OXFORD

Human Molecular Genetics, 2015, Vol. 24, No. 25 7445–7449

doi: 10.1093/hmg/ddv443
Advance Access Publication Date: 22 October 2015
Association Studies Article

ASSOCIATION STUDIES ARTICLE

Genome-wide genetic homogeneity between sexes and populations for human height and body mass index

Jian Yang[1,2,†,*], Andrew Bakshi[1], Zhihong Zhu[1], Gibran Hemani[1,3], Anna A.E. Vinkhuyzen[1], Ilja M. Nolte[5], Jana V. van Vliet-Ostaptchouk[5,6], Harold Snieder[5], The Lifelines Cohort Study[7], Tonu Esko[8,9,10,11], Lili Milani[8], Reedik Mägi[8], Andres Metspalu[8,12], Anders Hamsten[13], Patrik K.E. Magnusson[14], Nancy L. Pedersen[14], Erik Ingelsson[15,16] and Peter M. Visscher[1,2,†]

[1]Queensland Brain Institute, The University of Queensland, Brisbane, QLD 4072, Australia, [2]The University of Queensland Diamantina Institute, The Translation Research Institute, Brisbane, QLD 4102, Australia, [3]MRC Integrative Epidemiology Unit (IEU) at the University of Bristol, School of Social and Community Medicine, Bristol BS8 1TH, UK, [4]MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Institute of Metabolic Science, Box 285, Cambridge Biomedical Campus, Cambridge CB2 0QQ, UK, [5]Department of Epidemiology, [6]Department of Endocrinology and [7]The LifeLines Cohort Study, University of Groningen, University Medical Center Groningen, Groningen 9700 RB, The Netherlands, [8]Estonian Genome Center, University of Tartu, Tartu 51006, Estonia, [9]Division of Endocrinology, Boston Children's Hospital, Cambridge, MA 02141, USA, [10]Program in Medical and Populational Genetics, Broad Institute, Cambridge, MA 02242, USA, [11]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA, [12]Institute of Molecular and Cell Biology, University of Tartu, Tartu 51010, Estonia, [13]Cardiovascular Genetics and Genomics Group, Atherosclerosis Research Unit, Department of Medicine Solna and [14]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, 17177 Stockholm, Sweden, [15]Molecular Epidemiology and Science for Life Laboratory, Department of Medical Sciences, Uppsala University, Uppsala, Sweden and [16]Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA

To whom correspondence should be addressed at: Queensland Brain Institute, The University of Queensland, QBI Building (#79), St Lucia, QLD 4072, Australia. Tel: +61 7 3346 6396; Fax: +61 7 3346 6312; Email: jian.yang@uq.edu.au

Abstract

Sex-specific genetic effects have been proposed to be an important source of variation for human complex traits. Here we use two distinct genome-wide methods to estimate the autosomal genetic correlation (r_g) between men and women for human height and body mass index (BMI), using individual-level (n = ~44 000) and summary-level (n = ~133 000) data from genome-wide association studies. Results are consistent and show that the between-sex genetic correlation is not significantly different from unity for both traits. In contrast, we find evidence of genetic heterogeneity between sexes for waist–hip ratio (r_g = ~0.7) and

[†]These authors jointly supervised this work.

Received: August 10, 2015. **Revised and Accepted:** October 19, 2015

between populations for BMI (r_g = ~0.9 between Europe and the USA) but not for height. The lack of evidence for substantial genetic heterogeneity for body size is consistent with empirical findings across traits and species.

Introduction

There has been a long-standing question about the contribution of autosomal sex-specific genetic effects to complex trait variation in humans (1,2). For many traits, including height, body mass index (BMI), waist circumference, hip circumference and waist-to-hip ratio, twin studies find systematically lower correlations of opposite-sex twin pairs compared with same-sex dizygotic twin pairs (3). For example, Schousboe *et al.* (4) concluded from a large study of European twin pairs aged 20–40 years that 'the sets of genes contributing to variation in BMI are not identical among men and women', with an estimate of genetic correlation between men and women <0.5 (averaged across cohorts and age groups). However, genome-wide association studies (GWAS) using very large sample sizes have not identified any autosomal genetic variant that showed a significant sex-specific effect on height or BMI (5). In addition, pedigree studies in other species consistently report very large genetic correlations in body size between males and females (6–10), suggesting that little phenotypic variation is contributed by genotype–sex interactions. In this study, we address the question about genotype-by-sex and genotype-by-country interaction variation using a powerful population design in which unrelated individuals have genome-wide single nucleotide polymorphism (SNP) data and a phenotypic measure on height and BMI. We estimate the correlation in genome-wide additive effects that are captured by SNP data with high precision.

Results

Using the bivariate genomic restricted maximum likelihood (GREML) method (11) as implemented in the genetic complex trait analysis (GCTA) software (12), we estimated the genetic correlation between men and women for height and BMI in a combined GWAS data set comprising 44 126 unrelated individuals of European descent and ~1.2 M SNPs (see Material and Methods section). Since the bivariate GCTA-GREML method does not require traits measured on the same set of individuals, we regarded height (or BMI) in men as a distinct trait from that in women and used genome-wide SNP data to estimate genetic correlation between the 'two traits' (e.g. men's height versus women's height). The expected value of the genetic correlation (r_g) should, per definition, range from −1 or 1. In the analysis, we allowed the estimate to go beyond this range so that the estimate was unbiased (an estimate constrained between −1 and 1 is a biased estimated). The estimate of r_g was 1.02 for height with a standard error (SE) of 0.023, and 1.01 (SE = 0.064) for BMI (Table 1), both of which were not significantly different from 1 (Wald's test P = 0.48 for height and P = 0.86 for BMI).

A recently proposed alternative method, called bivariate linkage disequilibrium score (LDSC) regression analysis (13), is also able to estimate genetic correlations between traits measured on different samples. This method is built on the univariate LDSC regression analysis (14), which was developed to distinguish the effects due to population stratification from polygenic effects in GWAS. The LDSC regression method only requires GWAS summary data so that it can be applied to data with very large sample sizes. We used summary data from the sex-stratified Genetic Investigation of ANthropometric Traits (GIANT) meta-analyses (5) (up to 60 586 men and 73 137 women) for height and BMI (see Material and Methods section). The estimate of r_g from the bivariate LDSC regression analysis using summary data was 0.957 (SE = 0.023) for height and 0.879 (SE = 0.035) for BMI (Table 2). The estimate for BMI was significantly different from 1 (Wald's test P = 5.9e-4), which was inconsistent with the result from bivariate GREML analysis on individual-level data. We therefore investigated what could cause this apparent discrepancy.

Unlike the GREML analysis where there were roughly the same number of men and women in each cohort (Table 3), the bivariate LDSC regression analysis used summary data from a meta-analysis where there were cohorts that only contributed samples of one sex, e.g. all > 20 000 samples from the Women's Genome Health Study (WGHS) were women (see the Supplementary Material, Table S1 of Randall *et al.* (5)). Therefore, if there is genetic heterogeneity between cohorts (populations), the estimate of r_g between sexes from the bivariate LDSC regression analysis will be biased downwards, and this may explain the discrepancy between the estimates of r_g from bivariate GREML and LDSC regression. We therefore used the bivariate GREML analysis to test whether there is genetic heterogeneity between populations in the combined GWAS sample with individual-level data (n = 44 126). We stratified the data (from both sexes) into two groups, i.e. samples from the USA (ARIC, GENEVA-T2D, HRS) and Europe (TwinGene, Lifelines and EGCUT), and then used bivariate GCTA-GREML approach to estimate the genetic correlation between USA and Europe for height and BMI. The estimates of between-population r_g were 0.961 (SE = 0.030) for height and 0.894 (SE = 0.058) for BMI (Table 4). These estimates are remarkably consistent with the between-sex r_g estimates of 0.957 for height and 0.879 for BMI (Table 2) from bivariate LDSC regression using summary data. The results are consistent with the hypothesis that there is genetic heterogeneity between populations for BMI that confounds the estimate of r_g between sexes in the bivariate LDSC regression analysis. The European

Table 1. Estimates of genetic correlation between men and women from the bivariate GCTA-GREML analysis using individual-level data for five anthropometric traits.

Trait	Sample size (men versus women)	h^2_g (Men)		h^2_g (Women)		r_g		
		Est.	SE	Est.	SE	Est.	SE	$P(r_g = 1)$
Height	19 095 versus 24 504	0.447	0.018	0.431	0.015	1.022	0.031	0.483
BMI	19 016 versus 24 350	0.236	0.019	0.226	0.015	1.011	0.064	0.859
WCadjBMI	13 158 versus 15 874	0.167	0.026	0.174	0.022	0.774	0.119	0.057
HIPadjBMI	13 119 versus 15 854	0.231	0.026	0.185	0.022	0.855	0.101	0.149
WHRadjBMI	13 115 versus 15 846	0.159	0.026	0.182	0.022	0.607	0.112	4.4×10^{-4}

h^2_g = proportion of phenotypic variance explained by all SNPs used in the analysis. $P(r_g = 1)$: Wald's test P-value against $r_g = 1$.

Table 2. Estimates of genetic correlation between men and women from the LDSC regression analysis using summary data for five anthropometric traits.

Trait	Sample size (men versus women)	h^2_g (Men)		h^2_g (Women)		r_g		
		Est.	SE	Est.	Est.	Est.	SE	$P(r_g = 1)$
Height	60 505 versus 73 073	0.274	0.018	0.261	0.018	0.957	0.023	0.063
BMI	58 599 versus 67 935	0.167	0.012	0.186	0.010	0.879	0.035	5.9×10^{-4}
WCadjBMI	38 361 versus 42 727	0.143	0.014	0.110	0.013	0.780	0.071	1.9×10^{-3}
HIPadjBMI	32 920 versus 40 712	0.162	0.018	0.136	0.015	1.000	0.083	0.999
WHRadjBMI	34 594 versus 47 463	0.102	0.016	0.093	0.017	0.770	0.108	0.033

h^2_g = proportion of phenotypic variance explained by all SNPs used in the analysis. HIPadjBMI, BMI-adjusted hip circumference; WCadjBMI, BMI-adjusted waist circumference; WHRadjBMI, BMI-adjusted waist–hip ratio. The samples size shown in this table is the median of the per-SNP sample sizes reported in the summary data. $P(r_g = 1)$: Wald's test P-value against $r_g = 1$.

Table 3. Sample size and genotyping platform of each GWAS cohort with individual-level data.

Cohort	Men	Women	Genotyping array
ARIC	3567	4040	Affymetrix 6.0
GENEVA-T2D	2055	3190	Affymetrix 6.0
HRS	3531	4948	Illumina OmniExpress
TwinGene	3299	3197	Illumina Omni2.5
Lifelines	4230	5959	Illumina OmniExpress
EGCUT	2641	3469	Illumina Cyto SNP12 v2

More information about QC criteria of the genotype data can be found in the Supplementary Material, Figure S4 of Yang *et al.* (15).

samples (TwinGene, Lifelines and EGCUT) used in our bivariate GREML analyses are from Sweden, the Netherland and Estonia, which are not a perfect representative of the whole Europe, whereas the US samples (ARIC, GENEVA-T2D, HRS) are of descent of almost all Europeans (see the PCA plot in the Supplementary Material, Fig. S19 of Yang *et al.* (15)). Therefore, the between-population genetic heterogeneity for BMI could be due to genetic difference (e.g. difference in causal variants or difference in linkage disequilibrium between SNPs and causal variants) across European populations. Of course, we could not rule out the possibility that there is a genotype-by-environment interaction for BMI, where population (Europe or USA) is a proxy for environmental difference.

We noticed that the estimates of proportion of variance explained by all SNPs (h^2_g) from the LDSC regression analysis (Table 2) were much smaller than those from bivariate GREML analysis for both height and BMI (Table 1). The discrepancy could be partly attributed to the error of approximating the LDSC in the sample used for analysis by the estimate from the 1000 Genome Project (1KGP) data because the regression slope (from which h^2_g is estimated in the LDSC regression analysis) is inversely proportional to the variance of the estimated LDSCs, which consists of the true variation of LDSCs and the variation of estimation errors. This was demonstrated by the analysis of applying the LDSC regression analysis in the same data that were used in the GREML analysis (see Material and Methods section), where the estimates of h^2_g from LDSC regression (Supplementary Material, Table S1) were consistently smaller than those from GREML (Table 1). However, the estimates of r_g are generally consistent between the two methods with the standard errors from LDSC regression being about twice larger than those from GREML (Table 1 and Supplementary Material, Table S1). Another explanation of the discrepancy between estimate of h^2_g from GREML using individual-level data (Table 1) and that from LDSC regression using summary data (Table 2) is that the summary statistics available in the public domain had been corrected by the genomic control (GC) approach (16), i.e. reported SE equals to original SE multiplied by $\sqrt{\lambda_{GC}}$ in each cohort and further in the whole meta-analysis sample, where λ_{GC} is the genomic inflation factor (median of χ^2 statistics divided by 0.455). The LDSC regression analysis was based on z-statistics that are shrunk by GC correction, resulting in an underestimated h^2_g. This is demonstrated by the decrease in the estimate of h^2_g from LDSC regression as a result of GC correction in the combined GWAS sample with individual-level data (Supplementary Material, Table S1). Fortunately, GC correction did not affect the estimate of r_g because the genetic variance and the genetic covariance were scaled by the same factor, which cancels out when calculating r_g (Supplementary Material, Table S1). We cannot rule out the possibility that the deflated estimates of h^2_g from LDSC regression using meta-analysis summary data were partly due to technical factors [e.g. difference in design, genotyping platform and/or quality control (QC) criteria] that led to artifactual discrepancies in SNP calls. In addition, summary data from meta-analysis based on a large number of cohorts may suffer from more experimental noise, such that the implicit h^2_g is lower than that from an equivalent sample with individual-level data. In addition, we used an IMPUTE-INFO threshold of 0.3 for SNP inclusion in the analysis of the combined GWAS data. We demonstrated by additional analysis that the GREML estimates using an IMPUTE-INFO threshold of 0.6 (Supplementary Material, Table S2) were highly consistent with those using 0.3 (Table 1).

We further estimated the between-sex genetic correlation for BMI-adjusted waist circumference (WCadjBMI), hip circumference (HIPadjBMI) and waist–hip ratio (WHRadjBMI) using individual-level data in the combined GWAS sample (bivariate GREML analysis, n = up to 29 091) and using summary-level data from the sex-specific GIANT meta-analysis (bivariate LDSC regression analysis, n = up to 82 057) (see Material and Methods section). The waist circumference, hip circumference and waist–hip ratio (WHR) phenotypes were adjusted for BMI to capture body fat distribution independent of overall adiposity (5). Although the standard errors of the estimates were relatively large, the results from both analyses suggested that there was a significant between-sex genetic heterogeneity for WHRadjBMI, $r_g = 0.607$ (SE = 0.112) from GREML analysis and $r_g = 0.770$ (SE = 0.108) from LDSC regression, consistent with WHR being a sexually dimorphic trait under differential selection for men versus women (17,18). Our results also seem to suggest that there is between-sex genetic heterogeneity for WCadjBMI but not for HIPadjBMI (Tables 1 and 2). These estimates are consistent with the GWAS results reported in the Randall *et al.* study that out of 9 traits analysed, only WHRadjBMI and WCadjBMI had significant sex-specific effects (detected at 7 loci) (5).

Table 4. Bivariate GCTA-GREML estimates of genetic correlation between samples from the USA and Europe for height and BMI.

Trait	Sample size (men versus women)	h^2_g (USA)		h^2_g (Europe)		r_g		
		Est.	SE	Est.	Est.	Est.	SE	$P(r_g = 1)$
Height	21 006 versus 22 593	0.419	0.017	0.475	0.015	0.961	0.030	0.195
BMI	20 904 versus 22 462	0.240	0.017	0.248	0.016	0.894	0.058	0.068

h^2_g = proportion of phenotypic variance explained by all SNPs used in the analysis. $P(r_g = 1)$: Wald's test P-value against $r_g = 1$.

Discussion

The search for genetic variants with sex-specific effects and thereby sex-genotype interaction variation is often motivated by large differences in average phenotypes (as in height) or biological processes that are known to differ between the sexes (e.g. hormonal effects). However, mean differences do not imply that sex-by-genotype interactions exist. For example, a recent study (19) reported a mean difference in genetic profile for height across European nations, consistent with the observed difference in mean height phenotype. This finding does not conflict with the result from a previous study (20) that there is no evidence for a difference in heritability between European populations or result from our study (Table 4) that the between-population genetic correlation for height is not significantly different from unity. One caveat of our study is that we estimate the genetic correlation from effects captured by common SNPs. It is possible that rare variants of large effects individually have a different effect in men and women. Nevertheless, it is difficult to see how the missing part of genetic (co)variance would cumulatively lead to a much lower overall genome-wide correlation, since common SNPs imputed to the 1KGP reference capture the majority of genetic variation for height and BMI (15).

In conclusion, we provide precise estimates of the genetic correlation between men and women for height and BMI using very large samples of unrelated individuals. The results are consistent with the hypothesis that there is no between-sex genetic heterogeneity for both height and BMI and that there is some between-population genetic heterogeneity for BMI, suggesting that obesity in men and women is influenced by the same set of autosomal genes in the same population, and the effect of these genes may vary across populations. The analyses performed in this study in principle can be applied to detect genome-wide heterogeneity between sexes, populations or environments for other complex traits and diseases in humans and even in other species. Results from these kinds of analyses provide important information and a baseline for investigators seeking to identify individual genetic variants or genes with sex-, population- and/or environment-specific effects for complex traits.

Material and Methods

Bivariate GCTA-GREML analysis

We used individual-level data from seven GWAS cohorts, i.e. ARIC, NHS, HPFS, TwinGene, HRS, EGCUT and Lifelines. Informed consent was obtained from all subjects. QCs of the genotype data and imputation have been detailed elsewhere (15). In brief, we performed QCs of the genotype data in each cohort following the criteria as listed in the Supplementary Material, Table S4 of Yang *et al.* (15), and imputed the genotype data to the 1KGP reference panels (21) using IMPUTE2 (22). After imputation, we further excluded SNPs with Hardy–Weinberg equilibrium (HWE) test $P < 1e\text{-}6$, minor allele count < 3 or IMPUTE-INFO (the metric outputted from IMPUTE2 software as a measure of imputation accuracy) < 0.3. We used an IMPUTE-INFO threshold of 0.3 because a recent study (15) suggests that the proportion of variance explained by imputed variants decreases substantially when filtering variants with an IMPUTE-INFO threshold of >0.3. We included only in the analysis the variants with MAF ≥ 0.01 and those in common with HapMap Project Phase 3 reference panels (HM3) because the HM3 SNPs were optimized to capture common genetic variation (23). We used the GCTA (12,24) to estimate the genetic relatedness between all possible pairs of samples from SNP data (~1.2 M HapMap3 SNPs passed QCs), removed one of each pair of individuals with estimated genetic relatedness > 0.05 and retained 44 126 unrelated individuals (19 323 men and 24 803 women). All the individuals are of European descent as demonstrated by the principal component analysis in the Supplementary Material, Figure S19 of Yang *et al.* (15). We adjusted the phenotypes for age using linear regression and standardized the residuals to z-score in each gender group of each cohort to remove differences in the mean or variance between gender groups or cohorts. We then used the bivariate GCTA-GREML approach (11,12) to estimate the genetic correlation between men and women for height and BMI. We also performed the bivariate GREML analysis for WCadjBMI, HIPadjBMI and WHRadjBMI, where WC, HIP and WHR phenotypes were adjusted by BMI in each sex group of each cohort by linear regression to capture body fat distribution independent of overall adiposity (5). The sample sizes for the analyses of WC, HIP and WHR were smaller than those for height and BMI (Table 1) because the WC and HIP phenotypes in the HRS and GENEVA-T2D cohorts were not available to us.

Bivariate LDSC regression analysis

We accessed the summary data from GWAS meta-analyses for height, BMI, WCadjBMI, HIPadjBMI and WHRadjBMI (n = up to 133 723). These analyses were stratified by sex, and the summary data are available in the public domain (see the URLs section). We used the bivariate LDSC regression method (13), which is able to estimate genetic correlation between two traits using GWAS summary data. The method is implemented in LDSC software (see the URLs section). The LDSC data (provided in accompany with the LDSC regression software tool) were estimated from the 1KGP data with a window size of ±1 Mb. For consistency, we included in the bivariate LDSC analysis the SNPs in common with those used in the bivariate GCTA-GREML analysis. We further removed SNPs with reported sample size < 10 000 and those that did not exist in the LDSC file provided by LDSC software. We retained ~0.95 M SNPs for LDSC regression analysis. In addition, for method comparison, we also performed the bivariate LDSC regression analysis in the combined GWAS data (n = 44 126), with or without GC correction of the summary statistics, for height, BMI, WCadjBMI, HIPadjBMI and WHRadjBMI.

URLs

GIANT summary data: https://www.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files;

GCTA software: http://cnsgenomics.com/software/gcta/index.html;

LDSC software: https://github.com/bulik/ldsc

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

We thank William G. Hill for helpful discussion. This study makes use of data from the database of Genotypes and Phenotypes (dbGaP) under accessions phs000090.v1.p1, phs000091.v2.p1 and phs000428.v1.p1, and the EGCUT, Lifelines and TwinGene (see the Supplementary Material for the full set of acknowledgements for these data).

Conflict of Interest statement. None declared.

Funding

This research was supported by the Australian National Health and Medical Research Council (APP1078037), the Australian Research Council (DP130102666) and the Sylvia & Charles Viertel Charitable Foundation.

References

1. Ober, C., Loisel, D.A. and Gilad, Y. (2008) Sex-specific genetic architecture of human disease. *Nat. Rev. Genet.*, **9**, 911–922.
2. Weiss, L.A., Pan, L., Abney, M. and Ober, C. (2006) The sex-specific genetic architecture of quantitative traits in humans. *Nat. Genet.*, **38**, 218–222.
3. Polderman, T.J., Benyamin, B., de Leeuw, C.A., Sullivan, P.F., van Bochoven, A., Visscher, P.M. and Posthuma, D. (2015) Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat. Genet.*, **47**, 702–709.
4. Schousboe, K., Willemsen, G., Kyvik, K.O., Mortensen, J., Boomsma, D.I., Cornes, B.K., Davis, C.J., Fagnani, C., Hjelmborg, J., Kaprio, J. *et al.* (2003) Sex differences in heritability of BMI: a comparative study of results from twin studies in eight countries. *Twin. Res.*, **6**, 409–421.
5. Randall, J.C., Winkler, T.W., Kutalik, Z., Berndt, S.I., Jackson, A. U., Monda, K.L., Kilpelainen, T.O., Esko, T., Magi, R., Li, S. *et al.* (2013) Sex-stratified genome-wide association studies including 270,000 individuals show sexual dimorphism in genetic loci for anthropometric traits. *PLoS Genet.*, **9**, e1003500.
6. Hanrahan, J.P. and Eisen, E.J. (1973) Sexual dimorphism and direct and maternal genetic effects on body weight in mice. *Theor. Appl. Genet.*, **43**, 39–45.
7. Lynch, M. and Walsh, B. (1998) *Genetics and Analysis of Quantitative Traits*. Sunderland, MA: Sinauer Associates.
8. Splan, R.K., Cundiff, L.V. and Van Vleck, L.D. (1998) Genetic parameters for sex-specific traits in beef cattle. *J. Anim. Sci.*, **76**, 2272–2278.
9. Poissant, J., Wilson, A.J. and Coltman, D.W. (2010) Sex-specific genetic variance and the evolution of sexual dimorphism: a systematic review of cross-sex genetic correlations. *Evolution*, **64**, 97–107.
10. Maniatis, G., Demiris, N., Kranis, A., Banos, G. and Kominakis, A. (2013) Genetic analysis of sexual dimorphism of body weight in broilers. *J. Appl. Genet.*, **54**, 61–70.
11. Lee, S.H., Yang, J., Goddard, M.E., Visscher, P.M. and Wray, N. R. (2012) Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics*, **28**, 2540–2542.
12. Yang, J., Lee, S.H., Goddard, M.E. and Visscher, P.M. (2011) GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.*, **88**, 76–82.
13. Bulik-Sullivan, B., Finucane, H.K., Anttila, V., Gusev, A., Day, F. R., Loh, P.R., ReproGen, C. and Psychiatric Genomics, C., Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control, C., Duncan, L. *et al.* (2015) An atlas of genetic correlations across human diseases and traits. *Nat. Genet.*, doi: 10.1038/ng.3406.
14. Bulik-Sullivan, B.K., Loh, P.R., Finucane, H.K., Ripke, S. and Yang, J., Schizophrenia Working Group of the Psychiatric Genomics, C., Patterson, N., Daly, M.J., Price, A.L. and Neale, B.M. (2015) LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.*, **47**, 291–295.
15. Yang, J., Bakshi, A., Zhu, Z., Hemani, G., Vinkhuyzen, A.A., Lee, S.H., Robinson, M.R., Perry, J.R., Nolte, I.M., van Vliet-Ostaptchouk, J.V. *et al.* (2015) Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat. Genet.*, **47**, 1114–1120.
16. Devlin, B. and Roeder, K. (1999) Genomic control for association studies. *Biometrics*, **55**, 997–1004.
17. Singh, D. (1993) Adaptive significance of female physical attractiveness: role of waist-to-hip ratio. *J. Pers. Soc. Psychol.*, **65**, 293–307.
18. Leong, A. (2006) Sexual dimorphism of the pelvic architecture: a struggling response to destructive and parsimonious forces by natural & mate selection. *McGill J. Med.*, **9**, 61–66.
19. Robinson, M.R., Hemani, G., Medina-Gomez, C., Mezzavilla, M., Esko, T., Shakhbazov, K., Powell, J.E., Vinkhuyzen, A., Berndt, S.I., Gustafsson, S. *et al.* (2015) Population genetic differentiation of height and body mass index across Europe. *Nat. Genet.*, doi: 10.1038/ng.3401.
20. Wood, A.R., Esko, T., Yang, J., Vedantam, S., Pers, T.H., Gustafsson, S., Chu, A.Y., Estrada, K., Luan, J., Kutalik, Z. *et al.* (2014) Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.*, **46**, 1173–1186.
21. The 1000 Genomes Project Consortium. (2010) A map of human genome variation from population-scale sequencing. *Nature*, **467**, 1061–1073.
22. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. and Abecasis, G.R. (2012) Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat. Genet.*, **44**, 955–959.
23. Altshuler, D.M., Gibbs, R.A., Peltonen, L., Altshuler, D.M., Gibbs, R.A., Peltonen, L., Dermitzakis, E., Schaffner, S.F., Yu, F., Peltonen, L. *et al.* (2010) Integrating common and rare genetic variation in diverse human populations. *Nature*, **467**, 52–58.
24. Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W. *et al.* (2010) Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.*, **42**, 565–569.