genetically regulated (29). Therefore, it is possible that part of the genetic variation may be mediated by epigenetic effects.

We found some evidence for sex-specific genetic contributions to BMI. The lowest genetic correlation within opposite-sex dizygotic pairs was found at 13 y of age, probably coinciding with the onset of puberty. However, a sex-specific genetic contribution was also clear after puberty, which probably reflects the increasing differences in body composition between boys and girls with age (30). This is consistent with the sex-specific genetic contribution in adult BMI found in a study of twin cohorts with opposite-sex twins from 7 countries (31). However, it is noteworthy that lower genetic correlations for opposite-sex pairs were found even in infancy, indicating that a partly different set of genes regulates BMI before the major hormonal changes that occur during puberty. This suggests some caution when interpreting results from genetic studies that have relied on BMI pooling of boys and girls, even while focusing on prepubertal children. Otherwise, there were relatively minor differences in the genetic architecture of BMI between boys and girls, and the general age patterns were largely similar.

When comparing regions, North American and Australian children and adolescents presented greater means and larger total variation of BMI than their European and East Asian peers. The relative proportions of genetic and environmental sources of variations were, however, roughly similar in these 3 regions. These results are consistent with those of a previous international twin study showing larger mean and variance of BMI yet similar heritability estimates in Caucasian and in East Asian populations in adolescence (8). Thus, higher mean BMI was associated with greater variation in BMI, whereas the proportions of genetic variations were still largely similar. This suggests that genetic factors have an important role in individual differences in BMI in various populations differing in ethnicity and environmental exposures, as well as in their possible interactions. These results are consistent with studies in Denmark (32) and Sweden (33), suggesting that both total and genetic variation of BMI increased during the obesity epidemic. It is, however, noteworthy that we limited our East Asian cohorts to affluent populations, including the affluent Shandong and Guangdong provinces but excluding poorer areas of China. As reported previously, the heritability estimates of BMI were much lower and common environmental estimates higher in other areas of China (11), which may indicate larger differences between families in nutritional status. This emphasizes the importance of collecting data on twins living under different environmental exposures. Our study cannot reveal whether genetic or environmental factors are behind the differences in mean BMI between the regions. However, a recent study found that genetic factors explained a part of differences in mean BMI between European populations (34), and thus genetic factors may contribute to BMI differences also in our study cohorts in addition to environmental factors.

The data used in this study have both strengths and weaknesses. The main strength is the very large sample size, allowing an investigation of the change of the genetic and environmental contributions to individual differences in BMI in much more detail than in previous studies. We also have twin participants from different countries, thereby making it possible to stratify the analyses by regions of various ethnicities and obesogenic environments. Individual-based data also have many advantages over literature-based meta-analyses, such as better opportunities for
statistical modeling and lack of publication bias. However, even when the large majority of the twin cohorts in the world participated in this project, our data still had only limited power for East Asia, especially in adolescence. Another important limitation is that there were only few data sets available from the Middle East and Africa and a lack of data from South America. This underlines the need for new data collection in these geographic regions. There were some violations of the assumptions of twin modeling because of the larger variation in dizygotic twins than in monozygotic twins at some ages (14). The differences in the variation are, however, small and become statistically significant because of the very large sample size of our data. Finally, we did not have any area-level indicators and classified the cultural-geographic areas as less or more obesogenic based on the prevalence of adult obesity (10). Conceptualizing obesogenic environments is difficult, but it has been suggested that both micro- and macro-level environmental factors affect both food intake and physical exercise (35). More detailed measurements of the physical environment are thus needed to analyze the factors in the environment that potentially modify genetic influences on the development of obesity. Because several cohorts were available from one country, it is possible that one individual may have participated in more than one twin cohort. It is also a clear limitation that only a fraction of BMI values was based on clinical measures, and for most of them we needed to rely on parental or self-reported values.

In conclusion, we found evidence that environmental factors shared by co-twins contribute to BMI variation in early childhood and during puberty, but their role disappeared before the onset of adulthood. Heritability increased from midchildhood to the onset of adulthood, which may indicate gene–environment correlation processes, whereby an increasing independence of children from their parents led them to express their behaviors according to their genetic background. Genes affecting BMI were partly sex specific, even in infancy, with their contribution becoming more prominent during and after puberty. Obesogenic environment is associated with greater variation of BMI in North America and Australia than in East Asia, but the relative proportions of genetic and environmental variations were roughly similar. Our results suggest that, despite different ethnicities and environmental exposures, genetic factors play a major role in the variation of BMI in adolescence in affluent societies.

The authors’ responsibilities were as follows—KS, Y-MH, YY, KOK, FR, TIAS, DIB, and JK: planned the study design of the CODATwins project; RS und, Y-MH, YY, CH, JvBH, SM, SO, SA, FJ, FN, ZP, AB, CK, KJS, KLY, WC, AEH, TMM, WG, CY, LL, RPC, BMH, KC, AS, KOK, CAD, RRV, RJOFL, KH, JW, CHL, AF, TAM, TCE, AMG, MH, XD, MB-A, HB-N, MS, ADT, DLT, MAS, CF, CD, AK-N, DM, LA, SAB, KLLK, JLS, LJE, HMM, RFK, M McGue, SP, MG, DAB, M Bartels, TCEvB, JMC, R Saffery, DLF,.

**FIGURE 3** Proportions of natural logarithm of BMI variation with 95% CIs based on maximum likelihood estimation explained by additive genetic factors by age, sex, and region for European boys (A), European girls (B), North American and Australian boys (C), North American and Australian girls (D), East Asian boys (E), and East Asian girls (F). The number of twin pairs varied from 1107 at 6 y of age to 13,855 at 3 y of age in Europe, from 565 at 1 y of age to 5064 at 18 y of age in North America and Australia, and from 111 at 17 y of age to 2284 at 1 y of age in East Asia.
REFERENCES


