nature portfolio

Peer Review File

Genome-Wide Association Studies for Pelvic Organ Prolapse in the Japanese Population

Reviewers' comments:

Reviewer #1 (Remarks to the Author):

The manuscript "Genome-Wide Association Studies for Pelvic Organ Prolapse in the Japanese Population" by Matsunami et al presents a GWAS for pelvic organ prolapse (POP) in the Japanese population, and a subsequent multi-ancestry GWAS meta-analysis with a previously published European-ancestry study. The authors report one genome-wide significant locus in the Japanese analysis, and an additional novel locus in the multi-ancestry analysis. The authors also find the majority of hits from the previous European ancestry study replicate in the Japanese analysis with at least a nominal p-value, indicating that most of the susceptibility loci for POP are shared across the Japanese and European populations.

Overall, the study is a straightforward GWAS analysis with standard in silico follow-up analyses (eQTL lookup, gene-based analyses, tissue & pathway enrichment, and SMR). The manuscript is well-written, and although the Japanese ancestry analysis has a quite modest sample size for cases (~700), the nominal replication of the majority of EUR ancestry findings is encouraging.

I have only a few minor comments/questions:
1) The prevalence of POP is rather low in the analysed sample, stemming mostly from the fact that in the Biobank Japan sample, only ~400 cases were identified. Is the condition underdiagnosed or what could be the reason for such low prevalence in this dataset? Perhaps it would be good to add a few sentences about potential heterogeneity (or lack of) between the BBJ and OBi datasets as well?
2) Please add CI-s to OR estimate in the abstract
3) On page 8 it is stated "We further identified nine loci showing suggestive evidences for the association with POP in the Japanese ($5 \times 10\text{-}6 > P \geq 5 \times 10\text{-}8$, Supp Table.1).". Do these suggestive signals overlap with known POP loci or are these novel?
4) Please add distance between rs7072877 and FGFR2 in the text (page 9)
5) There's a typo in the title of Supplementary Table 11

Reviewer #2 (Remarks to the Author):

Matsunami et al report the first GWAS of Pelvic Organ Prolapse (POP) in a Japanese sample, in which they identify a known POP locus (WT1) at genome-wide significance. A

subsequent meta-analysis of this Japanese sample with a much larger European cohort identified a potentially new POP loci (FGFR2). Additional analyses attempt to establish functional roles for POP implicated SNPs.

The manuscript is well written, the methods are appropriate and well-described, and results are easy to track. Although the study is underpowered for novel variant discovery for this trait, it still provides some note-worthy advances: the first GWAS of POP in a Japanese cohort, evidence for cross-ancestry overlap of POP loci, and the Japanese cohort boosting the FGFR2 SNP over the standard 5e-8 genome-wide significance threshold in the Japanese-European meta-analysis. However I believe the interpretations of a couple key results must be carefully considered (see below).

Major items:

1. The author's claim that the effect size for the WT1 SNP is "significantly larger" in the Japanese population than the European population (lines 107-108, 230-233). This is possible, but also I suspect that there is a Winner's Curse bias on the effect estimates of replicated and nominally replicated loci in the Japanese dataset. Typically we think of the Winner's Curse impacting estimates from the original discovery sample. However in this situation, it is the sample size of the Japanese dataset that is underpowered as compared to the substantially larger discovery European dataset. Therefore to achieve genome-wide significance in the Japanese data, one requires a favorable draw in the form of an overestimate of the true effect size, as is likely observed for rs10742277. Figure S5 nicely demonstrates this phenomenon by showing that each of the loci for which this study was able to replicate/nominally replicate have estimated effect sizes that are larger in the Japanese cohort than the European cohort. Because these Japanese effect sizes likely contain an upward bias, it is not clear if they are truly larger than the effect size in Europeans.
Numerous methods have been proposed to correct for the Winner's Curse effect (e.g. https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1010546). Ideally the authors can compute a corrected effect size for rs10742277, to which they can compare to the reported European effect size. At the least, the authors should scale back the claim on the effect size being "significantly larger" in the Japanese and mention the potential for a Winner's Curse effect on their Japanese effect size estimates.

2. The authors perform an SMR/HEIDI analysis using the cross-ancestry POP GWAS summary statistics and eQTL/sQTL summary statistics from GTEx. Based on my

understanding of SMR/HEIDI from reading the original Zhu et al manuscript, the SMR (summary data Mendelian Randomization) test is first applied to identify associations between gene/transcript expression and the trait (POP). Genes passing the SMR test are then subject to the HEIDI (heterogeneity in dependent instruments) test to distinguish between a "model of pleiotropy" (defined as either a single SNP influencing both expression level and trait or a causal effect in which transcription is a mediator between the SNP and the trait) and linkage (two distinct causal SNPs, one influencing expression and the other influencing the trait). Importantly, the HEIDI test cannot distinguish between pleiotropy and causality/mediation. From the discussion section of the Zhu et al paper: "to avoid any potential misinterpretation of our method and results on causality, we describe our method as an approach to identify pleiotropic associations between gene expression and a complex trait and interpret all the results under a model of pleiotropy".

Therefore the authors need to be careful with their interpretation of the SMR/HEIDI results. Although I agree that the presented results support that "POP susceptibility and FGFR2 expression were likely affected by the same causal variant" (line 152-153), the SMR/HEIDI analysis cannot distinguish "whether susceptibility to POP was mediated by a decrease in FGFR2 expression" (line 154). Moreover, the p-values for the SMR test P-SMR should be presented before those of the HEIDI test and should not be used to argue that the risk allele might contribute to susceptibility to POP by lowering FGFR2 expression (lines 157-159).
In addition, the interpretation of SMR/HEIDI results should be similarly reframed to avoid causality/mediation when discussing the WT1 locus (lines 169-170) and altered splicing of the WT1-AS locus (lines 185-187).

It is fair to say (potentially in the Discussion) that causality/mediation is a possible model (along with pleiotropy) that is consistent with the observed SMR/HEIDI results.

3. The authors report that the gene-based association analysis using MAGMA identified WT1 (lines 190-192). Is this test simply capturing the effect of the genome-wide significant SNP rs10742277 or are there instead multiple additional low effect variants as well? Given that gene-based tests are intended to identify loci by combining signals that cannot be identified with standard single marker tests, it is important to distinguish between these scenarios. Is rs10742277 included among the 178 SNPs in the MAGMA test for WT1? Assuming that it is, can you eliminate its effect by either removing it (and all LD proxies) from the SNP list for WT1 or include genotype for rs10742277 as a covariate. You could then report MAGMA results with and without the effect of rs10742277 and interpret accordingly.

4. The authors report that in GTEx data, the SNP identified in the Japanese-European meta-analysis (rs7072877) was associated with decreased FGFR2 expression in three vascular tissues (lines 144-147). Were these the strongest and/or only tissues for which an association was observed for rs7072877, or were these chosen from a longer list because of their potential function?

In addition, the Pujol-Gualdo et al paper states in the Discussion section that "many GWAS associations we report point towards a link between metabolic and cardiovascular health and POP (KLF13, DUSP16, MAFF, VCL, and LDAH), which is a particularly interesting genetic association unraveled by our study". The connection between the novel FGFR2 association and vascular tissue would seem to be consistent with the above statement and perhaps worth mentioning.

5. Were the same criteria used to identify POP cases in the Okinawa Bioinformation Bank (OBi) and Biobank Japan? According to the methods section, an interview, clinical examination and Pelvic Organ Prolapse Quantification (POP-Q) assessment were used in the OBi (lines 297-299), and a "medical history of POP" was used for BioBank Japan samples. Can you please elaborate on what is meant by the medical history of POP. Is this, for example, based on ICD codes? If different criteria are used between the two cohorts, this should be acknowledged in the Discussion. I actually think that it is a strength that the Japanese meta-analysis was successful even if different clinical data were used for phenotyping in each cohort. (As an aside, I believe the European GWAS was entirely based on ICD codes for POP definition and that may impact direct comparisons of effect size as well.)

6. I was surprised that the test of heterogeneity of effect sizes in the meta-analyses was often not significant despite seemingly large differences in the estimated effect sizes between the Japanese and European cohorts. In particular, for the FGFR2 results from the cross-ancestry meta-analysis (Supp Table 3), the p-value for heterogeneity is reported as p=0.27 but several of the Japanese effect sizes (e.g. beta=1.41, 1.73) are quite different than the European effect size (beta=1.06). Are you also surprised that this SNP was not flagged for heterogeneity of effect sizes? Do you have any intuition why it was not? Finally, would this SNP still achieve genome-wide significance if a random-effect meta-analysis were used?

7. It is reported that nine loci showed suggestive evidence for association in the Japanese-only meta-analysis (lines 115-117, Suppl Table 1). Are any of these sites nearby loci previously identified in the European cohort? Would be valuable to simply clarify that.

Minor items:

8. Please report the genomic control values in the main text (and cite Suppl Figure S3) for both the Japanese meta-analysis and the Japanese+European meta-analysis prior to presenting GWAS results. As a reader I wanted confirmation that each GWAS was well-controlled before reading about findings.

9. I recommend starting new paragraphs at line 150 (Using summary data…) and line 170 (In addition…) to break up long blocks of text and separate distinct analyses.

Reviewer 1

We sincerely appreciate the reviewer's positive and valuable comments.
Following Reviewer 1's comments, we revised the paper as shown below.

【Reviewer 1's comment #1】

1) ***The prevalence of POP is rather low in the analysed sample, stemming mostly from the fact that in the Biobank Japan sample, only ~400 cases were identified. Is the condition underdiagnosed or what could be the reason for such low prevalence in this dataset? Perhaps it would be good to add a few sentences about potential heterogeneity (or lack of) between the BBJ and OBi datasets as well?***

【Reply】

We would like to thank Reviewer 1 for pointing out potential issues regarding case selection in our Japanese cohorts, Biobank Japan (BBJ) and Okinawa Bioinformation Bank (OBi).
BBJ is a hospital-based disease cohort that covers 47 target diseases other than POP. Since POP is not included in the 47 target diseases, we searched for the clinical information, including comorbidities and past medical history of each participant registered by doctors-in-charge. Then, we defined POP cases as follows.1) patients with cystocele, uterine prolapse, or vaginal prolapse as comorbidities or 2) patients having past medical history for cystocele, uterine prolapse, or vaginal prolapse. The information of comorbidity and past medical history was based on self-reports or medical records for underwent surgeries or other interventions (i.e., pessary). Therefore, several POP cases might be misclassified as controls, especially early-stage or asymptomatic cases.
In the Okinawa Bioinformation Bank (OBi), all POP cases were recruited by medical specialists for POP, and diagnosis of POP was made based on POP Quantification (POP-Q) system. As a result, most of the cases were surgical cases (95.4 %) and symptomatic POP cases with stage 3 or higher in POP-Q system (91.8 %).
Because relatively advanced stages of POP patients were analyzed as cases in both cohorts, we believe that considerable phenotypic heterogeneity between the two cohorts did not exist in this study.

We have added the sentences below in the discussion section.
Pages 19-20, Lines 319-328 in the manuscript with Track Change (TC), Pages18-19, Lines 301-309 in the clean manuscript.

BBJ cases were selected from participants with POP or a past history of POP based on self-reports or medical records for underwent surgeries or other interventions (i.e., pessary). Therefore, several POP cases might be misclassified as controls, especially early-stage or asymptomatic cases. In the OBi, all POP cases were recruited by medical specialists for POP, and diagnosis of POP was made based on POP Quantification (POP-Q) system. As a result, most of the cases were surgical cases (95.4 %) and symptomatic POP cases with stage 3 or higher in POP-Q system (91.8 %). Because relatively advanced stages of POP patients were analyzed as cases in both cohorts, it is suggested that considerable phenotypic heterogeneity between the two cohorts did not exist in this study.

In addition, we added the sentences below in the Method section.
Page 21, Lines 347-348 in the manuscript with Track Change (TC), Page 20 Line 328 in the clean manuscript.

Each subject participated in an interview and underwent a clinical examination and Pelvic Organ Prolapse Quantification (POP-Q) system assessment (stage 0: 0%, stage 1: 0%, stage 2: 8.2%, stage 3: 77.5%, stage 4: 14.3%)

Page 23, Lines 373-380 in the manuscript with Track Change (TC), Page 22, Lines 354-361 in the clean manuscript. Changes are indicated by underlining.

The BBJ is a hospital-based disease cohort which consists of DNA samples and clinical data from patients with 47 target diseases [48]. Clinical information, including comorbidities and past medical history, of each participant was registered in BBJ by doctors-in-charge [48]. All individuals in BBJ were genotyped using Illumina Infinium Omni Express, Human Exome, Infinium Omni Express Exome v1.0, or Infinium Omni Express Exome v1.2 (Illumina Inc). We defined POP cases as follows.1) patients with cystocele, uterine prolapse, or vaginal prolapse as comorbidities or 2) patients having past medical history for cystocele, uterine prolapse, or vaginal prolapse. ~~We selected cases from female participants or with a medical history of POP (cystocele, uterine prolapse, and vaginal prolapse), documented by the doctors- in- charge as previously described [48].~~

【**Reviewer 1's comment #2**】

2) Please add CI-s to OR estimate in the abstract

【**Reply**】

We have added CI-s to OR estimate in the abstract as per the reviewer's suggestion.
Page 4 Lines 49 and 52 in the manuscript with Track Change (TC), Page 4 Lines 49 and 52 in the clean manuscript. Changes are indicated by underlining.

WT1 locus with POP in the Japanese population; rs10742277; odds ratio (OR) = 1.48, 95% confidence interval (CI), 1.29 – 1.68, $P = 6.72 \times 10^{-9}$. Subsequent cross-ancestry GWAS meta-analysis combining the Japanese data and previously reported European data, including 28,857 cases and 622,916 controls, identified *FGFR2* locus as a novel susceptibility locus to POP (rs7072877; OR = 1.06, 95% CI, 1.04 - 1.08, $P = 4.11 \times 10^{-8}$).

【**Reviewer 1's comment #3**】

3) On page 8 it is stated "We further identified nine loci showing suggestive evidences for the association with POP in the Japanese ($5 \times 10^{-6} > P \geq 5 \times 10^{-8}$, Supp Table.1)." . Do these suggestive signals overlap with known POP loci or are these novel?

【**Reply**】

None of the nine suggestive loci overlap with previously reported POP loci. We have clarified this point in the main text.
Page 8 Lines 119-120 in the manuscript with Track Change (TC), Page 8 Lines 119-120 in the clean manuscript. Changes are indicated by underlining.

We further identified nine loci showing suggestive evidences for the association with POP in the Japanese ($5 \times 10^{-6} > P \geq 5 \times 10^{-8}$, Supp Table.1), none of these overlapped with loci previously identified in the European GWAS [12, 13].

【**Reviewer 1's comment #4**】

4) Please add distance between rs7072877 and FGFR2 in the text (page 9)

【**Reply**】

We have added distance between rs7072877 and *FGFR2* in the text.
Page 9 Lines 139-140 in the manuscript with Track Change (TC), Page 9 Lines 139-140 in the clean manuscript. Changes are indicated by underlining.

A GWAS meta-analysis with a sample size of 651,773 identified 25 genome-wide significant associations with POP, including the *FGFR2* locus which has not been previously reported as a POP susceptibility locus (rs7072877 located approximately 9 kbp downstream of ~~near~~ *FGFR2; P* = 4.11 × 10^{-8}, OR = 1.06, 95% CI 1.04−1.08, Table 2, Supp Table 3, Figure 1B).

【Reviewer 1's comment #5】

There's a typo in the title of Supplementary Table 11

【Reply】

We appreciate reviewer for pointing out the typo in the title of Supplementary Table 11. We have corrected the typo.

[Revised version]

Supplementary Table 11 MAGMA gene set analysis using POP_GWAS meta-analysis data using Japanese and European studies ($p < 5 \times 10^{-4}$)

Reviewer 2

We sincerely appreciate the reviewer's thoughtful peer review, positive comments and valuable suggestions.

Following Reviewer 2's comments, we revised the paper as shown below.

【Reviewer 2's comment #1】

1. The author's claim that the effect size for the WT1 SNP is"significantly larger" in the Japanese population than the European population (lines 107-108, 230-233). This is possible, but also I suspect that there is a Winner's Curse bias on the effect estimates of replicated and nominally replicated loci in the Japanese dataset. Typically we think of the Winner's Curse impacting estimates from the original discovery sample.

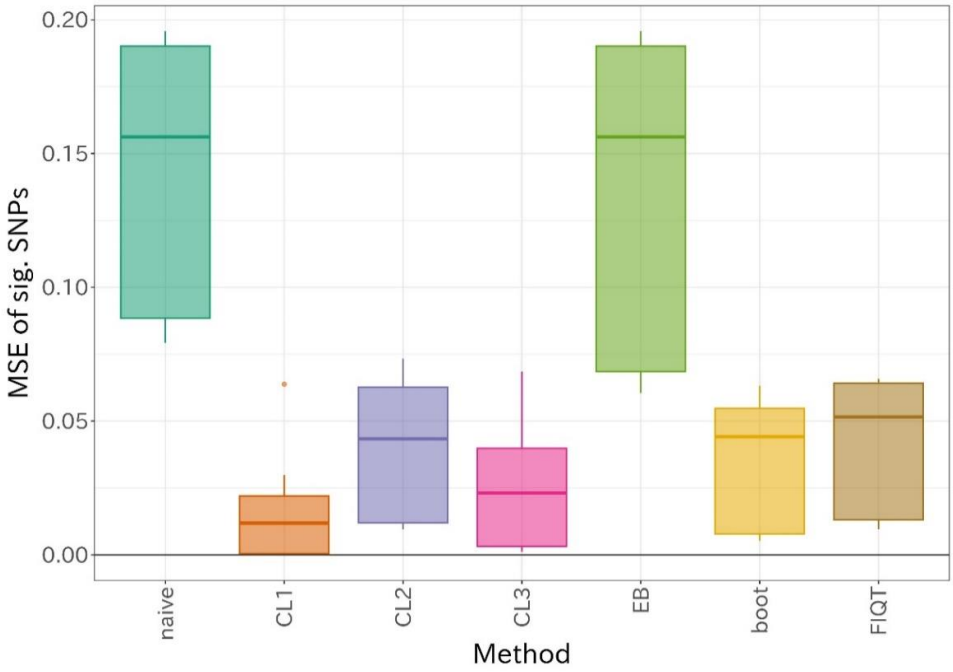
However in this situation, it is the sample size of the Japanese dataset that is underpowered as compared to the substantially larger discovery European dataset. Therefore to achieve genome-wide significance in the Japanese data, one requires a favorable draw in the form of an overestimate of the true effect size, as is likely observed for rs10742277. Figure S5 nicely demonstrates this phenomenon by showing that each of the loci for which this study was able to replicate/nominally replicate have estimated effect sizes that are larger in the Japanese cohort than the European cohort. Because these Japanese effect sizes likely contain an upward bias, it is not clear if they are truly larger than the effect size in Europeans. Numerous methods have been proposed to correct for the Winner's Curse effect (e.g.https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.10105 46). Ideally the authors can compute a corrected effect size for rs10742277, to which they can compare to the reported European effect size. At the least, the authors should scale back the claim on the effect size being "significantly larger" in the Japanese and mention the potential for a Winner's Curse effect on their Japanese effect size estimates.

【Reply】

We thank the reviewer for your thoughtful comments and specific instructions.

We applied four correcting methods：Conditional Likelihood(CL), empirical Bayes (EB), bootstrap methods (boot), and FDR Inverse Quantile Transformation (FIQT) method implemented in the R package "winnerscurse", according to your suggestion. For these analyses, effect sizes in European GWAS were set as replication. As a result, adjusted effect sizes of rs10742277 were become smaller than the unadjusted effect size for all

methods. The effect size of rs10742277 is estimated to be 0.2145901in the conditional likelihood method (CL) with the smallest estimated mean square error (MSE), suggesting this correction is most reliable (see below Figure and Table)**.**



We also estimated adjusted standard error (SE) or confidential interval (CI) for each estimated effect size; "se_adjust" function implemented in "winnerscurse" only estimate se for beta obtained by EB, boot, and FIQT, therefore we estimate CI for beta estimated by CL using "cl_interval" function. We have added the results as Supplementary Table 13.

Supplementary Table 13: Association data of rs10742277 in *WT1* after adjustement for potential Winner's curse effect

Method		MSE	adjusted beta for rs10742277	adjusted SE for rs10742277	adjusted 95%CI for rs10742277
unadjusted		0.13789	0.38940	0.06720	
CL	CL1	**0.00984**	0.21306	NA	
	CL2	0.03473	0.21612	NA	
	CL3	0.01972	**0.21459**	NA	**0.01226, 0.45842**
EB		0.12905	0.35235	0.04210	
boot		0.03466	0.19098	0.05382	
FIQT		0.03613	0.19096	0.04764	

MSE mean square error, SE standard error, CI confidence interval

CL conditional likelihood, EB empirical Bayes, boot bootstrap methods,

FIQT FDR Inverse Quantile Transformation.

Bold letters indicate most reliavle estimated association data with minimum MSE values

NA:SEwas not avairable in CL mehtod

Unfortunately, we were not able to analyze heterogeneity of the effect of rs10742277 between European and Japanese using adjusted beta-estimates by CL, because SE was not available in this method. When we used adjusted association data by three other methods, EB, boot, FIQT, heterogeneity of the effects of rs10742277 between European and Japanese is still significant as shown below.

	heterogeneity test vs European			
	Het I^2	Het ChiSq	Het Df	Het P-Value
unadjusted	95.8	23.866	1	1.03E-06
CL	NA	NA	NA	NA
	NA	NA	NA	NA
	NA	NA	NA	NA
EB	97.8	46.498	1	9.17E-12
boot	83.2	5.94	1	0.0148
FIQT	86.7	7.514	1	0.006123

We also added the following description in the discussion section.

Page 16, Lines 250-261 in the manuscript with Track Change (TC), Page 15, Lines 238-249 in the clean manuscript. Changes are indicated by underlining.

[Revised version]

The effect size of rs10742277-C in the Japanese population was significantly larger than that in the European population, suggesting that the *WT1* locus is ~~particularly~~ more relevant for Japanese women. Since the sample size in this study is smaller than those in the European studies especially for case groups, the possibility for Winners Curse effect might exist [27]. Then, we performed four analyses for correcting the Winners Curse effect; Conditional Likelihood (CL), empirical Bayes (EB), bootstrap methods (boot), and FDR Inverse Quantile Transformation (FIQT) method implemented in the R package "winnerscurse" [27]. The effect sizes (β-estimates) of rs10742277 in the Japanese were decreased after the corrections (Supplementary Table 13); therefore, the effect size of this locus in the Japanese might be overestimated in the original GWAS, although the corrected effect sizes are considered still larger than European populations.

We added Methods accordingly

Pages 28-29, Lines 467-481 in the manuscript with Track Change (TC), Page 27, Lines 445-459 in the clean manuscript.

[Revised version]

Correction for an effect size using the Winner's Curse methods

To evaluate the possible bias on effect size of rs10742277 in Japanese GWAS, we calculated corrected effect size using the Winner's Curse methods. We applied four correcting methods implemented in the R package “winnerscurse” [27]. The package includes Conditional Likelihood (CL) methods [52], Empirical Bayes (EB) method [53], FDR Inverse Quantile Transformation (FIQT) method [54], and Bootstrap (boot) method [27] (Forde et al. 2023). For analysis using CL methods, we used 25 variants with genome-wide significant ($p < 5 \times 10^{-8}$) association in Japanese POP GWAS. For other three methods, EB, FIQT and boot, we used 7,496,192 SNPs which association data was available from all of four Japanese sub-group. Effect sizes in European GWAS were used as “replication set”. We calculated adjusted beta as well as adjusted standard error with 100 bootstrap replicates for EB, FIQT and boot. Since standard error estimation using bootstrap was not available for adjusted beta estimated by CL method, we calculated adjusted 95% CI instead of standard error. The estimated mean square error (MSE) of SNPs were calculated to evaluate the estimation of each method.

【Reviewer 2’s comment #2】

2. The authors perform an SMR/HEIDI analysis using the cross-ancestry POP GWAS summary statistics and eQTL/sQTL summary statistics from GTEx. Based on my understanding of SMR/HEIDI from reading the original Zhu et al manuscript, the SMR (summary data Mendelian Randomization) test is first applied to identify associations between gene/transcript expression and the trait (POP). Genes passing the SMR test are then subject to the HEIDI (heterogeneity in dependent instruments) test to distinguish between a “model of pleiotropy” (defined as either a single SNP influencing both expression level and trait or a causal effect in which transcription is a mediator between the SNP and the trait) and linkage (two distinct causal SNPs, one influencing expression and the other influencing the trait).

Importantly, the HEIDI test cannot distinguish between pleiotropy and causality/mediation. From the discussion section of the Zhu et al paper:“to avoid any potential misinterpretation of our method and results on causality, we describe our method as an approach to identify pleiotropic associations between gene expression and a complex trait and interpret all the results under a model of pleiotropy”. Therefore the authors need to be careful with their interpretation of the SMR/HEIDI results. Although I agree that the presented results support that “POP susceptibility and FGFR2 expression were likely affected by the same causal variant” (line 152-153), the SMR/HEIDI analysis cannot distinguish “whether susceptibility to POP was mediated by a decrease in FGFR2 expression” (line 154). Moreover, the p-values

for the SMR test P-SMR should be presented before those of the HEIDI test and should not be used to argue that the risk allele might contribute to susceptibility to POP by lowering FGFR2 expression (lines 157-159).

In addition, the interpretation of SMR/HEIDI results should be similarly reframed to avoid causality/mediation when discussing the WT1 locus (lines 169-170) and altered splicing of the WT1-AS locus (lines 185-187).

It is fair to say (potentially in the Discussion) that causality/mediation is a possible model (along with pleiotropy) that is consistent with the observed SMR/HEIDI results.

【Reply】

We sincerely appreciated the reviewer's careful peer review and thoughtful comments. We have revised the description of SMR/HEIDI in the results, discussion, and supplementary note as follows.

Pages 10-11 Lines 156-164 in the manuscript with Track Change (TC), Pages 10-11 Lines 156-164 in the clean manuscript. Changes are indicated by underlining.

Using summary data-based Mendelian randomization (SMR) analysis and the heterogeneity in dependent instruments (HEIDI) test [19], we intended to identify associations between *FGFR2* expression and POP using summary data from eQTL studies and GWAS, followed by a heterogeneity test to see if *FGFR2* expression and the POP were affected by the same underlying causal variant [19]. The results of the SMR/HEIDI test suggested significant association between *FGFR2* expression and POP (*P*-SMR = 2.8×10^{-6} in the tibial artery, *P*-SMR = 8.6×10^{-6} in the aorta, *P*-SMR = 4.5×10^{-4} in the coronary artery, Supp Table 5), and POP susceptibility and *FGFR2* expression were likely affected by the same causal variant (*P*-value of HEIDI test ≥ 0.05, Supp Table 5). ~~we tested whether these two traits, POP susceptibility and *FGFR2* expression, were affected by the same causal variant, and whether susceptibility to POP was mediated by a decrease in *FGFR2* expression. The results of the SMR/HEIDI test suggested that POP susceptibility and *FGFR2* expression were likely affected by the same causal variant (*P*-value of HEIDI test ≥ 0.05, Supp Table 5), and carrying the risk allele (rs7072877-C) might contribute to susceptibility to POP by lowering the *FGFR2* expression (*P*-SMR = 2.8×10^{-6} in the tibial artery, *P*-SMR = 8.6×10^{-6} in the aorta, *P*-SMR = 4.5×10^{-4} in the coronary artery, Supp Table 5).~~

Page 12, Line 182 in the manuscript with Track Change (TC), Page 11, Lines 174-175 in the clean manuscript.

The lead SNPs for the eQTL in the heart (rs3858454) and GWAS (rs10219425) were in high LD (r^2 = 0.82 in 1KG CEU) (Supp Figure 8), and the result of the SMR/HEIDI test suggested their co-localization ~~and causality~~ (Supp Table 5).

Pages 12-13, Lines 196-199 in the manuscript with Track Change (TC), Page 12, Lines 189-192 in the clean manuscript. Changes are indicated by underlining.

The results of the SMR/HEIDI test indicated not all but some of sQTL signals had significant association with POP (p-SMR < 0.05, Supp Table 7), and POP susceptibility and splicing at *WT1-AS* were likely affected by the same causal variant (*P*-value of HEIDI test ≥ 0.05, Supp Table 7). ~~Results of SMR/HEIDI test suggested that not all but some of sQTL signals and POP susceptibility were affected by the same causal variant (p-HEIDI ≥ 0.05, Supp Table 7) and the effect of the risk allele (rs11031796-G: proxy of rs10219425-A) on POP susceptibility was mediated by altered splicing at *WT1-AS* (p-SMR < 0.05, Supp Table 7).~~

Page 16, Lines 263-266 in the manuscript with Track Change (TC), Page 16, Lines 251-254 in the clean manuscript. Changes are indicated by underlining.

Although the involvement of *WT1* antisense RNA (*WT1-AS*) in POP susceptibility has not yet been reported, our results of eQTL/sQTL analyses suggest a possible model that the effect of causal variant in *WT1* on POP susceptibility is mediated by the increased expression of *WT1-AS* and altered relative abundance of *WT1-AS* splicing isoforms, that is consistent with the observed SMR/HEIDI results [19]. ~~our eQTL/sQTL and subsequent SMR/HEIDI analyses suggest that the effect of causal variants in *WT1* on POP susceptibility is likely to be mediated by the increased expression of *WT1-AS* and altered relative abundance of *WT1-AS* splicing isoforms.~~

Pages 17-18, Lines 282-285 in the manuscript with Track Change (TC), Page 17, Lines 267-270 in the clean manuscript. Changes are indicated by underlining.

The results of eQTL analysis suggested a possible model that the effect of causal variant in *FGFR2* on susceptibility to POP was mediated by reduced *FGFR2* expression, that was consistent with the observed SMR/HEIDI results. ~~The results of eQTL analysis and subsequent SMR/HEIDI analysis suggested that the effect of causal variants in *FGFR2* on susceptibility to POP was mediated by reduced *FGFR2* expression.~~

Supplementary Note Page 1, Bottom of the "**eQTL analysis in *WT1* locus**" section . Changes are indicated by underlining.

The results of the SMR/HEIDI test suggested significant association between *WT1-AS* expression and POP (*P*-SMR = 2.08×10^{-4}), and *WT1-AS* expression and POP susceptibility and were likely affected by the same causal variant (*P*- HEIDI ≥ 0.05, Supp Table 5) in Heart_Atrial_Appendage. ~~The results of SMR/HEIDI test suggested co-localization of POP susceptibility and *WT1-AS* expression signals in Heart_Atrial_Appendage (*P*-HEIDI > 0.05, Supp Table 5) and the effect of the risk allele (rs11031796-G) on POP susceptibility was mediated by increased expression of *WT1-AS* (*P*-SMR = 2.08×10^{-4}, Supp Table 5).~~

Supplementary Note Page 2, Bottom of the " **sQTL analysis in *WT1* locus**" section. Changes are indicated by underlining.

Among these 16 sQTL signals, 7 signals in total; 4 in ovary, 2 in uterus, 1 in visceral adipose tissue, had significant association with POP (p-SMR < 0.05, Supp Table 7), and those 7 signals and POP susceptibility were likely affected by the same causal variant (*P*- HEIDI ≥ 0.05, Supp Table 7). ~~were suggested to be co-localized with GWAS signal (rs11031796) (*P*-HEIDI > 0.05) and causal of POP susceptibility (*P*- SMR < 0.05) (Supp Table 6).~~

【Reviewer 2's comment #3】

3. The authors report that the gene-based association analysis using MAGMA identified WT1 (lines 190-192). Is this test simply capturing the effect of the genome-wide significant SNP rs10742277 or are there instead multiple additional low effect variants as well? Given that gene-based tests are intended to identify loci by combining signals that cannot be identified with standard single marker tests, it is important to distinguish between these scenarios. Is rs10742277 included among the 178 SNPs in the MAGMA test for WT1? Assuming that it is, can you eliminate its effect by either removing it (and all LD proxies) from the SNP list for WT1 or include genotype for rs10742277 as a covariate. You could then report MAGMA results with and without the effect of rs10742277 and interpret accordingly.

【Reply】

We would like to thank the reviewer for raising this point. We agree with you and have incorporated this suggestion.

We performed additional GWAS for POP conditioning on rs10742277, including it as a covariate, and subsequently run a MAGMA gene-based analysis. As a result, association between WT1 gene and POP was disappeared.

WT1 with the effect of rs10742277, Z-value 4.8301 P =6.82E-07

WT1 without the effect of rs10742277, Z -value 0.7159 P =0.24

We have shown this result in Supplementary Table 8, and added the following description in the main text.

Page 13, Lines 210-213 in the manuscript with Track Change (TC), Page 13, Lines 196-201 in the clean manuscript. Changes are indicated by underlining.

[Revised version]

The analysis identified that *WT1* was significantly associated with POP ($P < 2.55 \times 10^{-6}$ = 0.05/19,589, Supp Table 8). Because the gene-based association between the *WT1* gene and POP was no longer significant after including the effect of the lead SNP (rs10742277) as covariate in the GWAS (P = 0.24, Supplementary Table 8), the significant association of *WT1* with POP may simply reflect the effect of rs10742277.

【Reviewer 2's comment #4】

4. The authors report that in GTEx data, the SNP identified in the Japanese-European meta-analysis (rs7072877) was associated with decreased FGFR2 expression in three vascular tissues (lines 144-147). Were these the strongest and/or only tissues for which an association was observed for rs7072877, or were these chosen from a longer list because of their potential function?

In addition, the Pujol-Gualdo et al paper states in the Discussion section that "many GWAS associations we report point towards a link between metabolic and cardiovascular health and POP (KLF13, DUSP16, MAFF, VCL, and LDAH), which is a particularly interesting genetic association unraveled by our study". The connection between the novel FGFR2 association and vascular tissue would seem to be consistent with the above statement and perhaps worth mentioning.

【Reply】

Thank you very much for your very valuable suggestions. The three vascular tissues were the only tissues in which a significant correlation (m-value > 0.9, m-value: Posterior Probability from METASOFT) between rs7072877 and FGFR2 expression was confirmed in GTex, and not that these chosen from a longer list because of their potential function.

We added the following sentence in the discussion.

Page 19, Lines 305-310 in the manuscript with Track Change (TC), Page 18, Lines 288-292 in the clean manuscript.

[Revised version]

Notably, the GTEx eQTL analysis identified significant correlation between rs7072877 and *FGFR2* expression only in three vascular tissues out of 49 tissues, which is in line with the statement by Pujol-Gualdo et al [13] that many GWAS associations previously reported by European study (*WT1*, *KLF13*, *DUSP16*, *MAFF*, *VCL*, and *LDAH*) suggested a link between metabolic and cardiovascular health and POP [13].

【Reviewer 2's comment #5】

5. Were the same criteria used to identify POP cases in the Okinawa Bioinformation Bank (OBi) and Biobank Japan? According to the methods section, an interview, clinical examination and Pelvic Organ Prolapse Quantification (POP-Q) assessment were used in the OBi (lines 297-299), and a "medical history of POP" was used for BioBank Japan samples. Can you please elaborate on what is meant by the medical history of POP. Is this, for example, based on ICD codes? If different criteria are used between the two cohorts, this should be acknowledged in the Discussion. I actually think that it is a strength that the Japanese meta-analysis was successful even if different clinical data were used for phenotyping in each cohort. (As an aside, I believe the European GWAS was entirely based on ICD codes for POP definition and that may impact direct comparisons of effect size as well.)

【Reply】

Thank you very much for the reviewer's comment on the details of selection procedures for POP cases in the two Japanese cohorts.

The selection criteria for cases of BBJ and OBi were basically the same： female participants with age over 40 years and diagnosed with POP (cystocele, uterine prolapse, vaginal prolapse).

BBJ is a hospital-based disease cohort that covers 47 target diseases other than POP. Clinical information, including comorbidities and past medical history, of each participant was registered in BBJ by doctors-in-charge. Then, we defined POP cases as follows 1) patients with cystocele, uterine prolapse, or vaginal prolapse as comorbidities or 2) patients having past medical history for cystocele, uterine prolapse, or vaginal prolapse. POP cases were searched by two medical doctors based on the disease name described as clinical information by free text, not as ICD-10 codes. The information of POP as comorbidity and past medical history was based on self-reports or medical records for underwent surgeries or other interventions (i.e., pessary). Therefore, several

POP cases might be misclassified as controls, especially early-stage or asymptomatic cases.
In the Okinawa Bioinformation Bank (OBi), all POP cases were recruited by medical specialists for POP, and diagnosis of POP was made based on POP Quantification (POP-Q) system. As a result, most of the cases were surgical cases and symptomatic POP cases with stage 3 or higher in POP-Q system.
Because relatively advanced stages of POP patients were analyzed as cases in both cohorts, we believe that considerable phenotypic heterogeneity between the two cohorts did not exist in this study.

In the revised version, we add more details about case selection. In addition, we have described our thoughts on potential phenotypic heterogeneity between Japanese cohorts in Discussion section.
Page 21, Lines 347-348 in the manuscript with Track Change (TC), Page 20 Line 328 in the clean manuscript.

Each subject participated in an interview and underwent a clinical examination and Pelvic Organ Prolapse Quantification (POP-Q) system assessment (stage 0: 0%, stage 1: 0%, stage 2: 8.2%, stage 3: 77.5%, stage 4: 14.3%)

Page 23, Lines 373-380 in the manuscript with Track Change (TC), Page 22, Lines 354-361 in the clean manuscript. Changes are indicated by underlining.

The BBJ is a hospital-based disease cohort which consists of DNA samples and clinical data from patients with 47 target diseases [48]. Clinical information, including comorbidities and past medical history, of each participant was registered in BBJ by doctors-in-charge [48]. All individuals in BBJ were genotyped using Illumina Infinium Omni Express, Human Exome, Infinium Omni Express Exome v1.0, or Infinium Omni Express Exome v1.2 (Illumina Inc). We defined POP cases as follows.1) patients with cystocele, uterine prolapse, or vaginal prolapse as comorbidities or 2) patients having past medical history for cystocele, uterine prolapse, or vaginal prolapse. ~~We selected cases from female participants or with a medical history of POP (cystocele, uterine prolapse, and vaginal prolapse), documented by the doctors- in- charge as previously described [48].~~

We have added the sentences below in the discussion section.
Pages 19-20, Lines 319-328 in the manuscript with Track Change (TC), Pages18-19, Lines 301-309 in the clean manuscript.

BBJ cases were selected from participants with POP or a past history of POP based on self-reports or medical records for underwent surgeries or other interventions (i.e., pessary). Therefore, several POP cases might be misclassified as controls, especially early-stage or asymptomatic cases. In the OBi, all POP cases were recruited by medical specialists for POP, and diagnosis of POP was made based on POP Quantification (POP-Q) system. As a result, most of the cases were surgical cases (95.4 %) and symptomatic POP cases with stage 3 or higher in POP-Q system (91.8 %). Because relatively advanced stages of POP patients were analyzed as cases in both cohorts, it is suggested that considerable phenotypic heterogeneity between the two cohorts did not exist in this study.

【Reviewer 2's comment #6】

6. I was surprised that the test of heterogeneity of effect sizes in the meta-analyses was often not significant despite seemingly large differences in the estimated effect sizes between the Japanese and European cohorts. In particular, for the FGFR2 results from the cross-ancestry meta-analysis (Supp Table 3), the p-value for heterogeneity is reported as p=0.27 but several of the Japanese effect sizes (e.g. beta=1.41, 1.73) are quite different than the European effect size (beta=1.06). Are you also surprised that this SNP was not flagged for heterogeneity of effect sizes? Do you have any intuition why it was not? Finally, would this SNP still achieve genome-wide significance if a random-effect meta-analysis were used?

【Reply】

We appreciate for reviewer's careful peer review.
We used cochran's Q test implemented in METAL https://github.com/statgen/METAL to access heterogeneity among the studies. Two researchers independently re-performed the heterogeneity test, and obtained the completely consistent results with the original result, therefore we believe there was no technical error during our analysis procedures. Although we don't have clear answer to explain the reason why this SNP was not flagged for heterogeneity of effect sizes in this study, small sample sizes and large variances (standard error) of Japanese data might lead to miss-evaluation of heterogeneity of effect sizes between the Japanese and European populations.

To answer reviewer's questions, we conducted two random model meta-analysis, additive random effects model (known as DerSimonian-Laird estimator) and multiplicative random effect model, implemented in METAL (https://github.com/explodecomputer/random-metal). The results are shown below.

MarkerName chr10:123228660 Effect/non-effect C/T									
Effect/non-effect C/T	beta	se	OR		95%CI				P-value for association
Fixed_effect	0.059289	0.010807	1.06	(	1.04	-	1.07	)	4.11E-08
Random_effect_additive_model	0.094085	0.043368	1.10	(	1.01	-	1.15	)	0.03
Random_effect_multiplicative_model	0.059289	0.013854	1.06	(	1.03	-	1.08	)	1.87E-05
P-value for heterogeneity test	0.27								
I^2	22								
Tau square	0.002649								

The association of rs7072877 with POP was significant also in the random effect model, although the association did not reach genome wide significance level.

【Reviewer 2's comment #7】

7. It is reported that nine loci showed suggestive evidence for association in the Japanese-only meta-analysis (lines 115-117, Suppl Table 1). Are any of these sites nearby loci previously identified in the European cohort? Would be valuable to simply clarify that.

【Reply】

None of these nine suggestive loci were nearby loci with previously reported POP loci. We have clarified this point in the main text.

Page 8 Lines 119-120 in the manuscript with Track Change (TC), Page 8 Lines 119-120 in the clean manuscript. Changes are indicated by underlining.

We further identified nine loci showing suggestive evidences for the association with POP in the Japanese ($5 \times 10^{-6} > P \geq 5 \times 10^{-8}$, Supp Table.1), none of these overlapped with loci previously identified in the European GWAS [12, 13].

【Reviewer 2's comment #8】

8. Please report the genomic control values in the main text (and cite Suppl Figure S3) for both the Japanese meta-analysis and the Japanese+European meta-analysis prior to presenting GWAS results. As a reader I wanted confirmation that each GWAS was well-controlled before reading about findings.

【Reply】

We have added the genomic control values in the main text as below

Page 7 Lines 102-103 in the manuscript with Track Change (TC), Page 7 Lines 102 in the clean manuscript.

[Revised version]

No significant inflation was observed (λ_{GC} = 1.01, $\lambda_{GC}1000$ = 1.007, Supp Figure 3).

Page 9 Lines 136-137 in the manuscript with Track Change (TC), Page 9 Lines 136-137 in the clean manuscript.

[Revised version]

No significant inflation was observed (λ_{GC} = 1.120, $\lambda_{GC}1000$ = 1.002, Supp Figure 3).

【Reviewer 2's comment #9】

9. I recommend starting new paragraphs at line 150 (Using summary data…) and line 170 (In addition…) to break up long blocks of text and separate distinct analyses.

【Reply】

We have modified the main text by starting new paragraphs as reviewer 2 suggested.

REVIEWERS' COMMENTS:

Reviewer #2 (Remarks to the Author):

The authors have been responsive to my comments and I have no further questions of suggestions.

Reviewer #3 (Remarks to the Author):

Thank you for addressing my concerns.