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.

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, which has been 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 hernia repair surgery and 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 the pathophysiology of inguinal 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.

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 weakness in the transversalis fascia, or indirect, in which abdominal contents protrude through a congenital defect in the inguinal ring via a patent processus vaginalis. Inguinal hernias. To date, only a small number of candidate genes have been investigated.

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.

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 lambda 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 × 10^{-8}) in the regions of EFEMP1 (rs2009262, odds ratio (OR) = 1.23, 95% confidence interval (CI): 96.9–99.9%) were confirmed to be associated with a higher risk of 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% CI: 99.1–99.9%) were confirmed to be inguinal hernia repairs. We reviewed 230 patient charts to validate the accuracy of inguinal hernia diagnoses, and, of those, 228 (99.1%, 95% CI: 99.1–99.9%) were confirmed to be inguinal hernia repairs. We reviewed 230 patient charts to validate the accuracy of inguinal hernia diagnoses, and, of those, 228 (99.1%, 95% CI: 99.1–99.9%) were confirmed to be inguinal hernia repairs. We reviewed 230 patient charts to validate the accuracy of inguinal hernia diagnoses, and, of those, 228 (99.1%, 95% CI: 99.1–99.9%) were confirmed to be inguinal hernia repairs.
rs10746560 (instead of rs6991952) in EBF2
EFEMP1 in the region were more strongly associated with direct inguinal hernia, but for three of the four loci, other SNPs in inguinal hernia were also the most strongly associated SNPs with indirect hernia in men. The four top SNPs associated may underlie the different subtypes of inguinal hernia.

This indicates that multiple variants within these risk loci (Fig. 2). This suggests that multiple variants within these risk loci

**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>0.81 (1.09-1.19)</td>
<td>0.81 (1.02-1.09)</td>
<td>0.81 (1.00-1.12)</td>
<td>0.73 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>0.43 (1.01-1.14)</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 (1.13-1.23)</td>
<td>0.62 (1.03-1.10)</td>
<td>0.62 (1.00-1.14)</td>
<td>3.69 x 10^-14</td>
</tr>
</tbody>
</table>

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

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, Adtm at a moderate level and Ef2 and Adams6 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.
three replicates were analysed and normalized gene expression values, FPKM, were obtained for each replicate using Cufflinks2. Confirmed those associations in an independent cohort. All four proteins, collagen and elastin. Protein expression was measured in mouse connective tissue. ADAMTS6 near Ebf2. Dysregulation of collagen homeostasis is thought to play an important role in the development of inguinal hernias. 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. Three novel inguinal hernia genetic susceptibility loci near WT1, 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 homeostasis is thought to play an important role in the development of inguinal hernias. 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 2 | Sex-stratified analysis of direct and indirect inguinal hernia among GERA discovery cohort.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Inguinal Hernia Type</th>
<th>Men OR (95% CI)</th>
<th>P-value</th>
<th>Women OR (95% CI)</th>
<th>P-value</th>
<th>Combined OR</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2009262</td>
<td>Direct</td>
<td>1.25 (1.16–1.36)</td>
<td>0.003</td>
<td>1.26 (1.17–1.36)</td>
<td>0.009</td>
<td>1.26 (1.17–1.36)</td>
<td>0.009</td>
</tr>
<tr>
<td>rs370763</td>
<td>Direct</td>
<td>1.14 (1.06–1.22)</td>
<td>0.133</td>
<td>1.14 (1.07–1.22)</td>
<td>0.117</td>
<td>1.14 (1.07–1.22)</td>
<td>0.117</td>
</tr>
<tr>
<td>rs6991952</td>
<td>Direct</td>
<td>1.15 (1.08–1.23)</td>
<td>0.322</td>
<td>1.15 (1.08–1.23)</td>
<td>0.322</td>
<td>1.15 (1.08–1.23)</td>
<td>0.322</td>
</tr>
<tr>
<td>rs3809060</td>
<td>Direct</td>
<td>1.17 (1.09–1.24)</td>
<td>0.55</td>
<td>1.17 (1.09–1.24)</td>
<td>0.55</td>
<td>1.17 (1.09–1.24)</td>
<td>0.55</td>
</tr>
</tbody>
</table>

CI, confidence interval; GERA, Genetic Epidemiology Research in Adult Health and Aging; I2, heterogeneity index; ORF, odds ratio from fixed effects model; ORR, odds ratio from random effects model; Phet, P value for heterogeneity from Cochran’s Q test; SNP, single-nucleotide polymorphism.

Figure 2 | Expression analysis of Efemp1, Wt1, Ebf2 and Adams6 by qRT-PCR (a) and RNA-seq (b). Efemp1 is expressed at a high level, Wt1 at a moderate level and Ebf2 and Adams6 at low levels in mouse connective tissue compared to a connective tissue expressed gene Coll2a1 (positive control) and Oct4 that is not expressed in this tissue (negative control). Data are represented as mean ± s.d. for the qRT-PCR and ± s.e.m. for the RNA-seq (n = 12). For qRT-PCR three samples were analysed in three replicates of each reaction and relative expression levels calculated by the ΔΔCT method. For RNA-seq, three replicates were analysed and normalized gene expression values, FPKM, were obtained for each replicate using Cufflinks2.

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 Adams6 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 WT1 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 homeostasis. 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 intricate balance between ECM degrading and protecting enzymes and may thus perturb connective tissue homeostasis. Combined, our IPA analysis suggests that EFEMP1 and WT1 play a role in connective tissue maintenance/homeostasis through their action on collagen and/or elastin.

Discussion
We identified four novel inguinal hernia genetic susceptibility loci near the genes WT1, 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 homeostasis 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 etiology between inguinal hernia risk and cancer, lung function and anthropomorphic traits. Given previous observational associations between inguinal

Figure 3 | Ingenuity Pathway Analysis outlines potential regulatory networks around EFEMP1 and WT1. Network Analysis for EFEMP1 and WT1 was carried out using the RNA-seq FPKM > 30 gene list (see Methods section). WT1 regulates many extracellular matrix genes, including MMP2 (matrix metalloproteinase-2) and CTGF (connective tissue growth factor). EFEMP1 directly interacts with ELASTIN, a component of elastic fibres in the ECM. TIMP3 (tissue inhibitor of metalloproteinase-3), which is activated by WT1 and interacts with EFEMP1 and was found to connect between the two networks.
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 immunoprecipitation 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.

The incidence of hernia susceptibility in humans peaks at birth and late adulthood. It is possible, and perhaps likely, that factors influencing both the development of fascia and their maintenance affect inguinal hernia susceptibility. Our discovery sample focused on surgically confirmed adult-onset inguinal hernia, with an average age of 66.2 years at diagnosis. It is unclear how the inguinal hernia risk loci identified here influence the risk of childhood-onset hernias, which are always of the indirect type and related to congenital persistence of the processus vaginalis. Furthermore, we confirmed our findings in subjects with self-reported hernia repair surgery, which likely represents a mix of hernia subtypes, with inguinal hernias being the most common. First-degree relatives of inguinal hernia patients have a greater risk of femoral, incisional, epigastric and umbilical hernias indicating a common metabolic aetiology and a shared genetic basis across different hernia subtypes, which may, in part, explain why we observe similar signals across the two cohorts. Future research should examine how the loci identified here contribute to the risk of other types of hernias and the extent to which the mechanisms underlying inguinal hernia development are common to other hernia subtypes and other connective tissue disorders.

In conclusion, our study identified four novel loci underlying the risk of adult-onset inguinal hernia. Our findings suggest a role for the regulation of both collagen homeostasis and elastin maintenance in the development of inguinal hernias, which appear to also influence anthropomorphic traits, the risk of cancer and lung function. Further research into the precise mechanisms through which these loci act may improve our understanding of hernia formation and point the way to more effective preventative, operative and non-surgical treatments of this common disorder.

Methods

Setting. KPNC is an integrated healthcare delivery organization, which has an active membership of 3.5 million people. It is the largest healthcare provider in Northern California. Approximately, one third of the Northern California population is enrolled in the KPNC health plan. Comparisons with the general population have shown that KPNC membership is representative of the population of Northern California, with the exception of extremes of the socioeconomic spectrum. In 1995, KPNC instituted a comprehensive EHR system, which records physician diagnoses, prescriptions and lab results from all inpatient and outpatient encounters. KPNC has high membership retention, with over 90% of those over age 65, and 66% of all active members as of June 2012, having five or more years of retrospective membership.

The GERA cohort. The GERA cohort is comprised of 110,266 adult men and women members of the KPNC Medical Care Plan. It is a component of the KPNC Research Program on Genes, Environment and Health. The detailed description of the cohort and study design can be found in dbGaP (Study Accession: phs000674.v1.p1). Briefly, participants were enrolled through participation in a mailed survey of all adult members of KPNC (~ 1.9 million) conducted in 2007. A total of 38,983 members consented to participate, which included information on demographic factors, behaviours and self-reported health. Beginning in July 2008, respondents to the survey were asked to sign and return a consent form authorizing use of biospecimens, survey data and data from participants’ EHR for use in studies of genetic and environmental influences on health and disease. Respondents who completed consent forms were mailed (Organix) saliva collection kits. A total of 110,266 participant samples were selected for genome-wide genotyping and telomere length measurement to ensure that at least 100,000 were successfully assayed (102,998 samples passed genotyping quality control). The average age of the participants was 62.9 years old (s.d. = 12.3 years); 69,987 participants (63%) were aged 60 years and older, and over 12,000 were aged 80 years and older. The sample is ethnically diverse, generally well-educated, with above average incomes. Length of membership in KPNC averaged 23.5 years, indicating the stability of KPNC membership and the length of medical history that is recorded for cohort members. All study procedures were approved by the Institutional Review Board of the Kaiser Foundation Research Institute.

23andMe cohort. Study participants in the replication cohort were drawn from the customer base of 23andMe, which has been previously described in detail. All individuals provided informed consent and answered surveys online according to the 23andMe human subjects protocol, which was reviewed and approved by Ethical and Independent Review Services, a private institutional review board (http://www.eandireview.com).

Phenotype definition. Hernia cases in the GERA cohort were identified from clinical diagnoses and surgical procedures captured in the EHR system. Hernia diagnoses were typically associated with pre-operative and post-operative diagnoses and a detailed operative report; hernias found or repaired among inpatients also resulted in hospital discharge diagnoses. The operative reports were reviewed by hospital coders so that the corrected discharge diagnosis and procedure codes were assigned. These procedure codes usually indicated whether an inguinal hernia repair was for a direct or indirect hernia based on inguinal hernia risk alleles. Furthermore, we confirmed our findings in subjects with self-reported hernia repair surgery, which likely represents a mix of hernia subtypes, with inguinal hernias being the most common. First-degree relatives of inguinal hernia patients have a greater risk of femoral, incisional, epigastric and umbilical hernias indicating a common metabolic aetiology and a shared genetic basis across different hernia subtypes, which may, in part, explain why we observe similar signals across the two cohorts. Future research should examine how the loci identified here contribute to the risk of other types of hernias and the extent to which the mechanisms underlying inguinal hernia development are common to other hernia subtypes and other connective tissue disorders.

In conclusion, our study identified four novel loci underlying the risk of adult-onset inguinal hernia. Our findings suggest a role for the regulation of both collagen homeostasis and elastin maintenance in the development of inguinal hernias, which appear to also influence anthropomorphic traits, the risk of cancer and lung function. Further research into the precise mechanisms through which these loci act may improve our understanding of hernia formation and point the way to more effective preventative, operative and non-surgical treatments of this common disorder.

The GER...
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.

rs3809060: 0.976; rs6991952: 0.999; and rs370763: 0.991)

The statistical association of the meta-analysis results was calculated using the MAGENTA software (http://www.broadinstitution.org/mpg/magenta/). To do this, we input a ranked list of our meta-analysis association results. We evaluated the results of the gene set enrichment analysis by nominal P value and FDR to control for multiple testing (10). We considered an FDR < 0.05 as significant.

Causal Network Analysis was carried out using IPA (Qiagen). To narrow down the number of genes for this analysis, we ranked results from the RNA-seq experiment by FPKM value and used an arbitrary cutoff of FPKM 30 (2,059 genes). Of the four genes, two exceeded the cutoff values, Efemp1 (FPKM: 621, rank: 99) and Wt1 (FPKM: 35, rank: 1,767), but not Efz2 (FPKM: 4,8, rank: 8,723) or Adams6 (FPKM: 1.2 rank: 12,467).

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.L.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/