Author(s): Sarah E. Medland, Tetyana Zayats, Beate Glaser, Dale R. Nyholt, Scott D. Gordon, Margaret J. Wright, Grant W. Montgomery, Megan J. Campbell, Anjali K. Henders, Nicholas J. Timpson, Leena Peltonen, Dieter Wolke, Susan M. Ring, Panos Deloukas, Nicholas G. Martin, George Davey Smith and David M. Evans,
Article Title: A Variant in LIN28B Is Associated with 2D:4D Finger-Length Ratio, a Putative Retrospective Biomarker of Prenatal Testosterone Exposure
Year of publication: 2010
Link to published article:
http://dx.doi.org/ 10.1016/j.ajhg.2010.02.017

Publisher statement: S. E. Medland et al. (2010). : A Variant in LIN28B Is Associated with 2D:4D Finger-Length Ratio, a Putative Retrospective Biomarker of Prenatal Testosterone Exposure. The American Journal of Human Genetics, Vol. 86(4), pp. 519-525

A variant in *LIN28B* is associated with 2D:4D finger length ratio a putative retrospective biomarker of prenatal testosterone exposure

Sarah E. Medland[1†], Tetyana Zayats[2,3†], Beate Glaser[2,3], Dale R. Nyholt[1], Scott D. Gordon[1], Margaret J. Wright[1], Grant W. Montgomery[1], Megan J. Campbell[1], Anjali K. Henders[1], Nicholas J. Timpson[2,3], Leena Peltonen[4,5,6,7], Dieter Wolke[8], Susan M. Ring[3], Panos Deloukas[4], Nicholas G. Martin[1], George Davey Smith[2,3], David M Evans[2,3*]

[1]Genetic Epidemiology, Queensland Institute of Medical Research, Australia

[2]MRC Centre for Causal Analyses in Translational Epidemiology, University of Bristol, United Kingdom

[3]Department of Social Medicine, University of Bristol, Bristol, United Kingdom

[4]Wellcome Trust Sanger, Institute, Cambridge CB10 1SA, UK

[5]Biomedicum Helsinki, Research Program in Molecular Medicine, University of Helsinki, Finland

[6]Department of Molecular Medicine, National Public Health Institute, Helsinki, Finland

[7]The Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

[8]Department of Psychology and Health Sciences, University of Warwick, Coventry, UK

Running Title: A LIN28B variant associates with 2D:4D

[†]These authors contributed equally to this work

[*]Address Correspondence to: David M. Evans. MRC Centre for Causal Analyses in Translational Epidemiology, Department of Social Medicine, University of Bristol, Oakfield House, Oakfield Grove, Bristol, BS8 2BN, United Kingdom. Tel: +44 (0)117 3310094, Fax: +44 (0)117 3310123. Email: dave.evans@bristol.ac.uk.

Abstract

The ratio of the lengths of an individual's second to fourth digit (2D:4D) is commonly used as a non-invasive retrospective biomarker for prenatal androgen exposure. In order to identify the genetic determinants of 2D:4D, we applied a genome-wide association approach to 1507 eleven year old children from the Avon Longitudinal Study of Parents and Children in whom 2D:4D ratio had been measured, as well as a sample of 1382, 12 to 16 year olds from the Brisbane Adolescent twin study. A meta-analysis of the two scans identified a single variant in the *LIN28B* gene that was strongly associated with 2D:4D (rs314277: $p = 4.1 \times 10^{-8}$), and which was subsequently independently replicated in a further 3659 children from the ALSPAC cohort ($p = 1.53 \times 10^{-6}$). The minor allele of the rs314277 variant has previously been linked to increased height and delayed age at menarche, but in our study was associated with increased 2D:4D in the opposite direction to previous reports on the correlation between 2D:4D and age at menarche. Our findings call into question the validity of 2D:4D as a simplistic retrospective biomarker for prenatal testosterone exposure.

Introduction

The ratio of the lengths of the second to fourth digits (2D:4D) is a sexually dimorphic trait that is on average a quarter of a standard deviation lower in males than females.[1] First identified by Ecker in 1875;[2] the measure was rediscovered by Wilson in the early 1980s,[3] and subsequently, Manning who hypothesized that the ratio reflected prenatal androgen exposure.[4,5] Consistent with this theory, sex differences in 2D:4D develop prenatally[6] and remain relatively stable across the lifespan[7]. Given the practical and ethical difficulties inherent in measuring testosterone exposure in the developing fetus, many researchers have adopted 2D:4D as a non-invasive retrospective biomarker for prenatal androgen exposure, although its use as such is controversial (see McIntyre 2006 for a review).[7] Despite this controversy, over 300 papers have been published using this measure in the last 10 years,[8] and 2D:4D has been shown to correlate with a wide range of diseases and physiological and psychological traits including autism,[9] attention deficit disorder, [10] fertility,[11] myocardial infarction,[12] visuo-spatial ability,[13] homosexuality,[14,15] athletic performance[16] and age at menarche.[17]

2D:4D is highly heritable with additive genetic effects explaining ~60% of the phenotypic variance,[18-21]. Although the results from one twin study suggested that female twins exhibit higher heritabilities than males for left hand 2D:4D,[18] a larger study failed to find any significant sex limitation or differences in the magnitude of heritability between the sexes for 2D:4D traits.[19] There has also been little progress identifying the individual variants underlying this genetic variation. Following the hypothesis that 2D:4D reflects prenatal exposure to testosterone, Manning et al[22] examined the association between the number of CAG repeats at the Androgen Receptor locus and 2D:4D in males (N=51) and reported a significant correlation between CAGn and 2D:4D for the right (r = 0.29) but not the left hand

(r = 0.005). To the best of our knowledge, no other genetic association studies have been performed using this trait to date.

In order to identify genetic determinants underlying variation in 2D:4D, we applied a genome-wide association approach to 1507 children from the Avon Longitudinal Study of Parents and Children (ALSPAC), a large population-based cohort in which 2D:4D had been measured at ~11 years of age,[23] as well as a sample of 1382 twelve, fourteen and sixteen year old twins and their singleton siblings from the Queensland Institute of Medical Research, Brisbane Adolescent twin study (QIMR, Australia).[19,24]

Subjects and Methods

Participants

ALSPAC is a population-based birth cohort study consisting of over 13,000 women and their children recruited from the county of Avon, UK in the early 1990s[23]. Both mothers and children have been extensively followed from the 8th gestational week onwards using a combination of self-reported questionnaires, medical records and physical examinations. Biological samples including DNA have been collected for 10,121 of the children from this cohort. The discovery sample reported in this study concerns 1507 children who had their 2D:4D measured at 11 years of age (mean = 11.75 years), and for whom genome-wide SNP typing had been performed.[25] The replication sample consisted of a further 5129 children from ALSPAC who were not part of the initial discovery cohort.

Participants in the QIMR Brisbane Adolescent Twin study were recruited from the general population, in the context of ongoing studies of melanoma risk factors and studies of cognition.[26] Twins and their singleton siblings were enlisted by contacting the principals of primary schools in the greater Brisbane area, media appeals and by word of mouth. It is estimated that approximately 50% of the eligible birth cohort were recruited into the study, which began in 1992. Digit ratios were available for 1382 individuals with genome-wide association data from 671 families (comprising 169 singletons; 332 sibling-pairs; 135 sibling-trios; 31 sibling-quads and 4 sibling-quins). Age range for the sample was 11- 24 (mean = 15.46, SD = 3.27).

In both samples participants' hands were photocopied during a clinical visit, and measurements of the second and fourth fingers were taken from the photocopies using digital calipers (accurate to .1 mm). The 2D:4D was calculated as the length of the second digit

divided by the length of the fourth digit multiplied by 100 so as to avoid computational difficulties due to the low variance of the trait. In both samples, the measure was normally distributed so no further transformation was required.

In ALSPAC, a random sample of 57 right and 48 left hands were measured in vivo to establish the validity of using a photocopy measurement to assess 2D:4D. Similarly in the QIMR cohort, 680 hands were measured twice from the same hand photocopy, once by hand using digitial calipers and once using a computer assisted measurement program and the reliability between measurement occasions calculated.

Childrens' standing height in ALSPAC was measured using a Harpenden Stadiometer. Both studies were performed with the approval of the appropriate ethics committees and informed consent of all participants and their parents.

Genotyping

One thousand five hundred and forty-three ALSPAC children were initially genotyped at 317,504 SNPs on the Illumina HumanHap317K SNP chip. Individuals exhibiting cryptic relatedness, non-European ancestry, high genome-wide heterozygosity and/or missing rates were removed from analyses as described previously,[25] leaving 1507 individuals in the analysis who had been measured for 2D:4D. Markers with minor allele frequency <1%, SNPs with >5% missing genotypes and any marker that failed an exact test of Hardy–Weinberg equilibrium ($P < 5x10^{-7}$) were excluded from further analyses leaving 310,613 SNPs that passed quality control.

The QIMR participants analyzed here were genotyped on the Illumina Human610-Quad SNP chip. These samples were genotyped in the context of a larger genome-wide association study (GWAS) which resulted in the genotyping of 16,140 individuals[26] using the Illumina 317, 370 and 610 SNP chips. Genotype data were screened for genotyping quality (GenCall < 0.7), SNP and individual call rates (< 0.95), HWE failure ($P < 10^{-6}$) and MAF (< 0.01). As these samples were genotyped in the context of a larger project the data were integrated with the larger QIMR genotype project and the data were checked for pedigree, sex and Mendelian errors and for non-European ancestry. As the QIMR genotyping project included data from the 317, 370 and 610 chip sets, to avoid introducing bias to the imputed data, a set of SNPs common to the three genotyping platforms was used for imputation (N = 274,604).

Follow-up genotyping of two SNPs in the *LIN28B* gene (rs314277 and rs314276) was carried out in a further 5129 individuals from the ALSPAC cohort by K-Biosciences, who employ a novel form of competitive allele specific PCR (KASPar) and TaqmanTM system for genotyping. The rs314277 SNP was chosen for replication because it showed maximum

association in the discovery cohort, whilst rs314276 was chosen because it had previously shown association with traits correlated with 2D:4D like age of menarche.[27]

Statistical Analyses

As ALSPAC and QIMR samples were genotyped on different arrays, consensus autosomal genotypic data were imputed using Markov Chain Haplotyping software (MaCH) with phased data from CEU individuals from release 22 of the HapMap project as the reference set of haplotypes. Only SNPs that could be imputed with relatively high confidence ($R^2 > 0.3$) and had a MAF >1% were used in subsequent analyses. In the ALSPAC cohort, association analysis of imputed SNPs was performed assuming an underlying additive model using the software package MACH2QTL which accounts for uncertainty in prediction of the imputed data by weighting genotypes by their estimated posterior probabilities. In the QIMR twins study, the most likely genotypes were imputed at each locus and these genotypes were subsequently analyzed using MERLIN.[28] Markers at physically genotyped loci on the X chromosome were analyzed using PLINK in the ALSPAC sample[29] and MINX in the QIMR cohort.[28] SNPs were tested for association with right 2D:4D, left 2D:4D and the mean of left and right. All analyses included sex as a covariate. Results for the two cohorts were then combined using fixed effects inverse variance meta-analysis. We also performed a chi-square test for heterogeneity to test whether the regression coefficients differed significantly between males and females for the regression of 2D:4D for all SNPs in the discovery GWAS (i.e. a test for additive genotype x sex interaction). P values for the tests of heterogeneity were combined across ALSPAC and QIMR cohorts using the software package METAL to produce an overall level of significance. Association in the replication sample was performed in PLINK via linear regression assuming an underlying additive genetic model.

Results

The Pearson's product moment correlation between the in vivo and photocopied measurements of the length of the second and fourth digits was high for the right and left hands (All r > 0.97). The correlations between finger lengths using digital calipers and using a computer assisted measurement program were similarly high (All r > .96) suggesting that our measurements have a high degree of repeatability.

Table 1 displays the mean and standard deviation of the 2D:4D measurements for the left, right and mean of the hands in the ALSPAC and QIMR discovery cohorts, and also in the ALSPAC replication sample. As expected, across samples and hands, mean 2D:4D was higher for females than males. QQ plots for the genome-wide association scan of ALSPAC individuals, the GWAS of the QIMR twins, and the combined meta-analyses are presented in Supplementary Figure 1. Both plots indicate that the observed GWAS test statistics lie close to expectation and suggest that potential technical and stratification artifacts had negligible impact on the results. Consistent with this interpretation, the genomic inflation factors in both the ALSPAC ($\lambda = 1.01$), QIMR ($\lambda = 1.01$) and meta-analyzed samples ($\lambda = 1.00$) indicate little inflation of the association test statistics. We also checked whether p values derived from markers on the X chromosome might exhibit a more dramatic deviation from the null hypothesis of no association than the genome as a whole. QQ plots of the combined meta-analyses for left, right and mean 2D:4D showed that there was little evidence that this was the case, although a few markers on the X chromosome did exceed null expectations for left 2D:4D (Supplementary Figure 2).

The genome-wide association results for the combined meta-analysis of left, right and mean 2D:4D are presented in Supplementary Figures 3 through 5. A single SNP, rs314277, in the *LIN28B* gene (Figure 1) reached genome-wide significance for mean 2D:4D (p = 4.1 x

10^{-8}) as well as suggestive significance for left ($p = 1.5 \times 10^{-6}$) and right 2D:4D (8.2×10^{-7}). Each copy of the minor allele was associated with a 0.6 increase in mean 2D:4D. No other imputed or genotyped SNP in this region met the criterion for genome-wide significance, although several SNPs including rs314276, which had shown association with pubertal development and height in previous studies,[27] showed nominal evidence of association with 2D:4D (left: $p = 7.9 \times 10^{-5}$; right: $p = 1 \times 10^{-3}$; mean: $p = 5.4 \times 10^{-5}$). Interestingly, although the recombination rate 50kB either side of rs314277 was low, there were few SNPs in appreciable r^2 with the marker, which may explain why the next most associated SNP had p values at least two orders of magnitude lower. Conditioning on rs314277 in both the ALSPAC and QIMR datasets reduced the signal at the surrounding loci but did not completely abolish all evidence of association (ALSPAC best p = 0.0021; QIMR best p = 0.0069), suggesting either the existence of a more strongly associated variant in the region that had not been imputed, or perhaps a smaller second signal independent of rs314277 associated with 2D:4D in these data.

There was close correspondence in the top most hits between left and right 2D:4D, which is not surprising given the moderate to high phenotypic correlation between these variables (ALSPAC: r = 0.69; QIMR r = 0.56). We therefore only present the results for mean 2D:4D in the main text. Supplementary Tables 1 to 3 list all SNPs with a combined p value of less than 1×10^{-5} for the meta-analysis of left, right and mean 2D:4D. These included SNPs in the genes *SMOC1*, *SOX7*, *SORBS2*, *GLIS1*, *EFNA1*, *ZNF695*, *VAV3* and *NEDD4L*. Of particular note were SNPs on chromosome 8 in the *SOX7* gene which were strongly associated with 2D:4D in the ALSPAC dataset, but not in the QIMR cohort. Four SNPs located within the androgen receptor gene and six SNPs in high linkage disequilibrium (LD) with these displayed little evidence of association with the 2D:4D measures (best p = 0.04 for rs4456006). Similarly, there was no strong evidence for association between SNPs in the

aromatase gene (*CYP19A1*) and any of the 2D:4D measures (all $p > 0.01$). Additionally, there were no large signals present in the *HOXD* cluster of genes (best: rs2857533 $p = 0.0015$), nor in the *HOXA* cluster (all $p > .05$), which play important roles in limb development.

We also performed a chi-square test for heterogeneity to test whether the regression coefficients differed significantly between males and females. No variants exhibited significant differences in the magnitude of the regression coefficients between the sexes, although this may be partially a consequence of the low power of these tests to detect interactions. A complete list of variants with p values $< 10^{-5}$ is displayed in Supplementary Tables 4 to 6.

We attempted to replicate the *LIN28B* association by genotyping both rs314277 and rs314276 in a further 5129 children from the ALSPAC cohort (Table 2). Whilst rs314277 was strongly associated with 2D:4D in the replication cohort ($\beta = 0.44$; 95% CI: 0.26 – 0.62; $p = 1.53 \times 10^{-6}$), rs314276 only showed nominal association ($p = 2.26 \times 10^{-4}$). Indeed conditioning on rs314277 suggested that the association at rs314276 could be entirely explained by the signal at rs314277 ($p = 0.202$). The association between rs314277 and 2D:4D also remained after conditioning on height ($p = 1.64 \times 10^{-6}$). There was no evidence for interaction between genotype at rs314277 and gender in the replication cohort ($p = 0.31$).

As 2D:4D has been hypothesized to reflect testosterone exposure *in utero*, we investigated a possible relationship between mother's genotype at rs314277 and offspring 2D:4D. Whilst there was some evidence for a relationship between maternal genotype and mean 2D:4D in offspring ($p = 0.05$), any hint of association disappeared after conditioning on the child's genotype ($p = 0.99$) suggesting that mother's genotype at rs314277 did not directly influence child 2D:4D.

Finally, given the previous association between rs314277 and height, we explored whether any of the confirmed variants from eleven different genome-wide association studies

of height and four genome-wide studies of age of menarche were associated with 2D:4D from our meta-analysis using the online GWAS catalogue (see URL). Apart from SNPs located at 6q21, the only other variants which showed nominal association with 2D:4D were rs4932217 near the gene *POLG* ($p = 0.01$), and rs2292303 in the gene *NUP37* ($p = 0.05$). Looking across the loci from all these studies, there did not appear to be an overlap in the direction of effect (i.e. variants associated with increased height did not generally appear to be associated with variants that increased (or conversely decreased) 2D:4D). Similarly, SNPs in the 9q31 region which have previously been associated with age of menarche[30] did not show association with mean 2D:4D either (all $p > 0.05$). A complete list of variants previously associated with height and age at menarche are presented in Supplementary Table 7 along with results from the current combined meta-analysis of mean 2D:4D.

Discussion

The ratio of an individual's second to fourth digit (2D:4D) is sexually dimorphic and has been frequently used as a non-invasive retrospective biomarker for prenatal androgen exposure.[8] 2D:4D has been shown to correlate with a wide range of diseases and physiological and psychological traits including autism,[9] attention deficit disorder, [10] fertility,[11] myocardial infarction,[12] visuo-spatial ability,[13] homosexuality,[14,15] athletic performance[16] and age at menarche.[17]In this study we identified a variant, rs314277, within intron two of the *LIN28B* gene that was robustly associated with 2D:4D.

LIN28B is the human orthologue of a gene that regulates developmental timing in *Caenorhabditis elegans*. Its product, LIN28-B, is an RNA binding protein that interacts directly with let-7 precursors preventing their processing to become mature miRNAs.[31] Polymorphisms within *LIN28B* have previously been associated with height[32] and age at menarche in girls.[27,33,34] Most recently, Viswanathan et al. demonstrated that activation of *LIN28B* promotes neoplastic transformation and is associated with aggressive forms of human malignancy.[35] Although it is unclear at present how variation in *LIN28B* might influence 2D:4D, it is noteworthy that the gene is highly expressed in the testis, placenta and fetal liver, [33,36] suggesting that *LIN28B* might influence 2D:4D early in development.

The association between rs314277 in *LIN28B* and 2D:4D is particularly interesting because the minor allele has previously been associated with increased height[32] and delayed menarche in females.[34] The direction of effect in these studies is consistent with the overall correlation between the variables, since girls who experience menarche later tend to be taller as adults than girls who reach puberty earlier.[37] Similarly, a recent study of 2D:4D and age at menarche found that girls with low right-hand 2D:4D ratios (although not left-hand 2D:4D) tended to experience delayed menarche.[17] In the present study, however, the minor allele at

rs314277 was associated with *increased* 2D:4D ratio, which is opposite to the direction predicted by these earlier reports. It is unclear why this might be the case, but suggests that the relationship between 2D:4D, age at menarche and height is complex. Given that the association between rs314277 and 2D:4D remained after conditioning on height, and given that no other SNPs which displayed association with height or age of menarche showed convincing evidence of association with 2D:4D, it is unlikely that the effect of rs314277 is mediated through these other variables. It is interesting to note that LIN28B has two isoforms distinguished by the presence or truncation of a conserved cold-shock domain which influences protein function.[38] It is possible that different isoforms of LIN28B and hence different biological pathways might influence 2D:4D, height and age at menarche helping explain the interesting pattern of correlations.

Our study is not the only report to have questioned the specificity of 2D:4D as a proxy for androgen exposure. In fact the relationship between 2D:4D and prenatal androgen levels has only been directly examined in one small sample (N=33), which found no significant relationship between testosterone levels within the amniotic fluid (at ~17 weeks gestation) and 2D:4D (at 24 months), but did find a correlation between 2D:4D and the ratio of fetal testosterone to estradiol,[4,39] suggesting that the relationship between 2D:4D and circulating sex steroids might be more complex than previously hypothesized. In this study we did not find evidence of a relationship between genetic variants in or near the androgen receptor and 2D:4D although it is unclear how well the SNPs in this study might tag CAG repeat expansion at the androgen receptor.

In conclusion, we have demonstrated that a variant within the *LIN28B* gene is associated with 2D:4D and important in the early development of the hands. The same variant is associated with height and age at menarche, but the direction of the association with 2D:4D was in the opposite direction to that predicted from an earlier study of 2D:4D and age

at menarche. Our result suggests that the relationship between 2D:4D, early development and fetal androgen exposure is likely to be more complex than previously appreciated.

Acknowledgments

We are extremely grateful to all the families who took part in the ALSPAC study, the midwives for their help in recruiting them and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. We thank John Manning for advice with the set up of the 2D:4D measurement, and Jeremy Horwood and Andrea Waylen for help with implementation in the assessment clinics. The UK Medical Research Council, the Wellcome Trust and the University of Bristol provide core support for ALSPAC and the MRC CAiTE (MRC 90600705). We thank the Sample Logistics and Genotyping Facilities at the Wellcome Trust Sanger Institute for generating the ALSPAC GWA data. DME and this work were supported by a Medical Research Council New Investigator Award (MRC G0800582 to D.M.E).

We thank the twins and their families from the QIMR Brisbane Adolescent Twin study for their participation; Dixie Statham, Ann Eldridge, Marlene Grace, Kerrie McAloney (sample collection); Lisa Bowdler, Sara Smith, Steven Crooks (DNA processing); David Smyth, Harry Beeby, Daniel Park (IT support). We thank the Australian National Health and Medical Research Council (NHMRC, grants 241944, 339462, 389927, 389875, 389891, 389892, 389938, 443036, 442915, 442981, 496739, 552485, 552498), and Australian Research Council (A7960034, A79906588, A79801419, DP0212016, DP0343921) for funding the QIMR Brisbane Adolescent Twin study. Statistical analyses were carried out on the Genetic Cluster Computer which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003). SEM, DRN and GWM are supported by the NHMRC Fellowship Scheme. This publication is the work of the authors and they will serve as guarantors for the contents of this paper.

Web Resources

The URLs for data presented herein are as follows:

Genetic Cluster Computer, http://www.geneticcluster.org

KBIOSCIENCES, http://www.kbioscience.co.uk

MACH, http://www.sph.umich.edu/csg/abecasis/MACH/

MACH2QTL, http://www.sph.umich.edu/csg /abecasis/MACH/

METAL, http://www.sph.umich.edu/csg/abecasis/metal/

Online GWAS Catalogue, http://www.genome.gov/gwastudies/. Accessed 15 January 2010.

References

1. Phelps VR (1952) Relative finger lengths as a sex-influenced trait in man. Amer J Hum Genet 4:72-89

2. Ecker A (1875) Einige Bemerkungen über einen schwankenden charakter in der hand des menschen [Some remarks on a varying property of the human hand]. Archiv fur Anthroplogie 8:67-74

3 Wilson G.D. (1983) Finger length as an index of assertiveness in women. Personal. Individ. Diff. 4:111-112

4. Manning JT (2002) Digit Ratio:A pointer to fertility, behavior, and health. Rutgers University Press, New Brunswick

5. Manning JT, Scutt D, Wilson J, Lewis-Jones DI (1998) The ratio of 2nd to 4th digit length: a predictor of sperm numbers and concentrations of testosterone, luteinizing hormone and oestrogen. Hum Reprod 13:3000-3004

6. Malas MA, Dogan S, Evcil EH, Desdicioglu K (2006) Fetal development of the hand, digits and digit ratio (2D:4D). Early Hum Dev 82:469-475

7. McIntyre MH, Ellison PT, Lieberman DE, Demerath E, Towne B (2005) The development of sex differences in digital formula from infancy in the Fels Longitudinal Study. Proc

Biol Sci 272:1473-1479

8. Voracek M, Loibl LM (2009) Scientometric analysis and bibliography of digit ratio (2D:4D) research, 1998-2008. Psychol Rep 104:922-956

9. Manning JT, Baron-Cohen S, Wheelwright S, Sanders G (2001) The 2nd to 4th digit ratio and autism. Dev Med Child Neurol, 43:160-4

10. Stevenson JC, Everson PM, Williams DC, Hipskind G, Grimes M, Mahoney ER. (2007) Attention deficit/hyperactivity disorder (ADHD) symptoms and digit ratios in a college sample. Am J Hum Biol 19:41-50

11. Bang AK, Carlsen E, Holm M, Petersen JH, Skakkebaek NE, Jørgensen N (2005) A study of finger lengths, semen quality and sex hormones in 360 young men from the general Danish population. Hum Reprod, 20:3109-13

12 Manning JT and Bundred PE. (2000) The ratio of 2nd–4th digit length: a new predictor of disease predisposition? Med Hypotheses 54:855–7

13. Collaer ML, Reimers S, Manning JT (2007) Visuospatial performance on an internet line judgment task and potential hormonal markers: sex, sexual orientation, and 2D:4D. Arch Sex Behav 36:177-192

14. Manning JT, Robinson SJ (2003) 2nd to 4th digit ratio and a universal mean for prenatal testosterone in homosexual men. Med Hypotheses 61:303-306

15. Rahman Q and Wilson GD (2003) Sexual orientation and the 2^{nd} to 4^{th} finger length ratio: evidence for organising effects of sex hormones or developmental instability? Psychoneuroendocrinology 28:288-303

16. Honekopp J, Manning JT, Muller C (2006) Digit ratio (2D:4D) and physical fitness in males and females: Evidence for effects of prenatal androgens on sexually selected traits. Horm Behav 49:545-549

17. Matchock RL (2008) Low digit ratio (2D:4D) is associated with delayed menarche. Am J

Hum Biol 20:487-489

18. Gobrogge KL, Breedlove SM, Klump KL (2008) Genetic and environmental influences on 2D:4D finger length ratios: a study of monozygotic and dizygotic male and female twins. Arch Sex Behav 37:112-118

19. Medland SE, Loehlin JC (2008) Multivariate genetic analyses of the 2D:4D ratio: examining the effects of hand and measurement technique in data from 757 twin families. Twin Res Hum Genet 11:335-341

20. Paul SN, Kato BS, Cherkas LF, Andrew T, Spector TD (2006) Heritability of the second to fourth digit ratio (2D:4D): A twin study. Twin Research and Human Genetics 9:215-219

21. Voracek M, Dressler SG (2007) Digit ratio (2D:4D) in twins: Heritability estimates and evidence for a masculinized trait expression in women from opposite-sex pairs. Psychological Reports 100:115-126

22. Manning JT, Bundred PE, Newton DJ, Flanagan BF (2003) The 2nd to 4th digit ratio and variation in the androgen receptor gene. Evolution and Human Behavior 24:399-405

23. Golding J, Pembrey M, Jones R (2001) ALSPAC--the Avon Longitudinal Study of Parents and Children. I. Study methodology. Paediatr Perinat Epidemiol 15:74-87

24. Wright MJ, Martin N (2004) Brisbane Adolescent Twin Study: outline of study methods and research projects. Australian Journal of Psychology 56:65-78

25. Timpson NJ, Tobias JH, Richards JB, Soranzo N, Duncan EL, Sims AM, Whittaker P, Kumanduri V, Zhai G, Glaser B, et al (2009) Common variants in the region around Osterix are associated with bone mineral density and growth in childhood. Hum Mol Genet 18:1510-1517

26. Medland S, Nyholt D, Painter J, McEvoy B, McRae A, Zhu G, Gordon S, Wray N, Ferreira M, Wright M, et al (Submitted) Common variants in the Trichohyalin gene

are associated with straight hair in Europeans. American Journal of Human Genetics.

27. Ong KK, Elks CE, Li S, Zhao JH, Luan J, Andersen LB, Bingham SA, Brage S, Smith GD, Ekelund U, et al (2009) Genetic variation in LIN28B is associated with the timing of puberty. Nat Genet 41:729-733

28. Abecasis GR, Cherny SS, Cookson WO, and Cardon LR (2002) Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 30:97-101

29. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, et al (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559-575

30. Perry JR, Stolk L, Franceschini N, Lunetta KL, Zhai G, McArdle PF, Smith AV, Aspelund T, Bandinelli S, Boerwinkle E, et al (2009) Meta-analysis of genome-wide association data identifies two loci influencing age at menarche. Nat Genet 41:648-650

31. Viswanathan SR, Daley GQ, Gregory RI (2008) Selective blockade of microRNA processing by Lin28. Science 320:97-100

32. Lettre G, Jackson AU, Gieger C, Schumacher FR, Berndt SI, Sanna S, Eyheramendy S, Voight BF, Butler JL, Guiducci C, et al (2008) Identification of ten loci associated with height highlights new biological pathways in human growth. Nat Genet 40:584-591

33. Sulem P, Gudbjartsson DF, Rafnar T, Holm H, Olafsdottir EJ, Olafsdottir GH, Jonsson T, Alexandersen P, Feenstra B, Boyd HA, et al (2009) Genome-wide association study identifies sequence variants on 6q21 associated with age at menarche. Nat Genet 41:734-738

34. He C, Kraft P, Chen C, Buring JE, Pare G, Hankinson SE, Chanock SJ, Ridker PM, Hunter DJ, Chasman DI (2009) Genome-wide association studies identify loci associated with age at menarche and age at natural menopause. Nat Genet 41:724-728

35. Viswanathan SR, Powers JT, Einhorn W, Hoshida Y, Ng TL, Toffanin S, O'Sullivan M, Lu J, Phillips LA, Lockhart VL, et al (2009) Lin28 promotes transformation and is associated with advanced human malignancies. Nat Genet 41:843-848

36. Guo Y, Chen Y, Ito H, Watanabe A, Ge X, Kodama T, Aburatani H (2006) Identification and characterization of lin-28 homolog B (LIN28B) in human hepatocellular carcinoma. Gene 384:51-61

37. Onland-Moret NC, Peeters PH, van Gils CH, Clavel-Chapelon F, Key T, Tjonneland A, Trichopoulou A, Kaaks R, Manjer J, Panico S, et al (2005) Age at menarche in relation to adult height: the EPIC study. Am J Epidemiol 162:623-632

38. Guo Y, Chen Y, Ito H, Watanabe A, Ge X, Kodama T, Auratani H (2006) Identification and characterization of lin-28 homolog B (LIN28B) in human hepatocellular carcinoma. Gene 384:51-61

39. Lutchmaya S, Baron-Cohen S, Raggatt P, Knickmeyer R, Manning JT (2004) 2nd to 4th digit ratios, fetal testosterone and estradiol. Early Human Development 77:23-82

Figure 1 Plot of the locus surrounding *LIN28B*. SNPs are plotted by position on chromosome 6 (Build 36) against GWAS association p values for average 2D:4D. SNP rs314277 is shown in dark red labeled with its P value in the discovery meta-analysis. Estimated recombination rates from HapMap are plotted in light blue to reflect local LD structure. The directly genotyped or imputed SNPs surrounding rs314277 are color coded to reflect their LD with rs314277 (according to pair-wise r^2 values from the HapMap CEU database). Genes and their direction of transcription are labeled at the bottom of the plot.

Table 1 Mean and standard deviation for 2D:4D in the discovery and replication cohorts. Females tend to have higher ratios than males.

2D:4D ratio		Mean (SD) 2D:4D in original GWAS sample		Mean (SD) 2D:4D in replication ALSPAC cohort
		QIMR	**ALSPAC**	
Right	Females	97.93 (3.18)	96.67 (3.25)	96.73 (3.17)
	Males	96.16 (3.30)	95.85 (3.19)	95.73 (3.19)
Left	Females	98.43 (3.46)	96.80 (3.34)	96.88 (3.16)
	Males	96.84 (3.43)	96.02 (3.23)	95.93 (3.15)
Average	Females	98.18 (2.91)	96.74 (3.04)	96.82 (2.90)
	Males	96.84 (3.44)	95.94 (2.93)	95.84 (2.85)

Table 2 Association results for the SNP markers rs314277 and rs314276 in the ALSPAC Discovery GWAS, the QIMR Discovery GWAS, the combined meta-analysis, and the replication data. In each case the minor allele is listed first in the table. Beta coefficients indicate the expected change in 2D:4D per additional minor allele.

		Alleles	**MAF**	**Left 2D:4D**			**Right 2D:4D**			**Mean 2D:4D**		
				Beta	**SE**	**P-value**	**Beta**	**SE**	**P-value**	**Beta**	**SE**	**P-value**
rs314277	ALSPAC Discovery	A/C	0.14	0.54	0.17	1.90×10^{-3}	0.58	0.17	4.90×10^{-4}	0.56	0.16	3.17×10^{-4}
	QIMR Discovery		0.14	0.74	0.19	1.50×10^{-4}	0.65	0.18	4.30×10^{-4}	0.70	0.17	2.40×10^{-5}
	Combined			0.63	0.13	1.50×10^{-6}	0.61	0.12	8.20×10^{-7}	0.63	0.11	4.10×10^{-8}
	Replication		0.15	0.42	0.01	2.75×10^{-5}	0.47	0.10	5.15×10^{-6}	0.44	0.09	1.53×10^{-6}
rs314276	ALSPAC Discovery	A/C	0.34	0.37	0.13	4.05×10^{-3}	0.27	0.12	0.027	0.32	0.12	5.30×10^{-3}
	QIMR Discovery		0.34	0.39	0.14	6.00×10^{-3}	0.33	0.13	0.014	0.36	0.12	3.00×10^{-3}
	Combined			0.38	0.09	7.92×10^{-5}	0.30	0.09	1.01×10^{-3}	0.34	0.08	5.42×10^{-5}
	Replication		0.34	0.21	0.07	5.34×10^{-3}	0.30	0.08	1.12×10^{-4}	0.26	0.07	2.26×10^{-4}

Supplementary Table 1 SNPs with a combined p value of less than $p < 1 \times 10^{-5}$ in the meta-analyses of the left 2D:4D ratio. SNPs located within genes (as defined by Ensembl annotation which includes transcribed as well as regulatory regions) are labelled in the column "Gene". "Meta-analysis effect size" and "meta-analysis p value" refer to the effect size and p value obtained from an inverse variance meta-analysis of the ALSPAC and QIMR datasets. "QIMR p value" and "ALSPAC p value" refer to the p values for the test of association in the individual cohorts.

Chromosome	SNP	Gene	Allele1	Allele2	Meta-analysis Effect size for Allele1 (SE)	Meta-analysis p value	QIMR p-value	ALSPAC p-value
1	rs12127718		A	T	-0.43 (0.09)	9.4×10^{-6}	9.7×10^{-4}	2.7×10^{-3}
2	rs4666630		A	G	1.18 (0.26)	6.1×10^{-6}	0.014	1.0×10^{-4}
6	rs314277	*LIN28B*	A	C	0.63 (0.13)	1.5×10^{-6}	1.5×10^{-4}	1.9×10^{-3}
8	rs11995447		A	T	-0.50 (0.09)	2.5×10^{-7}	3.0×10^{-4}	1.7×10^{-4}
8	rs4392860		A	G	0.54 (0.11)	3.6×10^{-7}	1.5×10^{-4}	4.7×10^{-4}
8	rs11775128		A	C	0.54 (0.11)	3.6×10^{-7}	1.5×10^{-4}	4.7×10^{-4}
8	rs4361726		A	G	-0.54 (0.11)	3.8×10^{-7}	1.5×10^{-4}	4.8×10^{-4}
8	rs4380891		T	C	0.54 (0.11)	5.8×10^{-7}	2.7×10^{-4}	4.6×10^{-4}
8	rs4523215		A	G	0.54 (0.11)	5.8×10^{-7}	2.7×10^{-4}	4.6×10^{-4}
8	rs13279329		T	G	0.48 (0.09)	7.3×10^{-7}	3.1×10^{-4}	4.8×10^{-4}
8	rs11984796		C	G	-0.53 (0.11)	6.7×10^{-7}	3.4×10^{-4}	4.6×10^{-4}
8	rs6984615	*UNQ9391*	A	G	0.53 (0.11)	6.5×10^{-7}	1.2×10^{-3}	1.3×10^{-4}
8	rs12543028		T	C	-0.53 (0.11)	7.3×10^{-7}	3.4×10^{-4}	5.0×10^{-4}
8	rs11250020		C	G	-0.53 (0.11)	7.3×10^{-7}	3.4×10^{-4}	5.0×10^{-4}
8	rs11987962		C	G	-0.53 (0.11)	7.9×10^{-7}	4.1×10^{-4}	4.4×10^{-4}
8	rs4532561		T	G	0.53 (0.11)	7.9×10^{-7}	4.1×10^{-4}	4.4×10^{-4}
8	rs4295624		C	G	-0.53 (0.11)	7.9×10^{-7}	4.1×10^{-4}	4.5×10^{-4}
8	rs7837520		A	G	-0.51 (0.10)	8.0×10^{-7}	6.6×10^{-4}	3.0×10^{-4}
8	rs11783249		A	T	0.51 (0.10)	8.7×10^{-7}	1.1×10^{-3}	2.0×10^{-4}
8	rs10086950		A	C	0.48 (0.10)	9.5×10^{-7}	6.5×10^{-4}	3.6×10^{-4}
8	rs11250024		T	C	0.52 (0.11)	9.8×10^{-7}	1.3×10^{-3}	1.9×10^{-4}
8	rs11250010		C	G	0.51 (0.10)	9.9×10^{-7}	1.3×10^{-3}	1.9×10^{-4}
8	rs11250013		C	G	0.51 (0.10)	9.9×10^{-7}	1.3×10^{-3}	1.9×10^{-4}
8	rs4841363		A	T	-0.52 (0.11)	9.8×10^{-7}	1.3×10^{-3}	1.9×10^{-4}
8	rs7842925		T	C	0.52 (0.11)	1.1×10^{-6}	5.2×10^{-4}	5.1×10^{-4}
8	rs10102650		T	C	0.48 (0.10)	9.5×10^{-7}	3.0×10^{-3}	7.8×10^{-5}
8	rs10090800		T	C	-0.49 (0.10)	1.1×10^{-6}	1.2×10^{-3}	2.2×10^{-4}
8	rs11250019		A	G	-0.52 (0.11)	1.2×10^{-6}	5.8×10^{-4}	5.1×10^{-4}
8	rs7821273		A	G	-0.52 (0.11)	1.2×10^{-6}	5.8×10^{-4}	5.2×10^{-4}
8	rs13282174		T	C	0.52 (0.11)	1.2×10^{-6}	5.8×10^{-4}	5.3×10^{-4}
8	rs11780245		T	G	0.52 (0.11)	1.2×10^{-6}	5.8×10^{-4}	5.3×10^{-4}
8	rs13266986	*MSRA*	A	G	-0.51 (0.10)	1.2×10^{-6}	2.4×10^{-3}	1.2×10^{-4}
8	rs13257718		T	G	0.50 (0.10)	1.7×10^{-6}	1.5×10^{-3}	3.0×10^{-4}
8	rs4471098		T	G	-0.50 (0.10)	2.0×10^{-6}	2.4×10^{-3}	2.2×10^{-4}
8	rs10103190		T	C	0.67 (0.14)	3.8×10^{-6}	1.6×10^{-3}	6.5×10^{-4}
8	rs17063833		T	C	-0.66 (0.14)	5.5×10^{-6}	2.8×10^{-3}	5.4×10^{-4}
8	rs6994475		T	G	-0.65 (0.14)	6.3×10^{-6}	2.6×10^{-3}	6.5×10^{-4}
8	rs6601465		A	G	-0.44 (0.10)	8.1×10^{-6}	0.016	1.1×10^{-4}

11	rs11024832		T	C	1.24 (0.26)	2.5×10^{-6}	0.046	6.1×10^{-6}
14	rs11621436	*SMOC1*	T	C	0.46 (0.10)	3.3×10^{-6}	1.5×10^{-3}	5.8×10^{-4}
14	rs11158820	*SMOC1*	A	G	-0.53 (0.11)	1.3×10^{-6}	3.2×10^{-3}	8.4×10^{-5}
18	rs892579		A	G	0.47 (0.10)	7.6×10^{-6}	1.5×10^{-3}	1.4×10^{-3}

Supplementary Table 2 SNPs with a combined p value of less than $p < 1 \times 10^{-5}$ in the meta-analyses of the right 2D:4D ratio. . SNPs located within genes (as defined by Ensembl annotation which includes transcribed as well as regulatory regions) are labelled in the column "Gene". "Meta-analysis effect size" and "meta-analysis p value" refer to the effect size and p value obtained from an inverse variance meta-analysis of the ALSPAC and QIMR datasets. "QIMR p value" and "ALSPAC p value" refer to the p values for the test of association in the individual cohorts.

Chromosome	SNP	Gene	Allele1	Allele2	Meta-analysis Effect size for Allele1 (SE)	Meta-analysis p value	QIMR p-value	ALSPAC p-value
1	rs12137291	*GLIS1*	T	G	-0.46 (0.09)	7.1×10^{-7}	6.1×10^{-5}	1.9×10^{-3}
1	rs4634849	*GLIS1*	T	C	-0.46 (0.09)	6.9×10^{-7}	6.1×10^{-5}	1.9×10^{-3}
1	rs12140369	*GLIS1*	A	T	0.45 (0.09)	7.5×10^{-7}	6.3×10^{-5}	1.9×10^{-3}
1	rs2950252	*GLIS1*	C	G	0.45 (0.09)	8.2×10^{-7}	6.3×10^{-5}	2.1×10^{-3}
1	rs2950250	*GLIS1*	T	G	0.46 (0.09)	9.5×10^{-7}	6.1×10^{-5}	2.7×10^{-3}
1	rs2948045	*GLIS1*	T	C	-0.45 (0.09)	1.2×10^{-6}	6.3×10^{-5}	2.9×10^{-3}
1	rs6588482	*GLIS1*	A	T	0.44 (0.09)	1.3×10^{-6}	3.7×10^{-5}	4.5×10^{-3}
1	rs4927011	*GLIS1*	T	C	-0.44 (0.09)	1.3×10^{-6}	3.6×10^{-5}	4.7×10^{-3}
1	rs4926604	*GLIS1*	T	C	0.44 (0.09)	1.3×10^{-6}	4.4×10^{-5}	4.1×10^{-3}
1	rs1879735	*GLIS1*	A	G	-0.44 (0.09)	1.4×10^{-6}	4.1×10^{-5}	4.5×10^{-3}
1	rs2948053	*GLIS1*	A	G	-0.43 (0.09)	2.8×10^{-6}	5.3×10^{-5}	6.6×10^{-3}
1	rs3013754	*GLIS1*	A	C	-0.43 (0.09)	2.8×10^{-6}	5.3×10^{-5}	6.6×10^{-3}
1	rs7551844	*GLIS1*	T	C	0.43 (0.09)	2.9×10^{-6}	5.3×10^{-5}	6.8×10^{-3}
1	rs10788958	*GLIS1*	C	G	0.47 (0.10)	2.5×10^{-6}	8.8×10^{-5}	6.0×10^{-3}
1	rs17382457	*GLIS1*	A	G	0.42 (0.09)	4.5×10^{-6}	4.3×10^{-5}	0.011
1	rs12563871	*GLIS1*	A	G	0.42 (0.09)	4.5×10^{-6}	4.3×10^{-5}	0.011
1	rs7542387	*GLIS1*	A	G	0.41 (0.09)	5.8×10^{-6}	5.7×10^{-5}	0.012
1	rs11801290	*GLIS1*	A	G	-0.41 (0.09)	6.8×10^{-6}	7.0×10^{-5}	0.012
1	rs1572703		A	G	0.43 (0.09)	2.8×10^{-6}	3.6×10^{-4}	2.1×10^{-4}
2	rs7556683	*ENSG205086*	T	C	0.40 (0.08)	7.5×10^{-6}	0.017	8.9×10^{-5}
2	rs6724513	*C2orf43*	A	G	0.56 (0.11)	2.8×10^{-7}	3.2×10^{-4}	2.2×10^{-4}
2	rs340600	*C2orf43*	T	G	-0.55 (0.11)	3.1×10^{-7}	3.2×10^{-4}	2.4×10^{-4}
4	rs4241809	*SORBS2*	T	C	0.52 (0.11)	4.4×10^{-6}	1.5×10^{-5}	0.031
6	rs314277	*LIN28B*	A	C	0.61 (0.12)	8.2×10^{-7}	4.3×10^{-4}	4.9×10^{-4}
8	rs11250064	*SOX7*	A	C	0.43 (0.09)	9.2×10^{-7}	0.36	1.0×10^{-9}
8	rs4503064	*SOX7*	A	G	0.41 (0.09)	3.0×10^{-6}	0.70	1.7×10^{-9}
8	rs10097478		A	G	0.53 (0.10)	1.4×10^{-7}	5.4×10^{-3}	2.5×10^{-6}
8	rs2733162		A	G	0.42 (0.09)	7.0×10^{-6}	2.1×10^{-4}	5.9×10^{-3}
18	rs4941373	*NEDD4L*	C	G	-0.51 (0.11)	8.5×10^{-6}	2.6×10^{-4}	6.1×10^{-3}
18	rs4058295	*NEDD4L*	A	C	-0.51 (0.11)	8.3×10^{-6}	5.5×10^{-4}	3.5×10^{-3}
18	rs2288775	*NEDD4L*	A	G	0.51 (0.11)	5.5×10^{-6}	3.7×10^{-4}	3.0×10^{-3}

Supplementary Table 3 SNPs with a combined p value of less than $p < 1 \times 10^{-5}$ in the meta-analyses of the Mean 2D:4D ratio. . SNPs located within genes (as defined by Ensembl annotation which includes transcribed as well as regulatory regions) are labelled in the column "Gene". "Meta-analysis effect size" and "meta-analysis p value" refer to the effect size and p value obtained from an inverse variance meta-analysis of the ALSPAC and QIMR datasets. "QIMR p value" and "ALSPAC p value" refer to the p values for the test of association in the individual cohorts.

Chromosome	SNP	Gene	Allele1	Allele2	Meta-analysis Effect size for Allele1 (SE)	Meta-analysis p value	QIMR p-value	ALSPAC p-value
1	rs4927011	*GLIS1*	T	C	-0.38 (0.08)	6.1×10^{-6}	3.9×10^{-4}	3.4×10^{-3}
1	rs11264329	*EFNA1*	A	G	0.37 (0.08)	4.4×10^{-6}	0.023	3.41×10^{-5}
1	rs6588482	*GLIS1*	A	T	0.38 (0.08)	6.4×10^{-6}	4.4×10^{-4}	3.4×10^{-3}
1	rs1879735	*GLIS1*	A	G	-0.38 (0.08)	6.7×10^{-6}	4.5×10^{-4}	3.4×10^{-3}
1	rs4926604	*GLIS1*	T	C	0.38 (0.08)	7.2×10^{-6}	4.9×10^{-4}	3.2×10^{-3}
1	rs17382220	*GLIS1*	A	T	-0.56 (0.12)	6.2×10^{-6}	1.4×10^{-4}	9.4×10^{-3}
1	rs4927012	*GLIS1*	T	C	-0.56 (0.12)	6.2×10^{-6}	1.4×10^{-4}	9.4×10^{-3}
1	rs4926603	*GLIS1*	A	G	0.58 (0.13)	6.0×10^{-6}	1.4×10^{-4}	9.5×10^{-3}
1	rs12140369	*GLIS1*	A	T	0.38 (0.08)	8.2×10^{-6}	7.9×10^{-4}	2.6×10^{-3}
1	rs3108391	*GLIS1*	T	C	-0.57 (0.13)	6.4×10^{-6}	1.4×10^{-4}	0.010
1	rs2950252	*GLIS1*	C	G	0.38 (0.08)	8.2×10^{-6}	7.9×10^{-4}	2.7×10^{-3}
1	rs11585273	*GLIS1*	A	G	0.57 (0.13)	6.8×10^{-6}	1.4×10^{-4}	0.010
1	rs11585344	*GLIS1*	A	G	0.57 (0.13)	6.7×10^{-6}	1.4×10^{-4}	0.010
1	rs12089978	*GLIS1*	T	C	-0.57 (0.13)	6.8×10^{-6}	1.4×10^{-4}	0.010
1	rs6588483	*GLIS1*	A	G	0.60 (0.13)	6.8×10^{-6}	2.6×10^{-4}	6.8×10^{-3}
1	rs12084713	*GLIS1*	A	G	0.57 (0.13)	7.1×10^{-6}	1.4×10^{-4}	0.011
1	rs17386087	*GLIS1*	T	C	-0.57 (0.13)	7.0×10^{-6}	1.4×10^{-4}	0.011
1	rs17386108	*GLIS1*	A	T	0.57 (0.13)	7.1×10^{-6}	1.4×10^{-4}	0.011
1	rs2948053	*GLIS1*	A	G	-0.37 (0.08)	9.0×10^{-6}	5.3×10^{-4}	4.0×10^{-3}
1	rs3013754	*GLIS1*	A	C	-0.37 (0.08)	9.0×10^{-6}	5.3×10^{-4}	4.0×10^{-3}
1	rs10888795	*GLIS1*	A	G	0.37 (0.08)	8.7×10^{-6}	5.3×10^{-4}	4.0×10^{-3}
1	rs7551844	*GLIS1*	T	C	0.37 (0.08)	8.7×10^{-6}	5.3×10^{-4}	4.0×10^{-3}
1	rs12137291	*GLIS1*	T	G	-0.37 (0.08)	8.5×10^{-6}	8.7×10^{-4}	2.7×10^{-3}
1	rs3006884	*GLIS1*	A	G	-0.57 (0.13)	7.0×10^{-6}	1.4×10^{-4}	0.011
1	rs2948045	*GLIS1*	T	C	-0.37 (0.08)	9.5×10^{-6}	7.9×10^{-3}	2.9×10^{-3}
1	rs4634849	*GLIS1*	T	C	-0.38 (0.08)	8.4×10^{-6}	8.7×10^{-4}	2.7×10^{-3}
1	rs11185151	*VAV3*	T	C	-0.36 (0.08)	9.5×10^{-6}	7.8×10^{-4}	3.3×10^{-3}
1	rs4914952	*VAV3*	A	C	-0.36 (0.08)	9.6×10^{-6}	7.1×10^{-4}	3.6×10^{-3}
1	rs1691228	*ZNF695*	T	C	-0.86 (0.19)	8.1×10^{-6}	3.3×10^{-3}	6.8×10^{-4}
2	rs6724513	*C2orf43*	A	G	0.52 (0.10)	1.6×10^{-7}	1.7×10^{-4}	2.2×10^{-4}
2	rs340600	*C2orf43*	T	G	-0.52 (0.10)	1.8×10^{-7}	1.7×10^{-4}	2.5×10^{-5}
4	rs4241809	*SORBS2*	T	C	0.51 (0.10)	1.1×10^{-6}	1.4×10^{-4}	1.9×10^{-3}
6	rs314277	*LIN28B*	A	C	0.63 (0.11)	4.1×10^{-8}	2.4×10^{-5}	3.2×10^{-4}
8	rs11250064	*SOX7*	A	C	0.43 (0.09)	9.2×10^{-7}	0.26	2.74×10^{-9}
8	rs4392860		A	G	0.44 (0.09)	3.1×10^{-6}	2.9×10^{-3}	2.6×10^{-4}
8	rs11775128		A	C	0.44 (0.09)	3.1×10^{-6}	2.9×10^{-3}	2.6×10^{-4}
8	rs4361726		A	G	-0.44 (0.09)	3.2×10^{-6}	2.9×10^{-3}	2.7×10^{-4}

8	rs4380891		T	C	0.44 (0.09)	2.9×10^{-6}	3.4×10^{-3}	2.3×10^{-4}
8	rs4523215		A	G	0.44 (0.09)	3.0×10^{-6}	3.4×10^{-3}	2.3×10^{-4}
8	rs11984796		C	G	-0.43 (0.09)	4.2×10^{-6}	5.4×10^{-3}	2.0×10^{-4}
8	rs7837520		A	G	-0.42 (0.09)	4.3×10^{-6}	9.5×10^{-3}	9.76×10^{-5}
8	rs10097478		A	G	0.43 (0.09)	4.5×10^{-6}	5.3×10^{-5}	0.013
8	rs12543028		T	C	-0.43 (0.09)	4.4×10^{-6}	5.4×10^{-3}	2.2×10^{-4}
8	rs11250020		C	G	-0.43 (0.09)	4.4×10^{-6}	5.4×10^{-3}	2.2×10^{-4}
8	rs11783249		A	T	0.42 (0.09)	4.6×10^{-6}	0.011	8.61×10^{-5}
8	rs13266986	*MSRA*	A	G	-0.09 (0.09)	5.0×10^{-6}	0.018	4.5×10^{-5}
8	rs11987962		C	G	-0.43 (0.09)	5.1×10^{-6}	6.6×10^{-3}	1.9×10^{-4}
8	rs4532561		T	G	0.43 (0.09)	5.1×10^{-6}	6.6×10^{-3}	1.9×10^{-4}
8	rs4295624		C	G	-0.43 (0.09)	5.1×10^{-6}	6.6×10^{-3}	1.9×10^{-4}
8	rs10090800		T	C	-0.40 (0.09)	5.3×10^{-6}	0.015	6.5×10^{-5}
8	rs11250010		C	G	0.41 (0.09)	5.2×10^{-6}	0.013	8.1×10^{-5}
8	rs11250013		C	G	0.41 (0.09)	5.2×10^{-6}	0.013	8.2×10^{-5}
8	rs6984615	*UNQ9391*	A	G	0.42 (0.09)	5.6×10^{-6}	0.017	6.2×10^{-5}
8	rs7842925		T	C	0.43 (0.09)	6.0×10^{-6}	7.0×10^{-3}	2.2×10^{-4}
8	rs11250019		A	G	-0.42 (0.09)	6.4×10^{-6}	7.3×10^{-3}	2.2×10^{-4}
8	rs7821273		A	G	-0.42 (0.09)	6.4×10^{-6}	7.3×10^{-3}	2.3×10^{-4}
8	rs7821267		A	C	-0.42 (0.09)	6.6×10^{-6}	7.3×10^{-3}	2.3×10^{-4}
8	rs13282174		T	C	0.42 (0.09)	6.6×10^{-6}	7.3×10^{-3}	2.3×10^{-4}
8	rs11780245		T	G	0.42 (0.09)	6.6×10^{-6}	7.3×10^{-3}	2.3×10^{-4}
8	rs11250024		T	C	0.42 (0.09)	7.5×10^{-6}	0.013	1.2×10^{-4}
8	rs4841363		A	T	-0.42 (0.09)	7.7×10^{-6}	0.013	1.2×10^{-4}
8	rs11995447		A	T	-0.38 (0.08)	8.4×10^{-6}	6.8×10^{-3}	2.9×10^{-4}
8	rs4503064	*SOX7*	A	G	0.37 (0.08)	7.4×10^{-6}	0.640	9.7×10^{-9}
11	rs11024832		T	C	1.05 (0.23)	5.4×10^{-6}	0.140	1.3×10^{-6}
11	rs4755664		A	G	0.38 (0.08)	9.5×10^{-6}	4.0×10^{-3}	6.0×10^{-4}
14	rs11621436	*SMOC1*	T	C	0.39 (0.09)	9.0×10^{-6}	8.9×10^{-4}	2.8×10^{-3}
20	rs6138060		C	G	0.59 (0.13)	4.5×10^{-6}	4.8×10^{-4}	2.6×10^{-3}

Supplementary Table 4 SNPs with a combined p value of less than $p < 1 \times 10^{-5}$ in the meta-analysis of the chi-square tests of heterogeneity between the sexes for left 2D:4D ratio. SNPs located within genes (as defined by Ensembl annotation which includes transcribed as well as regulatory regions) are labelled in the column "Gene".

Chromosome	SNP	Gene	P-value
2	rs13391185		2.59×10^{-6}
2	rs7584991		9.72×10^{-6}
3	rs6445404		9.75×10^{-6}
3	rs7651603		7.75×10^{-6}
5	rs381575	*NDUFS4*	9.82×10^{-6}
5	rs2398587		7.84×10^{-6}
7	rs4724644	*PKD1L1*	1.04×10^{-6}
7	rs1551276	*PKD1L1*	1.05×10^{-6}
7	rs1551277	*PKD1L1*	1.05×10^{-6}
7	rs9719534	*FLJ21075*	5.71×10^{-6}
7	rs965143	*FLJ21075*	7.17×10^{-6}
7	rs12702390	*FLJ21075*	6.24×10^{-6}
8	rs7002691	*COL14A1*	5.66×10^{-6}
8	rs16893630	*COL14A1*	5.79×10^{-6}
8	rs961223	*COL14A1*	5.95×10^{-6}
9	rs1874109	*FREM1*	4.98×10^{-6}
10	rs1047468	*VTI1A*	4.92×10^{-6}
10	rs10885989	*PNLIPRP1*	6.42×10^{-6}
10	rs2915753	*PNLIPRP1*	6.51×10^{-6}
11	rs645359		3.01×10^{-6}
11	rs11228791		8.58×10^{-6}
11	rs12362194		9.10×10^{-6}
12	rs11834880	*ERC1*	4.69×10^{-6}
12	rs7488441	*ERC1*	6.98×10^{-6}
13	rs2874196		2.53×10^{-6}
16	rs8055579	*CHD13*	8.74×10^{-6}
16	rs11861722	*CDH13*	7.56×10^{-6}
X	rs4830407		5.69×10^{-6}

Supplementary Table 5 SNPs with a combined p value of less than $p < 1 \times 10^{-5}$ in the meta-analysis of the chi-square tests of heterogeneity between the sexes for right 2D:4D ratio. SNPs located within genes (as defined by Ensembl annotation which includes transcribed as well as regulatory regions) are labelled in the column "Gene".

Chromosome	SNP	Gene	P-value
7	rs4724644	*PKD1L1*	4.71×10^{-6}
7	rs1551276	*PKD1L1*	4.71×10^{-6}
7	rs1551277	*PKD1L1*	4.71×10^{-6}
7	rs9719534	*PKD1L1*	6.87×10^{-6}
7	rs965143	*PKD1L1*	7.63×10^{-6}
7	rs12702390	*PKD1L1*	4.28×10^{-6}
19	rs7255485	*ZNF615*	9.66×10^{-6}
19	rs11881700	*ZNF615*	9.75×10^{-6}
19	rs7253318	*ZNF615*	7.72×10^{-6}
19	rs11882305	*ZNF615*	7.72×10^{-6}
19	rs11879112	*ZNF615*	7.72×10^{-6}
19	rs7251200	*ZNF615*	7.79×10^{-6}
19	rs11878981	*ZNF615*	9.98×10^{-6}
19	rs16983430	*ZNF615*	9.62×10^{-6}
19	rs7248935	*ZNF615*	9.71×10^{-6}
19	rs7249005	*ZNF615*	9.89×10^{-6}
19	rs7250212	*ZNF615*	9.98×10^{-6}
20	rs550408		9.39×10^{-6}

Supplementary Table 6 SNPs with a combined p value of less than $p < 1 \times 10^{-5}$ in the meta-analysis of the chi-square tests of heterogeneity between the sexes for mean 2D:4D ratio. SNPs located within genes (as defined by Ensembl annotation which includes transcribed as well as regulatory regions) are labelled in the column "Gene".

Chromosome	SNP	Gene	P-value
3	rs7625907	*KCNMB2*	2.87×10^{-6}
3	rs7620381	*KCNMB2*	6.72×10^{-6}
3	rs6768608	*KCNMB2*	3.02×10^{-6}
3	rs6802875	*KCNMB2*	3.18×10^{-6}
3	rs9847663	*KCNMB2*	3.48×10^{-6}
3	rs6414483	*KCNMB2*	4.19×10^{-6}
3	rs7637074	*KCNMB2*	5.57×10^{-6}
5	rs12514182	*NDUFS4*	9.33×10^{-6}
7	rs6956567		9.82×10^{-6}
7	rs4724644	*PKD1L1*	1.75×10^{-7}
7	rs1551276	*PKD1L1*	1.75×10^{-7}
7	rs1551277	*PKD1L1*	1.75×10^{-7}
7	rs9719534	*FLJ21075*	5.85×10^{-7}
7	rs12702390	*FLJ21075*	4.58×10^{-7}
7	rs965143	*FLJ21075*	7.10×10^{-7}
9	rs1874109	*FREM1*	3.58×10^{-6}
16	rs4782812	*CDH13*	6.64×10^{-6}
16	rs8055579	*CDH13*	5.30×10^{-6}
16	rs11861722	*CDH13*	4.45×10^{-6}
17	rs4791811	*NTN1*	8.45×10^{-6}
19	rs7255485	*ZNF614*	1.02×10^{-6}
19	rs11881700	*ZNF432*	9.95×10^{-7}
19	rs7253318	*ZNF432*	8.61×10^{-7}
19	rs11882305	*ZNF432*	8.61×10^{-7}
19	rs11879112	*ZNF432*	8.61×10^{-7}
19	rs7251200	*ZNF432*	8.71×10^{-7}
19	rs11878981	*ZNF432*	1.03×10^{-6}
19	rs2043296		1.21×10^{-6}
19	rs16983412	*ZNF841*	1.21×10^{-6}
19	rs16983414	*ZNF841*	1.21×10^{-6}
19	rs16983416		1.08×10^{-6}
19	rs7359836		1.10×10^{-6}
19	rs6509621		1.10×10^{-6}
19	rs16983430		1.18×10^{-6}
19	rs7248935		1.22×10^{-6}
19	rs7249005		1.24×10^{-6}
19	rs7250212		1.26×10^{-6}
19	rs16983438		1.27×10^{-6}
19	rs7255079		1.30×10^{-6}
19	rs8112628		1.34×10^{-6}
19	rs8100579		1.76×10^{-6}
20	rs6079727	*MACROD2*	6.47×10^{-6}

Supplementary Table 7 Confirmed variants from eleven different genome-wide association studies of height and four genome-wide studies of age of menarche and their association with mean 2D:4D in the current study. SNPs are listed that had a p value < 1 x 10^{-5} in the discovery GWAS of height or age of menarche. The Effect Column indicates whether the risk allele is associated with an increase or decrease in height, 2D:4D or in age of menarche.

							Previous Study (height)		Current Study (2D:4D)	
Previous Study (Height)	**Sample Size**	**Ethnicity**	**Reported SNP**	**Chromosome**	**Related Gene**	**Risk Allele**	**Effect**	**P-value**	**Effect**	**P-value**
Kim at al. Identification of 15 loci influencing height in a Korean population. J Hum Genet 2009							positive	2 x 10^{-8}	negative	0.88
			rs10513137	3	*ZBTB38, ACPL2*	A	positive	8 x 10^{-8}	positive	0.54
			rs10961780	9	*FREM1*	G	negative	2 x 10^{-6}	positive	0.37
			rs2079795	17	*C17orf82, TBX2, TBX4*	T	positive	3 x 10^{-6}	positive	0.23
			rs13273123	8	*PLAG1*	C	negative	3 x 10^{-6}	positive	0.68
			rs16910061	9	*FBP2*	T	negative	3 x 10^{-6}	positive	0.53
			rs7032940	9	*PALM2-AKAP2*	A	positive	3 x 10^{-6}	negative	0.88
			rs3791675	2	*EFEMP1*	G	positive	4 x 10^{-6}	positive	0.15
			rs4811971	20	*ANKRD60*	C	positive	6 x 10^{-6}	negative	0.58
			rs11989122	8	*EXT1*	T	negative	6 x 10^{-6}	positive	0.13
			rs2292303	**12**	***NUP37, C12orf48, PMCH***	**C**	**negative**	**8 x 10^{-6}**	**negative**	**0.05**
			rs2315504	17	*KRT23, KRT20*	C	positive	8 x 10^{-6}	negative	0.73
			rs10948197	6	*SUPT3H*	C	negative	8 x 10^{-6}	negative	0.28
			rs1569019	12	*GPR133*	A	positive	5 x 10^{-6}	positive	0.22
Tonjes at al. Genetic variation in GPR133 is associated with height: genome wide association study in the self-contained population of Sorbs. Hum Mol Genet 2009	3916	Caucasian	rs6717918	2	*DIS3L2, ALPP, NPPC*	T	positive	3 x 10^{-9}	positive	0.47
Estrada et al. A genome-wide association study of northwestern Europeans involves the C-type natriuretic peptide signaling pathway in the etiology of human height variation. Hum Mol Genet 2009	10074	Caucasian	rs139909	22	*TNRC6B, ADSL*	T	positive	2 x 10^{-7}	negative	0.94
			rs10472828	5	*C5orf23, NPR3*	C	positive	3 x 10^{-7}	positive	0.16

Supplementary Table 7 cont.

							Previous Study (height)		Current Study (2D:4D)	
Previous Study (Height)	**Sample Size**	**Ethnicity**	**Reported SNP**	**Chromosome**	**Related Gene**	**Risk Allele**	**Effect**	**P-value**	**Effect**	**P-value**
							positive	6×10^{-12}	positive	0.54
			rs13273123	8	*PLAG1*	G	negative	1×10^{-9}	positive	0.68
			rs3791675	2	*EFEMP1*	G	positive	2×10^{-9}	positive	0.15
			rs6918981	6	*HMGA1*	G	positive	3×10^{-8}	negative	0.88
			rs17038182	1	*Intergenic*	C	negative	5×10^{-7}	positive	0.37
Soranzo et al. Meta-analysis of genome-wide scans for human adult stature identifies novel Loci and associations with measures of skeletal frame size. PLoS Genet 2009	12611	Caucasian	rs8756	12	*HMGA2*	A	negative	5×10^{-14}	negative	0.91
			rs6088813	20	*UQCC*	A	negative	1×10^{-13}	negative	0.63
			rs6763931	3	*ZBTB38*	A	positive	3×10^{-12}	positive	0.38
			rs10946808	6	*HIST1H1D*	A	positive	6×10^{-12}	negative	0.57
			rs849141	7	*JAZF1*	A	positive	3×10^{-11}	positive	0.99
			rs6570507	6	*GPR126*	T	negative	4×10^{-11}	negative	0.30
			rs1776897	6	*HMGA1, C6orf106*	T	positive	8×10^{-11}	positive	0.59
			rs3118914	13	*DLEU7*	A	negative	4×10^{-10}	positive	0.13
			rs1182188	7	*GNA12*	A	positive	3×10^{-9}	negative	0.95
			rs6830062	4	*LCORL*	A	positive	5×10^{-9}	negative	0.59
			rs2282978	7	*CDK6*	A	negative	1×10^{-8}	positive	0.42
			rs710841	4	*PRKG2*	T	positive	2×10^{-8}	negative	0.21
			rs4842838	15	*ADAMTSL3*	A	negative	3×10^{-8}	negative	0.83
			rs13437082	6	*HLA-B*	A	negative	5×10^{-8}	positive	0.30
			rs11809207	1	*CATSPER4*	A	positive	6×10^{-8}	negative	0.50
			rs910316	14	*TMED10*	A	positive	1×10^{-7}	positive	0.18
			rs10472828	5	*NPR3*	T	negative	3×10^{-7}	negative	0.16
			rs7871764	9	*WDR40A*	T	positive	2×10^{-6}	negative	0.90
Soranzo et al. Meta-analysis of genome-wide scans for human adult stature identifies novel Loci and associations with measures of skeletal frame size. PLoS Genet 2009	12611	Caucasian	rs1812175	4	*HHIP*	A	negative	4×10^{-6}	positive	0.72
			rs7815788	8	*PLAG1*	A	negative	5×10^{-6}	negative	0.84
Johansson et al. Common variants in the JAZF1 gene associated with height identified by linkage and genome-wide association analysis. Hum Mol Genet 2009	3925	Caucasian	rs1635852	7	*JAZF1*	A	positive	9×10^{-10}	positive	0.34

Supplementary Table 7 cont.

							Previous Study (height)		Current Study (2D:4D)	
Previous Study (Height)	**Sample Size**	**Ethnicity**	**Reported SNP**	**Chromosome**	**Related Gene**	**Risk Allele**	**Effect**	**P-value**	**Effect**	**P-value**
Gudbjartsson et al. Many sequence variants affecting diversity of adult human height. Nat Genet 2008							positive	1×10^{-27}	positive	0.38
			rs8756	12	*HMGA2*	C	positive	2×10^{-16}	positive	0.91
			rs798544	7	*GNA12*	G	positive	6×10^{-15}	negative	0.80
			rs3748069	6	*GPR126*	A	positive	4×10^{-14}	positive	0.27
			rs1812175	4	*HHIP*	C	positive	1×10^{-11}	negative	0.72
			rs12198986	6	*BMP6*	A	positive	2×10^{-11}	positive	0.94
			rs3791679	2	*EFEMP1, PNPT1*	T	positive	6×10^{-11}	positive	0.14
			rs7153027	14	*TRIP11, FBLN5, ATXN3, CPSF2*	A	positive	1×10^{-10}	positive	0.38
			rs11205277	1	*Histone class 2A, MTMR11, SV2A*	G	positive	1×10^{-10}	negative	0.81
			rs6830062	4	*LCORL, NCAPG*	T	positive	1×10^{-10}	negative	0.59
			rs10946808	6	*Histone calss 1, Butyrophilin genes*	A	positive	6×10^{-10}	negative	0.57
			rs3760318	17	*CRLF3, ATAD5, CENTA2, EHF135*	C	positive	2×10^{-9}	negative	0.71
			rs4800148	18	*CABLES1, RBBP8, C18orf45*	A	positive	4×10^{-9}	negative	0.94
			rs2274432	1	*C1orf19, GLT25D2*	T	positive	8×10^{-9}	negative	0.57
			rs1776897	6	*HMGA1, LBH*	C	positive	1×10^{-8}	negative	0.59
			rs2282978	7	*CDK6, PEX1, GATAD1, ERVWE1*	C	positive	1×10^{-8}	negative	0.42
			rs967417	20	*BMP2*	C	positive	1×10^{-8}	negative	0.67
			rs4743034	9	*ZNF462*	A	positive	2×10^{-8}	positive	0.44
			rs4533267	15	*ADAMTS17*	A	positive	3×10^{-8}	negative	0.68
			rs678962	1	*DNM3*	G	positive	3×10^{-8}	positive	0.20
			rs185819	6	*HLA class III*	T	positive	3×10^{-8}	negative	0.85

Supplementary Table 7 cont.

							Previous Study (height)		Current Study (2D:4D)	
Previous Study (Height)	**Sample Size**	**Ethnicity**	**Reported SNP**	**Chromosome**	**Related Gene**	**Risk Allele**	**Effect**	**P-value**	**Effect**	**P-value**
Gudbjartsson et al. Many sequence variants affecting diversity of adult human height. Nat Genet 2008							positive	3×10^{-8}	positive	0.07
			rs7846385	8	*PXMP3, ZFHX4*	C	positive	5×10^{-8}	negative	0.69
			rs757608	17	*BCAS3, NACA2, TBX2, TBX4*	T	positive	6×10^{-8}	positive	0.23
			rs10958476	8	*PLAG1, MOS, CHCHD7, RDHE2, RPS20, LYN, TGS1, PENK*	C	positive	7×10^{-8}	negative	1.00
			rs4794665	17	*NOG, DGKE, TRIM25, COIL, RISK*	C	positive	1×10^{-7}	negative	0.57
			rs3825199	12	*SOCS2, MRPL42, CRADD, UBE2N*	C	positive	2×10^{-7}	negative	0.11
			rs946053	9	*COL27A1*	T	positive	2×10^{-7}	negative	0.42
			rs1490388	6	*C6orf173*	T	positive	6×10^{-7}	positive	0.15
			rs7209435	17	*MAP3K3, WDR68, LYK5, MT1F*	C	positive	7×10^{-7}	positive	0.42
			rs12199222	6	*NUP153, CAP2, KIFI13A*	T	positive	7×10^{-7}	negative	0.34
			rs2326458	16	*ZDHHC7, CRISPLD2, USP10*	C	positive	8×10^{-7}	negative	0.69
			rs6088792	20	*UQCC, GDF5, CEP250, EIF6, MMP24*	T	positive	8×10^{-7}	positive	0.40
			rs6733301	2	*ADCY3, RBJ, POMC, DNMT3A, DTNB*	G	positive	8×10^{-7}	positive	0.87

Supplementary Table 7 cont.

							Previous Study (height)		Current Study (2D:4D)	
Previous Study (Height)	**Sample Size**	**Ethnicity**	**Reported SNP**	**Chromosome**	**Related Gene**	**Risk Allele**	**Effect**	**P-value**	**Effect**	**P-value**
Gudbjartsson et al. Many sequence variants affecting diversity of adult human height. Nat Genet 2008							**positive**	**8 x 10^{-7}**	**positive**	**5 x 10^{-5}**
			rs2554380	15	*ADAMTSL3, SH3GL3*	T	positive	9 x 10^{-7}	negative	0.87
			rs2814828	9	*SPIN1, CCRK*	T	positive	9 x 10^{-7}	positive	0.65
			rs7249094	19	*ADAMTS10, MYO1F, PRAM1, OR2Z1*	G	positive	1 x 10^{-6}	negative	0.17
			rs1052483	2	*IHH, CRYBA2, FEV, SLC23A3, TUBA1*	C	positive	1 x 10^{-6}	negative	0.84
			rs749052	2	*NPPC, DIS3L2, COPS7B, PDE6D, PTMA*	A	positive	1 x 10^{-6}	negative	0.56
			rs11611208	12	*PDE3A, SLCO1C1, SLCO1B3*	A	positive	2 x 10^{-6}	negative	0.24
			rs2187642	12	*ETV6*	A	positive	2 x 10^{-6}	negative	0.53
			rs710841	4	*BMP3, PRKG2, RASGEF1B*	A	positive	2 x 10^{-6}	negative	0.21
			rs11177669	12	*LYZ, YEATS4, FRS2, CPSF6, CCT2, LRRC10*	A	positive	3 x 10^{-6}	positive	0.64
			rs1474563	X	*ITM2A*	T	positive	3 x 10^{-6}	negative	0.27
			rs9487094	6	*PPIL6, CD164, SMPD2, MNICAL1, ZBTB24*	G	positive	4 x 10^{-6}	positive	0.38
			rs5751614	22	BCR, GNAZ, RTDR1, IGLL1	A	positive	6 x 10^{-6}	positive	0.26

Supplementary Table 7 cont.

							Previous Study (height)		Current Study (2D:4D)	
Previous Study (Height)	**Sample Size**	**Ethnicity**	**Reported SNP**	**Chromosome**	**Related Gene**	**Risk Allele**	**Effect**	**P-value**	**Effect**	**P-value**
Gudbjartsson et al. Many sequence variants affecting diversity of adult human height. Nat Genet 2008							positive	6×10^{-6}	positive	0.90
			rs4345115	3	*GOLIM4, SERPINI1*	T	positive	7×10^{-6}	positive	0.72
			rs1239947	13	*DLEU7*	G	positive	8×10^{-6}	negative	0.47
			rs31198	5	*PIXT1, PCBD2, CATSPER3, TXNDC15, DDX46, CAMLG*	T	positive	8×10^{-6}	negative	0.84
			rs9395066	6	*SUPT3H, RUNX2*	C	positive	8×10^{-6}	positive	0.15
Lettre et al. Identification of ten loci associated with height highlights new biological pathways in human growth. Nat Genet 2008	15821	Caucasian	rs724016	3	*ZBTB38*	G	positive	8×10^{-22}	positive	0.38
			rs1042725	12	*HMGA2*	T	negative	3×10^{-20}	positive	0.94
			rs4896582	6	*GPR126*	A	negative	2×10^{-18}	negative	0.33
			rs10946808	6	*HIST1H1D*	G	negative	4×10^{-17}	negative	0.57
			rs6060369	20	*GDF5, UQCC*	C	positive	1×10^{-16}	positive	0.52
			rs1492820	4	*HHIP*	G	negative	1×10^{-11}	negative	0.90
			rs8007661	14	*TRIP11, ATXN3*	T	negative	5×10^{-10}	negative	0.33
			rs314277	**6**	***LIN28B***	**A**	**positive**	**1×10^{-8}**	**positive**	**4×10^{-8}**
			rs12986413	19	*DOT1L*	T	positive	3×10^{-8}	NA	NA
			rs2562784	15	*SH3GL3, ADAMTSL3*	G	positive	6×10^{-8}	negative	0.32
			rs2730245	7	*WDR60*	G	positive	3×10^{-7}	positive	0.33
			rs2040494	7	*CDK6*	C	negative	4×10^{-7}	negative	0.87
			rs9650315	8	*CHCHD7, RDHE2*	T	negative	4×10^{-7}	positive	0.93
			rs7466269	9	*FUBP3*	G	negative	7×10^{-7}	negative	0.42
			rs7869550	9	*PAPPA*	G	negative	1×10^{-6}	negative	0.57
			rs12449568	17	*ANKFN1*	C	positive	2×10^{-6}	negative	0.86
			rs763014	16	*RAB40C*	C	positive	5×10^{-6}	negative	0.30
			rs17104630	14	*NKX2-1*	G	negative	8×10^{-6}	positive	0.50

Supplementary Table 7 cont.

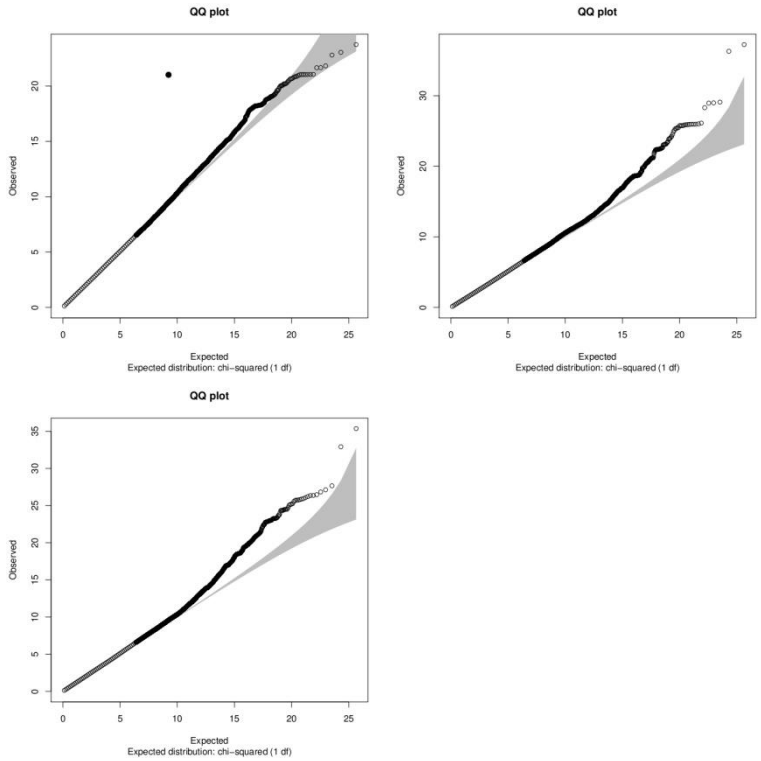
							Previous Study (height)		Current Study (2D:4D)	
Previous Study (Height)	**Sample Size**	**Ethnicity**	**Reported SNP**	**Chromosome**	**Related Gene**	**Risk Allele**	**Effect**	**P-value**	**Effect**	**P-value**
Weedon et al. Genome-wide association analysis identifies 20 loci that influence adult height. Nat Genet 2008			rs6440003	3	*ZBTB38*	A	positive	2×10^{-24}	positive	0.31
			rs2282978	7	*CDK*	C	positive	8×10^{-23}	negative	0.42
			rs1042725	12	*HMGA2*	C	positive	3×10^{-18}	negative	0.94
			rs6060373	20	*GDF*	G	negative	2×10^{-17}	positive	0.53
			rs16896068	4	*LCORL*	A	negative	2×10^{-13}	positive	0.58
			rs4549631	6	*LOC387103*	C	positive	5×10^{-13}	positive	0.25
			rs3791675	2	*EFEMP1*	C	positive	2×10^{-12}	positive	0.15
			rs2814993	6	*C6orf106*	A	positive	4×10^{-12}	positive	0.31
			rs12735613	1	*SPAG17*	A	negative	4×10^{-11}	positive	0.46
			rs10512248	9	*PTCH1*	G	positive	4×10^{-11}	positive	0.15
			rs11107116	12	*SOCS2*	G	negative	6×10^{-10}	positive	0.11
			rs6854783	4	*HHIP*	A	positive	2×10^{-9}	negative	0.84
			rs1390401	1	*ZNF678*	A	positive	5×10^{-9}	positive	0.32
			rs3116602	13	*DLEU7*	G	negative	7×10^{-9}	positive	0.13
			rs10906982	15	*ADAMTSL3*	A	positive	2×10^{-8}	negative	0.83
			rs6686842	1	*SCMH1*	C	negative	2×10^{-8}	positive	0.74
			rs6724465	2	*IHH*	A	negative	2×10^{-8}	positive	0.87
			rs10935120	3	*ANAPC13, CEP63*	A	negative	7×10^{-8}	negative	0.37
			rs8041863	15	*ACAN*	A	positive	8×10^{-8}	positive	0.76
			rs8099594	18	*DYM*	A	positive	3×10^{-7}	negative	0.55
Sanna et al. Common variants in the GDF5-UQCC region are associated with variation in human height. Nat Genet 2008	6669	Caucasian	rs6060369	20	*BFZB*	C	positive	2×10^{-16}	positive	0.52
			rs17690232	4	*PDGFRA*	C	positive	4×10^{-7}	positive	0.64
			rs4932217	**15**	***POLG***	**A**	positive	$\mathbf{8 \times 10^{-7}}$	**negative**	**0.01**
			rs724016	3	*ZBTB38*	G	positive	1×10^{-6}	positive	0.38
			rs10078095	5	*HOMER1*	C	positive	3×10^{-6}	positive	0.49
Weedon et al. A common variant of HMGA2 is associated with adult and childhood height in the general population. Nat Genet 2007	4921	Caucasian	rs1042725	12	*HMGA2*	C	positive	6×10^{-6}	negative	0.94

Supplementary Table 7 cont.

							Previous Study (menarche)		Current Study (2D:4D)	
Previous Study (Menarche)	**Sample Size**	**Ethnicity**	**Reported SNP**	**Chromosome**	**Related Gene**	**Risk Allele**	**Effect**	**P-value**	**Effect**	**P-value**
He et al. Genome-wide association studies identify loci associated with age at menarche and age at natural menopause. Nat Genet 2009							**positive**	**3×10^{-13}**	**positive**	**4×10^{-8}**
			rs314263	**6**	**LIN28B**	**C**	**positive**	**3×10^{-13}**	**positive**	**2×10^{-5}**
			rs369065	**6**	**LIN28B**	**C**	**positive**	**2×10^{-11}**	**positive**	**5×10^{-5}**
			rs7861820	9	TMEM38B	C	negative	3×10^{-9}	negative	0.10
			rs314280	**6**	**LIN28B**	**T**	**positive**	**2×10^{-8}**	**positive**	**3×10^{-4}**
			rs4946651	**6**	**LIN28B**	**A**	**positive**	**3×10^{-8}**	**positive**	**3×10^{-4}**
			rs12684013	9	TMEM38B	T	negative	4×10^{-8}	negative	0.11
			rs4452860	9	TMEM38B	G	negative	8×10^{-8}	negative	0.19
			rs7028916	9	TMEM38B	A	negative	1×10^{-7}	negative	0.22
			rs314262	**6**	**LIN28B**	**C**	**positive**	**1×10^{-7}**	**positive**	**3×10^{-4}**
Ong et al. Genetic variation in LIN28B is associated with the timing of puberty. Nat Genet 2009	4714	Caucasian	**rs314276**	**6**	**LIN28B**	**C**	**negative**	**4×10^{-16}**	**negative**	**5×10^{-5}**
Perry et al. Meta-analysis of genome-wide association data identifies two loci influencing age at menarche. Nat Genet 2009	17510	Caucasian	rs2090409	9	TMEM38B, SLC44A1, FKTN, FSD1L, TAL2, ANF462	A	negative	2×10^{-9}	negative	0.41
			rs7759938	**6**	**LIN28B**	**C**	**positive**	**7×10^{-9}**	**positive**	**1×10^{-5}**
Sulem et al. Genome-wide association study identifies sequence variants on 6q21 associated with age at menarche. Nat Genet 2009	477	Caucasian	**rs314280**	**6**	**LIN28B, HACE1, E3 ubiquitin protein ligase 1, BVES, POPDC3**	**T**	**positive**	**2×10^{-14}**	**positive**	**3×10^{-4}**

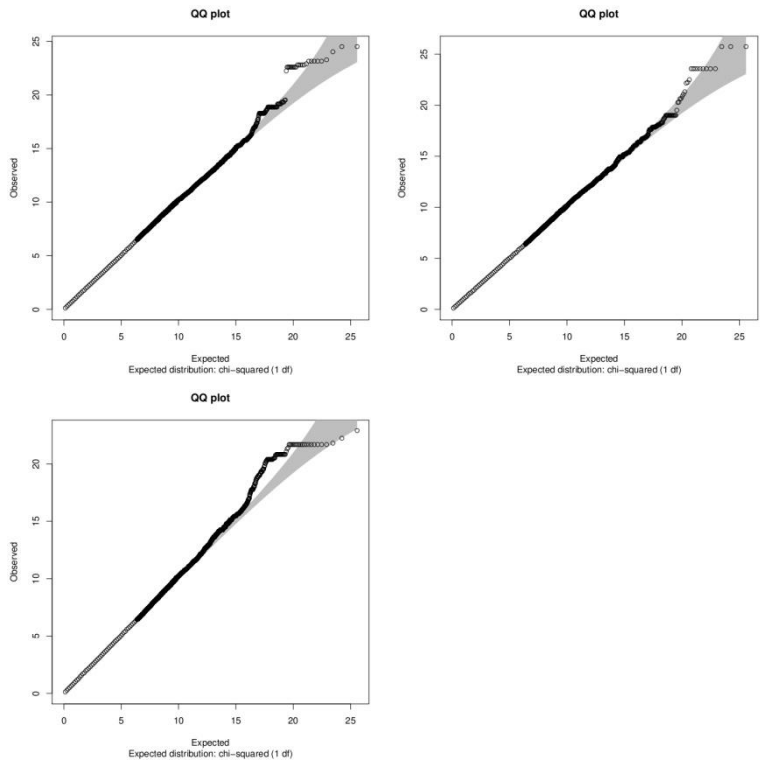
(A) ALSPAC:

Left 2D:4D B.) Right 2D:4D C.) Mean 2D:4D



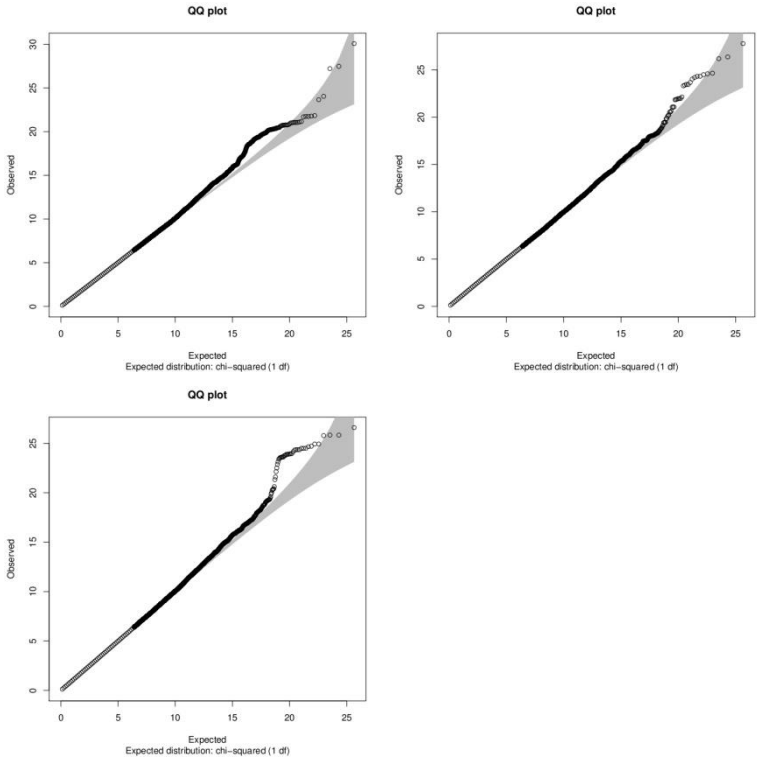
(B) QIMR:

A.) Left 2D:4D B.) Right 2D:4D C.) Mean 2D:4D



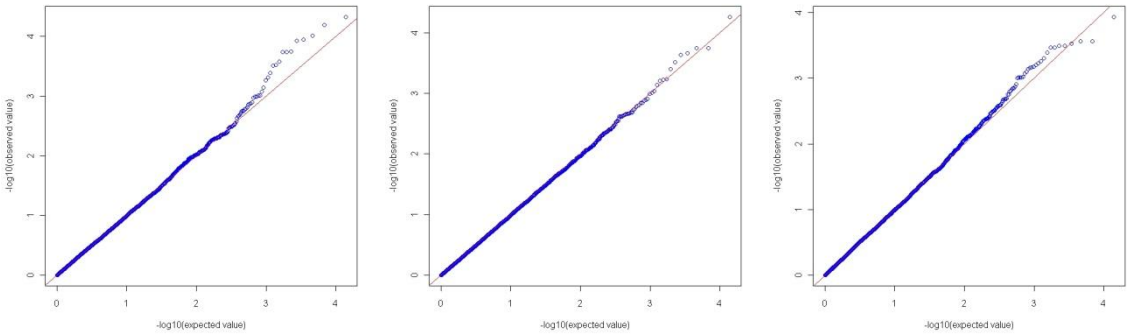
(C) Meta-analysis:

A.) Left 2D:4D B.) Right 2D:4D C.) Mean 2D:4D

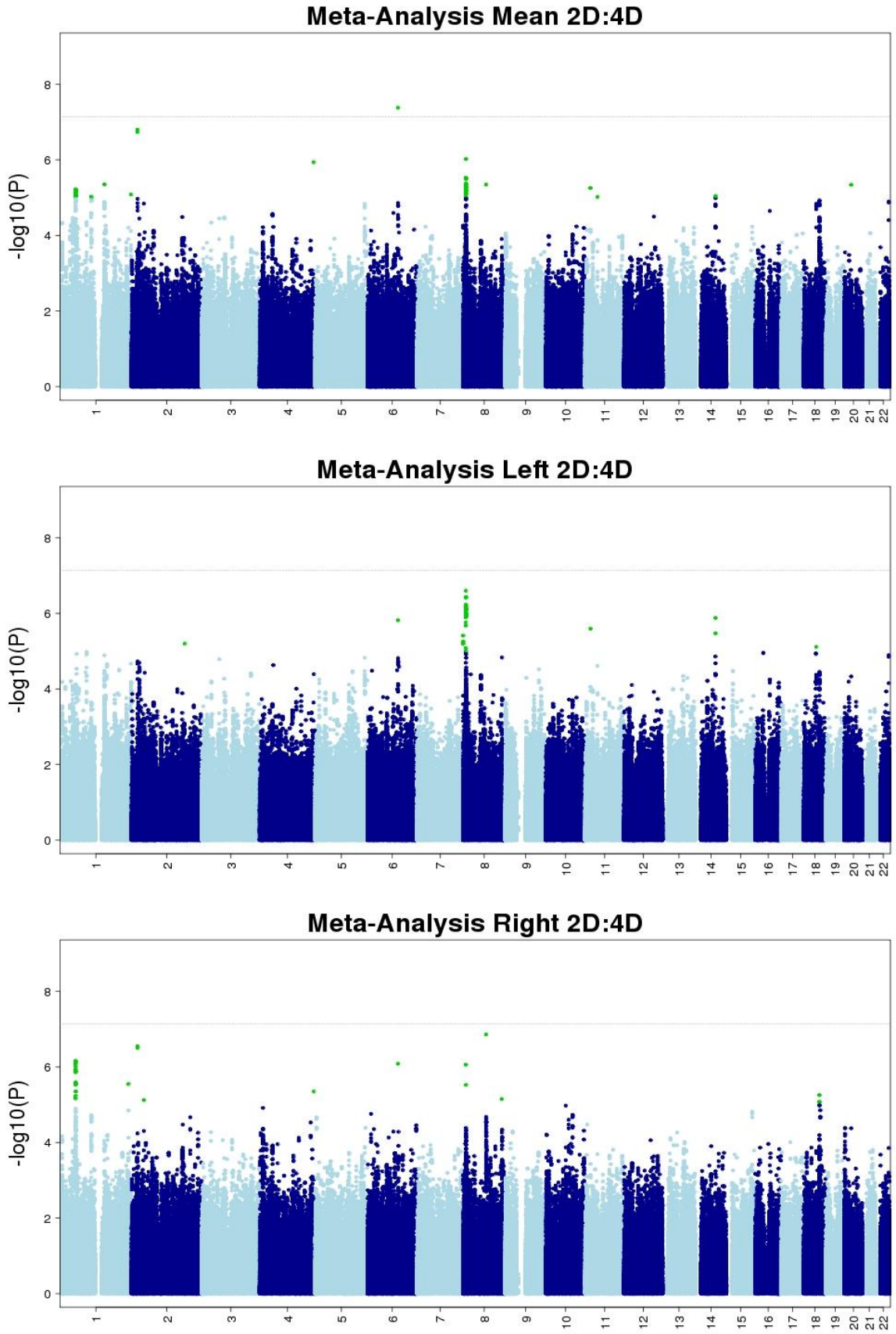


Supplementary Figure 1. QQ plots for analysis of 2D:4D in (A) ALSPAC, (B) QIMR and (C) a meta-analysis of both cohorts. Expected chi-square value under the global null hypothesis of no association is displayed on the x axis. Observed chi-square value is displayed on the y axis. The plots show little evidence of stratification, but suggest the existence of loci of small effect.

A.) Left 2D:4D B.) Right 2D:4D C.) Mean 2D:4D

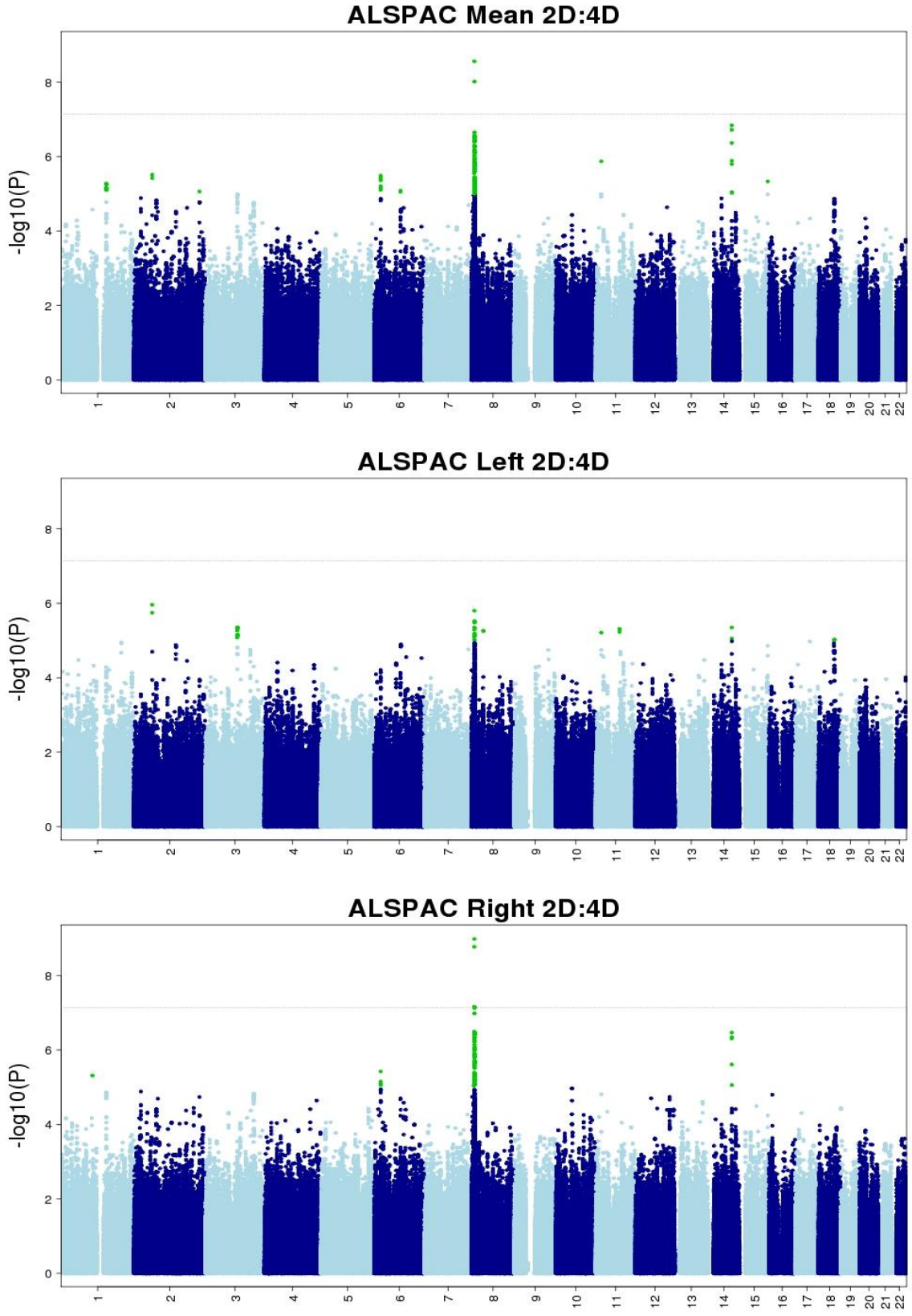


Supplementary Figure 2. QQ plots for analysis of X chromosome markers for (A) left, (B) right and (C) mean 2D:4D. Expected chi-square value under the global null hypothesis of no association is displayed on the x axis. Observed chi-square value is displayed on the y axis.



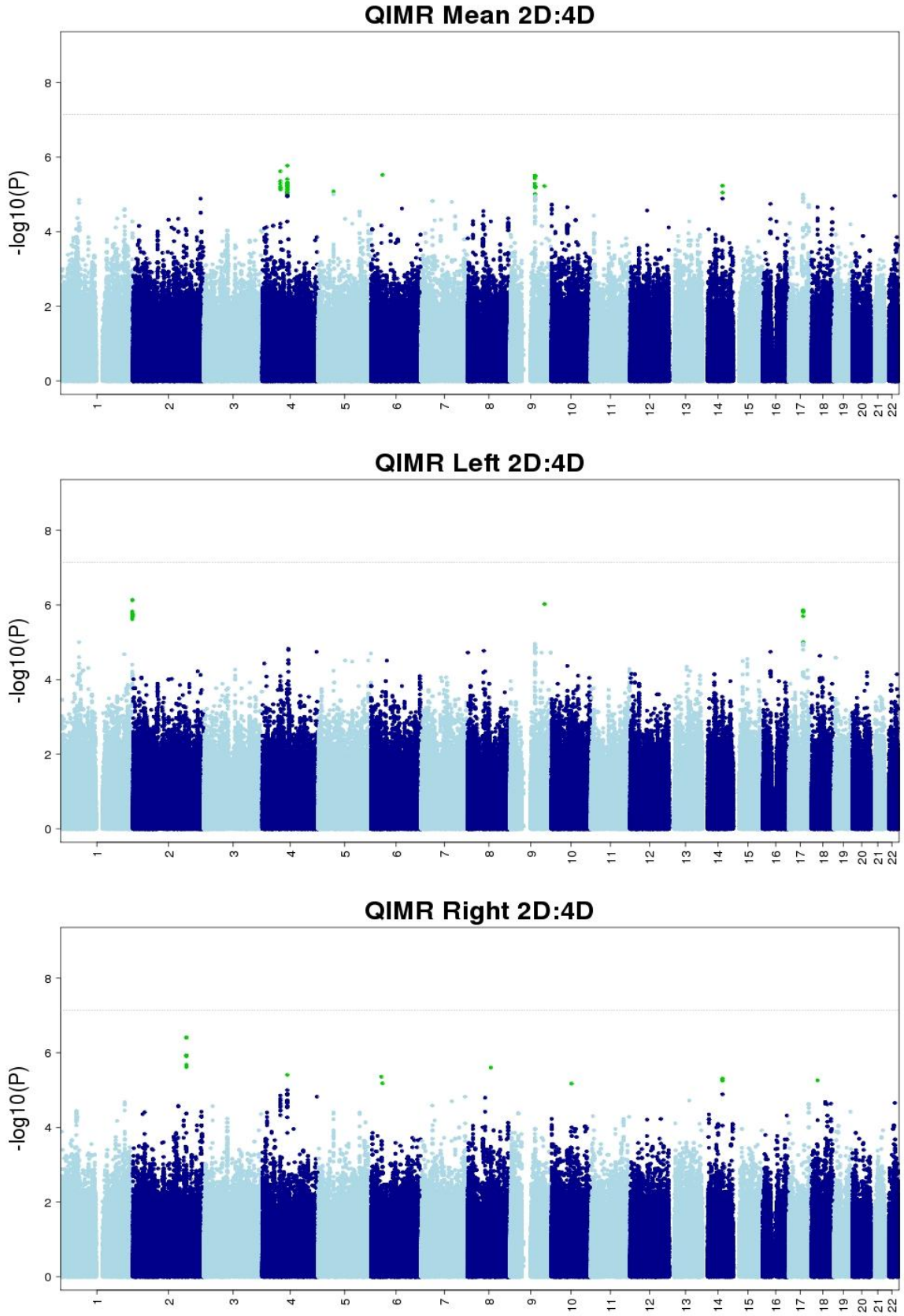
Supplementary Figure 3. Manhattan plots showing genome-wide association results of meta-analyses for the mean, left and right 2D:4D ratios. Green dots represent SNPs that have p values < 10^{-5}. The rs314277 SNP on chromosome 6 is the only one to reach a genome-wide

significance level of $p = 5 \times 10^{-8}$.



Supplementary Figure 4. Manhattan plots showing genome-wide association results for the mean, left and right 2D:4D ratios from the ALSPAC sample. Green dots represent SNPs that have p values < 10^{-5}. There was a large signal on chromosome 8p, but this was not replicated

in the QIMR dataset.



Supplementary Figure 5. Manhattan plots showing genome-wide association results for the mean, left and right 2D:4D ratios from the QIMR sample. Green dots represent SNPs that

have p values < 10^{-5}. No SNPs reached the threshold for genome-wide significance ($p = 5 \times 10^{-8}$).