A genome-wide association study identifies four novel susceptibility loci underlying inguinal hernia

Eric Jorgenson1,*, Nadja Makki2,3,*, Ling Shen1, David C. Chen4, Chao Tian5, Walter L. Eckalbar2,3, David Hinds5, Nadav Ahituv2,3 & Andrew Avins1

Inguinal hernia repair is one of the most commonly performed operations in the world, yet little is known about the genetic mechanisms that predispose individuals to develop inguinal hernias. We perform a genome-wide association analysis of surgically confirmed inguinal hernias in 72,805 subjects (5,295 cases and 67,510 controls) and confirm top associations in an independent cohort of 92,444 subjects with self-reported hernia repair surgeries (9,701 cases and 82,743 controls). We identify four novel inguinal hernia susceptibility loci in the regions of EFEMP1, WT1, EBF2 and ADAMTS6. Moreover, we observe expression of all four genes in mouse connective tissue and network analyses show an important role for two of these genes (EFEMP1 and WT1) in connective tissue maintenance/homoeostasis. Our findings provide insight into the aetiology of hernia development and highlight genetic pathways for studies of hernia development and its treatment.
inguinal hernias are amongst the most frequently diagnosed conditions in clinical practice and have a lifetime prevalence in the range of 20–27% in men and 3–6% in women. They can be classified as either direct, which occur through an acquired weakness in the transversalis fascia, connective tissue that comprises the floor of the inguinal canal, or indirect, in which abdominal contents protrude through a congenital defect in the inguinal ring via a patent processus vaginalis. Inguinal hernia repair is one of the most common surgical procedures, with more than 750,000 performed annually in the United States, and is associated with substantial costs. Inguinal hernias can lead to serious medical morbidity such as bowel incarceration and strangulation, and emergency hernia surgery to treat these conditions is associated with a substantial mortality risk. A subset of patients experience hernia recurrence after surgery and chronic pain affects over 6% of patients, highlighting the need for a better understanding of hernia aetiology, which could, in turn, lead to new approaches to therapy and improved treatment outcomes.

Several risk factors underlying the development of inguinal hernia in adults have been identified, including male sex, older age, chronic obstructive pulmonary disease, lower body mass index and family history. The risk of inguinal hernia is increased among first-degree relatives of individuals with a history of inguinal hernia, suggesting that there likely exist identifiable genetic risk factors responsible for many inguinal hernias. In addition, individuals with certain genetic syndromes, including cutis laxa, Marfan syndrome and Ehlers-Danlos syndrome, have a greater risk of developing inguinal hernias. To date, only a small number of candidate genes have been investigated. As a result, little is currently known about the specific genes that play a role in the pathophysiology of inguinal hernia.

To address this question, we conduct the first large-scale genome-wide association study (GWAS) of surgically confirmed inguinal hernia. We utilize information from participants in the Genetic Epidemiology Research in Adult Health and Aging (GERA) cohort, nested in the Kaiser Permanente integrated health plan in Northern California (KPNC). We confirm top associations in a large independent sample of research participants with self-reported information on history of hernia repair surgery from 23andMe. We then examine patterns of expression of genes in the associated regions in mouse connective tissue equivalent to human transversalis fascia and find that all four genes are expressed in this tissue, supporting their role in hernia development.

**Results**

Using information extracted from KPNC electronic health records (EHR), we identified hernia cases and controls among non-Hispanic white GERA participants and validated a subset of cases through chart review. In total, we identified 5,295 surgically confirmed inguinal hernia cases with male predominance (90.2%) and 67,510 controls with no known surgical or medical history of inguinal or other abdominal hernia in the GERA discovery cohort (Supplementary Table S1). Hernia repair discharge procedure codes indicated that 2,335 inguinal hernia cases had direct inguinal hernia repairs and 2,647 had indirect inguinal hernia repairs. We reviewed 230 patient charts to validate the accuracy of inguinal hernia diagnoses, and, of those, 228 (99.1%, 95% confidence interval (CI): 96.9–99.9%) were confirmed to be designated correctly as inguinal hernias of any type. Of the 118 charts reviewed specifically for accuracy of the diagnosis of direct inguinal hernia, 113 (95.7%, 95% CI: 90.3–98.6%) were found to be correctly identified. For indirect inguinal hernias, 110 of 112 chart diagnoses were found to be supported by the clinical data (98.2%, 95% CI: 93.7–99.8%). Thus, the positive predictive value of our algorithm for identifying hernia cases, as well as hernia type, was very high in this sample.

**Genetic association analysis of inguinal hernia.** We conducted a sex-stratified GWAS analysis of inguinal hernia in the GERA cohort, adjusting for age and the first 10 ancestry principal components. The genomic control λ values were 1.022 for the analysis of men and 1.021 for the analysis of women. We identified four loci that exceeded genome-wide significance ($P < 5 	imes 10^{-5}$) in the regions of EFEMP1 (rs2009262, odds ratio (OR) = 1.23, $P = 3.66 	imes 10^{-15}$), WT1 (rs3809060, OR = 1.18, $P = 4.69 	imes 10^{-14}$), EBF2 (rs6991952, OR = 1.14, $P = 1.17 	imes 10^{-10}$) and ADAMTS6 (rs370763, OR = 1.14, $P = 9.70 	imes 10^{-9}$) in the discovery cohort (Fig. 1, Table 1 and Supplementary Fig. 1). We confirmed these associations in 9,701 cases and 82,743 controls who were research participants from the 23andMe cohort with self-reported information on history of hernia repair surgery (Supplementary Table S2). In this replication cohort, we observed significant associations of all four single-nucleotide polymorphisms (SNPs; rs2009262, OR = 1.10, $P = 3.65 	imes 10^{-6}$, rs3809060, OR = 1.07, $P = 1.69 	imes 10^{-4}$, rs6991952, OR = 1.08, $P = 2.04 	imes 10^{-6}$ and rs370763, OR = 1.06, $P = 3.02 	imes 10^{-3}$, Table 1).

Each of these inguinal hernia risk genes has a plausible biological and pathophysiologic role in the development of hernias, which is known to have metabolic aetiology related to collagen subtype and maturation, elastin and matrix metalloproteinases, in addition to congenital and acquired factors. EFEMP1 knockout mice develop both direct and indirect inguinal hernias, have reduced elastic fibres in fascia and display signs of early aging. Nonsynonymous variants in WT1 have been identified in patients with Denys–Drash syndrome and Mechem syndrome with congenital diaphragmatic hernia. An antisense morpholino knockdown study of EBF2 resulted in defects in muscle development in Xenopus. ADAMTS6 is a member of a gene family that encode proteases that convert procollagen to collagen. Mutations in the gene family member ADAMTS2 have been associated with Ehlers–Danlos syndrome with congenital umbilical hernia.

To determine whether there were additional inguinal hernia risk alleles in the four inguinal hernia susceptibility loci, we repeated the GWA analysis in the GERA sample conditioning on the top associated SNPs at each of the four loci. We did not observe any other SNPs that were significantly associated with inguinal hernia in the conditional analysis. We then estimated the point prevalence of surgically confirmed inguinal hernia among non-Hispanic white KPNC members who were at least 50 years of age as of June 2013, which was 9.2% in men and 0.3% in women. These estimates are consistent with the lifetime prevalence of inguinal hernias previously reported in the literature, 27% for men, 6% for women, but lower due to the more stringent case definition and shorter observation time. Using both the point and lifetime prevalence estimates to provide a range, the four top SNPs explained 1.0–1.4% of the variation in the risk of inguinal hernia in men and 1.3–2.8% in women in our discovery sample. The narrow-sense heritability explained by common SNPs (minor allele frequency > 5%) ranged from 13.2 to 18.3% in men and 20.8 to 25.5% in women, suggesting that additional inguinal hernia susceptibility loci remain to be discovered.

**Direct and indirect inguinal hernia.** Inguinal hernias can be classified as direct, in which the abdominal contents herniate through the floor of the inguinal canal due to an acquired...
Four novel inguinal hernia risk loci with genome-wide significant associations were identified in the regions of EFEMP1 (chromosome 2), ADAMTS6 (chromosome 5), EBF2 (chromosome 8) and WT1 (chromosome 11). The dotted red line represents a significance threshold of $P = 5.0 \times 10^{-8}$.

### Table 1 | SNP associations reaching genome-wide significance in the combined analysis of discovery and replication cohorts.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr.</th>
<th>Position</th>
<th>Gene</th>
<th>Risk allele</th>
<th>Discovery (5,295 cases, 67,510 controls)</th>
<th>Replication (9,701 cases, 82,743 controls)</th>
<th>Combined OR (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2009262</td>
<td>2</td>
<td>56,012,214</td>
<td>EFEMP1</td>
<td>T</td>
<td>0.78 (1.17-1.30)</td>
<td>0.78 (1.06-1.15)</td>
<td>1.15 (1.11-1.19)</td>
<td>1.45 x 10^{-17}</td>
</tr>
<tr>
<td>rs370763</td>
<td>5</td>
<td>64,355,060</td>
<td>ADAMTS6</td>
<td>A</td>
<td>0.65 (1.09-1.19)</td>
<td>0.67 (1.02-1.09)</td>
<td>1.09 (1.06-1.12)</td>
<td>3.73 x 10^{-9}</td>
</tr>
<tr>
<td>rs6999952</td>
<td>8</td>
<td>25,707,412</td>
<td>EBF2</td>
<td>G</td>
<td>0.43 (1.10-1.19)</td>
<td>0.43 (1.05-1.12)</td>
<td>1.11 (1.08-1.14)</td>
<td>6.68 x 10^{-15}</td>
</tr>
<tr>
<td>rs3809060</td>
<td>11</td>
<td>32,458,807</td>
<td>WT1</td>
<td>G</td>
<td>0.62 (1.15-1.23)</td>
<td>0.63 (1.03-1.10)</td>
<td>1.19 (1.08-1.14)</td>
<td>3.69 x 10^{-14}</td>
</tr>
</tbody>
</table>

Chr., chromosome; CI, confidence interval; RAF, risk allele frequency; SNP, single-nucleotide polymorphism.

weakness in the transversalis fascia, or indirect, in which abdominal contents protrude through a congenital defect in the inguinal ring via enlargement of a patent processus vaginalis. We analysed the four inguinal hernia risk SNPs in GERA subjects with direct and indirect hernias separately to determine whether any of them predisposed subjects to a specific subtype of inguinal hernia. The ORs observed for direct inguinal hernia were slightly stronger for three of the four top SNPs in men than for indirect hernia (Table 2). In women, for whom there were fewer subjects with inguinal hernias (N = 549), only rs2009262 and rs3809060 were nominally associated with direct or indirect inguinal hernia ($P<0.05$), and both displayed larger effects for indirect compared with direct inguinal hernia.

We then examined the association of SNPs in the region (± 250 kb) of the four hernia susceptibility loci with direct and indirect hernia in men. The four top SNPs associated with inguinal hernia were also the most strongly associated SNPs with indirect inguinal hernia, but for three of the four loci, other SNPs in the region were more strongly associated with direct inguinal hernia, specifically rs11899888 (instead of rs2009262) in EFEMP1, rs12520760 (instead of rs370763) in ADAMTS6 and rs10946560 (instead of rs6991952) in EBF2 (Supplementary Fig. 2). This indicates that multiple variants within these risk loci may underlie the different subtypes of inguinal hernia.

To determine whether specific biological pathways or functions play a role in inguinal hernia development, we conducted a gene set enrichment analysis of our discovery cohort results using the program Meta-Analysis Gene-set Enrichment of variant Associations (MAGENTA)29. We identified four gene sets at a false discovery rate (FDR) $<0.05$: Jak Stat signalling, leukocyte extravasation signalling, actin cytoskeleton signalling and glycosaminoglycan biosynthesis chondroitin sulfate (Supplementary Table 3). We then used RegulomeDB to investigate the potential for SNPs in the identified inguinal hernia risk loci to influence the binding of transcription factors30. We identified 14 SNPs in the four regions that were classified as likely to affect transcription factor binding (Supplementary Table 4).

### Expression of inguinal hernia risk genes.

Using quantitative real-time PCR (qRT-PCR) and RNA sequencing (RNA-seq), we examined mRNA levels of the four genes in mouse connective tissue equivalent to human transversalis fascia (see Methods section). qRT-PCR found Efp1 to be expressed at a high level, Wt1 at a moderate level and Ebf2 and Adamts6 at low levels compared with a control connective tissue expressed gene (Col12a1; Fig. 2a). Our RNA-seq analysis showed comparable fragments per kilobase per million reads (FPKM) values, with all four genes correlating well with the relative expression levels determined by qRT-PCR (Fig. 2b). Combined, our results show that all four genes are expressed in connective tissue and could have a functional role in this tissue.
We next set out to characterize the gene regulatory networks associated with these genes. We carried out Causal Network Analysis on the highest expressing genes from our RNA-seq list (see Methods section) using the Ingenuity Pathway Analysis software (IPA, Qiagen). Since Ebf2 and Adamts6 were expressed at low levels, we only characterized interactions for Efemp1 and Wt1. We identified many interesting interactors for EFEMP1 including ELASTIN, a component of elastic fibres and COLLAGEN15A1, a component of collagen fibres (Fig. 3). The WTI network contained many extracellular matrix (ECM) proteins. These included MMP2 (matrix metalloproteinase-2), CTGF (connective tissue growth factor) and THBS1 (thrombospondin-1), all proteins known to play a role in connective tissue remodelling and homoeostasis. One common protein of interest between the two networks is TIMP3 (tissue inhibitor of metalloproteinase-3), which inhibits matrix metalloproteinases that degrade collagen and elastin. TIMP3 interacts with EFEMP1 and is thought to be activated by WT1. Changes in the expression levels of TIMP3 could shift the ratio of type I to type III collagen. The alteration of this ratio appears to be driven by greater expression of type III collagen mRNA in patients with inguinal hernias compared with controls. In addition, an imbalance in the activity of collagen degrading matrix metalloproteinases and their inhibitors (MMPs and TIMPs) has been reported in fibroblasts of patients with inguinal hernias. WT1 has been shown to inhibit MMP2 (ref. 36) and activate TIMP3 (ref. 37), which in turn inhibits MMPs. EFEMP1 interacts with TIMP3 and might thus augment the inhibitory role of WT1 on MMPs. In addition, ADAMTS family members are matrix metalloproteinases that convert procollagen to collagen. The association of genetic variants near ADAMTS6 supports the hypothesis that collagen dysregulation can influence the development of inguinal hernias. A GWAS of central corneal thickness (CCT) also identified the ADAMTS6 locus, along with an association with the

<table>
<thead>
<tr>
<th>SNP</th>
<th>Inguinal Hernia Type</th>
<th>Men (OR (95% CI), P-value)</th>
<th>Women (OR (95% CI), P-value)</th>
<th>Combined (OR, P&lt;sub&gt;R&lt;/sub&gt;, OR&lt;sub&gt;R&lt;/sub&gt;, P&lt;sub&gt;F&lt;/sub&gt;, I²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2009262</td>
<td>Direct</td>
<td>1.25 (1.16–1.36), 0.03</td>
<td>1.26 (1.17–1.36), 0.280</td>
<td>1.0 (95% CI), 0.0 (P&lt;sub&gt;R&lt;/sub&gt;)</td>
</tr>
<tr>
<td>rs2009262</td>
<td>Indirect</td>
<td>1.12 (1.13–1.31), 0.001</td>
<td>1.13 (1.14–1.29), 0.009</td>
<td>0.0 (95% CI), 0.0 (P&lt;sub&gt;R&lt;/sub&gt;)</td>
</tr>
<tr>
<td>rs370763</td>
<td>Direct</td>
<td>1.21 (1.06–1.22), 0.033</td>
<td>1.14 (1.07–1.22), 0.332</td>
<td>0.0 (95% CI), 0.0 (P&lt;sub&gt;R&lt;/sub&gt;)</td>
</tr>
<tr>
<td>rs6991952</td>
<td>Direct</td>
<td>1.15 (1.08–1.22), 0.110</td>
<td>1.14 (1.10–1.22), 0.228</td>
<td>0.0 (95% CI), 0.0 (P&lt;sub&gt;R&lt;/sub&gt;)</td>
</tr>
<tr>
<td>rs3809060</td>
<td>Direct</td>
<td>1.14 (1.08–1.21), 0.07</td>
<td>1.14 (1.07–1.20), 0.09</td>
<td>0.0 (95% CI), 0.0 (P&lt;sub&gt;R&lt;/sub&gt;)</td>
</tr>
</tbody>
</table>

**Discussion**

We identified four novel inguinal hernia genetic susceptibility loci near the genes WTI, EFEMP1, EBF2 and ADAMTS6, and confirmed those associations in an independent cohort. All four loci appear to be associated with both direct and indirect inguinal hernias. Each of these four genes is expressed in mouse connective tissue, with the expression of EFEMP1 being particularly high. Our IPA analysis suggests that WT1 and EFEMP1 might play a role in connective tissue maintenance/homoeostasis through their action on ECM enzymes including matrix metalloproteinases that degrade collagen and elastin fibres.

Dysregulation of collagen homoeostasis is thought to play an important role in the development of inguinal hernias. Collagen is the main structural protein of the abdominal fascia, and undergoes a continuous process of synthesis and degradation. Transversalis fascia samples from patients with indirect inguinal hernias were found to have lower levels of collagen compared with cadaver controls and showed a decreased ratio of type I to type III collagen. The alteration of this ratio appears to be driven by greater expression of type III collagen mRNA in patients with inguinal hernias compared with controls. In addition, an imbalance in the activity of collagen degrading matrix metalloproteinases and their inhibitors (MMPs and TIMPs) has been reported in fibroblasts of patients with inguinal hernias. WT1 has been shown to inhibit MMP2 (ref. 36) and activate TIMP3 (ref. 37), which in turn inhibits MMPs. EFEMP1 interacts with TIMP3 and might thus augment the inhibitory role of WT1 on MMPs. In addition, ADAMTS family members are matrix metalloproteinases that convert procollagen to collagen. The association of genetic variants near ADAMTS6 supports the hypothesis that collagen dysregulation can influence the development of inguinal hernias. A GWAS of central corneal thickness (CCT) also identified the ADAMTS6 locus, along with an association with the...
collagen gene COL5A1 (ref. 39), suggesting that ADAMTS6 may influence collagen homoeostasis in multiple tissues and disorders. Elastin is also a key component of transversalis fascia that complements the role of collagen by providing elasticity, which allows for the tissue to stretch and return to its original form. Mutations in the human elastin gene, ELN, cause cutis laxa40, which has been associated with an increased risk of inguinal hernias14 and supravalvular aortic stenosis41. In connective tissue, the integration of elastin to the microfibril scaffold is guided by fibulins42; EFEMP1 is a member of the fibulin gene family, and the EFEMP1 protein binds tropoelastin, the building block of the elastin protein43. EFEMP1 knockout mice have reduced elastic fibres in fascia and develop direct and indirect inguinal hernias22. Variants in the EFEMP1 locus have also been associated with a number of conditions and functional changes, including differences in forced vital capacity, a measure of lung function44. This shared association suggests that alterations in elastin maintenance may contribute to the development of both chronic obstructive pulmonary disease and inguinal hernia and may be the mechanism through which chronic obstructive pulmonary disease increases the risk of inguinal hernias. These alterations in elastin and connective tissues may act more generally to affect the risk of disorders of other elastic tissues, such as abdominal aortic aneurysm, for which inguinal hernia patients are at an increased risk32,45.

While this is the first study to identify inguinal hernia susceptibility loci, previous GWASs have identified these regions as influencing a number of human phenotypes, supporting a functional role for variation in inguinal hernia loci in human traits and diseases. WT1, so named for causing Wilm’s tumour46, has also been associated with tuberculosis 47. Variants in the EFEMP1 locus have been associated with height48 and forced vital capacity44, and its epigenetic silencing has been associated with multiple cancer types49,50. EBF2 has been associated with prostate cancer, though the variants identified were located proximal to those identified here51. SNPs in the ADAMTS6 region are associated with differences in CCT, an anthropomorphic measure of the eye, but not conditions associated with CCT, including keratoconus or primary open-angle glaucoma39. A second study also found suggestive evidence for association of this locus with osteosarcoma52. The pleiotropic effect of the loci identified in this study suggests a potential shared aetiology between inguinal hernia risk and cancer, lung function and anthropomorphic traits. Given previous observational associations between inguinal hernia...
hernia risk and body mass index and other connective tissue disorders, examining potential shared effects of genetic variation underlying these disorders may provide additional insight into hernia development.

Although these lines of evidence provide support for the role of these four genes in hernia development, further experiments are needed to demonstrate a causal role for these genes and specific SNPs in the gene regions. These experiments include examining epigenetic features by performing ChIP-seq (chromatin immuno-precipitation followed by deep sequencing) on fascia connective tissue and identifying SNPs that reside in putative gene regulatory regions that are also in linkage disequilibrium with SNPs associated with inguinal hernia risk. Complementary to this, differential enhancer assays can be carried out in human fibroblast cell lines to compare enhancer activity of the reference allele and the potential risk allele. Genome editing techniques, such as CRISPR/Cas9, can also be used to delete the regulatory region or to replace the reference allele with the risk allele, allowing for a more complete understanding of mechanisms through which the risk alleles act to influence the development of inguinal hernias.
Study Accession: phs000674.v1.p1). Since the principal component analysis was computationally intensive, it was run on a large set of individuals (N = 20,000) with the remaining individuals projected into the same space. These principal components were used in the GWAS to adjust for genetic ancestry.

23andMe. DNA extraction and genotyping were performed on saliva samples by National Genetics Institute, a CLIA-licensed clinical laboratory and a subsidiary of Laboratory Corporation of America. Samples were genotyped on four genotyping platforms. The V1 and V2 platforms were based on the Illumina HumanHap550 + BeadChip, including about 25,000 custom SNPs selected by 23andMe, with a total of about 560,000 SNPs. The V3 platform was based on the Illumina OmniExpress + BeadChip, with custom content to improve the overlap with the V2 array, with a total of about 950,000 SNPs. The V4 platform in current use is a fully custom array, including a lower redundancy subset of V2 and V3 SNPs with additional coverage of lower-frequency coding variation, and about 570,000 SNPs. Samples that failed to reach 98.5% call rate were reanalyzed.

Individuals whose analyses failed repeatedly were re-contacted by 23andMe customer service and offered additional assistance.

The subjects to be analysed were restricted to a set of individuals who have >97% European ancestry, as determined through an analysis of local ancestry41. Briefly, the algorithm first partitions phased genomic data into short windows of about 100 SNPs. Within each window, a support vector machine is used to classify individual haplotypes into one of 31 reference populations. The support vector machine classifications are then fed into a hidden Markov model that accounts for switch errors and incorrect assignments, and gives probabilities for each reference population in each window. Finally, simulated admixed individuals are used to recalculate the hidden Markov model probabilities so that the reported assignments are consistent with admixed admixture proportions.

For each genotyping platform was phased and imputed separately. Phasing was conducted using a phasing tool, Finch, developed at 23andMe, which implements the Beagle haplotype graph-based phasing algorithm42, modified to separate the haplotypes into separate haplotype blocks. Finch and phasing were used to accommodate genotyping error and recombination, to handle cases where there are no consistent paths through the haplotype graph for the individual being phased. Haplotype graphs for European and non-European samples were constructed on each 23andMe genotyping platform from a representative sample of genotyped individuals, and then performed out-of-sample phasing of all genotyped individuals against the appropriate graph.

For preparation for imputation, phased chromosomes were split into segments of no more than 10,000 genotyped SNPs, with overlaps of 200 SNPs. SNPs with Hardy–Weinberg equilibrium (HWE) P < 10^{-10}, call rate < 95%, or with large allele frequencies compared with Europeans were excluded from the analysis. SNPs with a minor allele frequency of 1% or less were excluded. Frequency discrepancies were identified by computing a 2 x 2 table of allele counts for European 100 Genomes samples and 2,000 randomly sampled 23andMe customers with European ancestry, and identifying SNPs with a P < 10^{-5}. Each phased segment was imputed against all-ethnicity 1000 Genomes haplotypes (excluding monomorphic and singleton sites) using Minimap2 (ref. 65) five using five rounds and 200 states for parameter estimation. The four SNPs reported here had high imputation r^2 values (rs2009262: 0.991; rs3809060: 0.976; rs6991952: 0.999; and rs3707636: 0.991)

**Statistical analysis.** GWA analysis. Analyses in the discovery cohort were conducted using PLINK v1.07 (http://pngu.mgh.harvard.edu/~purcell/plink) and R (http://www.r-project.org). We tested single-marker associations for men and women separately in a logistic regression model adjusted for age and the first 10 ancestry principal components using allele counts for typed SNPs and imputed dosages for the imputed SNPs and a log-additive genetic model. We then conducted meta-analyses across all SNPs. In the results section, we present top soruths that exceeded genome-wide significance (P < 5 x 10^{-8}) at novel loci. We examined the top associations by inspecting the cluster plots, call rates and HWE P values of the genotyped SNPs. To detect Hardy–Weinberg deviation due to genotyping error rather than population stratification, the HWE P values were calculated based on a subset of the homogeneous non-Hispanic white samples within the interquartile ranges of the first two principal components. The genomic control parameter λ was calculated for each analysis to assess inflation due to population stratification. To identify independent signals, we tested the genome-wide SNP associations with inguinal hernia by conditioning on the top SNPs from each of the four independent loci that exceeded genome-wide significance in the replication analyses. Results were conducted in men and women separately with covariates for age and the top five principal components. Results were combined in a fixed effects meta-analysis.

**References**


Acknowledgements

Development of the GERA cohort was supported by grant RC2 AG036607 from the National Institutes of Health, grants from the Robert Wood Johnson Foundation, the Ellison Medical Foundation, the Wayne and Gladys Valley Foundation and Kaiser Permanente. We are grateful to the KPNC members who have generously agreed to participate in the Kaiser Permanente Research Program on Genes, Environment and Health. We would like to thank the research participants and employees of 23andMe for making this work possible. This work was supported by the National Human Genome Research Institute of the National Institutes of Health (grant number R44HG006981). N.A. is supported in part by NIDDK award number 1R01DK090382, NINDS award number 1R01NS079231 and NCI award number 1R01CA197139. N.M. is supported by the Jane Coffin Childs Postdoctoral Fellowship.

Author contributions

E.J., L.S., N.A., N.M. and A.A. designed the study. E.J. drafted the manuscript with contributions from all other authors; L.S. performed statistical and bioinformatics analyses in the discovery cohort; C.T. performed statistical analyses and D.H. oversaw analyses of the replication cohort; N.M. ascertained samples and performed experimental work and analyses; W.E. performed computational analysis of RNA-seq data; N.A. oversaw all experimental work; D.C.C. advised on phenotypic characterization and clinical context; A.A. conducted chart review of cases samples; all authors contributed to the final paper.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Jorgenson, E. et al. A genome-wide association study identifies four novel susceptibility loci underlying inguinal hernia. Nat. Commun. 6:10130 doi: 10.1038/ncomms10130 (2015).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/