

Article

GWAS of Dizygotic Twinning in an Enlarged Australian Sample of Mothers of DZ Twins

Scott D. Gordon¹ , David L. Duffy¹, David C. Whiteman¹, Catherine M. Olsen¹, Kerrie McAloney¹, Jessica M. Adsett¹, Natalie A. Garden¹, Simone M. Cross¹, Susan E. List-Armitage¹, Joy Brown², Jeffrey J. Beck³, Hamdi Mbarek⁴, Sarah E. Medland¹, Grant W. Montgomery⁵, Nicholas G. Martin¹ 

¹QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia, ²Independent researcher, Invercargill, New Zealand, ³Avera Institute for Human Genetics, Avera McKennan Hospital and University Health Center, Sioux Falls, South Dakota, USA, ⁴Qatar Genome Institute, Doha and ⁵Institute of Molecular Bioscience, The University of Queensland, Brisbane, Queensland, Australia

Abstract

Female fertility is a complex trait with age-specific changes in spontaneous dizygotic (DZ) twinning and fertility. To elucidate factors regulating female fertility and infertility, we conducted a genome-wide association study (GWAS) on mothers of spontaneous DZ twins (MoDZT) versus controls (3273 cases, 24,009 controls). This is a follow-up study to the Australia/New Zealand (ANZ) component of that previously reported (Mbarek et al., 2016), with a sample size almost twice that of the entire discovery sample meta-analysed in the previous article (and five times the ANZ contribution to that), resulting from newly available additional genotyping and representing a significant increase in power. We compare analyses with and without male controls and show unequivocally that it is better to include male controls who have been screened for recent family history, than to use only female controls. Results from the SNP based GWAS identified four genomewide significant signals, including one novel region, *ZFPM1* (Zinc Finger Protein, FOG Family Member 1), on chromosome 16. Previous signals near *FSHB* (Follicle Stimulating Hormone beta subunit) and *SMAD3* (SMAD Family Member 3) were also replicated (Mbarek et al., 2016). We also ran the GWAS with a dominance model that identified a further locus *ADRB2* on chr 5. These results have been contributed to the International Twinning Genetics Consortium for inclusion in the next GWAS meta-analysis (Mbarek et al., in press).

Keywords: Genetics of twinning; Fertility

(Received 19 July 2023; accepted 18 September 2023; First Published online 23 November 2023)

Spontaneous dizygotic (DZ) twinning rates vary markedly between the major ethnicities (Bulmer, 1970; Monden et al., 2021; Sear et al., 2016). The DZ twinning rate can be a useful index of fertility in a population (Tong & Short, 1998), and age-specific changes in DZ twinning may reflect declining ovarian follicle reserve and individual fertility (Beemsterboer et al., 2006). On the other hand, twinning is associated with increased risks to maternal and infant health (Monden & Smits, 2017; Santana et al., 2016; Santana et al., 2018) and it is important to understand the factors regulating DZ twinning frequency and their relationship to fertility and reproductive aging. We previously reported the first two significant loci for being a mother of spontaneous DZ twins (MoDZT), based on a GWAMA of 1980 cases plus 12,953 controls (Mbarek et al., 2016). We found loci close to two genes, *FSHB*, the structural locus for follicle stimulating hormone (FSH) beta subunit and *SMAD3*, which regulates the response of the ovaries to FSH, and replicated

these findings in gene-based testing of ‘being a DZ twin’ in the UK Biobank. We have now expanded the Australian and New Zealand component of that study to 3273 MoDZT plus 24,009 controls of European ancestry.

Methods

The analysis was a case/control genome-wide association study (GWAS), where the cases were MoDZT and ‘controls’ were those not known to be cases and who have no self-reported closely related DZ twins (based on known pedigrees, or relatedness with cases above π -hat ~ 0.1 in an Identity By Descent (IBD) analysis; see later description under ‘Controls’). No phenotype information was used apart from zygosity of the twins, assisted reproductive technology (ART) status (see below) and family history of DZ twinning, as foregoing. All studies were approved by the QIMR Berghofer Human Research Ethics Committee.

Cases

Cases were genotyped MoDZT drawn from a twin family dataset collected and held by QIMR Berghofer (‘QIMR’). This primarily included (a) families with twins from various studies that had

Corresponding author: Nick Martin; Email: Nick.Martin@qimrberghofer.edu.au

Cite this article: Gordon SD, Duffy DL, Whiteman DC, Olsen CM, McAloney K, Adsett JM, Garden NA, Cross SM, List-Armitage SE, Brown J, Beck JJ, Mbarek H, Medland SE, Montgomery GW, Martin NG. (2023) GWAS of Dizygotic Twinning in an Enlarged Australian Sample of Mothers of DZ Twins. *Twin Research and Human Genetics* 26: 327–338, <https://doi.org/10.1017/thg.2023.45>

© The Author(s), 2023. Published by Cambridge University Press on behalf of International Society for Twin Studies. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.



recruited twins and members of their families since 1978 via the Australian Twin Registry and from whom DNA samples were subsequently collected (most of these twins were born before 1972 so ART is not an issue); (b) families with twins born after 1980 and enrolled in the Queensland Twin Registry (the Brisbane Adolescent Twin sample (BATS), Project 5 in Medland, Nyholt *et al.* (2009); see also Wright and Martin (2004) and Medland, Duffy *et al.* (2009)); (c) the QIMR Twinning Genetics (TG) study, which recruited MoDZT from multiplex families with closely related MoDZT, mainly sisters whom we had previously used in a sib-pair linkage study of DZ twinning. Participants were from Australia or New Zealand and ascertained through Mothers-of-Twins clubs and elsewhere (Painter *et al.*, 2010).

The two big issues with acceptability of a MoDZT case for our study are (1) that her twins are definitely DZ, and (2) that she conceived the twins spontaneously and did not use any ART. Both issues are addressed below. A schema summarizing our ascertainment procedure with regard to both ART and zygosity is shown in Appendix A.

Controls

Controls were drawn from two groups from the general Australian population, also filtered genetically to include only Europeans: (1) individuals from the studies supplying case groups (a) and (b) above where there were no known DZ twins in the family and genotyping was available; (2) genotyped individuals from the QIMR Berghofer QSkin Study (Olsen *et al.*, 2012) which is an unselected sample of Queenslanders aged 40–70 randomly contacted via the electoral roll (voter registration is compulsory in Australia and compliance is 97%). Around 44,000 completed an online survey on various aspects of health, but with a focus on skin cancer. All were invited to donate a saliva sample. QSkin subjects who were genotyped and had consented to re-contact were contacted by email and asked to complete a brief online screening questionnaire (Appendix B) about their family history of twinning including (female participants) whether they themselves had had twins and if so, their ART status and zygosity of the twins. Of the genotyped 17,642 QSkin subjects, we excluded 378 due to non-European ancestry, followed by 897 who were DZ twins themselves, a mother of twins, or had a close twin relative (leaving 16,367 at this point). For QSkin and elsewhere, we retained controls even if they were male, after running the final analysis twice: first with both male and female controls ($N = 24,009$, the main analysis), and second with only female controls ($N = 12,819$ controls). This showed that the resulting near doubling of sample size for controls outweighs the fact that male controls, by virtue of their inability to become pregnant, were only screened for a predisposition towards DZ twinning, based on limited family history information (see Results).

In the final GWAS analysis, tight relatedness exclusions were applied to controls who had an IBD-based π -hat ≥ 0.1 with any DZ twin or case/mother of DZ twin (even unselected cases) in a GRM including all individuals with available genotyping. In these instances the controls were removed, as they greatly outnumbered cases. Genotyped controls (both sexes) remaining after all QC, comprised QSkin ($N = 14,577$) and other controls as in (1) above, of which $N = 658$ were from the same batch of GSA genotyping as QSkin, $N = 3209$ typed on Core+Exome, PsychArray, Omni2.5, OmniExpress chips; $N = 2442$ typed on the 610K chip, and $N = 3123$ typed on 317K or 370K chips for a total of total $N = 24,009$ controls (12,820 female, 11,189 male).

Zygosity – Ensuring Twins Were DZ

For opposite-sex pairs (about one third of spontaneous European twins), zygosity is unequivocally DZ, but for same-sex pairs, zygosity was initially based on responses to standard questions regarding similarity of appearance. Later this was refined by genetic testing, initially using the Profiler[®] microsatellite set but more recently using genomewide genotyping with a SNP array so that in most cases final zygosity is based on direct genetic evidence. Direct genotyping of the twins and a variety of clinical or laboratory tests performed at and after recruitment already made the zygosity of twins for whom DNA was available quite certain. However, in many cases, the twins of a mother selected as a potential case had not been genotyped so we could not genetically confirm zygosity. In these cases, zygosity was established by questionnaires containing standard zygosity items and, in cases of remaining doubt, by telephone interview. If, after these steps, zygosity was still in doubt then the mother was excluded from the MoDZT case sample.

Screening Out MoDZT Who Had Used ART

There has been a large increase in DZ twinning following the widespread adoption from the 1980s on of ART, such as in vitro fertilization (IVF); before 1980 use of ART was uncommon (Monden *et al.*, 2021). In our previous work (Mbarek *et al.*, 2016) we showed that although ART accounts for only a small fraction of twin pregnancies, the sensitivity of the phenotype definition to contamination with DZ twins arising after ART is high, with as few as 10% ART cases entirely ablating any association signals. In addition to dilution of a genetic signal due to phenocopies, it is possible that some women requiring ART to conceive are genetically predisposed to low fertility and their inclusion would therefore tend to cancel the very genetic signals we are looking for.

For mothers contemplated for inclusion as MoDZT cases, where twins were born after 1980 and we had no existing ART screening information, we therefore made a substantial effort to re-contact the mother by telephone or email to confirm their ART status (see questionnaire in Appendix C) and at the same time to confirm twins' zygosity. Phase 1 of this effort involved attempting to contact 755 mothers of twins in the youngest (mostly QTwin-derived) cohort with 397 (53%) completing screening questions. Phase 2 involved contacting other mothers targeted for new GSA genotyping outside of the original Twinning Genetics Study; with 484 mothers contacted successfully; 81 of whom failed the ART exclusion and/or zygosity check; with another 427 uncontactable. Those failing or uncontactable were excluded from further genotyping, with an initial highest-priority candidate list of 793 individuals being reduced by 78 (9.8%) due to screening questions and 177 (22.3%) by lack of contact (538 remaining). An overall list of known ART cases from both phases was also excluded from the analysis itself. The same questionnaire also confirmed that pre-1980 use of ART was indeed rare, which gives some confidence in assuming lack of ART in that age group.

These efforts, along with existing screening data for other individuals, will have removed the great majority of potential cases where twin pregnancies were due to ART or twins were not DZ. However, there remain some mothers with pre-existing genotyping who could not be contacted and have not been screened for ART or re-contacted for zygosity. We chose to retain these uncontactable cases and so not filter them out explicitly, on the basis that (1) the great majority of these births were pre-1980 and unlikely to use ART; (2) zygosity error is known to be quite low.

Genotyping

These data are held as an integrated dataset (R10) expanded (added 660K, Omni family and GSA genotyping including QSkin) and reimputed from that used in Medland, Nyholt, et al. (2009); also used in Wood et al. (2014) and elsewhere. Genotyping was taken as a subset of the imputed version of an existing dataset of ~51,834 individuals. These were genotyped in batches of up to ~22,000 on successive generations of Illumina SNP arrays from 2006 onward. Genotyping as a result was spread across three families/generations of Illumina chip (listed chronologically): (1) HapMap-derived 317K/370K/610K; and (2) 1000 Genomes-derived, Omni family (Core+Exome, PsychChip, Omni 2.5, Omni Express). (3) Global Screening Array (GSA v1); this latter includes Qskin control samples which were typed on the Illumina Global Screening Array (model GSAMD-24v1-0_20011747_A1) array at Erasmus University Medical Centre's Human Genomics Facility (HuGe-F, André Uitterlinden director). It also includes a batch of 1922 samples (1806 after QC), including 775 MoDZT sister pairs, 329 MoDZT from parent-child trios, 309 singleton MoDZTs and 509 MoDZT from other studies, which were genotyped at the Avera Institute of Human Genetics, Sioux Falls SD, at Avera Health using the Avera-NTR GSA array (model GSA-24v1-0_A1_avera_20170513f, a semi-custom SNP array made by Illumina (Beck et al., 2019).

The entire dataset contains $N = 19,428$; 14,210; 23,821 individuals from HapMap-derived Illumina arrays; 1000G-derived Illumina arrays; and GSA v1 respectively. Analyzed individuals are a subset of this with $N = 6322$; 3922; 17,038 individuals respectively.

Individual batches of genotyping were called and QC'd as per standard protocols (within and post GenomeStudio, including dropping individuals <97%, and SNPs with GenTrain Score < 0.6, $MAF < 1\%$, $p(HWE) < 10^{-6}$, $mean(GC) < 0.7$, call rate <95%) before being integrated into an overall dataset. Further Mendelian and relatedness checks were performed by Identity By Descent (IBD) calculation run on the overall dataset. Batch effects at the genotyping stage are not significant after QC; they were checked by various means including allele frequency comparisons and repeat genotyping of the same samples. Family-based data cleaning prior to GSA genotyping allowed most sample issues to be identified and resolved. Other misidentified or wrong sex samples were removed. The X chromosome data were available post-QC.

Data Cleaning and Imputation

A Principal Components Analysis (PCA) was run using *smartpca* (in EigenSoft 7.2.0; original version described in Price et al., 2006) with samples from HapMap Phase 3 and Genome-EUTWIN populations used to define the PC axes; all QIMR genotyping was projected onto those axes. About 3% of individuals (from the overall dataset) with PC1 and/or PC2 >6sd from the mean of European reference populations were excluded as ancestry outliers. In the association analysis, PC1-PC4 were also used as covariates in all analyses, although from past experience, there is minimal population stratification after the stated exclusions.

Imputation was run on the Michigan Imputation Server using SHAPEIT (phasing), minimac3 (imputation) and the HRCr1.1 reference panel, for each of the three SNP-array families. Each run used only individuals genotyped with those arrays, and observed markers passing QC in all relevant batches (19,428 individuals, 280,280 markers HapMap-derived; 14,210 individuals, 238,591 markers 1000G-derived; 23,821 individuals, 481,926 markers

GSA) and merged post-imputation, preferring HapMap-derived; 1000G-derived; GSA in that order for each individual. Two binary analysis covariates record that choice of imputation run, to model any differences in imputation between the chip families.

Association Analysis (GWAS)

Although controls closely related to cases ($\pi\text{-hat} \geq 0.1$) were removed to improve the quality of controls, cases are drawn primarily from family-based studies, as are the non-QSkin controls, so substantial relatedness still exists within both the case and control groups. Reducing the analysis to an 'unrelateds only' analysis would have considerably reduced the sample size and power of the analysis. Hence the association tests were run using the R package SAIGE (v 0.36.3.3) (Zhou et al., 2018), which uses a Generalised Linear Mixed Model (GLMM) to model relatedness as a part of the initial stage of its analysis, and also has a high tolerance to unbalanced case control ratios as applies here. SAIGE 'Step 1' (the null GLMM fit) was run with, as input, the phenotype and covariates plus a genome-wide set of observed genotypes for ~39,500 markers available in all batches of genotyping. Phenotype was the case/control variable coded 1 = case ($N = 3273$), 0 = control ($N = 24,009$). Covariates were the first 4 genetic Principal Components (PCs) from the *smartpca* ancestry PCA run (PC1-PC4), plus two binary covariates encoding which of the three imputation runs was used for the individual in question. SAIGE 'Step 2' was then run in parallel (for speed) across blocks of ~50,000 markers with as input the Variance Ratio and GMMAT model estimate files from 'Step 1' plus imputed genotypes (against the HRC r1.1 reference panel; or TOPMed r2, females-only, for the X chromosome) for that block of markers. Step 2 used a minor allele frequency (MAF) cutoff of 0.001 and minimum minor allele count (MAC) cutoff of 5. The results for these blocks were then concatenated and filtered based on imputation metadata. For the purposes of results presented here, only markers with $MAF \geq 0.01$ (1%) and $r^2 \geq 0.6$ (the lowest, that is, worst r^2 across the three imputation runs) were retained.

Results

The results of the SAIGE SNP-based GWAS analysis are summarized in Figure 1 (a Manhattan Plot, i.e., p value vs. genomic position) and Figure 2 (the corresponding Quantile-Quantile or Q-Q Plot), after the chosen filter of $MAF \geq 0.01$ (1%) and $R_{sq} \geq 0.6$. These show four unique association peaks below the conventional genome-wide significance threshold for GWAS results, $p = 5 \times 10^{-8}$, along with a number of narrowly below significant associated regions that are not explored further here. The significant regions are listed in Table 1. The Genomic Inflation Factor (λ) is ~1.125 (1.131 for autosomes). No significant peaks are seen on the X-chromosome.

Figure 3 explores the question of *whether to include male controls in the analysis*; cases are, by definition, female, so it might be intuited that one should only use female controls. The equivalent GWAS was run for the autosomes using only female controls ($N = 12,819$; λ for this female-only analysis is ~1.085) to compare with the analysis using controls of both sexes ($N = 24,009$); both analyses used $N = 3273$ cases. The two sets of (QC'd) p values are plotted against each other and show that the analysis using female controls only consistently results in less significant p values for markers below, or close to, the genomewide significance threshold (5×10^{-8}). This is particularly notable for the

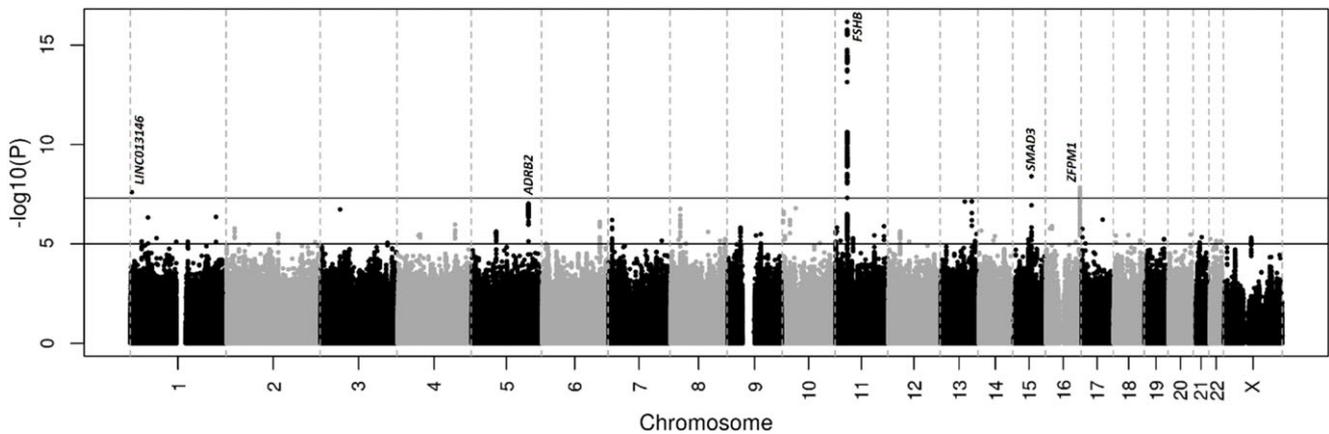


Figure 1. Manhattan plot for additive model (MoDZT vs. controls). Shows association p values calculated by SAIGE ($-\log_{10}$ transformed) after removing markers with MAF < 1% or $R_{sq} < 0.6$. Genome-wide significance threshold $p = 5 \times 10^{-8}$ is the upper line. Suggestive threshold 10^{-5} is the lower line. Horizontal axis is genomic position, measured in basepairs, increasing toward the right (hg19). Labels match the genes quoted for the adjacent peak in Table 1.

Note: MoDZT, mothers of dizygotic twins; MAF, minor allele frequency.

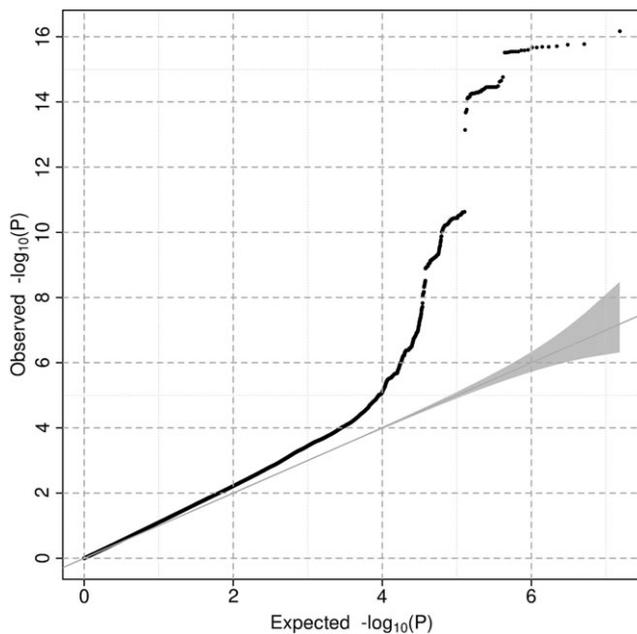


Figure 2. Q-Q plot for additive model (MoDZT vs. controls). Shows association p values calculated by SAIGE ($-\log_{10}$ transformed, vertical axis) after removing markers with MAF < 1% or $R_{sq} < 0.6$; against p value expected at that rank. The red line and shaded area mark the 95% confidence interval under the null ($\lambda = 1$). Observed $\lambda = 1.131$.

Note: MoDZT, mothers of dizygotic twins; MAF, minor allele frequency.

most-associated *FSHB* and *SMAD3* regions; however, there is no pervasive shift from the $y = x$ diagonal for nonassociated high p markers. We take this as our justification for including male controls in our GWAS analyses to maximize power. This is to be expected for a low prevalence ($\sim 1\%$) trait like DZ twinning where screening out ‘affecteds’ (which we could not do perfectly) makes little difference.

MAGMA v1.0.9a (de Leeuw *et al.*, 2015) was used to perform gene-based association tests using the supplied 1000 European reference genome for LD information. A Manhattan plot showing the resulting p values for the 19,104 genes returning a result is

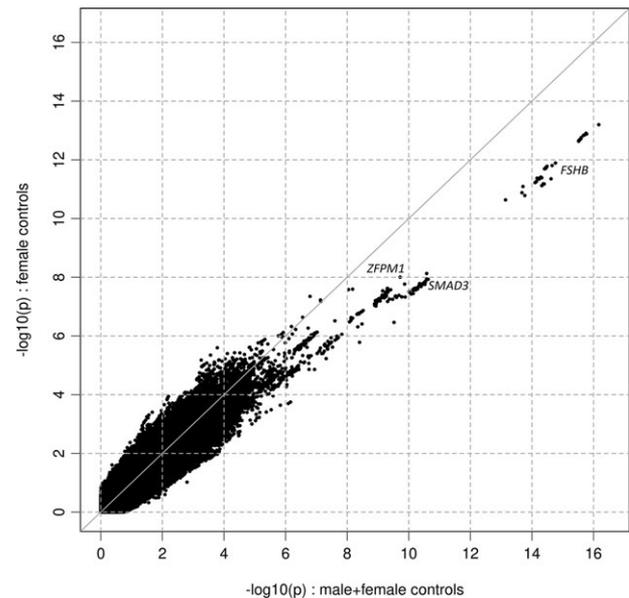


Figure 3. A comparison plot, marker by marker, between p values for that marker in the main analysis with both male and female controls (horizontal axis; 3273 cases, 24,009 controls) and an alternative analysis with only female controls but otherwise equivalent (vertical axis; 3273 cases, 12,819 controls). The diagonal shows $y = x$. Labels show indicative locations of the three strongest association peaks in the main analysis.

shown in Figure 4. The equivalent tests were also run using VEGAS v2.2 (Mishra *et al.*, 2015) with broadly consistent results for the most associated genes (not presented here). Both programs produce results that are broadly consistent with the strongest SNP associations (Table 1) although some regions are significant at the gene-based level but not at the individual SNP level.

The four per-SNP association peaks are shown as regional association plots, with LD information, as Figure 5 (a–d). For comparison purposes, Figure 5(f) also shows a version of panel (a) referenced to the previously published ‘peak’ SNP rs11031006 for the *FSHB* peak; and panel (g) shows the remaining ‘peak’ SNP rs12064669 from Table 3 of Mbarek *et al.* (2016), which no longer appears associated and is likely to have been a false signal.

Table 1. Genomewide significant genes and loci for additive model (MoDZT vs. control), after removing markers with MAF < 1% or R_{sq} < 0.6. Positions are for NCBI Human Genome Build 37 (also known as hg19), as chromosome:basepair_position. Only the most associated SNP for each gene/region is shown; each association peak has up to 221 genomewide significant SNPs (for *FSHB*) or 19 (for *SMAD3* case).

Gene	Gene-based test (MAGMA v1.09a)				Top SNP in gene						
	CHR	START bp	STOP bp	p	rsID	Bp position	Ref/eff alleles	Eff allele frequency	Beta (SE)	p	
<i>FSHB</i>	11	30202563	30306824	6.24×10^{-21}	rs74485684	30242287	T/C	0.1542	-0.3568 (0.0427)	6.78×10^{-17}	
<i>ARL14EP</i>	11	30294649	30409165	1.60×10^{-20}	rs75525300	30307989	A/G	0.1540	-0.3517 (0.0428)	1.95×10^{-16}	
<i>MPPED2</i>	11	30356040	30658260	8.75×10^{-15}	rs12294104	30382899	C/T	0.1665	-0.3255 (0.0411)	2.40×10^{-15}	
<i>ADRB2</i>	5	148156156	148304628	2.02×10^{-8}	rs4705276	148249230	A/G	0.1489	0.2245 (0.0421)	9.46×10^{-8}	
<i>ZNF469</i>	16	88443879	88557165	2.14×10^{-8}	rs4426338	88528348	T/C	0.3086	-0.1865 (0.0329)	1.49×10^{-8}	
<i>ZFPM1</i>	16	88470014	88651574	5.26×10^{-8}	rs4426338	88528348	T/C	0.3086	-0.1865 (0.0329)	1.49×10^{-8}	
<i>SMAD3</i>	15	67308036	67537533	4.71×10^{-7}	rs17293443	67437863	T/C	0.2270	0.2144 (0.0364)	3.97×10^{-9}	
<i>CLYBL</i>	13	100208919	100599388	2.43×10^{-6}	rs2038695	100559123	C/A	0.5408	-0.1676 (0.0311)	7.16×10^{-8}	
<i>IPO8</i>	12	30731915	30898929	3.22×10^{-6}	rs7316477	30786098	A/G	0.5108	0.1416 (0.0300)	2.36×10^{-6}	
<i>CAPRIN2</i>	12	30812486	30957448	6.23×10^{-6}	rs6487927	30826335	T/C	0.5122	0.1405 (0.0299)	2.66×10^{-6}	
<i>ADM5</i>	19	50141942	50244247	4.28×10^{-6}	rs12980063	50196992	A/G	0.3846	0.1409 (0.0311)	5.76×10^{-6}	
<i>DOCK5</i>	8	24992287	25323227	8.91×10^{-6}	rs6185	25280800	C/G	0.2434	-0.1917 (0.0367)	1.73×10^{-7}	
<i>LINC01346</i>	1	3999474	4013839	N/A	rs6426385	4010889	C/T	0.6631	0.2202 (0.0395)	2.51×10^{-8}	
<i>RLIM</i>	X	73752811	73884461	2.24×10^{-6}	rs5981637	73764809	A/G	0.0320	0.2862 (0.0626)	4.87×10^{-6}	

Note: MoDZT, mothers of dizygotic twins; SNP, single nucleotide polymorphism.

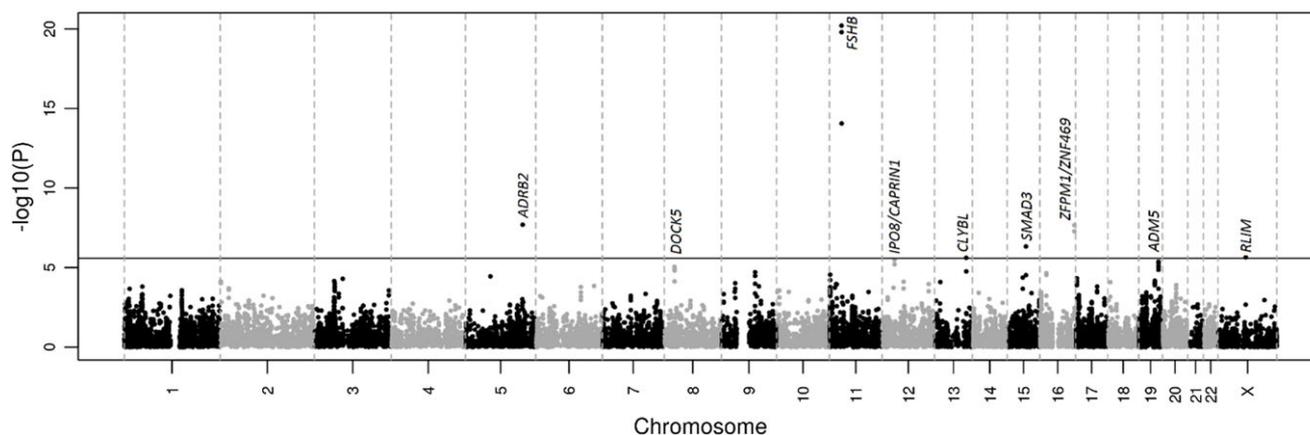


Figure 4. Manhattan plot for MAGMA gene-based test (showing field P_MULTI – composite gene-based p value. The horizontal line is at $p = 2.61 \times 10^{-6} = 0.05/N$ (N = number of tested genes = 19,154). Labels show the most-associated significant or near-significant gene(s) adjoining the label.

The strongest two associations are those on chromosomes 11 and 15, which reproduce the strongest two associations from our previous paper (Mbarek et al., 2016) with the expected higher statistical significance. We previously concluded that both relate to genes *FSHB* (Follicle Stimulating Hormone Beta subunit) and *SMAD3* (SMAD Family Member 3) respectively, which are both strong candidates as DZ twinning genes. The MAGMA gene-based tests strongly support both genes along with (at lower significance) the adjacent genes *ARL14EP* (ADP Ribosylation Factor Like GTPase 14 Effector Protein) and *MPPED3* in the case of *FSHB* (Table 1).

The most convincing of the two other associated regions is at rs4426338, a novel locus close to the gene *ZFPM1* (Zinc Finger Protein, FOG Family Member 1) on chromosome 16. The LD block also covers the smaller genes *ZNF469* and *MIR5189* (see

Figure 5c). In fact, the MAGMA gene-based p value for *ZNF469* is slightly lower (2.14×10^{-8} vs. 5.26×10^{-8} for *ZFPM1*; Table 1).

The remaining genomewide significant association at rs6426385 on chromosome 1 has no other marker (after our QC filter) with a p value below 2.85×10^{-4} , far higher than the genomewide threshold 5×10^{-8} (Figure 5d). The lack of other associated markers nearby suggests that this is a false-positive result and it is not considered further here. The only nearby gene is noncoding RNA gene *LINC01346*.

Also prominent and at or near significance in the gene-based results (Table 1; Figure 4) but not reaching genomewide significance for individual SNPs, are *ADRB2* on chromosome 5; *DOCK5/GNRH1/KCTD9* on chromosome 8; *IPO8/CAPRIN2* on chromosome 12; *CLYBL* on chromosome 13; *ADM5/PRMT1/IRF3* on chromosome 19; and *RLIM/SLC16A2* on chromosome X.

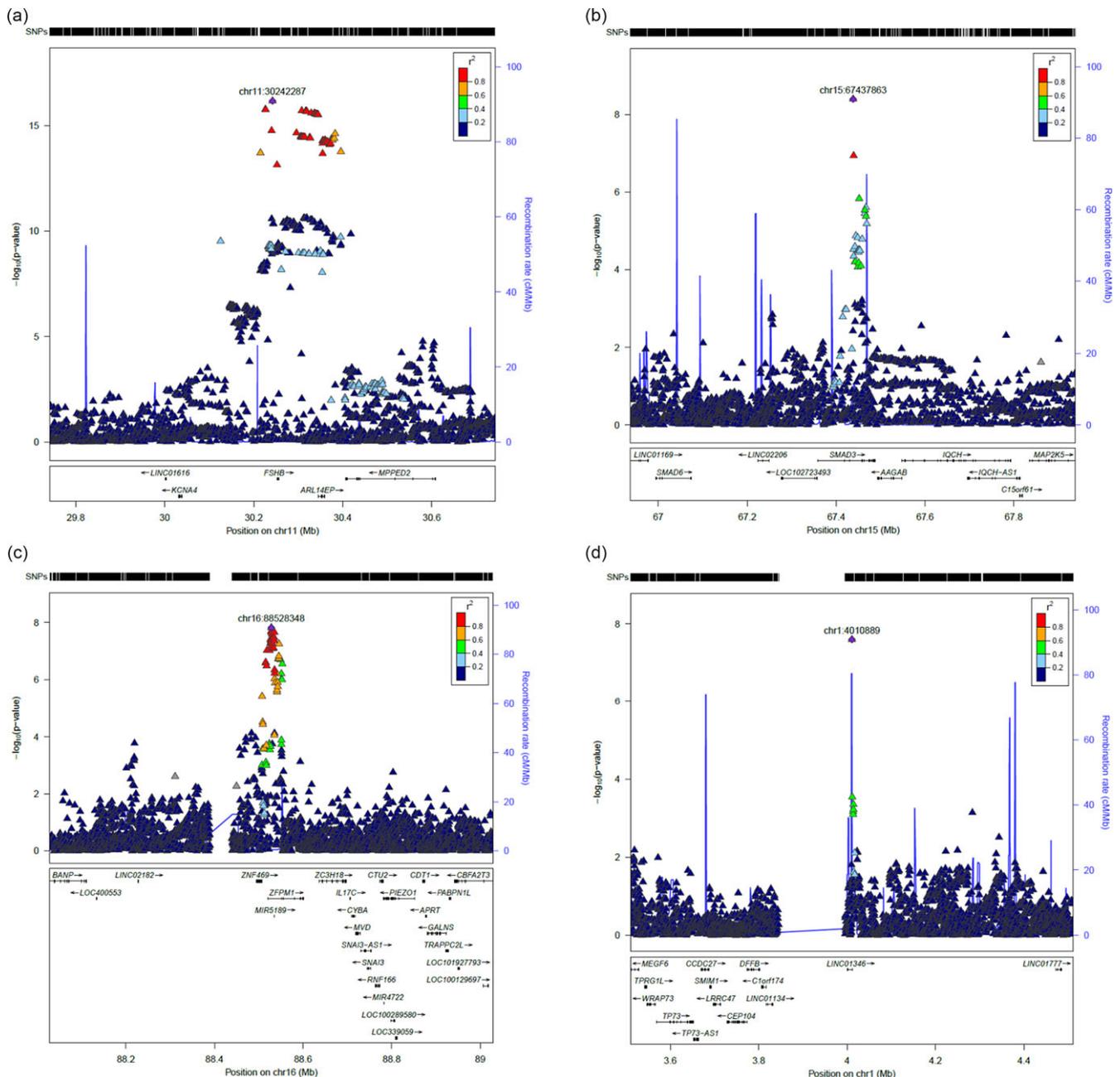


Figure 5. Regional association plots and gene annotation for association peaks. Panels (a)-(d) show the four dominant peaks in Table 1, and panel (e) the chr X peak; reported SNP shown as a purple diamond. Panels f-g shows the additional two peaks reported in the previous meta-analysis paper (Mbarek *et al.*, 2016); (f) rs11031006/11:30226528, which is a lesser-associated chromosome 11 SNP also visible in panel (a), but was the lead chromosome 11 marker reported in Table 3 of the previous paper); and (g) rs12064669/1:230688643 which was reported in that paper, but is not associated in the current analysis. Panels (h)-(i) show the peak surrounding *ADRB2* both in the additive model (h) and dominant model (i), referenced to the SNP most-associated in the additive model (rs4705276). All plots prepared using LocusZoom v1.4 with LD r^2 estimates derived from 1000 Genomes v3 European genotypes. Grey indicates lack of LD information versus the peak marker (chromosome X and some individual markers). Note: SNP, single nucleotide polymorphism.

ADRB2 also features in the Dominance Model results (see below). It and *CLYBL* are close to SNP-wise genomewide significance (Table 1).

Fitting a dominance model. Bulmer (1970) interpreted his mother-daughter recurrence data (tetrachoric correlation $r = .08$) being significantly less than that of sisters ($r = .14$, $p = .007$) as implying a strong nonadditive genetic contribution to DZ twinning. We therefore repeated our GWAS using a dominance model, by rerunning SAIGE for the purpose of this exploratory

study, with the input dosage values recoded from the imputed genotype probabilities to treat the heterozygote genotype as a double minor allele (i.e., a minor allele dosage of 2 rather than the normal 1 for heterozygotes). The results were effectively identical to the main additive model except for the region in/around the *ADRB2* gene (ADRenocceptor Beta 2) on chr 5, with the p value for lead SNP rs4705276 near this gene changing from 1.04×10^{-7} under an additive model to 2.02×10^{-8} under a dominance model (note: run against TOPMedr2 instead of HRCr1.1 used elsewhere).

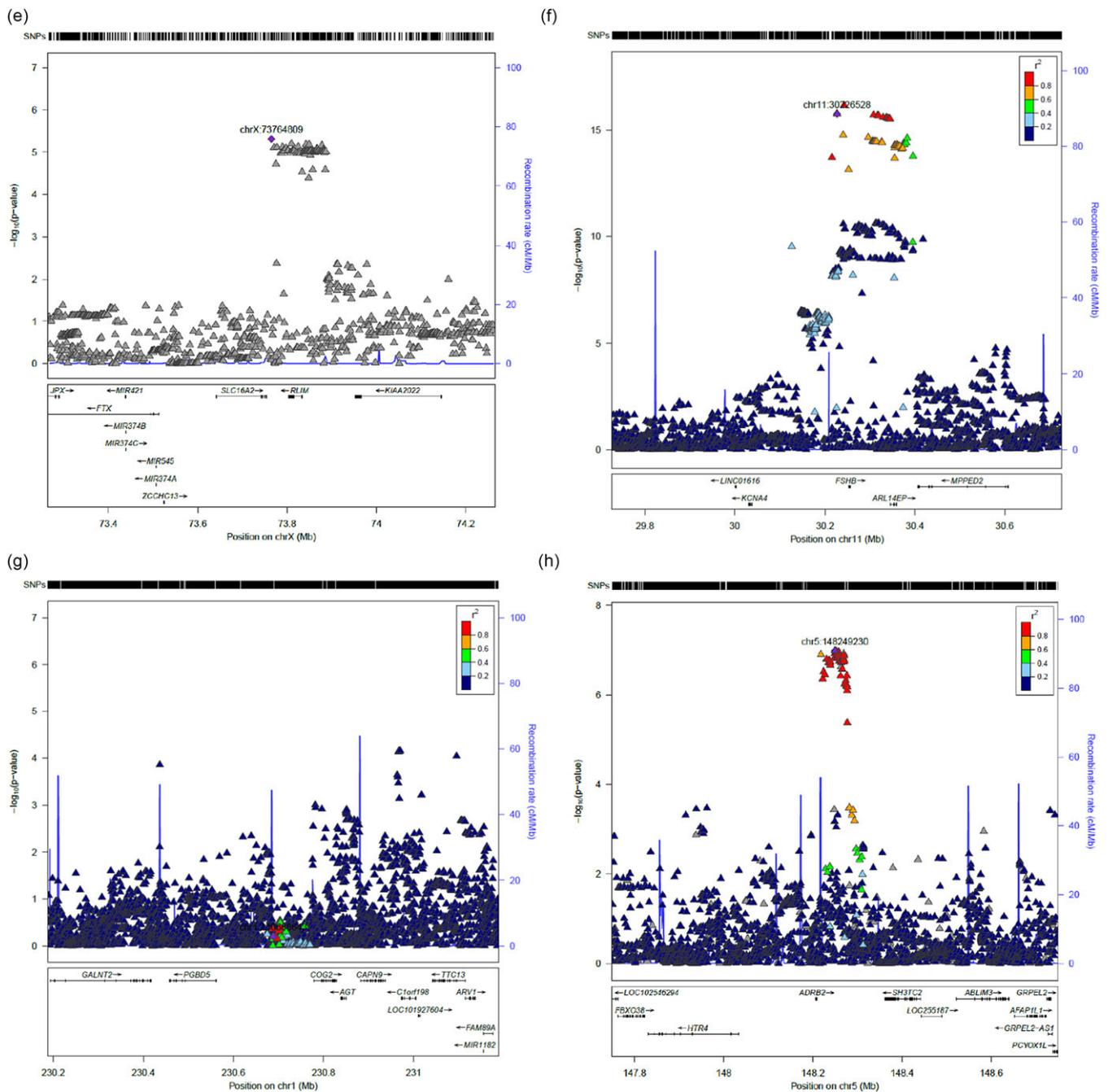


Figure 5b. (Continued)

Interestingly, *ADRB2* is also the most associated gene in the MAGMA gene-based results, after the chr 11 peak (gene-based $p \sim 2.02 \times 10^{-8}$). Refer also to Figure 5h-i.

Discussion

Our results, with almost double the number of MoDZT cases (3273 vs. 1980) over the previous Mbarek et al. (2016) article, have consolidated and expanded the list of genes involved in spontaneous DZ twinning, at least in Europeans. However, it should be noted that ~600 cases in the present study also contributed to the meta-analysis in the 2016 paper, so the results of the two studies are not wholly independent.

For the *FSHB* peak (Figure 5a), we find a different peak SNP (rs74485684), which is only ~16 kb away from our previously-reported rs11031006, and well within the same LD block, just ~10kb upstream of *FSHB*. *FSHB* is a strong DZ twinning candidate gene as FSH plays a central role in regulating ovarian follicle growth and ovulation (e.g., Trevisan et al., 2019). Previously we confirmed in an independent Icelandic population that rs11031006 is associated with changed FSH levels (Mbarek et al., 2016). However, a role for the other gene under the association peak, *ARL14EP*, cannot be ruled out.

For the *SMAD3* peak (Figure 5b), we reproduce the previously found peak SNP rs17293443, which is within that gene’s boundaries. *SMAD3* is suspected (based on mouse studies and

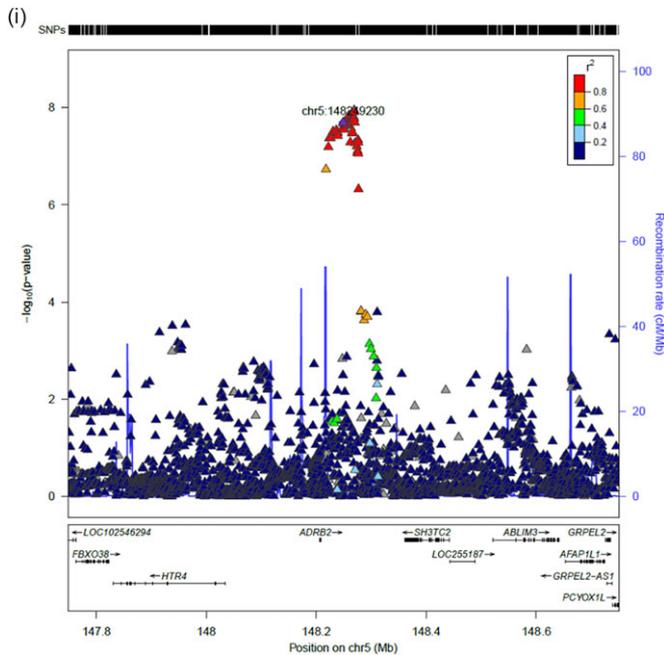


Figure 5c. (Continued)

experiments on human tissue in culture) to be strongly expressed in human ovarian luteinized granulosa cells; to stimulate production of estradiol and progesterone; with possible roles in regulating *FSHR* (Follicle Stimulating Hormone Receptor) and other genes (e.g., Li *et al.*, 2017; Liu *et al.*, 2014).

The lead SNP on chromosome 16 near *ZFPMI* (Figure 5c) rs4584807 is in high LD with nearby SNPs associated with age at menopause (Mbarek *et al.*, 2016; Mbarek *et al.*, 2019), sex hormone binding globulin (Ruth *et al.*, 2020), and FSH concentrations (Chen *et al.*, 2013; Mbarek *et al.*, 2016).

The chromosome X gene-based association peak is near *RLIM* (see also Figure 5e), known also as *RNF12*, which was first identified as an X-linked activator of X chromosome inactivation (Jonkers *et al.*, 2009). This E3 ubiquitin ligase is involved in the regulation of LIM-homeodomain transcriptions factors. It has been linked to a variety of biological activities, including the appropriate expression of GnRH receptor (GnRHR), which is necessary for the correct regulation of the gonadotropins, LH and FSH, by GnRH (Bach *et al.*, 1999; McGillivray *et al.*, 2005). Also nearby is gene *SLC16A2*, which is involved in transport of thyroid hormones.

Fitting a dominance model, we also find significant evidence for a fourth gene, *ADRB2* (Figure 5h-i), which is a plausible twinning candidate. *ADRB2* belongs to the family of the beta-adrenergic receptor. It has been shown that both alpha and beta-adrenergic receptors play a role in rodent ovulation (Kannisto *et al.*, 1985, p. 361: "The experiments indicate that both α - and β -adrenergic receptors mediate an increased rate of ovulation through an effect, on the one hand, at the level of the follicle wall and, on the other hand, by a humoral-type of ovarian mechanism."). Previously, evidence was found pointing to a role for *ADRB2* specifically in rat ovulation (Ratner *et al.*, 1980). *ADRB2* is expressed in human and monkey ovaries and might stimulate both growth of small follicles and induce *FSHR*, contributing to follicular development (Föhr *et al.*, 1993; Mayerhofer, *et al.*, 1997; Merz *et al.*, 2015).

Our article also addresses the perennial question of whether, for a sex-limited trait like DZT, it is better to use same-sex only controls, or controls of both sexes. In accord with expectations for a low prevalence trait, our analyses show unequivocally that we obtain more power using both sexes as controls, even though males cannot be screened for the trait.

Our results are presented here in bare outline and only briefly, because our immediate intention is to contribute them to a large new meta-analysis, to be published shortly (Mbarek *et al.*, *in press*).

Acknowledgments. The authors wish to acknowledge the valuable contribution of Joy Brown, a community volunteer (and a mother of DZ twins herself) who carried out most of the work identifying and recruiting mothers of twins from New Zealand, in her own time.

Author contribution. NGM, DLD, GWM designed the study; NGM, SDG, SEM, GWM analysed data; NGM, SDG, HM drafted the manuscript; NGM and JB contributed to recruitment; DCW, CMO, KM, JMA, NAG, SMC, SEL, JJB, GWM contributed to data and sample collection and/or processing.

Financial support. None.

Competing interests. None.

Ethical standard. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

References

- Bach, I., Rodriguez-Esteban, C., Carrière, C., Bhushan, A., Krones, A., Rose, D. W., Glass, C. K., Andersen, B., Izpisua Belmonte, J. C., & Rosenfeld, M. H. (1999). RLIM inhibits functional activity of LIM homeodomain transcription factors via recruitment of the histone deacetylase complex. *Nature Genetics*, 22, 394–399. <https://doi.org/10.1038/11970>
- Beck, J. J., Hottenga, J. J., Mbarek, H., Finnicum, C. T., Ehli, E. A., Hur, Y.-M., Martin, N. G., de Geus, E. J. C., Boomsma, D. I., & Davies, G. E. (2019). Genetic similarity assessment of twin-family populations by custom-designed genotyping array. *Twin Research and Human Genetics*, 22, 210–219. <https://doi.org/10.1017/thg.2019.41>
- Beemsterboer, S. N., Homburg, R., Gorter, N. A., Schats, R., Hompes, P. G. A., & Lambalk, C. B. (2006). The paradox of declining fertility but increasing twinning rates with advancing maternal age. *Human Reproduction*, 21, 1531–1532. <https://doi.org/10.1093/humrep/del009>
- Bulmer, M. G. (1970). *The biology of twinning in man*. Clarendon Press.
- Chen, Z., Tao, S., Gao, Y., Zhang, J., Hu, Y., Mo, L., Kim, S.-T., Yang, X., Tan, A., Zhang, H., Qin, X., Li, L., Wu, Y., Zhang, S., Zheng, S. L., Xu, J., Mo, Z., & Sun, J. (2013). Genome-wide association study of sex hormones, gonadotropins and sex hormone-binding protein in Chinese men. *Journal of Medical Genetics*, 50, 794–801. <https://doi.org/10.1136/jmedgenet-2013-101705>
- de Leeuw, C. A., Mooij, J. M., Heskes, T., & Posthuma, D. (2015). MAGMA: generalized gene-set analysis of GWAS data. *PLOS Computational Biology*, 11, e1004219. <https://doi.org/10.1371/journal.pcbi.1004219>. See <https://ctg.cncr.nl/software/magma> for update to version 1.08.
- Föhr, K. J., Mayerhofer, A., Sterzik, K., Rudolf, M., Rosenbusch, B., & Gatzl, M. (1993). Concerted action of human chorionic gonadotropin and norepinephrine on intracellular-free calcium in human granulosa-lutein cells: evidence for the presence of a functional alpha-adrenergic receptor. *Journal of Clinical Endocrinology and Metabolism*, 76, 367–373. <https://doi.org/10.1210/jcem.76.2.8381798>
- Jonkers, I., Barakat, T. S., Achame, E. M., Monkhorst, K., Kenter, A., Rentmeester, E., Grosveld, F., Grootegoed, J. A., & Gribnau, J. (2009). RNF12 is an X-Encoded dose-dependent activator of X chromosome inactivation. *Cell*, 139, 999–1011. <https://doi.org/10.1016/j.cell.2009.10.034>
- Kannisto, P., Owan, C., & Walles, B. (1985). Involvement of local adrenergic receptors in the process of ovulation in gonadotrophin-primed immature

- rats. *Journal of Reproduction and Fertility*, 75, 357–362. <https://doi.org/10.1530/jrf.0.0750357>
- Li, Y., Schang, G., Boehm, U., Deng, C. X., Graff, J., & Bernard, D. J. (2017). SMAD3 regulates follicle-stimulating hormone synthesis by pituitary gonadotrope cells in vivo. *Journal of Biological Chemistry*, 292, 2301–2314. <https://doi.org/10.1074/jbc.M116.759167>
- Liu, Y., Chen, X., Xue, X., Shen, C., Shi, C., Dong, J., Zhang, H., Liang, R., Li, S., & Xu, J. (2014). Effects of Smad3 on the proliferation and steroidogenesis in human ovarian luteinized granulosa cells. *IUBMB*, 66, 424–437. <https://doi.org/10.1002/iub.1280>
- Mayerhofer, A., Dissen, G. A., Costa, M. E., & Ojeda, S. R. (1997). A role for neurotransmitters in early follicular development: induction of functional follicle-stimulating hormone receptors in newly formed follicles of the rat ovary. *Endocrinology*, 138, 3320–3329. <https://doi.org/10.1210/endo.138.8.5335>
- Mbarek, H., Gordon, S.D., Mortlock, S., Duffy, D. L., et al. (in press). GWAS meta-analysis of dizygotic twinning illuminates genetic regulation of female fecundity. *Human Reproduction*. <https://doi.org/10.1093/humrep/dead247>
- Mbarek, H., Steinberg, S., Nyholt, D. R., Gordon, S. D., Miller, M. B., McRae, A. F., Hottenga, J. J., Day, F. R., Willemsen, G., de Geus, E. J., Davies, G. E., Martin, H. C., Penninx, B. W., Jansen, R., McAloney, K., Vink, J. M., Kaprio, J., Plomin, R., Spector, T. D., ... Boomsma, D. I. (2016). Identification of common genetic variants influencing spontaneous dizygotic twinning and female fertility. *American Journal of Human Genetics*, 98, 898–908. <https://doi.org/10.1016/j.ajhg.2016.03.008>
- Mbarek, H., van de Weijer, M. P., van der Zee, M. D., Ip, H. F., Beck, J. J., Abdellaoui, A., Ehli, E. A., Davies, G. E., Baselmans, B. M. L., Nivard, M. G., Bartels, M., de Geus, E. J., & Boomsma, D. I. (2019). Biological insights into multiple birth: Genetic findings from UK Biobank. *European Journal of Human Genetics*, 27, 970–979. <https://doi.org/10.1038/s41431-019-0355-z>
- McGillivray, S. M., Bailey, J. S., Ramezani, R., Kirkwood, B. J., & Mellon, P. L. (2005). Mouse GnRH receptor gene expression is mediated by the LHX3 homeodomain protein. *Endocrinology*, 146, 2180–2185. <https://doi.org/10.1210/en.2004-1566>
- Medland, S. E., Duffy, D. L., Wright, M. J., Geffen, G. M., Hay, D. A., Levy, F., van-Beijsterveldt, C. E., Willemsen, G., Townsend, G. C., White, V., Hewitt, A. W., Mackey, D. A., Bailey, J. M., Slutske, W. S., Nyholt, D. R., Treloar, S. A., Martin, N. G., & Boomsma, D. I. (2009). Genetic influences on handedness: data from 25,732 Australian and Dutch twin families. *Neuropsychologia*, 47, 330–337. <https://doi.org/10.1016/j.neuropsychologia.2008.09.005>
- Medland, S. E., Nyholt, D. R., Painter, J. N., McEvoy, B. P., McRae, A. F., Zhu, G., Gordon, S. D., Ferreira, M. A., Wright, M. J., Henders, A.K., Campbell, M. J., Duffy, D. L., Hansell, N. K., Macgregor, S., Slutske, W. S., Heath, A. C., Montgomery, G. W., & Martin, N. G. (2009). Common variants in the trichohyalin gene are associated with straight hair in Europeans. *American Journal of Human Genetics*, 85, 750–755. <https://doi.org/10.1016/j.ajhg.2009.10.009>
- Merz, C., Saller, S., Kunz, L., Xu, J., Yeoman, R.R., Ting, A. Y., Lawson, M. S., Stouffer, R. L., Hennebold, J. D., Pau, F., Dissen, G. A., Ojeda, S. R., Zelinski, M. B., & Mayerhofer, A. (2015). Expression of the beta-2 adrenergic receptor (ADRB-2) in human and monkey ovarian follicles: A marker of growing follicles? *Journal of Ovarian Research*, 8, 8. <https://doi.org/10.1186/s13048-015-0136-4>
- Mishra, A., & Macgregor, S. (2015). VEGAS2: Software for more flexible gene-based testing. *Twin Research and Human Genetics*, 18, 86–91. <https://doi.org/10.1017/thg.2014.79>
- Monden, C., Pison, G., & Smits, J. (2021). Twin Peaks: More twinning in humans than ever before. *Human Reproduction*, 36, 1666–1673. <https://doi.org/10.1093/humrep/deab029>
- Monden, C. W. S., & Smits, J. (2017). Mortality among twins and singletons in sub-Saharan Africa between 1995 and 2014: A pooled analysis of data from 90 demographic and health surveys in 30 countries. *Lancet Global Health*, 5, e673–e679. [https://doi.org/10.1016/S2214-109X\(17\)30197-3](https://doi.org/10.1016/S2214-109X(17)30197-3)
- Olsen, C. M., Green, A. C., Neale, R. E., Webb, P. M., Cicero, R. A., Jackman, L. M., O'Brien, S. M., Perry, S. L., Ranieri, B. A., & Whiteman, D. C. (2012). Cohort profile: The QSkin Sun and Health Study. *International Journal of Epidemiology*, 41, 929–929. <https://doi.org/10.1093/ije/dys10>
- Painter, J. N., Willemsen, G., Nyholt, D., Hoekstra, C., Duffy, D. L., Henders, A. K., Wallace, L., Healey, S., Cannon-Albright, L. A., Skolnick, M., Martin, N. G., Boomsma, D. I., & Montgomery, G. W. (2010). A genome wide linkage scan for dizygotic twinning in 525 families of mothers of dizygotic twins. *Human Reproduction*, 25, 1569–1580. <https://doi.org/10.1093/humrep/deq084>
- Price, A., Patterson, N., Plenge, R., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, 38, 904–909. <https://doi.org/10.1038/ng1847>
- Ratner, A., Weiss, G. K., & Sanborn, C. R. (1980). Stimulation by β -adrenergic receptors of the production of cyclic AMP and progesterone in rat ovarian tissue. *Journal of Endocrinology*, 87, 123–129. <https://doi.org/10.1677/joe.0.0870123>
- Ruth, K. S., Day, F. R., Tyrrell, J., Thompson, D. J., Wood, A. R., Mahajan, A., Beaumont, R. N., Wittemans, L., Martin, S., Busch, A. S., Erzurumluoglu, A. M., Hollis, B., O'Mara, T. A.; Endometrial Cancer Association Consortium; McCarthy, M. I., Langenberg, C., Easton, D. F., Wareham, N. J., Burgess, S., Murray, A., Ong, K. K., Frayling, T. M., & Perry, J. R. B. (2020). Using human genetics to understand the disease impacts of testosterone in men and women. *Nature Medicine*, 26, 252–258. <https://doi.org/10.1038/s41591-020-0751-5>
- Santana, D. S., Cecatti, J. G., Surita, F. G., Silveira, C., Costa, M. L., Souza, J. P., Mazhar, S. B., Jayaratne, K., Qureshi, Z., Sousa, M. H., Vogel, J. P.; WHO Multicountry Survey on Maternal and Newborn Health Research Network. (2016). Twin pregnancy and severe maternal outcomes: The World Health Organization Multicountry Survey on Maternal and Newborn Health. *Obstetrics & Gynecology*, 127, 631–641. <https://doi.org/10.1097/AOG.0000000000001338>
- Santana, D. S., Silveira, C., Costa, M. L., Souza, R. T., Surita, F.G., Souza, J. P., Mazhar, S. B., Jayaratne, K., Qureshi, Z., Sousa, M. H., Vogel, J. P., Cecatti, J. G.; WHO Multi-Country Survey on Maternal and Newborn Health Research Network. (2018). Perinatal outcomes in twin pregnancies complicated by maternal morbidity: Evidence from the WHO Multicountry Survey on Maternal and Newborn Health. *BMC Pregnancy Childbirth*, 18, 449. <https://doi.org/10.1186/s12884-018-2082-9>
- Sear, R., Lawson, D.W., Kaplan, H., & Shenk, M. K. (2016). Understanding variation in human fertility: what can we learn from evolutionary demography? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 371, 20150144. <https://doi.org/10.1098/rstb.2015.0144>
- Tong, S., & Short, R. V. (1998). Dizygotic twinning as a measure of human fertility. *Human Reproduction*, 13, 95–98. <https://doi.org/10.1093/humrep/13.1.95>
- Trevisan, C. M., de Oliveira, R., Christofolini, D. M., Barbosa, C. P., & Bianco, B. (2019). Effects of a polymorphism in the promoter region of the Follicle-Stimulating Hormone Subunit Beta (FSHB) gene on female reproductive outcomes. *Genetic Testing and Molecular Biomarkers*, 23, 39–44. <https://doi.org/10.1089/gtmb.2018.0182>
- Wood, A. R., Esko, T., Yang, J., Vedantam, S., Pers, T. H., Gustafsson, S., Chu, A. Y., Estrada, K., Luan, J., Kutalik, Z., Amin, N., Buchkovich, M. L., Croteau-Chonka, D. C., Day, F. R., Duan, Y., Fall, T., Fehrmann, R., Ferreira, T., Jackson, A. U., ... Frayling, T. M. (2014). Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature Genetics*, 46, 1173–86. <https://doi.org/10.1038/ng.3097>
- Wright, M. J., & Martin, N. G. (2004). Brisbane Adolescent Twin Study: Outline of study methods and research projects. *Australian Journal of Psychology*, 56, 65–78. <https://doi.org/10.1080/00049530410001734865>
- Zhou W., Nielsen, J. B., Fritsche, L. G., Dey, R., Gabrielsen, M. E., Wolford, B. N., LeFaive, J., VandeHaar, P., Gagliano, S. A., Gifford, A., Bastarache, L. A., Wei, W. Q., Denny, J. C., Lin, M., Hveem, K., Kang, H. M., Abecasis, G. R., Willer, C. J., & Lee, S. (2018). Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nature Genetics*, 50, 1335–1341. <https://doi.org/10.1038/s41588-018-0184-y>

Appendix A

Summary of criteria for case and control selection

Selection of Australian/New Zealand Cases

Ideally: use mothers of DZ twins (MoDZT)

[Also considered the second-best option : DZ twins as cases where their mother was ungenotyped – but only half the genetic power and difficult to analyses together with MoDZT cases, so not used here].

We scanned our database for genotyped MoDZT.

- If opposite-sex (50%) then definitely DZ → include *****
- If same sex then tell DZ from MZ using :
 - o Discordance in genotyping or blood groups *****
 - o Discordance in visible phenotype (hair, eye colour) ***
 - o Never mistaken for each other by teachers, grandparents **
- Checked zygosity with the mother if we also needed to ask about ART (below)
- Otherwise exclude

Exclude artificial twin pregnancies: did the mother use Assisted Reproductive Technology (ART) to conceive the twins?

- We contacted the mother (letter, phone) and she reported use of ART → exclude *****
- If no contact available and twins born <1980 then ART very unlikely → include **
- Strategic decision to retain remaining uncontactable cases as mostly pre-1980 and ART is rare, due to the large increase in power resulting.

Selection of Australian/New Zealand Controls

- From QIMR Genetic Epidemiology Group databases : start with fathers of twins, MZ twins, others with no twin relatives.
- From QSkin: emailed to screen for zygosity and related twins (thanks David Whiteman; Catherine Olsen) → exclude if a DZ twin or self-report related to DZ twin/twin of unspecified zygosity.
- Also exclude genetically based on $IBD \geq 0.1$ to known DZ twins/MoDZT.

Appendix B

The text of the screening self-report questionnaire for QSkin controls (reformatted)

Females Only (if Male, go to Question 4)

1. Are you the mother of twins? *No / Yes*

If you answered 'No', please go to Question 4.

2. Have your twins ever been genetically tested to determine whether they are identical or not?

No need to test, opposite sex – one twin male, one twin female / No / Yes

2a. If yes, what was the result: *Identical / Non-identical*

2b. If no: 'Non-identical twins are no more alike than ordinary brothers and sisters. Identical twins, on the other hand, have such a strong resemblance to each other in stature, skin tone, hair colour, facial features, etc., that people often mistake one for the other, or say they are "as alike as two peas in a pod".' Do you think your twins are: *Identical / Non-identical*

3. Were your twins conceived with the help of Assisted Reproductive Technology (Fertility) Treatment (e.g., hormone treatment or IVF)? *No / Yes*

Both Males and Females

4. Are you a twin or are there any twins in your extended biological family (i.e., blood relatives)?

No / Yes

If you answered 'No', end of questionnaire.

4a. If yes, what is their relationship to you (select all that apply):

*I am a twin / My Mother / My Father / My Brother/Sister / My Uncle/Aunt / My Nephew/Niece
My Cousin / My Grandparent / My Son/Daughter / Other*

4b. If yes, are any of the twins non-identical (select all that apply):

*Me and my twin / My Mother / My Father / My Brother/Sister / My Uncle/Aunt /
My Nephew/Niece / My Cousin / My Grandparent / My Son/Daughter / Other*

