Genetic and environmental influences on height from infancy to early adulthood: An individual-based pooled analysis of 45 twin cohorts

Aline Jelenkovic1,2, Reijo Sund1, Yoon-Mi Hur3, Yoshie Yokoyama4, Jacob v. B. Hjelmborg5, Sören Möller3, Chika Honda9, Patrik K. E. Magnusson17, Nancy L. Pedersen7, Syuichi Ooki8, Sari Aaltonen1,9, Maria A. Stazi10, Corrado Fagnani10, Cristina D’Ippolito10, Duarte L. Freitas11, José Antonio Maia12, Fuling Ji13, Feng Ning13, Zengchang Pang13, Esther Rebato2, Andreas Busjahn14, Christian Kandler15, Kimberly J. Saudino16, Kerry L. Jang17, Wendy Cozen18,19, Amie E. Hwang18, Thomas M. Mack18,19, Wenjing Gao20, Canqing Yu20, Liming Li20, Robin P. Corley21, Brooke M. Huibregtse21, Catherine A. Derom22,23, Robert F. Vlietinck22, Ruth J. F. Loos24, Kauko Heikkilä9, Jane Wardle25,26, Claire H. Llewellyn25, Abigail Fisher25, Thalia C. Eley26, Alice M. Gregory27, Mingguang He28,29, Xiaohu Ding28, Zengchang Pang30,31,32, Henning Beck-Nielsen32, Morten Sodemann33, Adam D. Tarnoki34,35, David A. Butler43, Meike Bartels44, Toos C. E. M. van Beijsterveldt45, Jeffrey M. Craig45,46, Richard Saffery45,46, Lise Dubois47, Michel Boivin48,49, Mara Brendgen50, Ginette Dionne48, Frank Vitaro51, Nicholas G. Martin52, Sarah E. Medland52, Grant W. Montgomery53, Gary E. Swan54, Ruth Krasnow55, Per Tynelius56, Paul Lichtenstein7, Claire M. A. Haworth57, Robert Plomin58, Gombojav Bayasgalan58, Danshiitsoodol Narandalai59,58, K. Paige Harden60, Elliot M. Tucker-Drob60, Timothy Spector61, Massimo Mangino61, Genevieve Lachance61, Laura A. Baker62, Catherine Tuvblad62,63, Glen E. Duncan63, Dedra Buchwald64, Gonneke Willemse65, Axel Skytthe6, Kirsten O. Kyvik66,67, Kaare Christensen5,67, Sevgi Y. Öncel68, Fazil Aliev69, Finn Rasmussen66, Jack H. Goldberg70,71, Thorkild I. A. Sørensen72,73, Dorret I. Boomsma74, Jaakko Kaprio9,73,74 & Karri Silventoinen1,6

Height variation is known to be determined by both genetic and environmental factors, but a systematic description of how their influences differ by sex, age and global regions is lacking. We conducted an individual-based pooled analysis of 45 twin cohorts from 20 countries, including 180,520 paired measurements at ages 1–19 years. The proportion of height variation explained by shared environmental factors was greatest in early childhood, but these effects remained present until early adulthood. Accordingly, the relative genetic contribution increased with age and was greatest in adolescence (up to 0.83 in boys and 0.76 in girls). Comparing geographic-cultural regions (Europe, North-America and Australia, and East-Asia), genetic variance was greatest in North-America and Australia and lowest in East-Asia, but the relative proportion of genetic variation was roughly similar across these regions. Our findings provide further insights into height variation during childhood and adolescence in populations representing different ethnicities and exposed to different environments.
Human height is a classic anthropometric quantitative trait for its ease of measurement, approximately normal distribution and relative stability in adulthood, and thus has been the target of extensive research across many fields of science. The study of height has a long standing tradition in genetics; in fact, the field of quantitative...
genetics was born out of studies of human height in the late 19th and early 20th centuries. Galton1 published data as early as 1886 on the relationship between parent and offspring height and inferred that "when dealing with the transmission of stature from parents to children, the average height of the two parents is, … all we need care to know about them". Later on, Pearson and Lee2 presented correlations of height between relatives, also providing evidence for the inheritance of height. In 1918 Fisher3 calculated the first heritability estimate of height, i.e., the proportion of total variation explained by genetic variation; in this seminal paper presenting the statistical principles of quantitative genetics, he demonstrated that continuous characters are caused by a combination of many genetic loci with small effects (polygenic inheritance), replacing the blending inheritance hypothesis proposed by Galton. Since then many lines of evidence such as twin, adoption and family studies have estimated the role of genetic factors in the determination of height, showing that it is one of the most heritable human quantitative phenotypes4. Interest in the genetic influences on height was renewed when genetic linkage studies enabled research into genetic effects over the whole genome5 and genome-wide association (GWA) studies allowed identification of loci consistently associated with height in populations of different ancestry6-10.

Beside the genetic factors, a multitude of environmental factors can affect height. They can operate during the whole growth period, but infancy is probably the most sensitive phase regarding external influences11,12. In the presence of adverse environmental conditions, the physical growth of children can decline and even adult height may be affected12-14. Nutrition and especially lack of dietary protein is universally the most important environmental factor influencing height, but also childhood diseases, in particular infections, can affect growth13. These and other proximate biological determinants are further associated with social and economic conditions manifesting as socio-economic differences in height both within and between populations15.

Although the heritable nature of height has been recognized for more than one hundred years, only a few studies have explored in detail the genetic variation of height during childhood and adolescence. Twin studies have consistently estimated that the heritability of height is lowest (0.2–0.5) in infancy15-17, rapidly increases in childhood with varying values15,16,18, and reaches estimates ranging from 0.70 to 0.90 in adolescence and adulthood15,17,19-21. However, these studies leave unclear whether environmental factors shared by co-twins, which are generally important in infancy and childhood, persist in adolescence or after the cessation of growth15–21. A study in four countries with over 12,000 twin pairs from birth to 19 years of age showed that the effect of shared environment remained up through 12 years, and was present again at 16 years15. Somewhat different results were observed in a longitudinal study of two Finnish twin cohorts, which found that common environmental factors affecting height at different ages in adolescence and early adulthood20.

Height is also a classic example of a sexually dimorphic trait; on average, men are taller than women in all human populations15. However, much less is known about sex-differences in genetic and environmental contributions to height variation. Greater heritability estimates for males than for females in childhood15 and adulthood21 have been reported. Also sex-specific genetic effects have been found for height, but the results are inconsistent across studies18-21,31.

Further, a greater mean height has been consistently observed in Western populations as compared with East-Asian populations15, but most studies on the genetic and environmental factors influencing height variation to date are based on Western populations. A multinational study on adolescent twins from eight countries showed that even when the total variation of height was higher in Western populations, the heritability estimates were largely similar between Western and East-Asian populations19. However, these studies did not address the possible differences in the genetic variation pattern between ethnic-cultural groups in childhood and late adolescence/early adulthood.

Using height measures obtained from 45 twin cohorts in 20 countries participating in the COllaborative project of Development of Anthropometrical measures in Twins (CODATwins), we conducted an individual-based analysis of pooled twin cohorts (i) to analyze the genetic and environmental contribution to variation of height from 1 to 19 years of age; (ii) to explore sex-differences in these contributions over each year of age; and (iii) to assess whether this age pattern varies by geographic-cultural region (Europe, North-America and Australia, and East-Asia).

Results

Descriptive statistics of height by age and sex for the pooled data (all cohorts together) and by geographic-cultural region are presented in Table 1. Mean height expectedly increased with age in both sexes with the exception of the slight decrease observed at 18/19 years of age, which reflects differences in the distribution of different cohorts within each age group. Mean height was greater in boys than in girls; only at the age of 11 and 12 years were girls slightly taller than boys, reflecting the earlier onset of pubertal growth in girls. The difference in mean height between consecutive age groups was very similar in boys and girls during childhood; these mean height differences started to decrease considerably from 12 years in girls and 14 years in boys. The variation of height increased with age and reached the peak at 12 years in girls and 13 in boys, and then decreased slightly. When comparing geographic-cultural regions, mean height was tallest in Europe, somewhat shorter in North-America and Australia and shortest in East-Asia at all ages in boys and girls. The variation of height showed a less clear pattern but was generally greatest in North-America and Australia and lowest in East-Asia.

Figure 1 presents the proportions of height variation explained by additive genetic, common (shared) environmental and unique environmental factors from 1 to 19 years of age in the pooled data (estimates with 95% confidence intervals (CIs) are available in Supplementary Table S1). The proportion of environmental variation shared by co-twins was greatest at age 1 (0.48 in boys and 0.49 in girls), decreased over childhood and stabilized in adolescence, remaining considerable until 19 years (except at ages 14 and 16 in boys) with values generally lower than 0.2. Accordingly, heritability was lowest at age 1 (0.40 in boys and 0.38 in girls) increased with age in early and middle childhood (~2–5 and 6–8 years of age, respectively) and was generally greater than 0.7 in late childhood (~9–11 years of age) and adolescence; the greatest heritability estimates were found for boys at ages 14 and
The proportion of height variation explained by environmental factors unique to each twin individual, which also includes measurement error, did not show any clear age pattern and was largely similar at all ages (0.05–0.14). In spite of the observed sex differences in the relative variance components at most ages (See Supplementary Table S2), the age pattern was generally similar in boys and girls; the biggest sex-differences were found in late adolescence when the heritability estimates were slightly greater in boys. The point estimates for the genetic correlations within opposite-sex DZ pairs were generally lower than 0.5 suggesting sex-specific genetic effects. There was a trend for these correlations to be lowest in adolescence, although the largest 95% confidence intervals were estimated at 14 and 16 years of age (Fig. 2).

Univariate models for height were then conducted separately in the three geographic-cultural regions. Only the estimates of additive genetic factors are presented in Fig. 3, but all estimates with 95% CIs are available in Supplementary Table S1. The three geographic-cultural regions showed the general trend of increasing proportion of additive genetic factors with age during childhood. Explained by its largest sample size, the pattern in Europe was practically the same to that observed for all cohorts together, but with slightly greater heritability estimates at most ages. In North-America and Australia and East-Asia, heritability estimates in childhood were

|          | All cohorts | Europe | North-America and Australia | East-Asia |
|----------|-------------|--------|-----------------------------|-----------|
| **Boys** |             |        |                             |           |
| Age 1    | 15512       | 74.3   | 12674                       | 75.3      |
|          |             | 4.3    |                             |           |
| Age 2    | 12667       | 87.0   | 10357                       | 87.7      |
|          |             | 4.3    |                             |           |
| Age 3    | 16840       | 96.3   | 13761                       | 96.8      |
|          |             | 4.4    |                             |           |
| Age 4    | 9719        | 102.4  | 7244                        | 103.0     |
|          |             | 5.3    |                             |           |
| Age 5    | 7987        | 111.5  | 6014                        | 112.6     |
|          |             | 6.0    |                             |           |
| Age 6    | 3167        | 115.5  | 1328                        | 118.3     |
|          |             | 6.7    |                             |           |
| Age 7    | 13396       | 124.3  | 11316                       | 125.3     |
|          |             | 6.6    |                             |           |
| Age 8    | 6298        | 128.8  | 4333                        | 130.3     |
|          |             | 6.5    |                             |           |
| Age 9    | 6870        | 134.2  | 4212                        | 135.8     |
|          |             | 7.0    |                             |           |
| Age 10   | 11228       | 140.9  | 8902                        | 142.1     |
|          |             | 7.2    |                             |           |
| Age 11   | 8839        | 144.6  | 6470                        | 145.6     |
|          |             | 7.4    |                             |           |
| Age 12   | 11599       | 151.9  | 8342                        | 152.7     |
|          |             | 8.1    |                             |           |
| Age 13   | 4745        | 158.9  | 3322                        | 159.7     |
|          |             | 9.3    |                             |           |
| Age 14   | 8896        | 165.9  | 6308                        | 166.0     |
|          |             | 8.9    |                             |           |
| Age 15   | 5307        | 171.8  | 3782                        | 172.3     |
|          |             | 8.7    |                             |           |
| Age 16   | 8013        | 175.7  | 5806                        | 176.2     |
|          |             | 7.5    |                             |           |
| Age 17   | 9829        | 177.0  | 6795                        | 178.2     |
|          |             | 7.4    |                             |           |
| Age 18   | 14722       | 176.1  | 6514                        | 179.6     |
|          |             | 7.6    |                             |           |
|          |             |        |                             |           |
| **Girls**|             |        |                             |           |
| Age 1    | 15520       | 72.9   | 12498                       | 74.1      |
|          |             | 4.3    |                             |           |
| Age 2    | 12223       | 85.9   | 9885                        | 86.7      |
|          |             | 4.3    |                             |           |
| Age 3    | 17216       | 95.3   | 13949                       | 95.8      |
|          |             | 4.5    |                             |           |
| Age 4    | 9571        | 101.3  | 7144                        | 101.8     |
|          |             | 5.2    |                             |           |
| Age 5    | 7817        | 110.6  | 5884                        | 111.9     |
|          |             | 6.2    |                             |           |
| Age 6    | 2807        | 114.6  | 886                         | 117.9     |
|          |             | 6.5    |                             |           |
| Age 7    | 13736       | 123.5  | 11368                       | 124.5     |
|          |             | 6.6    |                             |           |
| Age 8    | 6072        | 127.8  | 3979                        | 129.5     |
|          |             | 6.7    |                             |           |
| Age 9    | 6642        | 133.3  | 3868                        | 134.8     |
|          |             | 7.1    |                             |           |
| Age 10   | 11298       | 140.5  | 8882                        | 141.6     |
|          |             | 7.4    |                             |           |
| Age 11   | 8743        | 145.1  | 6288                        | 145.8     |
|          |             | 7.7    |                             |           |
| Age 12   | 11873       | 152.9  | 8376                        | 153.7     |
|          |             | 8.2    |                             |           |
| Age 13   | 4599        | 158.3  | 3078                        | 159.5     |
|          |             | 7.6    |                             |           |
| Age 14   | 9388        | 162.2  | 6644                        | 162.6     |
|          |             | 6.7    |                             |           |
| Age 15   | 5299        | 164.5  | 3786                        | 165.5     |
|          |             | 6.9    |                             |           |
| Age 16   | 8801        | 164.9  | 6516                        | 165.4     |
|          |             | 6.5    |                             |           |
| Age 17   | 9203        | 165.8  | 7523                        | 166.1     |
|          |             | 6.5    |                             |           |
| Age 18   | 7704        | 166.3  | 5504                        | 167.1     |
|          |             | 6.7    |                             |           |
| Age 19   | 8021        | 166.2  | 5582                        | 167.2     |
|          |             | 6.7    |                             |           |

Table 1. Descriptive statistics of height by age, sex and geographic-cultural region. N: number of twin individuals; SD: standard deviation.

16 (0.83 and 0.82, respectively). The proportion of height variation explained by environmental factors unique to each twin individual, which also includes measurement error, did not show any clear age pattern and was largely similar at all ages (0.05–0.14). In spite of the observed sex differences in the relative variance components at most of ages (See Supplementary Table S2), the age pattern was generally similar in boys and girls; the biggest sex-differences were found in late adolescence when the heritability estimates were slightly greater in boys. The point estimates for the genetic correlations within opposite-sex DZ pairs were generally lower than 0.5 suggesting sex-specific genetic effects. There was a trend for these correlations to be lowest in adolescence, although the largest 95% confidence intervals were estimated at 14 and 16 years of age (Fig. 2).

Univariate models for height were then conducted separately in the three geographic-cultural regions. Only the estimates of additive genetic factors are presented in Fig. 3, but all estimates with 95% CIs are available in Supplementary Table S1. The three geographic-cultural regions showed the general trend of increasing proportion of additive genetic factors with age during childhood. Explained by its largest sample size, the pattern in Europe was practically the same to that observed for all cohorts together, but with slightly greater heritability estimates at most ages. In North-America and Australia and East-Asia, heritability estimates in childhood were
generally somewhat lower than in Europe. In East-Asia the pattern in adolescence was not so clear because of the smaller sample size leading to wider 95% CIs. In spite of the roughly similar age patterns, the proportions of height variation explained by genetic and environmental factors were different between the geographic-cultural regions (See Supplementary Table S2). The Chinese National Twin Registry was excluded from these analyses because the heritability estimates in that cohort were substantially lower than in other East-Asian cohorts. When data from this cohort was included in the analyses for East-Asia, the proportion of genetic factors decreased and

---

**Figure 1.** Proportion of height variation explained by additive genetic (a,b), shared environmental (c,d) and specific environmental (e,f) factors with 95% confidence intervals by age and sex for all cohorts together.

**Figure 2.** Additive genetic correlations with 95% confidence intervals within opposite-sex DZ pairs by age for all cohorts together.
common environmental factors increased considerably; the change in heritability estimates was from 0.1 to 0.3 units depending on the age group (data available on request).

Finally, we studied how age modifies the genetic and environmental variances of height by using gene-age interaction analysis, with data pooled across all age groups. Figure 4 shows the change in the predicted raw genetic and environmental variances in height as a function of age (parameter estimates with 95% CIs are available in Supplementary Table S3). Additive genetic variation increased steadily from age 1, reached its peak at 14 years in boys and 13 years in girls and then decreased again in the pooled data; however, common and unique environmental variation were largely similar across ages. When stratified by geographic-cultural region, genetic variation was largest in North-America and Australia, somewhat lower in Europe, and lowest in East-Asia, particularly for boys. The pattern of genetic variance increasing to a maximum and thereafter decreasing was consistent across the regions. Also common environmental variation was greatest in North-America and Australia, somewhat lower in Europe, and lowest in East-Asia, particularly for boys. The pattern of genetic variance increasing to a maximum and thereafter decreasing was consistent across the regions. Also common environmental variation was greatest in North-America and Australia, reaching the peak at 10 years in boys and 7 years in girls, whereas in Europe and East-Asia it was similar across ages. Unique environmental variation showed a similar pattern and magnitude in the three geographic-cultural regions. The differences between the regions were highly statistically significant in boys [difference in $-2 \log$-likelihood values ($\Delta -2 LL = 1257$, difference in degrees of freedom ($\Delta d.f. = 30$, p-value $< 0.0001$)] and girls ($\Delta -2 LL = 1364$, $\Delta d.f. = 30$, p-value $< 0.0001$). When comparing sexes, in Europe and North-America and Australia there was a trend toward a greater genetic variation for boys than for girls, which increased with age. In East-Asia, however, genetic variation was slightly greater for girls until 14 years of age and for boys in late adolescence.

Discussion

The present study of 180,520 paired measurements from 86,037 complete twin pairs in 20 countries revealed that environmental factors shared by co-twins contribute to the inter-individual variation in height from infancy to early adulthood. The relative proportion of common environmental factors was greatest during the first years of life, representing almost half of the variation at age 1, and decreased over childhood and adolescence. The interpretation of these results, however, deserves some caution. It has been questioned whether twin studies are suitable for estimating heritability of height in infancy, since early growth patterns in twins differ considerably
from singleton growth patterns\textsuperscript{25}. Prenatal environmental factors can act very differently on MZ twins leading to differences in body size within pairs (the most extreme case is the twin-to-twin transfusion syndrome). This is an important issue because in the classical twin design heritability is estimated by comparing the resemblance of MZ and DZ twin pairs, and thus body size differences in MZ pairs will result in lower heritability estimates. Since children may take several years to fully catch-up after birth, the high proportion of height variation explained by the shared environment in infancy may still reflect these prenatal environmental factors. Among other possible explanations, it might be that the shared environment represents the effects of gestational age or the effects of the higher measurement error (correlated in twins) at earlier ages.

The influence of the shared environment on height variation up to 19 years, which is consistent with previous studies in adolescents\textsuperscript{19} and adults\textsuperscript{21} with enough statistical power to detect this component, suggests that adult height variation reflects childhood living conditions. Studies have shown that the secular trend in adult height occurs during the first two years of life mainly due to increases in leg length\textsuperscript{26}. A plausible explanation is that the period of most rapid growth, when the effect of an adverse environment is strongest, coincides with the period when most growth takes place in the long bones of the legs\textsuperscript{26}. Multinational studies analyzing the genetic and environmental influences on body length segments, particularly leg length, are thus needed to disentangle the aetiology of total height variation. The small but considerable effect of unique environment on height variation, very similar across ages, may partly be due to measurement error, which is modelled as part of unique environmental factors. However, it is likely that it also reflects real environmental factors, for example, different exposure to childhood diseases.

A recent and large meta-analysis of twin correlations and variance components for 17,804 traits carried out separately in four age groups (0–11, 12–17, 18–64 and 65+ years) showed that the heritability estimate of height at 12–17 years was considerably greater than at 0–11 years\textsuperscript{27}. Given the rapid growth that occurs in infancy,
childhood and adolescence, in this individual-based pooled analysis we analyzed the heritability of height in one year age groups. We found that genetic contributions increase over childhood with heritability estimates in the range of previous studies in children and adults\(^5,6,18,20,21\). GW A studies have identified many common genetic variants for adult height. The most recent GW A meta-analysis in 253,288 individuals of European ancestry identified 697 genome-wide significant SNPs in 423 loci that together explained one-fifth of the heritability for adult height\(^6,20\). Further, in a study using whole-genome sequencing data from 44,126 unrelated individuals, all imputed variants explained 56% of variance for height suggesting that missing heritability is negligible for human height\(^29\).

However, much less is known on the genetics of height in children. Van der Valk et al.\(^29\) found that polygenic scores based on 180 SNPs previously associated with adult height explained 2.95% of the variance of infant length, and that of 180 known adult height loci, only 11 were genome-wide significantly associated with infant length.

The pattern of total height variation across ages was largely driven by genetic variance. The most consistent result is the increasing genetic variance with age, reaching its peak at around 13 years in girls and 14 years in boys. After that point, even if mean height continued to increase, genetic variance started to decrease in such a way that in late adolescence the magnitude was similar to that before pubertal events start. Adolescence is characterized by the onset of puberty and the occurrence of growth spurts. Although a secular and population-dependent decline has been observed in the age at onset of pubertal growth spurt and peak height velocity since the mid 1900s\(^30,31\), the pubertal height spurt generally begins at age 10–11 years in girls and 11–13 years in boys and reaches peak height velocity at about 12 years and 14 years, respectively\(^13,30\). In this study, twins within age groups are at various stages of puberty. In addition to the substantial heritability reported for pubertal timing\(^32\), a genome-wide genetic correlation (0.13) between age at menarche and adult height has also been found\(^33\). In fact, a genome-wide association meta-analysis showed that five loci associated with pubertal timing impacted multiple aspects of growth, both before and during puberty\(^34\). Therefore, it is possible that some of the genetic variance in height at these ages is confounded with genetic variance in pubertal events.

In spite of the largely similar age patterns observed in boys and girls, boys showed somewhat greater heritability estimates and genetic variation, especially in late adolescence. Greater heritability estimates in boys than in girls have previously been reported from birth through 19 years\(^13\) and in adulthood\(^21\). Moreover, some studies have shown a sex-specific genetic effect on height variation in adolescents\(^19\) and adults\(^24\). It is clear that both of the sex chromosomes are implicated in determining mean height. Short stature has been demonstrated in females with Turner syndrome who have only one X chromosome\(^35\) and taller stature seen in XYY men compared with XY men\(^36\). However, sex chromosomes have also been associated with height variation; for example, Gudbjartsson et al.\(^37\) identified 27 regions of the genome including a locus on X chromosome that together explained around 3.7% of the population variation in height. In our multochondal data, the lowest genetic correlations within opposite-sex DZ pairs were found at 14–16 years of age and again at 18 years, suggesting that sex-specific genes have a role in the genetic variation of height not only during puberty, but also in late adolescence.

Comparison between geographic-cultural regions showed that mean height was greatest in Europe, somewhat shorter in North-America and Australia and shortest in East-Asia, but total variance was largest in North-America and Australia. Accordingly, genetic variation was also greatest in North-America and Australia and lowest in East-Asia. However, the relative proportions of additive and environmental variations were more similar in the different geographic-cultural regions. These results are consistent with a previous comparative twin study which found that the mean and variance of height were larger in Caucasian than in East-Asian populations in adolescence, but the heritability estimates were still at the same level\(^19\). An important proportion of the differences in total variances between geographic-cultural regions were attributable to genetic differences. It may be that allelic frequencies and effects of the genes involved in height vary between Europeans, North-Americans and Australians and East-Asians, leading to differences in genetic variation between the three population groups. A recent study across 14 European countries found that many independent loci contribute to population genetic differences in height, and estimated that these differences account for 24% of the captured additive genetic variance\(^38\). However, a major part of the differences in genetic variation may also be because of gene-environment interactions modelled as part of the additive genetic component in our model. That is, the higher genetic variation observed in Caucasians could arise because there is a set of genes expressed more strongly in Western environments. For example, a study of adults of Japanese descent living in the United States and native Japanese found that Japanese men and women were shorter than Japanese-Americans, suggesting that environmental factors play a role in physical growth\(^39\). Analyzing this question in detail would require collection of twins or GW A studies in unrelated individuals with East-Asian origin living in a Western environment.

The study in Caucasian and East-Asian populations showed that approximately 91% of the differences in the total variance between these two population groups was attributable to genetic variances\(^11\). However, our study found that shared environmental variance also differed between geographic-cultural regions. The lower shared environmental variance observed in East Asian girls and greater in North-America and Australia during childhood may reflect cultural differences in terms of nutrition and other environmental resources. It is also important to note that we limited our East-Asian cohorts to affluent East-Asian populations including the Shandong and Guangdong provinces but excluding poorer areas of China. As reported previously, the heritability estimates of height were considerably lower and common environmental estimates higher in the poorer areas\(^40\), which may indicate larger differences between families in nutrition and infection history in these areas of China. This emphasizes the need to collect data on twins living under different environmental exposures.

The main strength of the present study is the very large sample size of our multinational database of twin cohorts, with height data from 1 through 19 years of age, allowing a more detailed investigation of the genetic and environmental contributions to individual differences in height during childhood and adolescence than in the previous studies. Twin participants are from 20 different countries, thereby making it possible to stratify the analyses by regions representing different ethnicities and environments. Important advances of individual-based data are better opportunities for statistical modelling and lack of publication bias. However, our study also has
limitations. The equal-environment assumption, upon which twin methodology is based, assumes that MZ and DZ twins are equally exposed to environmental factors relevant to the outcome. If equal-environment assumption is violated, it should be seen as differences in variances between MZ and DZ twins, but we did not find such evidence. In the classical twin design phenotypic assortment increases DZ correlations and thus inflates the common environmental component when not accounted for in the modelling. Assortative mating is well recognized for height, and when the potential underestimation of heritability estimates was corrected using a sample of twins and their parents, these authors showed that doing so increased the heritability estimates from 0.75 to 0.85. In our database we do not have information on parental height and thus could not take into account assortative mating, which may thus explain part of the shared environmental variation. A recent study showed that increased homozygosity, which is influenced by inbreeding, was associated with decreased height and that the effect sizes were similar across different continental groups and populations with different degrees of genome-wide homozygosity. These authors thus suggested that homozygosity, rather than confounding as a result of environmental or additive genetic effects, directly contributes to phenotypic variance. Further, most of the height measurements were self-reported, which are prone to error and can bias our analyses toward lower heritability estimates and higher estimates of unique environmental effects. Finally, countries and/or ethnic-cultural regions are not equally represented, and the database is heavily weighted towards populations following the Westernized lifestyle; even when the large majority of the twin cohorts in the world participated in this project, our data still had limited power for East-Asia especially in adolescence. An even bigger problem is that there are few data available from South-Asia, Middle-East and Africa and no data from South-America. This demonstrates the need for new data collections in these regions.

Our findings provide further insights into height variation during childhood and adolescence in populations representing different ethnicities and exposed to different environments. Worthwhile objectives for future research are to study whether the same genetic and environmental factors contributing to height variation operate throughout time or new genes or new environmental factors start to operate at different ages, and to analyze the heritability of growth in height. Further, a major challenge in future studies with more information on birth and pregnancy related variables is to explore the reasons for the low heritability of height at young ages.

In conclusion, environmental factors shared by co-twins exert their strongest influence on height variation in childhood, but these effects remain until the onset of adulthood. Genetic variation in height increased steadily during childhood and reached its peak at around 13 years in girls and 14 years in boys, which may be confounded with genetic variation in pubertal events. Especially in adolescence, there was a trend toward somewhat greater genetic variation in boys than in girls, and part of the genetic variation of height was sex-specific. Genetic variation of height was larger in North-America and Australia and Europe compared with East-Asia, but the relative proportions of genetic and environmental variations between these three geographic-cultural regions were roughly similar. These findings suggest that, in spite of different ethnicities and environmental exposures, genetic factors play a major role on height variation in adolescence and early adulthood, but environmental factors shared by co-twins are also important.

**Methods**

**Ethics.** All participants were volunteers and gave their informed consent when participating in the study. No experimental data were asked and thus we did not ask ethical approval. Only a limited set of observational variables and anonymized data were delivered to the data management centre at University of Helsinki. The pooled analysis was approved by the ethical committee of Department of Public Health, University of Helsinki, and the methods were carried out in accordance with the approved guidelines.

**Sample.** This study is based on the data from the CODATwins project described elsewhere. Briefly, the CODATwins project was intended to recruit all twin projects in the world with information on height and weight measurements. For the present analyses, we selected height measurements at ages from 0.5 to 19.5 years (n = 420,707). Age was classified to single-year age groups (e.g., age 1 refers to 0.5–1.5 years range). Impossible values and outliers were checked by visual inspection of histograms for each age and sex group and were removed to obtain an approximately normal distribution (0.3% of the measurements). Since individuals in longitudinal studies have more than one measurement over time, analyses were restricted to one observation per individual in each age group. Analyses were additionally restricted to having at least 50 measurements per cohort. Finally we had data from 45 cohorts in 20 countries: one cohort from Africa (Guinea Bissau Twin Study), two cohorts from Australia (Peri/Postnatal Epigenetic Twins Study, and Queensland Twin Register), seven cohorts from East-Asia (Chinese National Twin Registry, Guangzhou Twin Eye Study, Japanese Twin Cohort, Mongolian Twin Registry, Qingdao Twin Registry of Children, South Korea Twin Registry, and West Japan Twins and Higher Order Multiple Births Registry), 20 cohorts from Europe (Adult Netherlands Twin Registry, Berlin Twin Register, Bielefeld Longitudinal Study of Adult Twins, Danish Twin Cohort, East Flanders Prospective Twin Survey, Finnish Older Twin Cohort, FinnTwin12, FinnTwin16, Gemini Study, Genesis 12–19 Study, Hungarian Twin Registry, Italian Twin Registry, Portugal Twin Cohort, Swedish Twin Cohorts, Swedish Young Male Twins Study of Adults, Swedish Young Male Twins Study of Children, TCHAD-study, Twins Early Developmental Study, TwinsUK, and Young Netherlands Twin Registry), two cohorts from Middle-East (Longitudinal Israeli Study of Twins and Turkish Twin Study) and 13 cohorts from North-America (Boston University Twin Project, California Twin Program, Colorado Twin Registry, Michigan Twins Study, Mid Atlantic Twin Registry, Minnesota Twin Family Study, NAS-NRC Twin Registry, Quebec Newborn Twin Study, SRI-international, Texas Twin Project, University of British Columbia Twin Project, University of Southern California Twin Study, and University of Washington Twin Registry). A more detailed description of the participating twin cohorts was presented previously.
During the course of the study, we found that the heritability estimates of height were substantially lower in the Chinese National Twin Registry than in other East-Asian cohorts, as also reported previously. Because of this heterogeneity, we did not include the data from the Chinese National Twin Registry in the reported analyses but tested how it would change the results in East-Asia. The final database comprised information on 86,037 different complete twin pairs with a total of 180,520 paired measurements (39% monozygotic (MZ), 34% same-sex dizygotic (SSDZ) and 27% opposite-sex dizygotic (OSDZ) twin pairs); that is, since some twin pairs have measurements at different ages, our database is based on measurement pairs. In order to analyze possible ethnic-cultural differences in the genetic and environmental contribution on height, cohorts were pooled in three groups according to their geographical and cultural characteristics: Europe (20 cohorts), North-America and Australia (15 cohorts) and East-Asia (6 cohorts) with 131,856, 29,856 and 17,924 paired measurements, respectively. In the additional analyses including the Chinese National Twin Registry, the number of pairs in East-Asia was 27,067. The cohort from Africa and the two from Middle-East were not included in these sub-analyses by geographic-cultural region because the data is too small to study these two areas separately.

Statistical analyses. To analyze genetic and environmental influences on the variation of height, we used classic twin modelling based on linear structural equations. Briefly, the analysis is based on the fact that MZ twins share the same gene sequence, whereas DZ twins share, on average, 50% of their genes identical-by-descent. On this basis, it is possible to divide the total variation of height into variance due to additive genetic effects (A: correlated 1.0 for MZ and 0.5 for DZ pairs), dominance genetic effects (D: 1.0 for MZ and 0.25 for DZ pairs), common (shared) environmental effects (C: by definition, correlated 1.0 for MZ and DZ pairs) and unique (non-shared) environmental effects (E: by definition, uncorrelated in MZ and DZ pairs). However, since our data included only twins reared together, we cannot simultaneously estimate common environmental and dominance genetic effects. All genetic models were fitted by the OpenMx package using the maximum likelihood method.

Prior to conducting the modelling, height values were adjusted using linear regressions and the resulting residuals were used as input phenotypes for the following analyses. Adjustment was carried out for birth year, exact age at the time of the measurement and study cohort within one-year and sex groups in univariate analyses, and for birth year and cohort in gene-environment interaction analyses.

The ACE sex-limitation model was selected as a starting point of the univariate modelling based on the following criteria: (i) MZ within-pair correlations were clearly higher than DZ correlations consistent with the influence of genetic effects, (ii) the magnitude of the difference between MZ and DZ correlations (rMZ > rDZ) indicated the presence of common environmental effects and (iii) the lower within-pair correlations for OSDZ than for SSDZ twins observed at most ages suggested the presence of sex-specific genetic effects (See Supplementary Table S4). Previous findings from this international database showed that both male and female DZ twins have greater height than MZ twins in these age groups, and thus different means for MZ and DZ twins were allowed. The fit of the univariate models for height at each one-year age group is shown in Supplementary Table S2. In the present study, the equal-environment assumption was tested by comparing the ACE model to the saturated model. The fit of the model after Bonferroni correction of multiple testing did not worsen at most ages, which suggested that the assumption of equality of variances between MZ and DZ twins was not violated. When fixing A, C and E parameters to be the same in boys and girls, the fit of the model was poorer at many ages, particularly from the beginning of adolescence, suggesting that these variance components differ between sexes. We additionally fitted a scale model allowing for different sizes of variance components but fixing the relative size of these components to be equal. Since this model also showed differences, we decided to present the results separately for boys and girls. Sex-specific genetic effects were considerable at some ages, and thus all modelling results are presented in sex-limited form for consistency. Finally, comparative model fitting revealed that the C parameter could be not excluded from the model without a significant (P < 0.001) deterioration in fit for all ages.

In order to study how age modifies the genetic and environmental variances of height, we additionally conducted gene-environment interaction modelling using age as an environmental modification factor. The advantage of the gene-environment interaction model, as compared to fitting a series of univariate models, is that it estimates a fewer number of parameters and thus has more statistical power to analyze the age patterns in genetic and environmental variances. The parameter estimates differed in boys and girls (Δ − 2LL = 407, Δd.f. = 15, p-value < 0.0001) and thus sex-specific models were conducted. Because the size of the sex-specific genetic effect varied according to age, only same-sex pairs were included in these age-moderation analyses. In addition to linear effects of age, quadratic age effects were also included on the variance components since they were highly statistically significant in boys (Δ − 2LL = 3298, Δd.f. = 3, p-value < 0.0001) and girls (Δ − 2LL = 4091, Δd.f. = 3, p-value < 0.0001).

References
1. Galton, F. Regression towards mediocrity in hereditary stature. *Journal of the Anthropological Institute* **15**, 246–262 (1886).
2. Pearson, K. & Lee, A. On the laws on inheritance in man. *Biometrika* **2**, 356–462 (1903).
3. Fisher, R. A. The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh* **52**, 399–433 (1918).
4. Silventoinen, K. Determinants of variation in adult body height. *J. Biosoc. Sci.* **35**, 263–285 (2003).
5. Perola, M. et al. Combined genome scan for body stature in 6,602 European twins: evidence for common Caucasian loci. *PloS Genet.* **3**, e97 (2007).
6. Cho, Y. S. et al. A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. *Nat. Genet.* **41**, 527–534 (2009).
7. Hao, Y. et al. Genome-wide association study in Han Chinese identifies three novel loci for human height. *Hum. Genet.* **132**, 681–689 (2013).
Acknowledgements

This study was conducted within the CODATwins project (Academy of Finland #266592). Support for participating twin projects: The Boston University Twin Project is funded by grants (#R01 HD066435 & #R01 MH062375) from the National Institutes of Health to K. Saudino. California Twin Program was supported by The California Tobacco-Related Disease Research Program (7RT-0134H, 8RT-0107H, 6RT-0354H) and the National Institutes of Health (1R01ES01510-01). Chinese National Twin Registry is funded by Special Fund for Health Scientific Research in the Public Welfare (Project No: 201502006), China. Colorado Twin Registry is funded by NIDA funded center grant DA011015, & Longitudinal Twin Study HD0333; Author HuiBregtse is supported by to 5T32DA017637-11. Danish Twin Registry is supported by the National Program for Research
Author Contributions

K.S., K.J., T.I.A.S., D.I.B., F.R., K.O.K., Y.-M.H. and Y.Y. planned the study design of the CODATwins project. A.B., C.K., K.J.S., K.I., W.C., A.E.H., T.M.M., W.G., C.Y., L.L., R.P.C., B.M.H., K.C., A.S., K.O.K., T.I.A.S., C.A.D., R.E.F., R.J.F.E.K., J.K., K.H., J.W., C.H.L., A.E., T.A.M., T.C.E., A.M.G., M.H., X.D., M.B.-A., H.B.-N., M.S., A.D.T., D.L.T., M.A.S., C.F., C.D., S.O., A.K.-N., D.M., L.A., S.A.B., K.L.K., H.H.M., I.J.E., L.J.S., R.F.K., M.M., S.P., M.G., D.A.B., D.I.B., G.W., T.C.E.M.v.B., M.B., C.H., J.M.C., R.S., D.L.F., J.A.M., F.J., F.N., Z.P., L.D., M.B., M.B., G.D., F.V., N.G.M., S.E.M., G.W.M., Y.-M.H., J.H.G., S.M., J.V.B.H., S.A., R.S., G.E.S., R.K., P.K.E.M., N.L.P., F.R., P.T., C.M.A.H., R.P., K.P.H., E.M.T.-D., S.Y.O., F.A., G.B., D.N., T.S., M.M., L.A.B., C.T., G.E.D., D.B. and Y.Y. collected the data used in this study. K.S. and A.J. were in charge of data management. A.J. conducted the analyses, wrote the first draft of the manuscript and has primary responsibility of the final content. All authors have commented the manuscript and read and approved the final version of the manuscript.
Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Jelenkovic, A. et al. Genetic and environmental influences on height from infancy to early adulthood: An individual-based pooled analysis of 45 twin cohorts. Sci. Rep. 6, 28496; doi: 10.1038/srep28496 (2016).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/