
















# Genome-wide association identifies seven loci for pelvic organ prolapse in Iceland and the UK Biobank

Thorhildur Olafsdottir<sup>1</sup>, Gudmar Thorleifsson<sup>1</sup>, Patrick Sulem<sup>1</sup>, Olafur A. Stefansson<sup>1</sup>, Helga Medek<sup>2</sup>, Karl Olafsson<sup>2</sup>, Orri Ingthorsson<sup>3</sup>, Valur Gudmundsson<sup>3</sup>, Ingileif Jonsdottir<sup>1,4,5</sup>, Gisli H. Halldorsson<sup>1</sup>, Ragnar P. Kristjansson<sup>1</sup>, Michael L. Frigge<sup>1</sup>, Lilja Stefansdottir<sup>1</sup>, Jon K. Sigurdsson<sup>1</sup>, Asmundur Oddsson<sup>1</sup>, Asgeir Sigurdsson<sup>1</sup>, Hannes P. Eggertsson<sup>1</sup>, Pall Melsted<sup>1,6</sup>, Bjarni V. Halldorsson<sup>1,7</sup>, Sigrun H. Lund<sup>1</sup>, Unnur Styrkarsdottir<sup>1</sup>, Valgerdur Steinthorsdottir<sup>1</sup>, Julius Gudmundsson<sup>1</sup>, Hilma Holm<sup>1</sup>, Vinicius Tragante<sup>1,8</sup>, Folkert W. Asselbergs<sup>8,9,10</sup>, Unnur Thorsteinsdottir<sup>1,4</sup>, Daniel F. Gudbjartsson<sup>1,6</sup>, Kristin Jonsdottir<sup>2</sup>, Thorunn Rafnar<sup>1</sup> & Kari Stefansson<sup>1,4</sup>

Pelvic organ prolapse (POP) is a downward descent of one or more of the pelvic organs, resulting in a protrusion of the vaginal wall and/or uterus. We performed a genome-wide association study of POP using data from Iceland and the UK Biobank, a total of 15,010 cases with hospital-based diagnosis code and 340,734 female controls, and found eight sequence variants at seven loci associating with POP ( $P < 5 \times 10^{-8}$ ); seven common (minor allele frequency  $> 5\%$ ) and one with minor allele frequency of 4.87%. Some of the variants associating with POP also associated with traits of similar pathophysiology. Of these, rs3820282, which may alter the estrogen-based regulation of *WNT4*, also associates with leiomyoma of uterus, gestational duration and endometriosis. Rs3791675 at *EFEMP1*, a gene involved in connective tissue homeostasis, also associates with hernias and carpal tunnel syndrome. Our results highlight the role of connective tissue metabolism and estrogen exposure in the etiology of POP.

<sup>1</sup>deCODE Genetics/Amgen, Sturlugata 8, 101 Reykjavik, Iceland. <sup>2</sup>Department of Obstetrics and Gynecology, Landspítali University Hospital, 101 Reykjavik, Iceland. <sup>3</sup>Department of Obstetrics and Gynecology, Akureyri Hospital, 600 Akureyri, Iceland. <sup>4</sup>Faculty of Medicine, School of Health Sciences, University of Iceland, 101 Reykjavik, Iceland. <sup>5</sup>Department of Immunology, Landspítali University Hospital, 101 Reykjavik, Iceland. <sup>6</sup>School of Engineering and Natural Sciences, University of Iceland, 101 Reykjavik, Iceland. <sup>7</sup>School of Science and Engineering, Reykjavik University, 101 Reykjavik, Iceland. <sup>8</sup>Department of Cardiology, Division Heart & Lungs, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands. <sup>9</sup>Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London, UK. <sup>10</sup>Health Data Research UK and Institute of Health Informatics, University College London, London, UK. ✉email: [thorhildur.olafsdottir@decode.is](mailto:thorhildur.olafsdottir@decode.is); [kstefans@decode.is](mailto:kstefans@decode.is)

**P**elvic organ prolapse (POP) refers to the condition where one or more of the pelvic organs herniate to or beyond the vaginal opening and can involve the bladder, uterus, rectum, or post-hysterectomy vaginal cuff<sup>1</sup>. Symptoms affect quality of life and include bothersome sense of vaginal bulge, urinary or bowel symptoms, and sexual dysfunction<sup>2–4</sup>. The five stages of POP as defined by Pelvic Organ Prolapse Quantification system (POPQ)<sup>5</sup> range from stage 0; no prolapse, to stage IV which is complete eversion of the total length of the lower genital tract. Prevalence estimates for all stages of POP are 30–50% among postmenopausal women diagnosed on physical examination<sup>6–8</sup>, but 7–26% if diagnosis is restricted to pelvic organs herniating beyond or at the hymenal remnant (stages II–IV)<sup>9,10</sup>. Symptom-based prevalence estimates are lower, or 3–6%<sup>11,12</sup> with incidence peaking between ages of 70 and 79<sup>7</sup>. Between 11 and 13% of women have had surgery for POP or related conditions by age 80 years<sup>13–15</sup>. POP is the leading indication for hysterectomy in postmenopausal women and accounts for 1 in 6 hysterectomies in all age-groups<sup>16</sup>.

Risk factors for POP vary depending on POP classification. For POP stages II–IV (a descent at or beyond the hymenal remnant), risk factors are number of children delivered, vaginal delivery, advancing age and BMI<sup>17</sup>, suggesting the role of tissue trauma, estrogen exposure and intra-abdominal pressure in the pathogenesis of POP. For all stages of POP, Hispanic ethnicity<sup>6</sup>, lack of pelvic floor muscle strength<sup>18</sup> and family history<sup>19</sup> are also reported as risk factors, and previous hysterectomy is associated with severe POP<sup>20</sup>. For symptomatic POP, poor health status, constipation, or irritable bowel syndrome are also among reported risk factors<sup>11</sup>.

Although the etiology of POP is not fully understood, and is likely multifactorial, abnormality in the connective tissue supporting the vagina and pelvic organs or in the muscles in the pelvic floor have been proposed to contribute to the pathophysiology of the condition<sup>21,22</sup>. The pelvic organs' support is dependent on interactions between the levator ani muscle and pelvic connective tissue, i.e. the ligaments holding the organs in alignment<sup>23</sup>. A higher prevalence and more recurrent POP is seen in women with joint hypermobility than in others<sup>24,25</sup>. Moreover,

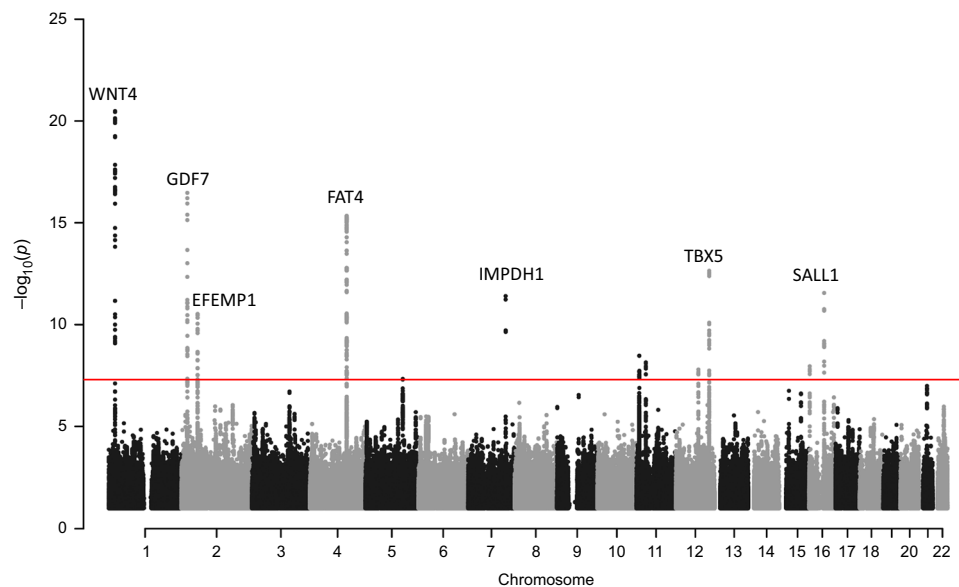
higher serum concentration of procollagen III<sup>25</sup> is found in women with joint hypermobility and in those with recurrent POP. It is, however, unclear whether such changes in collagen metabolism cause POP or are the result of a trauma<sup>26,27</sup>.

Genetic factors have been estimated to explain 43% of the variation in risk of POP in a twin study<sup>28</sup>. However, previous candidate gene-, linkage- and genome-wide-association studies (GWASs) on POP, generally of small sample sizes, have not yielded sequence variants that associate with POP<sup>26,29–31</sup>.

To provide insights into the etiology of POP, we performed a combined GWAS for POP in Iceland and the UK using data on 15,010 cases with hospital-based diagnosis code and 340,734 controls. We discovered eight variants at seven loci that associate with POP and point to a role of connective tissue metabolism and estrogen in the etiology of POP.

## Results

**Association analysis.** We performed a meta-analysis of two GWA-studies on POP, one from Iceland and the other from the UK (UK Biobank: UKB) with a combined sample of 15,010 cases and 340,734 controls of European ancestry (Supplementary Data 1). The Icelandic GWAS included 3409 cases and 131,444 controls and the corresponding numbers from UKB were 11,601 and 209,288. Cases were identified from hospital-based diagnosis records (International classification of disease (ICD) edition 10 code N81: Female genital prolapse) and controls were all females (see Methods for a detailed description of the Icelandic and UK datasets). To account for multiple testing, we used a weighted Bonferroni procedure based on sequence variant annotation<sup>32</sup> (Supplementary Data 2). A total of 245 variants at seven loci associate with POP at genome-wide significance (Fig. 1 and Supplementary Fig. 1, Supplementary Data 3). Conditional analysis at each locus identified a secondary signal at one (2p16.1) of the seven loci (Supplementary Data 4) resulting in eight distinct associations with POP (Table 1, Supplementary Fig. 2). All eight variants were nominally significant in both populations and accounting for multiple testing ( $P$ -value  $> 0.05/8 = 6.25 \times 10^{-3}$ ), there is no significant heterogeneity in the effect estimates from the two datasets (Table 1). Seven out of the eight variants are



**Fig. 1** Manhattan plot of association results between sequence variants from Iceland and UK Biobank and pelvic organ prolapse in the meta-analysis.

The significance of association for each variant ( $P$ -values on  $-\log_{10}$  scale) are plotted against the respective position on each chromosome. Red line indicates genome-wide significance level ( $5 \times 10^{-8}$ ). Genes closest to the 7 loci of interest are annotated in the Figure. The plot was created using qqman: an R package for visualizing GWAS results using Q-Q and Manhattan plots<sup>102</sup>. Variants with imputation information  $> 0.9$  are displayed.

**Table 1 Association results for lead variants at loci reaching genome-wide significance in a meta-analysis of pelvic organ prolapse.**

SNP	Pos hg38 <sup>a</sup> and cytoband	Annotation	EAF (%)	EA	OA	Gene <sup>b</sup>	LD class	Pop.	OR <sup>c</sup> (95% CI)	P	Info	P <sub>het</sub>
rs3820282	chr1:22,141,722 1p36.12	Intron	17.17	T	C	WNT4	23	Ice	0.83 (0.78, 0.90)	1.0 × 10 <sup>-6</sup>	1.00	
rs9306894	chr2:20,678,345 2p24.1	3' UTR	34.48	G	A	GDF7	5	UKB <i>Meta</i>	0.85 (0.82, 0.89) 0.85 (0.82, 0.88)	5.0 × 10 <sup>-16</sup> 3.3 × 10 <sup>-21</sup>	1.00	0.62
rs1430191	chr2:55,806,292 2p16.1	Intergenic	47.85	T	C	EFEMP1	3	Ice <i>Meta</i>	1.10 (1.04, 1.16) 1.12 (1.09, 1.15) 1.11 (1.09, 1.14)	1.7 × 10 <sup>-3</sup> 4.1 × 10 <sup>-15</sup> 3.4 × 10 <sup>-17</sup>	1.00	0.51
rs3791675	chr2:55,884,174 2p16.1	Intron	21.07	T	C	EFEMP1	30	Ice <i>Meta</i>	1.15 (1.08, 1.22) 1.07 (1.04, 1.10) 1.09 (1.06, 1.12)	5.4 × 10 <sup>-6</sup> 5.2 × 10 <sup>-6</sup> 1.0 × 10 <sup>-9</sup>	1.00	0.034
rs7682992	chr4:126,037,217 4q28.1	Intergenic	20.63	T	A	FAT4	78	UKB <i>Meta</i>	0.81 (0.74, 0.87) 0.88 (0.85, 0.92) 0.87 (0.84, 0.90)	9.3 × 10 <sup>-8</sup> 6.0 × 10 <sup>-12</sup> 2.7 × 10 <sup>-17</sup>	1.00	0.034
rs72624976	chr7:128,392,779 7q32.1	3' UTR	4.87	T	C	IMPDH1	5	UKB <i>Meta</i>	1.21 (1.13, 1.29) 1.11 (1.07, 1.15) 1.13 (1.10, 1.16)	1.0 × 10 <sup>-8</sup> 6.7 × 10 <sup>-10</sup> 4.5 × 10 <sup>-16</sup>	1.00	0.027
rs1247943	chr12:114,235,616 12q24.21	Intergenic	46.03	G	A	TBX5	10	Ice <i>Meta</i>	0.74 (0.65, 0.84) 0.82 (0.76, 0.88) 0.80 (0.75, 0.85)	3.4 × 10 <sup>-6</sup> 1.0 × 10 <sup>-7</sup> 4.0 × 10 <sup>-12</sup>	1.00	0.18
rs12325192	chr16:51,454,218 16q12.1	Intergenic	17.90	T	C	SALL1	20	UKB <i>Meta</i>	1.10 (1.05, 1.17) 1.09 (1.06, 1.12) 1.09 (1.07, 1.12)	3.3 × 10 <sup>-4</sup> 1.5 × 10 <sup>-10</sup> 2.3 × 10 <sup>-13</sup>	1.00	0.73
								Ice	0.93 (0.86, 1.00)	0.045	1.00	
								UKB <i>Meta</i>	0.89 (0.86, 0.92) 0.89 (0.87, 0.92)	1.3 × 10 <sup>-11</sup> 2.8 × 10 <sup>-12</sup>	1.00	0.32

P P-value, OR odds ratio, P<sub>het</sub> P-value for heterogeneity in the effect estimate between the Icelandic and UK Biobank data, EA effect allele frequency (average of ICE and UKB), EA effect allele, OA other allele, Pop. population, Info information score, LD linkage disequilibrium, Pop. population. Results of the meta-analysis of the GWA-studies from Iceland (Ice) and UKB are shown in italics.  
<sup>a</sup>Variant positions (Chromosome:position) are according to GRCh38/hg38.  
<sup>b</sup>For intergenic variants, nearest gene is reported.  
<sup>c</sup>Odds-ratios correspond to effect alleles.  
<sup>d</sup>Two distinct signals were detected at the 2p16.1 locus, rs3791675 and rs1430191. OR and P for rs3791675 are conditioned on rs1430191 and vice versa.

common, with minor allele frequency (MAF) range of 17–48%, and one is of low frequency (MAF = 4.87%). The number of variants that correlate with the lead variants ( $r^2 > 0.8$ ) at these loci ranges from 3 to 78 (Table 1, Supplementary Fig. 2). No previously suggested POP variants from candidate gene, linkage or GWAS studies<sup>26,29–31</sup> associate with POP (Supplementary Data 5, 6).

We assessed the robustness of the POP association results for the eight lead variants in a group of more stringently diagnosed POP cases that are defined by procedure codes specific to the treatment of POP in addition to the ICD10 code N81 (Supplementary Data 7). The statistical power for this analysis comes mainly from the UKB data because information on POP related surgeries was available for 61% of the ICD-coded POP cases in UK, but only 13% of the Icelandic POP cases. The effect sizes tended to be greater for the association using cases that had undergone surgery compared to those diagnosed only based on ICD code (Supplementary Data 8). This most likely reflects increased severity of POP among those that had undergone surgery or a greater number of falsely diagnosed POP cases among the ICD based cases. We conclude that our results based on ICD codes alone are robust as the effects are in the same direction and the effect sizes are not substantially different from those using surgically treated POP cases.

We used LD score regression to estimate the SNP heritability of POP<sup>33</sup>. Using LD scores for about 1.1 million variants found in European populations we estimated SNP heritability in the meta-analysis to be 12.4% (95% CI 9.9–14.8%).

### Functional annotation and biological inference of risk loci.

None of the lead POP variants are coding or in high LD ( $r^2 > 0.8$ ) with coding variants (Supplementary Data 3) and none was highly correlated with the top cis-eQTL for neighboring genes in any of the available tissues using mRNA expression data from the GTEx database and our own RNA-sequence data on Icelandic samples from blood (13,162 individuals) and adipose tissue (749 individuals).

To gain a further understanding of the variants associating with POP we explored how they correlate with other traits in combined data from Iceland and UKB – a database with approximately 800 traits ( $P$ -value threshold =  $0.05/800 = 6.3 \times 10^{-5}$ ) – in addition to looking up all variants and their correlates ( $r^2 > 0.5$ ) in the GWAS-catalog<sup>34</sup> (Supplementary Data 9, 10). As we observed overlaps of genetic association to different traits we conducted adjusted association analyses (co-localization analyses) at the POP loci to identify association results that are consistent with a single signal representation (see Methods).

Six of the eight POP variants also associate with a total of 13 other phenotypes. Three of those variants associate with six phenotypes with pathogenesis that may be the same as that of POP. Two POP variants associate with conditions related to estrogen exposure, rs3820282 at the *WNT4* locus (leiomyoma of uterus, gestational duration and endometriosis) and rs12325192 at the *SALL1* locus (leiomyoma). The third, rs3791675 at the *EFEMP1* locus, associates with conditions related to aberrant connective tissue function (hernias and carpal tunnel syndrome) (Table 2).

Secondary, non-symptomatic diagnosis of POP through conditions such as leiomyoma and endometriosis may introduce a confounding of or biases in the association results from our GWAS meta-analysis. We therefore tested the association of our eight POP variants with POP after excluding endometriosis and leiomyoma cases from the analysis. This did not affect effect or significance substantially (Supplementary Data 11) indicating that

the POP associations are independent of endometriosis and leiomyoma.

**POP associated loci.** The most significant POP association is with rs3820282–T ( $P = 3.3 \times 10^{-21}$ , OR = 0.85), located in intron 1 of *WNT4*, which encodes a protein involved in development of the female reproductive tract<sup>35</sup>. Loss of *WNT4* function in humans leads to underdevelopment and sometimes absence of the uterus and vagina<sup>36</sup>. Variants correlated ( $r^2 > 0.8$ ) with the POP-protecting allele rs3820282–T have been associated with increased risk of endometriosis<sup>37</sup>, leiomyoma<sup>38</sup>, increased gestational duration<sup>39</sup> and with decreased bone mineral density<sup>40</sup> (Supplementary Data 10), all of which we replicate (Supplementary Data 9). Rs3820282–T also associates with lower number of children in our combined data (Supplementary Data 9). Our co-localization analyses revealed that apart from the bone mineral density and number of children signals, the association results for leiomyoma, gestational duration, and endometriosis are likely to be the same signal as the POP association at the locus (Supplementary Data 12 and Table 2).

Rs12325192 is located near *SALL1* and the POP-protecting allele rs12325192–T is correlated ( $r^2 = 0.94$ ) with rs66998222, a variant we have previously reported to associate with leiomyoma<sup>38</sup>; a trait that is positively correlated with estrogen exposure (Table 2). Mutations in *SALL1* have been associated with Townes-Brocks syndrome, a condition associated with renal malformation, suggesting a role of *SALL1* in genitourinary development<sup>41</sup>.

Two distinct POP variants, rs3791675 and rs1430191, are located in and near *EFEMP1* (also known as *FBLN3*), a gene encoding fibulin-3. Fibulins are components of microfibrils, the building blocks of elastic fibers that are generated in fibroblasts<sup>42</sup> and provide tissue elasticity. *EFEMP1* is expressed in mouse connective tissue and network analyses suggest a role for *EFEMP1* in connective tissue maintenance/homoeostasis<sup>43</sup>. Rare mutations in *EFEMP1* have been found to cause autosomal dominant Doyne honeycomb degeneration of retina, characterized by yellow–white deposits known as drusen that accumulate beneath the retinal pigment epithelium<sup>44</sup>. The strongest POP association at the locus is with rs3791675–T ( $P = 2.7 \times 10^{-17}$ , OR = 0.87). Its correlates ( $r^2 > 0.8$ ) have been associated with various traits in Europeans, including less height<sup>45,46</sup>, smaller BMI-adjusted waist circumference<sup>47,48</sup> and protection against inguinal hernia<sup>43</sup> (Supplementary Data 10). We replicate these associations with our data (Table 2). In addition, rs3791675–T associates with protection against various hernias; inguinal hernia, femoral hernia, umbilical, and ventral hernia, consistent with the hypothesis of a similar proposed collagen pathophysiology for POP and hernias<sup>49–52</sup>. Rs3791675–T also associates with protection against diverticular disease, a disease in which abnormal collagen and decreased tensile strength of the colonic wall have been proposed to contribute to pathogenesis<sup>53</sup>. We furthermore found rs3791675–T to associate with increased risk of carpal tunnel syndrome – a condition also linked to connective tissue metabolism<sup>54</sup> – higher pulse pressure, and the lung function ratio forced expiratory volume (FEV1)/forced vital capacity (FVC) (Supplementary Data 9). According to our co-localization analyses, association results for inguinal hernia, height, FEV1/FVC, carpal tunnel syndrome, ventral hernia, and waist circumference (Supplementary Data 12 and Table 2), are consistent with a single signal origin.

Rs9306894, located in the 3'UTR of *GDF7* (also known as *BMP12*), associates with POP. *GDF7* encodes a secreted ligand of the TGF- $\beta$  superfamily of proteins and is thought to be involved in tendon and ligament formation and repair<sup>55,56</sup>. The

**Table 2 A summary of previously reported (\*) and novel associations of POP variants with other traits.**

<b>SNP [EA/OA]</b>	<b>Position hg38</b>	<b>EAf (%)</b>	<b>Gene</b>	<b>Trait</b>	<b>OR/<math>\beta</math></b>	<b>(95% CI)</b>	<b>P</b>	<b>r<sup>2</sup></b>	<b>Published*/Top SNP</b>	<b>PubmedID</b>
rs3820282 [T/C]	1:22141722	17.17	WNT4	POP	0.85	(0.82, 0.88)	3.3 × 10 <sup>-21</sup>	0.877	rs10917151*	30194396
				Leiomyoma	1.12	(1.09, 1.16)	5.3 × 10 <sup>-13</sup>	0.878	rs56318008*	28877031
				Gestational duration	0.058	(0.04, 0.08)	3.9 × 10 <sup>-11</sup>	0.945	rs12037376*	28537267 <sup>a</sup>
				Endometriosis	1.12	(1.06, 1.18)	1.1 × 10 <sup>-5</sup>			
rs9306894 [G/A]	2:20678345	34.48	GDF7	POP	1.11	(1.09, 1.14)	3.4 × 10 <sup>-17</sup>	0.996	rs9306895*	29892016 <sup>b</sup>
				Prostate cancer	1.07	(1.04, 1.11)	6.2 × 10 <sup>-6</sup>	0.620	rs7255*	27527254
				BE and EA	1.09	(1.05, 1.14)	4.2 × 10 <sup>-5</sup>			
rs3791675 [T/C]	2:55884174	21.07	EFEMP1	POP	0.87	(0.88, 0.93)	2.7 × 10 <sup>-17</sup>	0.960	rs59985551*	26686553
				Inguinal hernia <sup>c</sup>	0.82	(0.80, 0.84)	6.2 × 10 <sup>-44</sup>	0.950	rs3791679*	25282103
				Height	-0.081	(-0.07, -0.10)	6.7 × 10 <sup>-27</sup>	1	rs3791675	
				FEV1/FVC	0.033	(0.03, 0.04)	4.1 × 10 <sup>-24</sup>	0.960	rs191535629	
				Carpal tunnel s.	1.11	(1.08, 1.15)	9.0 × 10 <sup>-11</sup>	0.950	rs3791679	
				Ventral hernia	0.86	(0.81, 0.91)	5.9 × 10 <sup>-8</sup>	0.950	rs3791679*	
				Waist circumference	-0.036	(-0.04, -0.03)	1.8 × 10 <sup>-35</sup>	0.950	rs3791679*	25673412
rs7682992 [T/A]	4:126037217	20.63	FAT4	POP	1.13	(1.10, 1.16)	4.6 × 10 <sup>-16</sup>	0.640	rs2893158	
				Stress incontinence <sup>d</sup>	1.11	(1.06, 1.15)	6.4 × 10 <sup>-7</sup>			
rs1247943 [G/A]	12:114235616	46.03	TBX5	POP	1.09	(1.07, 1.12)	2.3 × 10 <sup>-13</sup>	0.990	rs2555019*	30410027 <sup>e</sup>
				BPH	0.92	(0.90, 0.94)	4.2 × 10 <sup>-12</sup>	0.814	rs1270884*	29892016 <sup>f</sup>
				Prostate cancer	1.05	(1.02, 1.08)	1.8 × 10 <sup>-3</sup>			
rs12325192 [T/C]	16:51454218	17.90	SALL1	POP	0.89	(0.87, 0.92)	2.8 × 10 <sup>-12</sup>	0.936	rs66998222*	30194396
				Leiomyoma	0.91	(0.88, 0.94)	4.0 × 10 <sup>-9</sup>			

P value for fixed effects meta-analysis, OR odds ratio,  $\beta$  effect estimate (italic), EAF effect allele frequency (average of ICE and UK data), r<sup>2</sup> linkage disequilibrium between POP index variant and the corresponding variant listed, BE and EA Barrett's oesophagus and esophageal adenocarcinoma combined, BPH benign prostatic hyperplasia, FEV1/FVC the ratio of forced expiratory volume and forced vital capacity.

Shown are association results consistent with a single signal origin, after performing conditional analysis using data from ICE and UKB (height and gestational duration is Iceland only, BE and EA is UKB only). Top variants are defined for each trait within  $\pm 500$  kb from the POP variant at each locus. Variant locations are according to GRCh38/hg38. P and OR for rs3791675 are adjusted for the secondary signal rs1430191 in the POP phenotype (see Supplementary Data 4).

<sup>a</sup>See also Rahmoghlu et al.<sup>65</sup>

<sup>b</sup>Originally reported in Amin Al Olama et al.<sup>103</sup>

<sup>c</sup>rs59985551 is top variant for inguinal hernia in the combined data from ICE and UKB. The previously reported variant at the locus, rs2009262, associates with POP with P = 1.63 × 10<sup>-36</sup> (OR = 0.83) and is in r<sup>2</sup> = 0.85 with rs3791675.

<sup>d</sup>319 of the cases diagnosed with stress incontinence (ICD10 N393) are also diagnosed with POP.

<sup>e</sup>Two independent signals were reported at the 12q24.21 locus, rs2555019 and rs8853. Only rs2555019 is correlated with the POP variant rs1247943 (r<sup>2</sup> = 0.002 for rs8853).

<sup>f</sup>Originally reported in Eeles et al.<sup>60</sup>

protein has been found to play a crucial role in tenogenesis of mesenchymal stem cells and is used in tissue-engineering to treat tendon injury<sup>57</sup>. A correlated variant near *GDF7* (rs2289081–C,  $r^2 = 0.85$ ) has been associated with decreased pulse pressure<sup>58</sup> and rs7255–T ( $r^2 = 0.62$  to rs9306894) has been associated with increased risk of Barrett's oesophagus and esophageal adenocarcinoma combined (BE and EA) (Supplementary Data 10), which we replicate (Supplementary Data 9) but co-localization analysis only supports a single signal origin for BE and AE combined (Supplementary Data 12). We replicate the previous association of a highly correlated variant rs9306895 ( $r^2 = 0.996$ ) with prostate cancer<sup>59</sup> ( $P = 6.15 \times 10^{-6}$ , OR = 1.07) (Table 2).

For the POP variant rs1247943, close to *TBX5* (distance: 172 Kb), we replicate the previous associations of two highly correlated variants with the POP risk-increasing allele; rs1270884 ( $r^2 = 0.81$ ) and rs2555019 ( $r^2 = 0.99$ ) with increased risk of prostate cancer<sup>60</sup> and lowered risk of benign prostatic hyperplasia<sup>61</sup> (Table 2).

Rs7682992–T, the POP variant close to the *FAT4* gene (distance: 0.5 Mb), associates with increased risk of stress incontinence (Table 2). *FAT4* is part of the planar cell polarity pathway that controls tissue organization; loss of *FAT4* in mice leads to cystic kidney disease<sup>62</sup>.

Rs72624976 is located in the 3'UTR of *IMPDH1*, one of two target genes for the immunosuppressant drug mycophenolic acid<sup>63</sup>. Mutations in *IMPDH1* have been associated with the eye disorders retinitis pigmentosa<sup>64</sup> and Leber's congenital amaurosis<sup>65</sup>.

**Comorbidity.** To test further for comorbidity between POP and other traits, we collected sequence variants reported to associate with traits that associate with one or more of our POP variants: leiomyoma ( $N = 32$ )<sup>38,66</sup>, gestational duration ( $N = 5$ )<sup>39</sup>, endometriosis ( $N = 27$ )<sup>67</sup>, BE and AE combined ( $N = 14$ )<sup>68</sup>, prostate cancer ( $N = 145$ )<sup>59</sup>, inguinal hernia ( $N = 4$ )<sup>43</sup>, height ( $N = 3,290$ )<sup>69</sup>, FEV1/FVC ( $N = 52$ )<sup>70</sup>, BMI-adjusted waist circumference ( $N = 70$ )<sup>48</sup> and benign prostatic hyperplasia ( $N = 23$ )<sup>61</sup>. Of the 3,652 unique variants collected, 17 were correlated within height ( $r^2 > 0.8$ ). We tested these 3,635 (3652–17) independent sequence variants for association with POP. As 44 variants were correlated across traits ( $r^2 > 0.8$ ), we used the significance threshold of  $P < 1.4 \times 10^{-5}$  (0.05/3591) (3652–17–44 = 3591) (Supplementary Data 13–22). We found variants at nine loci to associate with POP; five of which associate at genome-wide significance with POP, three that were previously associated with height ( $P < 1.4 \times 10^{-5}$ ) at *TXNDC5* (6p24.3), *SLC12A2* (5q23.3) and *LOXL1* (15q24.1) (Supplementary Data 19) and one at *WT1* that associates with inguinal hernia (11p13) (Supplementary Data 18).

We further used the variants collected for each of the ten traits as instruments to explore a possible causal relationship between each trait and POP<sup>71</sup>. By regressing the effect estimates of each group of variants (trait-specific) on POP on the effect estimates of those variants on the corresponding traits, using AF  $\times$  (1–AF) as a weight<sup>72</sup>, we did not observe a correlation between the effect estimates (Supplementary Fig. 3).

We also plotted the POP effects of the eight POP variants against their effects on the 13 traits listed in Table 2 (Iceland and UKB combined, apart from height and gestational duration which are data from Iceland) (Supplementary Fig. 4). It is apparent that although a handful of variants associate with POP and other phenotypes, there is little overlap between the variants associating with POP and the other phenotypes and little correlation between effect estimates. We are not aware of previous literature findings showing associations of these phenotypes within POP cases and their family members.

**BMI is not associated with POP at the genetic level.** We screened for evidence of a genetic relationship of POP and two of the most consistently reported risk factors for POP in addition to age; BMI and number of children. We verified that variants reported to associate with BMI<sup>73,74</sup> or number of children<sup>75</sup> do not associate with POP (Supplementary Data 23, 24) and that adding BMI or number of children separately as covariates in our models did not affect the effect sizes or significance of the eight lead variants when tested for association with POP (Supplementary Data 25). Furthermore, no correlation was found between effect estimates of BMI- or number of children sequence variants and their effects on POP and vice versa (Supplementary Figs. 5, 6). Using polygenic risk scores for POP, we saw little evidence of association with number of children ( $P$ -value =  $9.5 \times 10^{-5}$ , beta = 0.015) in Iceland only and none with BMI (Supplementary Data 26). For BMI, we found no evidence of causal relationship with POP at the genetic level. With that in mind, and since only one of the POP variants (rs3820282) associates with number of children we do not find strong support for a causal pathway or a common third factor affecting POP and those two traits.

## Discussion

We conducted a meta-analysis of two GWAS for POP and discovered eight variants at seven loci that associate with POP. Two sequence variants, rs3820282 at *WNT4* and rs12325192 near *SALL1* also associate with other traits that are strongly affected by estrogen exposure, i.e. leiomyoma (rs3820282 and rs12325192), gestational duration (rs3820282) and endometriosis (rs3820282). Of the 23 correlated variants at the 1p36.12 locus, rs3820282 has the best functional candidacy. Rs3820282–T resides in a region marked by mono-methylation of histone H3 at lysine residue K4 (H3K4me1), a characteristic of enhancers, and is predicted to alter a conserved binding site for the transcription factors (TFs) ESR1 and ESR2<sup>76</sup> supported by CHIP-seq data in breast and bone cell lines<sup>77</sup>. This suggests that the variant may alter the estrogen-based regulation of *WNT4* or adjacent genes<sup>39</sup>. A previous report showed chromatin looping between the region containing rs12038474 (LD to rs3820282;  $r^2 = 0.81$ ) and *CDC42* in endometrial adenocarcinoma<sup>78</sup>. Our analysis of available Hi-C data<sup>79</sup> along with enhancer-gene predictions<sup>80</sup> suggests *WNT4*, *HSPG2*, and *CDC42* as possible enhancer-gene targets, using relevant cell-types and tissues. Based on the role of *WNT4* in female sex organ development<sup>35</sup>, *WNT4* can be considered to be a strong candidate gene at the locus. One of the eight POP variants, rs3791675 at *EFEMP1*, associates with traits with proposed collagen pathophysiology, i.e. inguinal hernia, ventral hernia and carpal tunnel syndrome. Furthermore, two genes at POP associated loci, *GDF7* and *EFEMP1*, have functions related to connective tissue metabolism.

We find that the overlap between POP variants and other traits are limited to distinct signals and thus do not allow for strong inference regarding common pathways or causal relationships. However, the pleiotropy found at three of the POP loci, where the associating traits have similar proposed pathophysiology as POP, together with the known functions of the proposed target genes, may fuel further explorations of the possible role of estrogen exposure and connective tissue metabolism in the etiology of POP. Through analysis of variants that are reported to associate with conditions that associate with one or more of our POP variants, we identified four additional POP variants, three that associate with height and one with inguinal hernia.

In mice, a POP phenotype has been associated with knock-down of fibulin 3 (*EFEMP1*), lysyl oxidase like 1 (*LOXL1*), fibulin 5 (*FBLN5*), and homeobox A11 (*HOXA11*) as summarized in a

recent review on key genes and pathways for POP<sup>81</sup>. Of those four genes, variants at the *EFEMP1* locus associate with POP in our data and, through analysis of height variants, also a variant at the *LOXL1* locus; rs12440667-T ( $P = 3.6 \times 10^{-6}$ , OR = 0.94) (Supplementary Data 19, Supplementary Fig. 7). *EFEMP1* encodes fibulin-3 but members of the fibulin family of extracellular matrix (ECM) proteins have been found to be important in elastic fiber assembly<sup>82,83</sup>. The two *EFEMP1* mouse knockout studies reported to date<sup>84,85</sup> demonstrate an abnormality in the integrity of elastic fibers in fascia connective tissue and connective tissue of the vaginal wall with abnormal pelvic organ support in addition to increased activity of matrix metalloproteinase (MMPs) that degrade matrix collagen in connective tissue. These mice also displayed an early onset of aging-associated traits including decreased body mass, lordokyphosis, reduced hair growth, and generalized fat, muscle and organ atrophy, as well as reduced lifespan. The POP variants in and near to *EFEMP1* did not associate with traits likely to cause lordokyphosis (any fracture and osteoporosis in the ICE-UKB data, and BMD (hip and spine measures)), BMI (both sexes and females only) or lifespan (females) in the Icelandic data (Supplementary Data 27).

In this paper, we report the first set of genome-wide significant POP variants identified through GWAS. The genetic overlap observed between POP and several traits with similar pathophysiology point towards a role of estrogen exposure and connective tissue metabolism in the etiology of POP. The results provide new insights for future research to further the understanding of POP.

## Methods

**Datasets.** The meta-analysis combined the results of two GWA studies of pelvic organ prolapse (POP). The Icelandic data originates from Landspítali – The National University Hospital Inpatient Registry from January 1983 to August 2018. A total of 3,699 women had a POP diagnosis (ICD10 code N81 or ICD9 code 618). Controls were recruited through different genetic research projects at deCODE genetics. We had genotype information for 92% of the POP cases, so that the Icelandic dataset consisted of 3409 cases and 131,444 female controls. The UK Biobank data (UKB) consists of 11,601 cases and 209,228 controls of European ancestry (self-reported white British with similar genetic ancestry based on principal component analysis and with consistent reported and genetically determined gender)<sup>86</sup>, recruited between 2006 and 2010 aged 40–69, and with follow-up until 2016<sup>87</sup>. The UKB cases had ICD10 code N81 in hospital inpatient records, with data dating back to between 1981 and 1997, depending on the external source (Hospital Episode Statistics from England (89% of participants), Patient Episode Database for Wales (7%) and Scottish Morbidity Record (7%)) and include primary and secondary diagnosis of POP ([http://biobank.ctsu.ox.ac.uk/showcase/docs/inpatient\\_mapping.pdf](http://biobank.ctsu.ox.ac.uk/showcase/docs/inpatient_mapping.pdf)).

Hospital-based POP diagnosis are likely to be stage II or greater, according to the POPQ standardized grading system<sup>5</sup>. Furthermore, the ratio of cases to controls in the datasets (0.055 in UKB and 0.026 in ICE) are consistent with reported prevalence estimates that use POPQ stage II–IV as a criterion for diagnosis or with symptom-based prevalence estimates<sup>8–11</sup>. The 2-fold difference in the ratio of cases to controls of POP in UKB data compared to the data from Iceland is explained by the difference in age-distributions (Supplementary Data 1). To further compare the POP phenotype between Iceland and UKB, we compared the percentage of cases within each of the nine subgroups of ICD10 code N81. In both datasets the majority of cases are diagnosed with cystocele or rectocele and fewer with uterovaginal prolapse (Supplementary Data 28).

The leiomyoma<sup>38</sup>, benign prostatic hyperplasia<sup>61</sup>, prostate cancer<sup>61</sup>, endometriosis<sup>37</sup>, diverticular disease<sup>88</sup>, and lung function<sup>70</sup> datasets have been described previously. We obtained blood pressure measurements from Landspítali – The National University Hospital of Iceland, the Primary Health Care Clinics of the Reykjavík area and at recruitment for deCODE studies. Data on gestational duration (GD) (Iceland only) for the first available pregnancy was obtained from the National Birth Registry, excluding multiple pregnancies, stillbirths and pre and post term births (GD  $\leq 259$  days and GD  $\geq 301$  days) and adjusting for age of mother, sex, and year of birth of the child. Information on number of children was extracted from deCODE genealogy database, adjusting for year of birth and county. Information on BMI was corrected for year of birth, age and county, conditional on age >18. BMI values originate from measured and self-reported data and are mean values for multiple measures within individuals. Waist circumference measurement was adjusted for age, gender, and BMI, conditional on age > 18. The BMD values at head (DEXA, Hologic QDR4500A) were age, height and body mass index corrected (Iceland only). Quantitative traits were rank-based inverse normal transformed to a

standard normal distribution separately for each gender. Other datasets from Iceland originate from deCODE genetics phenotype database which contains extensive medical information on various diseases and other traits. Cases with hernias were identified by hospital-based ICD10 diagnosis codes: K40 (inguinal hernia), K41 (femoral hernia), K43 (ventral hernia) in both datasets; Iceland and UKB and the same applies for diverticular disease (K57), stress incontinence (N393), carpal tunnel syndrome (G560), Barrett's oesophagus and esophageal adenocarcinoma combined (K227 and C15, data in UKB only), and actinic keratosis (L57). For sex specific phenotypes in Iceland, only controls for the relevant sex were included in the analysis.

The study was approved by the Icelandic National Bioethics Committee (bioethics consent number VSN 18-067) in agreement with conditions issued by the Data Protection Authority of Iceland. Written informed consent was obtained from all genotyped subjects. Personal identities relating to participants' data and biological samples (i.e. blood samples, buccal samples, or adipose tissue samples) were encrypted by a third-party system (Identity Protection System), approved and monitored by the Data Protection Authority<sup>89</sup>. DNA was extracted from blood and buccal samples and RNA from blood and adipose tissue samples.

**Genotyping and imputation.** Details of genotyping and imputation methods in the Icelandic part of the study have been described<sup>90</sup>. In brief, the whole genomes of 15,220 Icelanders, participating in various disease projects at deCODE genetics, were sequenced to a mean depth of at least  $10\times$  (median  $32\times$ ) using Illumina technology. Genotypes of SNPs and indels were called using joint calling with the Genome Analysis Toolkit HaplotypeCaller (GATK version 3.4.07)<sup>91</sup>. As all of the sequenced individuals had also been chip-typed and long-range phased, information about haplotype sharing was used to improve genotype calls. In total, 151,677 Icelanders have been genotyped with various Illumina SNP chips, long-range phased and imputed based on the sequenced data set<sup>92</sup>. Using genealogical information, genotype probabilities for 282,894 untyped relatives of the genotyped individuals were calculated to further increase the sample size for association analysis and to increase the power to detect associations.

The informativeness of genotype imputation was estimated by the ratio of the variance of imputed expected allele counts and the variance of the actual allele counts:

$$\frac{\text{Var}(E(\theta|\text{chip data}))}{\text{Var}(\theta)}$$

where  $\theta$  is the allele count. Here,  $\text{Var}(E(\theta|\text{chip data}))$  is estimated by the observed variance in the imputed expected counts and  $\text{Var}(\theta)$  was estimated by  $p(1-p)$ , where  $p$  is the allele frequency. Variants were annotated using Ensembl release 80 and Variant Effect Predictor (VEP) version 2.8<sup>93</sup>.

Genotyping of UKB samples was performed using a custom-made Affymetrix chip, UK BiLEVE Axiom<sup>94</sup> in the first 50,000 participants and with the Affymetrix UK Biobank Axiom array<sup>95</sup> in the remaining participants with 95% of the signals on both chips. Imputation was performed by Wellcome Trust Centre for Human Genetics using a combination of the Haplotype Reference Consortium (HRC), 1000 Genomes phase 3 and the UK10K haplotype resources<sup>86</sup>. This yields a total of 92.6 million imputed variants, however, we used only markers that were imputed on the basis of the HRC panel owing to problems in the UK10K + 1000 Genomes panel imputation, resulting in 37.1 million imputed variants used in the current study.

**GWAS and meta-analysis.** We used logistic regression assuming an additive model in the case-control analysis to test for association between variants and disease, treating disease status as the response and expected genotype counts from imputation as covariates, and using likelihood ratio test to compute  $P$ -values. For the Icelandic data, the model also includes as nuisance variables available individual characteristics that correlate with phenotype status. In Iceland, these covariates are: county of birth, current age or age at death (first and second order terms included), availability of blood sample for the individual and an indicator function for the overlap of the lifetime of the individual with the time span of phenotype collection. The association analysis for both the Icelandic and UKB datasets was done using software developed at deCODE genetics<sup>90</sup>. For the association testing in UKB, 40 principal components were used to adjust for population substructure and age was included as a covariate in the logistic regression model.

To account for inflation in test statistics due to cryptic relatedness and stratification, we applied the method of linkage disequilibrium (LD) score regression<sup>33</sup>. LD scores were downloaded from an LD score database ([ftp://atguftp.mgh.harvard.edu/brendan/lk\\_eur\\_r2\\_hm3snps\\_se\\_weights.RDS](ftp://atguftp.mgh.harvard.edu/brendan/lk_eur_r2_hm3snps_se_weights.RDS); accessed 23 June 2015). With a set of 1.1 million variants, we regressed the  $\chi^2$  statistics from our GWAS scan on the LD scores and used the intercept as a correction factor.  $P$  values were then adjusted by dividing the corresponding  $\chi^2$  values by the correction factors. The estimated correction factor for POP based on LD score regression was 1.12 for the Icelandic and 1.05 for the UK datasets, respectively.

Variants in the UKB imputation dataset were mapped to National Center for Biotechnology Information (NCBI) Genome Reference Consortium Build 38 positions (GRCh38) and matched to the variants in the Icelandic dataset based on allele variation. Only sequence variants from the Haplotype Reference Consortium panel (HRC) were included in the meta-analysis. The results from the two datasets

were combined using a fixed effect model (a Mantel-Haenszel model)<sup>96</sup> in which the datasets were allowed to have different population frequencies for alleles and genotypes but were assumed to have a common OR and weighted with the inverse of the variance. We selected a threshold of 0.8 imputation info and MAF >0.01% for variants available in the Icelandic data set and/or the UKB data set. A total of 37,144,863 variants were used in the analysis. We tested for heterogeneity by comparing the null hypothesis of the effect being the same in both populations to the alternative hypothesis of each population having a different effect using a likelihood ratio test (Cochran's Q) reported as  $P_{\text{het}}$  (Table 1 for POP and Supplementary Data 29 for other traits).

We used the weighted Bonferroni method to account for all 37,144,863 variants being tested ( $P\text{-value} < (0.05 \times \text{weight})/37,144,863$ ). Using the weights given in Sveinbjornsson et al.<sup>32</sup>, this procedure controls the family-wise error rate at 0.05;  $P \leq 2.2 \times 10^{-7}$  for high-impact variants (including stop-gained, frameshift, splice acceptor or donor and initiator codon variants,  $N = 9,658$ ),  $P \leq 4.4 \times 10^{-8}$  for missense, splice-region variants and in-frame-indels ( $N = 180,803$ ),  $P \leq 4.0 \times 10^{-9}$  for low-impact variants (including synonymous, 3' and 5' UTR, and upstream and downstream variants,  $N = 2,653,622$ ),  $P \leq 6.7 \times 10^{-10}$  for intron and intergenic variants ( $N = 34,300,780$ ) (Supplementary Data 2). All  $P$ -values are two-sided.

**Conditional analysis.** We applied approximate conditional analyses, implemented in the GCTA software<sup>97</sup> to the meta-analysis summary statistics to look for additional association signals at each of the genome-wide significant loci. LD between variants was estimated using a set of 8700 whole-genome sequenced Icelandic individuals. The analysis was restricted to variants present in both the Icelandic and UKB datasets and within 1 Mb from the index variants. We tested 7 loci and about 100,000 variants in the conditional analysis and report one variant (rs1430191) with conditional  $P$  value  $< 5 \times 10^{-7}$ . The results from GCTA were verified by conditional analysis using genotype data in the Icelandic and UK datasets separately and results presented in Table 1 are obtained by meta-analyzing those results.

**Co-localization analysis.** The variants explored as potentially representing the same signal as the POP variants were found by two means. First, we looked up correlates of the POP variants in the GWAS-catalog ( $r^2 > 0.5$ ) (Supplementary Data 10) and extracted those with  $0.5 < r^2 < 0.9$  for the co-localization analysis because of the limited potential to distinguish between variants in LD of  $r^2 > 0.9$ . Second, for each of the traits associating with the POP variants in our data we extracted the strongest associating variant ( $r^2 < 0.9$ ) for the adjusted (conditional) analysis. For the tests performed in the co-localization analysis, we use previously reported variants as index variants for the secondary traits when available for a conditional analysis, given that the reported analyses are based on a similar sized or a larger sample than our combined data from Iceland and UKB. Otherwise we use the variant most strongly associating with the trait in our data. The results from the conditional analyses are consistent with a single signal representation if two conditions are met: First, the  $P$ -value of the index variant for the secondary trait is  $\geq 0.05$  after adjusting for the POP variant at the locus and second, the POP variant at the locus associates with the secondary trait at a  $P$ -value  $< 6.3 \times 10^{-5}$ . The latter condition holds for traits identified through the phenomescan (PheWAS) but a  $P$ -value  $< 0.05$  is required if a correlate of the POP variant is previously reported to associate with the trait in a considerably larger sample than in our data (Supplementary Data 12).

**RNA sequencing analysis.** Generation of poly(A)+ complementary DNA sequencing libraries, RNA sequencing, and data processing were carried out as described previously<sup>98,99</sup>. Two tissue types were available for this analysis: whole blood ( $N = 13,162$ ) and adipose tissue ( $N = 749$ ). We used generalized linear regression to test for association between sequence variants and rank-transformed gene expression estimates.

**Polygenic scores, heritability, and functional annotation.** We calculated two sets of polygenic risk scores (PRSs<sub>POP</sub>) both using Icelandic and UK data essentially as previously described<sup>100</sup>. Briefly, the PRSs were calculated using genotypes for about 630,000 autosomal markers included on the Illumina SNP chips to avoid uncertainty due to imputation quality. We estimated linkage disequilibrium (LD) between markers using 14,938 phased Icelandic samples and used this LD information to calculate adjusted effect estimates using LDpred<sup>100,101</sup>. To avoid over-fitting do to population substructure, the effect estimates calculated using the Icelandic data were used as weights when generating the weighted PRS (PRS<sub>POPICE</sub>) for testing in the UK, and the effect estimates generated from the UK data were used to derive the weighted PRS (PRS<sub>POPUKB</sub>) for testing in Iceland. We created several PRSs assuming different fractions of causal variants (the  $P$  parameter in LDpred), and selected the best PRS based on prediction of POP in the Icelandic and UK datasets (1% causal variants). The most predictive PRS<sub>POPICE</sub> was then used to calculate the correlation with selected phenotypes in the UKB data, and the most predictive PRS<sub>POPUKB</sub> was tested for correlation with the selected phenotypes in Iceland. The correlation between the PRS and phenotypes was calculated using logistic regression in R (v3.5) (<http://www.R-project.org>) adjusting for year of birth and principle components by including them as covariates in the analysis. We

summarize the correlation of the two sets of PRS scores calculated with the Icelandic and UKB phenotypes as weighted average of the effect estimates from both analyses.

Using precomputed LD scores for about 1.1 million variants found in European populations (downloaded from: [https://data.broadinstitute.org/alkesgroup/LDSCORE/eur\\_w\\_ld\\_chr.tar.bz2](https://data.broadinstitute.org/alkesgroup/LDSCORE/eur_w_ld_chr.tar.bz2)), we estimated SNP heritability with LD score regression<sup>33</sup>.

Chromatin interaction map data were derived from Hi-C sequencing for selected cell- and tissue types<sup>79</sup>. The data were downloaded from Omnibus, accession number GSE87112, in pre-processed format (Fit-Hi-C algorithm) representing false-discovery rates (FDR) for contact regions at 40 kb resolution. To define statistically significant contacts we used a threshold value of FDR  $< 10^{-6}$  in relevant cell-types and tissue (IMR90 fibroblasts, mesenchymal stem cells, and muscle tissue i.e. left ventricle, right ventricle, and psoas muscle)<sup>79</sup>. DNA contact regions containing the lead variant, or those containing variants in LD ( $r^2 > 0.8$ ) to the lead variant, were then identified to find interacting target genes. Using JEME (joint effect of multiple enhancers)<sup>80</sup>, we similarly looked for enhancer elements containing the lead variant (and those in LD) to find target genes using similar cell- and tissue types (IMR90 fibroblasts, mesenchymal stem cells, and muscle tissue i.e. colon, stomach, duodenum, and male skeletal muscle). The intersection of genes identified by Hi-C and JEME are then regarded as strong candidate gene targets.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The Icelandic population WGS data has been deposited at the European Variant Archive under accession code PRJEB15197. We declare that the data supporting the findings of this study are available within the article, its Supplementary Data files and upon reasonable request. A reporting summary for this Article is available as a Supplementary Information file. The UK Biobank data can be obtained upon application ([ukbiobank.ac.uk](http://ukbiobank.ac.uk)). The genome-wide association scan summary data will be made available at <http://www.decode.com/summarydata>.

## Code availability

We use the following publicly available software in conjunction with the above-described algorithms in the sequencing processing pipeline (whole-genome sequencing, association testing, RNA-sequencing mapping and analysis): BWA 0.7.10 mem, <https://github.com/lh3/bwa>; GenomeAnalysisTKLite 2.3.9, <https://github.com/broadgsa/gatk/>; Picard tools 1.117, <https://broadinstitute.github.io/picard/>; SAMtools 1.3, <http://samtools.github.io/>; Bedtools v2.25.0-76-g5e7c696z, <https://github.com/arq5x/bedtools2/>; Variant Effect Predictor, <https://github.com/Ensembl/ensembl-vep/>; BOLT-LMM, <https://data.broadinstitute.org/alkesgroup/BOLT-LMM/downloads/>; LDSC (LD Score), <https://github.com/bulik/ldsc>.

Received: 7 August 2019; Accepted: 25 February 2020;

Published online: 17 March 2020

## References

- Jelovsek, J. E., Maher, C. & Barber, M. D. Pelvic organ prolapse. *Lancet* **369**, 1027–1038 (2007).
- Digesu, G. A., Chaliha, C., Salvatore, S., Hutchings, A. & Khullar, V. The relationship of vaginal prolapse severity to symptoms and quality of life. *BJOG* **112**, 971–976 (2005).
- Swift, S. E., Tate, S. B. & Nicholas, J. Correlation of symptoms with degree of pelvic organ support in a general population of women: what is pelvic organ prolapse? *Am. J. Obstet. Gynecol.* **189**, 372–377 (2003).
- Jelovsek, J. E. & Barber, M. D. Women seeking treatment for advanced pelvic organ prolapse have decreased body image and quality of life. *Am. J. Obstet. Gynecol.* **194**, 1455–1461 (2006).
- Bump, R. C. et al. The standardization of terminology of female pelvic organ prolapse and pelvic floor dysfunction. *Am. J. Obstet. Gynecol.* **175**, 10–17 (1996).
- Hendrix, S. L. et al. Pelvic organ prolapse in the women's health initiative: gravity and gravidity. *Am. J. Obstet. Gynecol.* **186**, 1160–1166 (2002).
- Luber, K. M., Boero, S. & Choe, J. Y. The demographics of pelvic floor disorders: current observations and future projections. *Am. J. Obstet. Gynecol.* **184**, 1496–1501 (2001).
- Nygaard, I., Bradley, C. & Brandt, D. Pelvic organ prolapse in older women: prevalence and risk factors. *Obstet. Gynecol.* **104**, 489–497 (2004).
- Swift, S. et al. Pelvic Organ Support Study (POSS): the distribution, clinical definition, and epidemiologic condition of pelvic organ support defects. *Am. J. Obstet. Gynecol.* **192**, 795–806 (2005).



10. Gutman, R. E., Ford, D. E., Quiroz, L. H., Shippey, S. H. & Handa, V. L. Is there a pelvic organ prolapse threshold that predicts pelvic floor symptoms? *Am. J. Obstet. Gynecol.* **199**, 683.e681–683.e687 (2008).
11. Rortveit, G. et al. Symptomatic pelvic organ prolapse: prevalence and risk factors in a population-based, racially diverse cohort. *Obstet. Gynecol.* **109**, 1396–1403 (2007).
12. Nygaard, I. et al. Prevalence of symptomatic pelvic floor disorders in US women. *JAMA* **300**, 1311–1316 (2008).
13. Olsen, A. L., Smith, V. J., Bergstrom, J. O., Colling, J. C. & Clark, A. L. Epidemiology of surgically managed pelvic organ prolapse and urinary incontinence. *Obstet. Gynecol.* **89**, 501–506 (1997).
14. Wu, J. M., Matthews, C. A., Conover, M. M., Pate, V. & Jonsson Funk, M. Lifetime risk of stress urinary incontinence or pelvic organ prolapse surgery. *Obstet. Gynecol.* **123**, 1201–1206 (2014).
15. Kurkijärvi, K., Aaltonen, R., Gissler, M. & Mäkinen, J. Pelvic organ prolapse surgery in Finland from 1987 to 2009: a national register based study. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **214**, 71–77 (2017).
16. Keshavarz, H., Hillis, S. D., Kieke, B. A. & Marchbanks, P. A. Hysterectomy surveillance—United States 1994–1999. *MMWR Surveill. Summ.* **2002** **51**, 1–8 (2002).
17. Vergeldt, T. F., Weemhoff, M., Int’Hout, J. & Kluivers, K. B. Risk factors for pelvic organ prolapse and its recurrence: a systematic review. *Int. Urogynecol. J.* **26**, 1559–1573 (2015).
18. Samuelsson, E. C., Victor, F. T., Tibblin, G. & Svardsudd, K. F. Signs of genital prolapse in a Swedish population of women 20 to 59 years of age and possible related factors. *Am. J. Obstet. Gynecol.* **180**, 299–305 (1999).
19. Linde, S. L., van Kempen, L. C., Vierhout, M. E. & Kluivers, K. B. A systematic review of clinical studies on hereditary factors in pelvic organ prolapse. *Int. Urogynecol. J.* **23**, 1327–1336 (2012).
20. Swift, S. E., Pound, T. & Dias, J. K. Case-control study of etiologic factors in the development of severe pelvic organ prolapse. *Int. Urogynecol. J. Pelvic Floor Dysfunct.* **12**, 187–192 (2001).
21. Kerkhof, M. H., Hendriks, L. & Brolmann, H. A. Changes in connective tissue in patients with pelvic organ prolapse—a review of the current literature. *Int. Urogynecol. J. Pelvic Floor Dysfunct.* **20**, 461–474 (2009).
22. Mei, S. et al. The role of smooth muscle cells in the pathophysiology of pelvic organ prolapse. *Female Pelvic Med. Reconstr. Surg.* **19**, 254–259 (2013).
23. DeLancey, J. O. L. What’s new in the functional anatomy of pelvic organ prolapse? *Curr. Opin. Obstet. Gynecol.* **28**, 420–429 (2016).
24. Norton, P. A., Baker, J. E., Sharp, H. C. & Warenski, J. C. Genitourinary prolapse and joint hypermobility in women. *Obstet. Gynecol.* **85**, 225–228 (1995).
25. Knuuti, E., Kaupilla, S., Kotila, V., Risteli, J. & Nissi, R. Genitourinary prolapse and joint hypermobility are associated with altered type I and III collagen metabolism. *Arch. Gynecol. Obstet.* **283**, 1081–1085 (2011).
26. Ward, R. M. et al. Genetic epidemiology of pelvic organ prolapse: a systematic review. *Am. J. Obstet. Gynecol.* **211**, 326–335 (2014).
27. Lim, V. F., Khoo, J. K., Wong, V. & Moore, K. H. Recent studies of genetic dysfunction in pelvic organ prolapse: the role of collagen defects. *Aust. N. Z. J. Obstet. Gynaecol.* **54**, 198–205 (2014).
28. Altman, D., Forsman, M., Falconer, C. & Lichtenstein, P. Genetic influence on stress urinary incontinence and pelvic organ prolapse. *Eur. Urol.* **54**, 918–923 (2008).
29. Giri, A. et al. Genetic determinants of pelvic organ prolapse among African American and Hispanic women in the Women’s Health Initiative. *PLoS ONE* **10**, 141647 (2015).
30. Allen-Brady, K. et al. Identification of six loci associated with pelvic organ prolapse using genome-wide association analysis. *Obstet. Gynecol.* **118**, 1345–1353 (2011).
31. Cartwright, R. et al. Systematic review and metaanalysis of genetic association studies of urinary symptoms and prolapse in women. *Am. J. Obstet. Gynecol.* **212**, 199.e191–199.e124 (2015).
32. Sveinbjornsson, G. et al. Weighting sequence variants based on their annotation increases power of whole-genome association studies. *Nat. Genet.* **48**, 314–317 (2016).
33. Bulik-Sullivan, B. K. et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291 (2015).
34. MacArthur, J. et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.* **45**, D896–D901 (2017).
35. Kobayashi, A. & Behringer, R. R. Developmental genetics of the female reproductive tract in mammals. *Nat. Rev. Genet.* **4**, 969–980 (2003).
36. Ledig, S. & Wieacker, P. Clinical and genetic aspects of Mayer-Rokitansky-Küster-Hauser syndrome. *Medizinische Genetik* **30**, 3–11 (2018).
37. Sapkota, Y. et al. Meta-analysis identifies five novel loci associated with endometriosis highlighting key genes involved in hormone metabolism. *Nat. Commun.* **8**, 15539 (2017).
38. Rafnar, T. et al. Variants associating with uterine leiomyoma highlight genetic background shared by various cancers and hormone-related traits. *Nat. Commun.* **9**, 3636 (2018).
39. Zhang, G. et al. Genetic associations with gestational duration and spontaneous preterm birth. *N. Engl. J. Med.* **377**, 1156–1167 (2017).
40. Medina-Gomez, C. et al. Life-course genome-wide association study meta-analysis of total body BMD and assessment of age-specific effects. *Am. J. Hum. Genet.* **102**, 88–102 (2018).
41. Kohlhase, J., Wischermann, A., Reichenbach, H., Froster, U. & Engel, W. Mutations in the SALL1 putative transcription factor gene cause Townes-Brocks syndrome. *Nat. Genet.* **18**, 81–83 (1998).
42. Miner, J. H. in *Cell-Extracellular Matrix Interactions in Cancer* (eds Roy Zent & Ambra Pozzi) 1–17 (Springer, New York, 2010).
43. Jorgenson, E. et al. A genome-wide association study identifies four novel susceptibility loci underlying inguinal hernia. *Nat. Commun.* **6**, 10130 (2015).
44. Stone, E. M. et al. A single EFEMP1 mutation associated with both Malattia Leventinese and Doyme honeycomb retinal dystrophy. *Nat. Genet.* **22**, 199–202 (1999).
45. Lango Allen, H. et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832–838 (2010).
46. Wood, A. R. et al. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* **46**, 1173–1186 (2014).
47. Justice, A. E. et al. Genome-wide meta-analysis of 241,258 adults accounting for smoking behaviour identifies novel loci for obesity traits. *Nat. Commun.* **8**, 14977 (2017).
48. Shungin, D. et al. New genetic loci link adipose and insulin biology to body fat distribution. *Nature* **518**, 187–196 (2015).
49. Bendavid, R. The unified theory of hernia formation. *Hernia* **8**, 171–176 (2004).
50. Segev, Y. et al. Are women with pelvic organ prolapse at a higher risk of developing hernias? *Int. Urogynecol. J. Pelvic Floor Dysfunct.* **20**, 1451–1453 (2009).
51. Salnikova, L. E., Khadzhieva, M. B. & Kolobkov, D. S. Biological findings from the PheWAS catalog: focus on connective tissue-related disorders (pelvic floor dysfunction, abdominal hernia, varicose veins and hemorrhoids). *Hum. Genet.* **135**, 779–795 (2016).
52. Miedel, A., Tegerstedt, G., Maehle-Schmidt, M., Nyren, O. & Hammarstrom, M. Nonobstetric risk factors for symptomatic pelvic organ prolapse. *Obstet. Gynecol.* **113**, 1089–1097 (2009).
53. Hobson, K. G. & Roberts, P. L. Etiology and pathophysiology of diverticular disease. *Clin. Colon Rectal Surg.* **17**, 147–153 (2004).
54. Aboonq, M. S. Pathophysiology of carpal tunnel syndrome. *Neurosciences* **20**, 4–9 (2015).
55. Wolfman, N. M. et al. Ectopic induction of tendon and ligament in rats by growth and differentiation factors 5, 6, and 7, members of the TGF-beta gene family. *J. Clin. Investig.* **100**, 321–330 (1997).
56. Xu, K. et al. Synergistic promoting effects of bone morphogenetic protein 12/ connective tissue growth factor on functional differentiation of tendon derived stem cells and patellar tendon window defect regeneration. *J. Biomech.* **66**, 95–102 (2018).
57. Hou, Y. et al. Tenomodulin highly expressing MSCs as a better cell source for tendon injury healing. *Oncotarget* **8**, 77424–77435 (2017).
58. Warren, H. R. et al. Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nat. Genet.* **49**, 403–415 (2017).
59. Schumacher, F. R. et al. Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nat. Genet.* **50**, 928–936 (2018).
60. Eeles, R. A. et al. Identification of 23 new prostate cancer susceptibility loci using the iCOGS custom genotyping array. *Nat. Genet.* **45**, 385–391 (2013).
61. Gudmundsson, J. et al. Genome-wide associations for benign prostatic hyperplasia reveal a genetic correlation with serum levels of PSA. *Nat. Commun.* **9**, 4568 (2018).
62. Saburi, S. et al. Loss of Fat4 disrupts PCP signaling and oriented cell division and leads to cystic kidney disease. *Nat. Genet.* **40**, 1010 (2008).
63. Wu, T. Y. et al. Pharmacogenetics of the mycophenolic acid targets inosine monophosphate dehydrogenases IMPDH1 and IMPDH2: gene sequence variation and functional genomics. *Br. J. Pharmacol.* **161**, 1584–1598 (2010).
64. Bowne, S. J. et al. Mutations in the inosine monophosphate dehydrogenase 1 gene (IMPDH1) cause the RP10 form of autosomal dominant retinitis pigmentosa. *Hum. Mol. Genet.* **11**, 559–568 (2002).
65. Bowne, S. J. et al. Spectrum and frequency of mutations in IMPDH1 associated with autosomal dominant retinitis pigmentosa and leber congenital amaurosis. *Invest. Ophthalmol. Vis. Sci.* **47**, 34–42 (2006).
66. Välimäki, N. et al. Genetic predisposition to uterine leiomyoma is determined by loci for genitourinary development and genome stability. *eLife* **7**, e37110 (2018).

67. Rahmioglu, N. et al. Large-scale genome-wide association meta-analysis of endometriosis reveals 13 novel loci and genetically-associated comorbidity with other pain conditions. *bioRxiv*, 406967, (2018).
68. Gharahkhani, P. et al. Genome-wide association studies in oesophageal adenocarcinoma and Barrett's oesophagus: a large-scale meta-analysis. *Lancet Oncol.* **17**, 1363–1373 (2016).
69. Yengo, L. et al. Meta-analysis of genome-wide association studies for height and body mass index in approximately 700,000 individuals of European ancestry. *Hum. Mol. Genet.* **27**, 3641–3649 (2018).
70. Wain, L. V. et al. Genome-wide association analyses for lung function and chronic obstructive pulmonary disease identify new loci and potential druggable targets. *Nat. Genet.* **49**, 416–425 (2017).
71. Smith, G. D. & Ebrahim, S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.* **32**, 1–22 (2003).
72. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **44**, 512–525 (2015).
73. Locke, A. E. et al. Genetic studies of body mass index yield new insights into obesity biology. *Nature* **518**, 197–206 (2015).
74. Turcot, V. et al. Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity (vol 50, pg 26, 2018). *Nat. Genet.* **50**, 766–767 (2018).
75. Barban, N. et al. Genome-wide analysis identifies 12 loci influencing human reproductive behavior. *Nat. Genet.* **48**, 1462 (2016).
76. Luong, H. T. et al. Fine mapping of variants associated with endometriosis in the WNT4 region on chromosome 1p36. *Int. J. Mol. Epidemiol. Genet.* **4**, 193–206 (2013).
77. Yevshin, I., Sharipov, R., Valeev, T., Kel, A. & Kolpakov, F. GTRD: a database of transcription factor binding sites identified by ChIP-seq experiments. *Nucleic Acids Res.* **45**, D61–D67 (2017).
78. Powell, J. E. et al. Endometriosis risk alleles at 1p36.12 act through inverse regulation of CDC42 and LINC00339. *Hum. Mol. Genet.* **25**, 5046–5058 (2016).
79. Schmitt, A. D. et al. A compendium of chromatin contact maps reveals spatially active regions in the human. *Genome Cell Rep.* **17**, 2042–2059 (2016).
80. Cao, Q. et al. Reconstruction of enhancer–target networks in 935 samples of human primary cells, tissues and cell lines. *Nat. Genet.* **49**, 1428 (2017).
81. Zhou, Q., Hong, L. & Wang, J. Identification of key genes and pathways in pelvic organ prolapse based on gene expression profiling by bioinformatics analysis. *Arch. Gynecol. Obstet.* **297**, 1323–1332 (2018).
82. Giltay, R., Timpl, R. & Kostka, G. Sequence, recombinant expression and tissue localization of two novel extracellular matrix proteins, fibulin-3 and fibulin-4. *Matrix Biol.* **18**, 469–480 (1999).
83. Kobayashi, N. et al. A comparative analysis of the fibulin protein family. Biochemical characterization, binding interactions, and tissue localization. *J. Biol. Chem.* **282**, 11805–11816 (2007).
84. McLaughlin, P. J. et al. Lack of fibulin-3 causes early aging and herniation, but not macular degeneration in mice. *Hum. Mol. Genet.* **16**, 3059–3070 (2007).
85. Rahn, D. D. et al. Failure of pelvic organ support in mice deficient in fibulin-3. *Am. J. Pathol.* **174**, 206–215 (2009).
86. Bycroft, C. et al. Genome-wide genetic data on ~500,000 UK Biobank participants. *bioRxiv* <https://doi.org/10.1101/166298> (2017).
87. Sudlow, C. et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
88. Sigurdsson, S. et al. Sequence variants in ARHGAP15, COLQ and FAM155A associate with diverticular disease and diverticulitis. *Nat. Commun.* **8**, 15789 (2017).
89. Gulcher, J. R., Kristjansson, K., Gudbjartsson, H. & Stefansson, K. Protection of privacy by third-party encryption in genetic research in Iceland. *Eur. J. Hum. Genet.* **8**, 739–742 (2000).
90. Gudbjartsson, D. F. et al. Large-scale whole-genome sequencing of the Icelandic population. *Nat. Genet.* **47**, 435 (2015).
91. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
92. Kong, A. et al. Detection of sharing by descent, long-range phasing and haplotype imputation. *Nat. Genet.* **40**, 1068–1075 (2008).
93. McLaren, W. et al. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* **26**, 2069–2070 (2010).
94. Wain, L. V. et al. Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. *Lancet Respir. Med.* **3**, 769–781 (2015).
95. Welsh, S., Peakman, T., Sheard, S. & Almond, R. Comparison of DNA quantification methodology used in the DNA extraction protocol for the UK Biobank cohort. *BMC Genomics* **18**, 26 (2017).
96. Mantel, N. & Haenszel, W. Statistical aspects of the analysis of data from retrospective studies of disease. *J. Natl. Cancer Inst.* **22**, 719–748 (1959).
97. Yang, J. et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375 (2012).
98. Oskarsson, G. R. et al. A truncating mutation in EPOR leads to hyporesponsiveness to erythropoietin with normal haemoglobin. *Commun. Biol.* **1**, 49 (2018).
99. Styrkarsdottir, U. et al. Meta-analysis of Icelandic and UK data sets identifies missense variants in SMO, IL11, COL11A1 and 13 more new loci associated with osteoarthritis. *Nat. Genet.* **50**, 1681–1687 (2018).
100. Kong, A. et al. Selection against variants in the genome associated with educational attainment. *Proc. Natl. Acad. Sci. USA* **114**, E727–E732 (2017).
101. Vilhjalmsdottir, B. J. et al. Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *Am. J. Hum. Genet.* **97**, 576–592 (2015).
102. Turner, S. D. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *bioRxiv* <https://doi.org/10.1101/005165> (2014).
103. Amin Al Olama, A. et al. Multiple novel prostate cancer susceptibility signals identified by fine-mapping of known risk loci among Europeans. *Hum. Mol. Genet.* **24**, 5589–5602 (2015).

### Acknowledgements

This research has been conducted using the UK Biobank Resource under Application Number 24898 and 24711. We thank the individuals who participated in the study and whose contribution made this work possible. Folkert Asselbergs is supported by UCL Hospitals NIHR Biomedical Research Centre.

### Author contributions

T.O., T.R., G.T., D.F.G., P.S., U.T., and K.S. designed the study and interpreted the results. H.M., K.O., O.I., V.G., and K.J. carried out the subject ascertainment, recruitment, and collection of clinical data. U.S., V.S., J.G., I.J., H.H., V.T., F.W.A., A.O., A.S., and S.L. collected, processed, and analyzed the genotype and phenotype data. T.O., G.T., L.J., J.K.S., A.O., O.A.S., R.P.K., M.L.F., S.L., B.V.H., G.H.H., H.P.E., P.M., P.S., and D.F.G. performed the statistical and bioinformatics analyses. T.O., T.R., D.F.G., G.T., P.S., U.T., and K.S. drafted the manuscript. All authors contributed to the final version of the paper.

### Competing interests

The authors that are affiliated with deCODE are employees of deCODE genetics/Amgen. The remaining authors declare no competing interests.

### Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s42003-020-0857-9>.

Correspondence and requests for materials should be addressed to T.O. or K.S.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020