Enrichment of putative PAX8 target genes at serous epithelial ovarian cancer susceptibility loci

Siddhartha P Kar*,1, Emily Adler2, Jonathan Tyrer1,3, Dennis Hazelett4,5, Hoda Anton-Culver6, Elisa V Bandera7, Matthias W Beckmann8, Andrew Berchuck9, Natalia Bogdanova10, Louise Brinton11, Ralf Butzow12, Ian Campbell13,14, Karen Carty15, Jenny Chang-Claude16,17, Linda S Cook18, Daniel W Cramer19, Julie M Cunningham20, Agnieszka Dansonka-Mieszkowska21, Jennifer Anne Doherty22, Thilo Dörk23, Matthias Dürst24, Diana Eccles25, Peter A Fasching26, James Flanagan27, Aleksandra Gentry-Mahara28, Rosalind Glasspool29, Ellen L Goode30, Marc T Goodman31,32, Jacky Gronwald33, Florian Heitz34,35, Michelle A T Hildebrandt36, Estrid Høgdal37,38, Claus K Høgdal39, David G Hunstman40,41,42, Allan Jensen37, Beth Y Karlan43, Linda E Kelenen44, Lambertus A Kiemeney45, Susanne K Kjaer37,46, Jolanta Kupryjanczyk47, Diether Lambrechts48,49, Douglas A Levine50, Qian L Li51,52, Jolanta Lissowska53, Karen H Lu44, Jan Lubinski43, Leon F A G Massuger55, Valerie McGuire56, Iain McNeish57, Usha Menon28, Francesmary Modugno58,59,60, Alvaro N Monteiro51, Kirsten B Moysich62, Roberta B Ness63, Heli Nevanlinna64, James Paul65, Celeste L Pearce66, Tanja Pejovic67,68, Jennifer B Permutt61, Catherine Phelan61, Malcolm C Pike2,69, Elizabeth M Poole70, Susan J Ramus71, Harvey A Rischl72, Mary Anne Rossing73,74, Helga B Salvesen75,76, Joellen M Schildkraut77,78, Thomas A Sellers51, Mark Sherman11, Nadeem Siddiqui79, Weiva Sieh80, Honglin Song3, Melissa Southey80, Kathryn L Terry81,82, Shelley S Tworoger70,82, Christine Walsh43, Nicolas Wentzensen11, Alice S Whittemore66, Anna H Wu2, Hannah Yang11, Wei Zheng83, Argyrios Ziogas84, Matthew L Freedman85,86, Simon A Gayther5,87, Paul D P Pharoah1,3 and Kate Lawrenson5,88

Background: Genome-wide association studies (GWAS) have identified 18 loci associated with serous ovarian cancer (SOC) susceptibility, but the biological mechanisms driving these findings remain poorly characterised. Germline cancer risk loci may be enriched for target genes of transcription factors (TFs) critical to somatic tumorigenesis.

Methods: All 615 TF-target sets from the Molecular Signatures Database were evaluated using gene set enrichment analysis (GSEA) and three GWAS for SOC risk: discovery (2196 cases/4396 controls), replication (7035 cases/21693 controls; independent from discovery), and combined (9627 cases/30845 controls; including additional individuals).

Results: The PAX8-target gene set was ranked 1/615 in the discovery (P_{GSEA}<0.001; FDR=0.21), 7/615 in the replication (P_{GSEA}=0.004; FDR=0.37), and 1/615 in the combined (P_{GSEA}<0.001; FDR=0.21) studies. Adding other genes reported to interact with PAX8 in the literature to the PAX8-target set and applying an alternative to GSEA, interval enrichment, further confirmed this association (P=0.006). Fifteen of the 157 genes from this expanded PAX8 pathway were near eight loci associated with SOC risk at \(P<10^{-5}\) (including six with \(P<5\times10^{-8}\)). The pathway was also associated with differential gene expression after shRNA-mediated silencing of PAX8 in HeyA8 (P_{GSEA}=0.025) and IGROV1 (P_{GSEA}=0.004) SOC cells and several PAX8 targets near SOC risk loci demonstrated in vitro transcriptomic perturbation.

Conclusions: Putative PAX8 target genes are enriched for common SOC risk variants. This finding from our agnostic evaluation is of particular interest given that PAX8 is well-established as a specific marker for the cell of origin of SOC.
Epithelial ovarian cancer (OC) is the most common cause of gynaecological cancer death in the United Kingdom (Cancer Research UK, 2016). The high mortality associated with the disease is in part because it is often diagnosed at an advanced stage and a better understanding of germline genetic predisposition to OC may eventually lead to precision screening and earlier diagnosis (Bowtell et al., 2015). Genome-wide association studies (GWAS) have so far identified 18 loci associated with susceptibility to all invasive OC or to its most common histological subtype, serous OC (SOC), that accounts for ~70% of all cases (Song et al., 2009; Bolton et al., 2010; Goode et al., 2010; Bojesen et al., 2013; Couch et al., 2013; Permuth-Wey et al., 2013; Pharoah et al., 2013; Kuchenbaecker et al., 2015). Post-GWAS studies that integrate molecular phenotypes with GWAS findings are essential to elucidate the function of the known loci in SOC development and to unravel the potential role of loci that just fail to reach the threshold for genome-wide statistical significance (P < 5 × 10⁻⁸; Freedman et al., 2011; Kar et al., 2015; Lawrenson et al., 2015).

The vast majority of single-nucleotide polymorphisms (SNPs) associated with cancer susceptibility lie in non-coding regions of the genome and so do not have any impact on protein structure and function. A growing body of evidence suggests that many inherited common risk variants instead fall into non-coding regulatory elements, such as enhancers or transcription factor (TF)-binding sites (Sur et al., 2013). Different alleles of these SNPs impact the biological activity of the regulatory elements and thus modify expression of a local (cis-acting) target gene or genes.

Expression of many TFs occur in a tissue-specific manner, and binding sites and transcriptional target genes for such lineage-specific TF drivers of cancer can be enriched at risk loci, also in a tissue-specific manner. For example, breast cancer risk SNPs are enriched for binding sites of the TFs ESR1 and FOXA1 in breast cancer cells while prostate cancer risk variants are enriched for androgen receptor-binding sites in prostate cells (Cowper-Sal lari et al., 2012; Lu et al., 2012; Jiang et al., 2013; Chen et al., 2015). However, for SOC, similar links between TFs and genetic risk have not been evaluated. This is partly because the TF-target gene networks active in SOC and SOC precursor cells are poorly characterised. Moreover, genome-wide TF-binding sites have not been profiled by chromatin immunoprecipitation combined with sequencing (ChIP-Seq) in SOC precursor and SOC tissues by initiatives such as the Encyclopedia of DNA Elements and the Nuclear Receptor Cistrome projects that enabled the corresponding studies for breast and prostate cancers (Tang et al., 2011; ENCODE Project Consortium, 2012).

In the absence of such data, we searched for an in silico resource that would allow an agnostic assessment of association between putative target genes of many different TFs and susceptibility to SOC. The Molecular Signatures Database (MSigDB) is a compendium of annotated functional pathways that includes 615 TF-target gene sets (Subramanian et al., 2005). All genes in each set share the same upstream cis-regulatory motif that is a predicted binding site for a particular TF and they thus represent the inferred target genes of that TF. The motifs themselves are regulatory motifs of mammalian TFs derived from the TRANSFAC database (Matys et al., 2006). In this study, we undertook pathway analysis using gene set enrichment (Subramanian et al., 2005) to test for overrepresentation of signals associated with SOC risk in these 615 TF-target gene sets using the two largest SOC GWAS data sets currently available for discovery and for independent replication. We further confirmed our top replicated gene set – targets of the TF PAX8 – using an alternative pathway analysis approach and used in vitro transcriptomic modelling to demonstrate perturbation of this gene set in the cellular context of SOC.

### MATERIALS AND METHODS

**Discovery, replication, and combined study populations.** The discovery pathway analysis was performed on a meta-analysis of a North American and UK phase 1 GWAS of 2196 SOC cases and 4396 controls. The replication pathway analysis used data from 7035 SOC cases and 21,693 controls that were independent of the discovery participants and obtained from 43 case-control studies genotyped under the Collaborative Oncological Gene-environment Study (COGS) project. The two GWAS and the COGS studies have been described previously (Song et al., 2009; Permuth-Wey et al., 2011; Pharoah et al., 2013). The combined pathway analysis was based on a total of 9627 SOC cases and 30,845 controls from a meta-analysis that included the North American and UK GWAS, the COGS, and additional cases and controls from the Ovarian Cancer Association Consortium (OCAC) as reported previously (Kuchenbaecker et al., 2015). All participants were of European ancestry, provided informed consent, and had been recruited under protocols approved by a local ethics committee.

**Single-nucleotide polymorphism data.** The discovery, replication, and combined pathway analyses used summary findings (P-values) for association between SNP germline genotype and SOC susceptibility in the respective study populations. The discovery stage included 2,508,744 SNPs that had either been genotyped or imputed with imputation accuracy, r^2 > 0.3 and had a minor allele frequency (MAF) > 1% in both the North American and the UK GWAS. Samples were genotyped on Illumina (San Diego, CA, USA) platforms (317K/550K/610K) and imputed into the HapMap II (release 22) Utah residents with Northern and Western European ancestry (CEU) reference panel. As with most gene-based common variant association tests (Petersen et al., 2013), the gene-ranking procedure described below (Saccone et al., 2007; Christoforou et al., 2012) had been developed for HapMap-imputed GWAS and this guided our choice of HapMap-imputed SNP data over the more heavily correlated 1000 Genomes-imputed SNP data, which were also available. The replication stage was based on summary findings from COGS for a subset of 2,421,023 SNPs out of the ~2.5 million SNPs from the discovery stage that had either been genotyped on the Illumina iCOGS custom array or imputed into the 1000 Genomes (March 2012) European reference panel with r^2 > 0.3 and had a MAF > 1% in the COGS studies. The combined pathway scan was also based on data for the same subset of SNPs but from association analysis in the combined study population. Sample and genotyping quality control, imputation, association- and meta-analysis steps for generating these three data sets have been described previously (Song et al., 2009; Permuth-Wey et al., 2011; Pharoah et al., 2013; Kuchenbaecker et al., 2015).

**Gene-set enrichment analysis.** Pathway analysis was conducted using the Preranked tool in the GSEA software (version 2.2.1; Subramanian et al., 2005)) with default settings, 1000 permutations (unless otherwise specified), and no restrictions imposed on the size of gene sets that could be included. GSEA requires a list of genes ranked by any metric and a collection of annotated biological pathways or gene sets.

All 615 TF target gene sets (containing between 5 and 2657 genes; median = 219 genes) annotated in the Molecular Signatures Database (MSigDB version 5.0-C3; www.broadinstitute.org/gsea/msigdb) were tested in the GSEA. Each of these gene sets represents a group of genes that share a single TF-binding site motif defined in the TRANSFAC database (version 7.4; www.gene-regulation.com; Matys et al., 2006)). The gene sets are named after the corresponding TRANSFAC TF-binding site matrix identifier and additional details of their curation and nomenclature is available at MSigDB.
The ranked list of genes for GSEA was generated from the genome-wide SNP data by the following steps: (1) start and end positions for 23 161 genes that were unambiguously mapped were downloaded using R version 3.0.3 (Vienna, Austria; TxDb.Hsapiens.s.UCSC.hg19.knownGene: Annotation package for TxDb object(s). Version 2.8.0); (2) all SNPs that lay between these start and end positions were assigned to the corresponding genes; (3) the most significant P-value among all SNPs within the boundaries of each gene was adjusted for the number of SNPs in the gene by a modification of Sidak’s correction (Saccone et al, 2007; Christoforou et al, 2012) that has been shown to reduce the effect of gene size on the P-value and account for correlations due to linkage disequilibrium between SNPs (Segrè et al, 2010); (4) the genes were ranked in descending order of the negative logarithm (base 10) of the most significant P-value (after adjustment). These steps were performed on SNPs and their P-values from the discovery (19 540 genes containing ≥1 SNP), replication, and combined data (19 364 genes containing ≥1 SNP). Quantile-quantile plots of these gene-level P-values were plotted for each data set (Supplementary Figure S1).

PAX8 target genes and pathway. The MSigDB gene set for putative targets of the TF PAX8, termed VSPAX8_B, contained 106 genes (Supplementary Table S1). These genes were grouped together because their promoter regions (± 2 kb around transcription start site) contained at least one instance of the TRANSFAC motif NCNNTNNTGCRTGANNNN that matches annotation for a PAX8-binding site. Four of these genes were open reading frames and excluded from pathway analyses (Supplementary Table S1). We used Qiagen’s Ingenuity Target Explorer (targetexplorer.ingenity.com) to identify 55 additional genes (Supplementary Table S2) that were known to interact with PAX8 according to the literature, though not necessarily by binding PAX8 as a TF (description and citation for each interaction in Supplementary Table S2). Cells treated with PAX8 shRNA-1 and shRNA-2 and luciferase-labelled IGROV1 plus scrambled controls vs IGROV1 treated with PAX8 shRNA-3 and shRNA-4 (all experiments in triplicate). For microarray-GSEA, phenotype labels for these comparisons were permuted 1000 times and the standard signal-to-noise ratio was used to rank genes. Of the 157 PAX8 pathway genes, 154 were profiled on the Illumina HT-12 (three genes not profiled: LUC7L3, PKM, TMA16). Two-sided exact binomial test P-values calculated using the binom.test function in R version 3.0.3 were used to evaluate the proportion of PAX8 pathway genes that were differentially expressed.

RESULTS

We began by testing the association between each of the 615 TF-target gene sets in MSigDB/TRANSFAC and SOC susceptibility by GSEA using P-values for ~2.5 million SNPs from a meta-analysis of a North American and UK phase 1 GWAS of 2196 SOC cases and 4396 controls of European ancestry (Song et al, 2009;
Permuth-Wey et al., 2011). SNPs were mapped to genes and the most significant SNP P-value for association with SOC risk in each gene was used to rank genes genome-wide for the GSEA. In this discovery pathway scan, 57 of the 615 TF-target gene sets were associated with SOC risk at P_{GSEA}<0.05 (Supplementary Table S5) and putative target genes of the TF PAX8 (designated ‘VSPAX8_8’) emerged as the top ranked set (P<0.001; FDR=0.21; Table 1). Next, we sought to replicate these findings using genetic association results from independent samples not included in the discovery step. We performed a second GSEA for all 615 TF-target gene sets using association P-values for ~2.4 million SNPs genome-wide in 7035 SOC cases and 21,693 controls from the Collaborative Oncological Gene-environment Study. In this replication pathway scan, 54 of the 615 TF-target gene sets were associated with SOC risk at P_{GSEA}<0.05 (Supplementary Table S6), including 22 gene sets identified at the same significance level in the discovery analysis. The gene set containing targets of PAX8 was ranked 7th out of 615 (P=0.004; FDR=0.37; Table 1) and none of the six gene sets with a higher rank than the PAX8 P-value for association with SOC risk in this combined study population (P_{GSEA}<0.05; Supplementary Table S7), was also strongly associated with SOC risk in the combined study population (P_{GSEA}=4.10^{-4} after 10,000 permutations). Next, we confirmed the association between the PAX8 pathway and SOC susceptibility using an alternative pathway analysis method called interval-based enrichment (INRICH; (Lee et al., 2012)) and used it to pinpoint specific PAX8 target genes likely driving the pathway-level signal. We identified all uncorrelated SNPs (n=47; Supplementary Table S4) associated with SOC risk at P<10^{-5} in the combined study population, generated two-megabase-wide intervals centred on each SNP, and merged overlapping intervals to yield 45 intervals. Fifteen of the 157 genes from the PAX8 pathway were located in eight of these 45 intervals (P_{INRICH}=0.006 compared to 5000 permuted sets of 45 two-megabase-wide intervals). The P<10^{-5} index SNP at the center of five of these eight intervals was in fact genome-wide significant (P<5.9x10^{-8}; Table 2). The SNP anchoring a sixth interval (rs2268177), though just short of genome-wide significance in the combined study population (P=9.5x10^{-7}), has previously achieved this threshold in a meta-analysis of samples from OCAC with samples from the Consortium of Investigators of Modifiers of BRCA1/2 that were not included in this combined study population (Kuchenbaecker et al., 2015). Although we used a megabase flanking each P<10^{-5} SNP to define these intervals (to capture long-range SNP-gene cis-regulatory effects (van Heyningen and Bickmore, 2013)), in five of the eight intervals the nearest PAX8 pathway gene was less than 100 kb from the central SNP and only for two intervals did this distance extend beyond 200 kb (Table 2).

Finally, we tested whether the PAX8 pathway that had thus far been defined by combining annotations from MSigDB/TRANSFAC and curation of the published literature (that included experiments conducted in non-ovarian cell types) was cell- and cancer-type relevant. PAX8 expression was stably silenced by shRNAs in the ovarian cancer cell lines HeyA8 and IGROV1 (Supplementary Figure S2) and gene expression microarray profiling performed in knockdown and control cultures. The 157-gene PAX8 pathway, of which 154 genes were profiled on the microarray, was significantly associated with differential gene expression after PAX8 silencing in both cell line models (\( P_{\text{microarray-GSEA}}=0.03 \) for HeyA8; \( P_{\text{microarray-GSEA}}=0.004 \) for IGROV1; Supplementary Figure S3). In HeyA8 cells, 45 of these 154 genes from the PAX8 pathway were differentially expressed at \( P<0.05 \) (corresponding \( P_{\text{FDR}}<0.31; 14/154 at FDR<0.05; P_{\text{exact binomial}}<2.2x10^{-10} \) for 45/154 against 7.7/154 expected under the null hypothesis at the 5% level; Supplementary Table S8). In IGROV1 cells, 41 of the 154 PAX8 pathway genes were differentially expressed at \( P<0.05 \) (FDR<0.28; 17/154 at FDR<0.05; \( P_{\text{exact binomial}}<2.2x10^{-16} \); Supplementary Table S9). Overall, 69 of the 154 genes were differentially expressed at \( P<0.05 \) in at least one of, and 17 genes differentially expressed in both, the cell lines after PAX8 silencing (\( P_{\text{exact binomial}}=0.002 \) for 17/154 against 7.7/154 expected). On cross-examining results from the differential expression and INRICH analyses, we observed that of the 15 PAX8 pathway genes in eight intervals associated with SOC risk at \( P<10^{-5} \) (Table 2), BN2, TNF, and NCL were differentially expressed at \( P<0.05 \) in both cell lines, HOXB5, HOXB7, HOXB8, and SP6 in IGROV1 cells only, and TERT and OSBPL7 in HeyA8 cells only (Table 3). Notably, BN2 and HOXB7

### Table 1. Top 10 transcription factor target gene sets from pathway analysis (GSEA) in each study population

| Transcription factor target gene sets | Number of genes | P-value | FDR |
|--------------------------------------|-----------------|---------|-----|
| **Discovery study population (2196 cases/4396 controls)** | | | |
| VSPAX8_8 | 92 | 0.01 | 0.21 |
| GGAIRNKYCCA_UNKNOW | 70 | 0 | 0.21 |
| VSSOX5_01 | 241 | 0.01 | 0.31 |
| RRAGTGTG_UNKNOWN | 232 | 0.01 | 0.33 |
| VSSRY_02 | 230 | 0.01 | 0.35 |
| VSPAP1_02 | 251 | 0.01 | 0.38 |
| VSSRF_05_01 | 196 | 0.01 | 0.41 |
| VSSRF_02_02 | 199 | 0.01 | 0.42 |
| AANANWYTGC_UNKNOWN | 180 | 0.01 | 0.43 |
| YAATNRNNYNKNNNAT_UNKNOWN | 100 | 0.01 | 0.44 |
| **Replication study population (7035 cases/21693 controls)** | | | |
| CAGNYYNKAA_UNKNOWN | 68 | 0.01 | 0.15 |
| VSMAF_06 | 231 | 0.01 | 0.37 |
| VSEF_01 | 60 | 0.01 | 0.16 |
| VSY1_01 | 218 | 0.01 | 0.42 |
| VSCART1_01 | 202 | 0.01 | 0.38 |
| VSPARRA_01 | 36 | 0.01 | 0.24 |
| VSPAX8_8 | 91 | 0.01 | 0.37 |
| TCCATTKW_UNKNOWN | 210 | 0.01 | 0.43 |
| VSFIF2_01 | 108 | 0.01 | 0.43 |
| YKACATT_UNKNOWN | 258 | 0.01 | 0.52 |
| **Combined study population (9627 cases/30845 controls)** | | | |
| VSPAX8_8 | 91 | 0.01 | 0.21 |
| TTGCWCAAY_VSCEBB_02 | 55 | 0.01 | 0.24 |
| YAATNRNNYNKNNNAT_UNKNOWN | 98 | 0.01 | 0.27 |
| YCATTA_UNKNOWN | 501 | 0.01 | 0.46 |
| VSP02F32_02 | 69 | 0.01 | 0.19 |
| VSFIR2_01 | 108 | 0.01 | 0.28 |
| VSCART1_01 | 202 | 0.01 | 0.43 |
| VOCT1_04 | 201 | 0.01 | 0.44 |
| CAGNYNKAA_UNKNOWN | 68 | 0.01 | 0.17 |
| VSNK02S_01 | 112 | 0.02 | 0.28 |

Abbreviations: FDR = false discovery rate; GSEA = gene set enrichment analysis.
were PAX8 target genes within 200 kb of a genome-wide significant index SNP at the 9p22.2 and 17q21.32 SOC risk loci, respectively, and differentially expressed in IGROV1 cancer cells at FDR < 0.05 (Table 3). While we were completing our study, genome-wide ChIP-Seq maps of PAX8 binding in three FTSEC and three additional ovarian cancer cell lines were published (Elias et al., 2016). We intersected these PAX8-binding sites with all SNPs at \( P < 10^{-5} \) for association with SOC risk in the combined data set within the eight intervals identified by INRICH (+ / − 1 Mb of the eight unique index SNPs listed in Table 2). SNPs at the 9p22.2, 17q21.32, and 19p13.11 genome-wide significant SOC risk loci overlapped PAX8 ChIP-biding sites in at least two of the three ovarian cancer cell lines with the most significant binding site SNP at each locus having \( P < 7.1 \times 10^{-5} \) for association with SOC risk (Supplementary Table S10).

### DISCUSSION

This is the first analysis to demonstrate that genes potentially targeted by the TF PAX8 are enriched for common genetic variation associated with SOC risk, suggesting that PAX8 may be a master transcriptional regulator of susceptibility to SOC. The emergence of the PAX8 pathway from our agnostic genome-wide approach evaluating overrepresentation of SOC risk SNPs in 615 gene sets each containing putative targets of a different TF followed by replication in independent samples and methodological replication is highly significant given recent calls for an improved understanding of the role of PAX8 in SOC biology (Bowtell et al., 2015). PAX8 encodes a member of the paired box family of TFs that contains a partial homeodomain (Chi and Epstein, 2002). It is essential for normal embryonic development of the Müllerian ducts (Mittag et al., 2007). A systematic, genome-wide RNA interference screen that included 25 ovarian cancer cell lines previously demonstrated that PAX8 is an ovarian cancer lineage-specific dependency, that is, it was the only gene that met three criteria: (a) essential for the survival and proliferation of ovarian cancer cell lines; (b) focally amplified in primary high-grade serous ovarian tumours; and (c) differentially overexpressed in ovarian cancer cell lines (Cheung et al., 2011). PAX8 has also been shown to be a key player in the proliferation, migration, and invasion of ovarian cancer cells and silencing this gene significantly inhibited anchorage-independent growth in vitro and tumour formation in a nude mouse xenograft model in vivo (Di Palma et al., 2014). Furthermore, PAX8 has been shown to drive murine SOCs originating in the fallopian tube (Perets et al., 2013). PAX8 is routinely used clinically as an epithelial marker to identify primary or metastatic tumours of Mullerian origin.
PAX8 and ovarian cancer susceptibility

(Özcan et al., 2011a, b) because the TF is lineage-specific for FTSECs and ovarian surface epithelial cells (OSECs), both of which are contenders for being the cell of origin of SOC (Özcan et al., 2011b; Adler et al., 2015). The FDR for the PAX8-target gene set in the discovery, replication, and combined GSEA was 0.21, 0.37, and 0.21, respectively, suggesting that the result may be valid four out of five times or three out of five times. However, the convergence of orthogonal pieces of evidence: independent replication of the PAX8 GSEA findings across two of the largest genetic association studies and by INRICH, confirmation of the pathway-level association signal on including additional genes known to interact with PAX8 curated from published experiments, and significant perturbation of several of the putative PAX8 target genes near SOC risk loci in the transcriptomes of two ovarian cancer cell line models after abrogating PAX8 expression; all strongly support the observed association between this gene set and SOC risk. This convergence of evidence from integration of GWAS and cellular models also provides new insight into the potential transcriptional regulatory network of PAX8 and highlights 15 genes that are likely to be driving the association between the PAX8 pathway and risk of developing SOC (Table 3). Among these, BNC2 and HOXB7 are particularly noteworthy. BNC2 lies at 9p22.2, the first SOC risk locus to be identified (Song et al., 2009), and several functional SNPs in this region reside in enhancer elements and in a scaffold/matrix attachment region sequence that targets BNC2 (Buckley et al., under review). We have previously implicated HOXB7 at the 17q21.32 SOC risk locus in network analysis combining SOC GWAS data with SOC gene expression profiles from The Cancer Genome Atlas (Kar et al., 2015). HOXB7 overexpression in OSECS is associated with increased proliferation via fibroblast growth factor signalling (Naora overexpression in OSECs is associated with increased proliferation; all strongly support the observed association between this gene set and SOC risk. The identification of SOC risk SNPs in PAX8 ChIP-binding sites from The Cancer Genome Atlas (Kar et al., 2013). The FDR for the PAX8-target gene set in the discovery, replication, and combined GSEA was 0.21, 0.37, and 0.21, respectively, suggesting that the result may be valid four out of five times or three out of five times. However, the convergence of orthogonal pieces of evidence: independent replication of the PAX8 GSEA findings across two of the largest genetic association studies and by INRICH, confirmation of the pathway-level association signal on including additional genes known to interact with PAX8 curated from published experiments, and significant perturbation of several of the putative PAX8 target genes near SOC risk loci in the transcriptomes of two ovarian cancer cell line models after abrogating PAX8 expression; all strongly support the observed association between this gene set and SOC risk. This convergence of evidence from integration of GWAS and cellular models also provides new insight into the potential transcriptional regulatory network of PAX8 and highlights 15 genes that are likely to be driving the association between the PAX8 pathway and risk of developing SOC (Table 3). Among these, BNC2 and HOXB7 are particularly noteworthy. BNC2 lies at 9p22.2, the first SOC risk locus to be identified (Song et al., 2009), and several functional SNPs in this region reside in enhancer elements and in a scaffold/matrix attachment region sequence that targets BNC2 (Buckley et al., under review). We have previously implicated HOXB7 at the 17q21.32 SOC risk locus in network analysis combining SOC GWAS data with SOC gene expression profiles from The Cancer Genome Atlas (Kar et al., 2015). HOXB7 overexpression in OSECS is associated with increased proliferation via fibroblast growth factor signalling (Naora et al., 2001). Our identification of SOC risk SNPs in PAX8 ChIP-binding sites at the BNC2 and HOXB7 loci further support involvement of these two genes in SOC susceptibility potentially through a PAX8-regulated mechanism. A critical next step after this pathway-based study will involve systematic genome-wide discovery of additional specific SOC risk alleles in PAX8-binding sites and/or variants that are involved in the regulation of genes that are also regulated by PAX8 (Freedman et al., 2011; Sur et al., 2013). This link between SOC risk alleles and the PAX8-target gene network may be established through dynamic expression quantitative trait locus analyses undertaken against a background of PAX8 knockdown in the relevant cell types (Califano et al., 2012; Fletcher et al., 2013).

Overall, consistent with recent pathway-level results in other cancers (Huang et al., 2015; Qian et al., 2015), the present study suggests that the genetic architecture of SOC susceptibility may be underpinned by a complex interplay between genes at the known genome-wide significant risk loci and at as yet unidentified loci that just fail to reach genome-wide significance but are functionally related to the known loci. We have demonstrated that genes interacting up- and downstream of PAX8 harbour SNPs strongly associated with SOC risk and a more comprehensive exploration of these targets may eventually open up opportunities for rational pathway-guided biomarker and therapeutic development to combat this lethal disease in its earliest stages.

ACKNOWLEDGEMENTS

This work was in part supported by a Gates Cambridge Scholarship and a Junior Research Fellowship in Clinical Medicine from Homerton College, Cambridge (to SPK) and by a K99/R00 Pathway to Independence Award (R00CA184415) from the United States National Cancer Institute (to KL). The COGS project was funded by the European Commission’s Seventh Framework Programme under grant agreement HEALTH-F2-2009-223175. This project was also in part funded by the National Cancer Institute GAME-ON Post-GWAS Initiative U19-CA148112. The Ovarian Cancer Association Consortium (OCAC) is supported by a grant from the Ovarian Cancer Research Fund thanks to donations by the family and friends of Kathryn Sladek Smith (PPD/RPCI.07). Funding for the constituent studies of OCAC were provided by (these are listed by funding agency with grant numbers in parentheses): The American Cancer Society (CRTG-00-196-01-CCCE); the California Cancer Research Program (00-01389V-20170, N01-CN25403, 2II0200); the Canadian Institutes for Health Research (MOP-86727); Cancer Council Victoria; Cancer Council Queensland; Cancer Council New South Wales; Cancer Council South Australia; Cancer Council Tasmania; Cancer Foundation of Western Australia; the Cancer Institute of New Jersey; Cancer Research UK (C490/A6187, C490/A10119, C490/A10124, C536/A13086, C536/A6689); the Celma Mastry Ovarian Cancer Foundation; the Danish Cancer Society (94-222-52); the ELAN Program of the University of Erlangen-Nuremberg; the Eve Appeal; the Helsinki University Central Hospital Research Fund; Helse Vest; Imperial Experimental Cancer Research Centre (C1312/A15589); the Norwegian Cancer Society; the Norwegian Research Council; the Ovarian Cancer Research Fund; Naatsel Kankerplan of Belgium; Grant-in-Aid for the Third Term Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health Labour and Welfare of Japan; the L & S Milken Foundation; the Polish Ministry of Science and Higher Education (4 POSC 024 14, 2 POSA 068 27); Malaysian Ministry of Higher Education (UM.C/HIR/ MOHE/06) and Cancer Research Initiatives Foundation; the Roswell Park Cancer Institute Alliance Foundation; the US National Cancer Institute (K07-CA095666, K07-CA143047, K22-CA138563, N01-CN55424, N01-PC067010, N01-PC035137, P01-CA17054, P01-CA087696, P30-CA15883, P50-CA105009, P50-CA136395, R01-CA014089, R01-CA016056, R01-CA17054, R01-CA04949, R01-CA050385, R01-CA054419, R01-CA058598, R01-CA058860, R01-CA061107, R01-CA061132, R01-CA063682, R01-CA064277, R01-CA067626, R01-CA071766, R01-CA074850, R01-CA076016, R01-CA080742, R01-CA080978, R01-CA083918, R01-CA087538, R01-CA092044, R01-CA095023, R01-CA106414, R01-CA122443, R01-CA112523, R01-CA114343, R01-CA126841, R01-CA136924, R01-CA149429, R03-CA113148, R03-CA115195, R37-CA070867, R37-CA070867, U01-CA069471, U01-CA071766, K07-CA080668, P50-CA159981, R01CA095023, M01-RR000056, R01-CA063678, U1M-CA186107, P01-CA87969, UM1-CA176726, U1M-CA182910, and Intramural research funds); the US Army Medical Research and Material Command (DAMD17-98-1-8659, DAMD17-01-1-0729, DAMD17-02-1-0666, DAMD17-02-1-0669, W81XWH-10-1-0282); the National Health and Medical Research Council of Australia (199600 and 400282); the German Federal Ministry of Education and Research of Germany Programme of Clinical Biomedical Research (01 GB 94013); the state of Baden-Württemberg through Medical Faculty of the University of Ulm (P.685); the Minnesota Ovarian Cancer Alliance; the Mayo Foundation; the Fred C. and Katherine B Andersen Foundation; the Lon V. Smith Foundation (LVS-39420); the Oak Foundation; the OHSU Foundation; the Mermaid I project; the Rudolf-Bartling Foundation; the UK National Institute for Health Research Biomedical Research Centres at the University of Cambridge, Imperial College London, the Royal Marsden Hospital; WorkSafeBC. The Nurses’ Health Study and Nurses’ Health Study II would like to thank the participants for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA and WY.
CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Adler E, Mhawech-Fauceglia P, Gayther SA, Lawrenson K (2015) PAX8 expression in ovarian surface epithelial cells. Hum Pathol 46: 948–956.

Bailey TL, Machanick P (2012) Inferring direct DNA binding from ChIP-seq.

van der Schoot CE, Poole EM, Hogervorst FBL, Tworoger SS, Liu J, Silva IDS, Cramer DW, Gibson L, Terry KL, Fletcher O, Vitonis AF, Rudolph A, Weber RP, Flesch-Janys D, Iversen E, Nickels S, Schildkraut JM, Nordestgaard BG, Flyger H, Vachon C, Olson JE, Wang X, Levine DA, Hartman M, Yatabe Y, Hamann U, Karlan BY, Sangrajrang S, Kjaer SK, Edwards R, Meindl A, Heitz F, Matsuo K, du Bois A, Wu AH, Harter P, Leminen A, Anton-Culver H, Bunker CH, Kristensen V, Ness RB, Muir K, Jukkola-Vuorinen A, Rzepecka IK, Pylka¨s K, Bidzinski M, Kauppila S, Figueroa J, Kupryjanczyk J, Chanock SJ, Dansonka-Mieszkowska A, Menon U, Pearce CL, Bru ¨ning T, Pike MC, Ko Y-D, Lissowska J, Ziogas A, Ashworth A, Gentry-Maharaj A, Jones M, Ramus SJ, Orr N, Lux MP, Permuth-Wey J, Peissel B, Sellers TA, Ficarazzi F, Barile M, Flanagan JM, Kriege M, Paul J, Margolin S, Siddiqui N, Severi G, Knight JA, Høgdall CK, Apicella C, Gore M, Tsimiklis H, Song H, Le Marchand L, Yang HP, Mulligan AM, Glendon G, Engelholm SA, Balasubramanian SP, Kelemen LE, Schneeweiss A, Le ND, Sohn C, Tajima K, Tseng C-C, Stram DO, van den Berg D, Yip CH, Ikram MK, Stewart-Brown S, Ditsch N, Lichtner P, Schmutzler RK, Ito H, Iwata H, Shrubsole M, Deming-Halverson S, Lophatananon A, Siriwanarangsan P, Steveson-Brown S, Ditch S, Lichter P, Schmutzler RK, Ito H, Iwata H, Tajima K, Tseng C-C, Stram DO, van den Berg D, Yip CH, Iramak SM, Teh Y-C, Cai H, Wu, Signorello LB, Cai Q, Noh D-Y, Koo K-Y, Miao H, Iau PT-C, Teo YY, McKay J, Shapiro C, Ademuwujie F, Fountzilas G, Hsiung C-N, Yu J-C, Hou M-F, Healey CS, Lucarrini C, Pock S, Stopa-Lyonnat D, Petelon P, Rebbeck TR, Piedmonte M, Singer CF, Friedman E, Thomassen M, Offit K, Hansen TVO, Neuhausen SL, Szabo CI, Blanco I, Garber J, Narod SA, Weitzel JN, Montagna M, Olah E, Godwin AK, Yannoukakos D, Goldgar DE, Caldes T, Miyayntov EN, Tihomirov L, Arun BK, Campbell I, Mensenkamp AR, van Asperen CJ, van Roozendael KEP, Meijers-Heinboer H, Colleé MJ, Oosterwijk JC, Hoening MJ, Koopman RA, van der Luijt RB, Osen TAM, Evans DG, Frost D, Fineberg E, Barwell J, Walker I, Kennedy MJ, Platte R, Davidson R, Ellis SD, Cole T, Bresac-de Pailleters B, Buecher D, Bamiala F, Faire L, Fiere M, Sinilkovina OM, Caron O, Giraud M, Saizov M, Bonadonna V, Caux-Moncoutier V, Tolozcko-Grabarek A, Gronwald J, Brytski S, Spurdel AB, Bonanni B, Zaffaroni D, Giannini G, Bernard L, Dolcetti R, Manoukian S, Arnold N, Eng C, Deissler H, Rhien K, Niederacher D, Plendl H, Sutter C, Wappenscheidt B, Borg A, Melin B, Rantal A, Soller M, Nathanson KL, Domchek SM, Rodriguez GC, Salani R, Kaulig DG, Teas M-K, Salus LA, Laitman Y, Skytte A-B, Kruse TA, Jensen UB, Robson M, Gerdes A-M, Ejlertsen B, Foretova L, Savage SA, Lester J, Soucy P, Kuchenbaecker KB, Olswold C, Cunningham JM, Slafer S, Pankratz VS, Dics E, Lakhani SR, Couch FJ, Hall P, Monteiro ANA, Gayther SA, Pharoah PDP, Reddel RR, Goode EL, Greene MH, Easton DF, Berchuck A, Antoniou AC, Chenex-Trench G, Dunning AM (2013) Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. Nat Genet 45: 371–384; 3841–e2.

Bolton KI, Tyrer J, Song H, Rams SJ, Notaridou M, Jones C, Sher T, Gentry-Maharaj A, Wozniak E, Tsai Y-Y, Weidhaas J, Paik D, Van Den Berg DJ, Stram DO, Pearce CL, Wu AH, Brewster W, Anton-Culver H, Ziegas A, Narod SA, Levine DA, Kaye SB, Brown R, Paul J, Fanjia I, Sieh W, McGuire V, Whitemore AS, Campbell I, Gore ME, Lissowska J, Yang HP, Medrek E, Gronwald J, Lubinski J, Jakubowska A, Le ND, Cook LS, Kelemen LE, Brook-Wilson A, Massuger LFAG, Kiemeney LA, Aben KKH, van Altena AM, Houlston R, Tomlinson I, Palmieri RT, Moormann PG, Schildkraut J, Iversen ES, Phelan C, Varens K, Cunningham JM, Goode EL, Fridley BL, Kruger-Kjaer S, Blaeker J, Hogdall E, Hogdall C, Gross J, Karlan BY, Ness RB, Edwards RP, Onduski K, Moysich KB, Baker JA, Modugno F, Heikkinnen T, Butzow R, Nevanlinna H, Leimenen A, Bogdanova N, Antonenkovna N, Doerk T, Hillelmans P, Dirst M, Runnebaum I, Thompson PJ, Carney ME, Goodeman MT, Liu G, Wang-Gohrke S, Healey CS, Luccarini C, Peock S, Kuchenbaecker KB, Stefanos A, Birrer MJ, Dunning AM (2013) Multiple independent variants at the TERT locus are associated with susceptibility to ovarian cancer. PloS one 8(11): e78804.

Bowtell DD, Böhm S, Ahmed AA, Aspuria P-J, Bast RC, Beral V, Berek JS, Miller AB, Bo¨hm S, Ahmed AA, Aspuria P-J, Bast RC, Beral V, Berek JS, Ahmed S, Baynes C, Bolla MK, Wang Q, Dennis J, McGuffog L, Chen X, Weischer M, Nielsen SF, Maranian MJ, Ghoussaini M, Ahmed S, Baynes C, Bolla MK, Wang Q, Dennis J, McGuffog L, Barrowdale D, Lee A, Healey S, Lush M, Tessier DC, Vincent D, Baert F, Australian Cancer Study, Australian Ovarian Cancer Study, Kathleen Cunningham Foundation Consortium for Research into Familial Breast cancer (kConFab), Gene Environment Interaction and Breast Cancer (GENICA), Swedish Breast Cancer Study (SWE-BRCA), Hereditary Breast and Ovarian Cancer Research Group Netherlands (HEBON), Epidemiological Study of BRCA1 & BRCA2 Mutation Carriers (EMBRACE), Genetic Modifiers of Cancer Risk in BRCA1/2 Mutation Carriers (GEMO), SIBL, The Lyon, Albert Einstein, Boston, Cambridge, Harvard, Massachusetts, and World Health Organization Collaborative Group (2012) Inferring direct DNA binding from ChIP-seq. Nat Genet 44: 880–884.

Bowtell DD, Böhm S, Ahmed AA, Aspuria P-J, Bast RC, Beral V, Berek JS, Miller AB, Bo¨hm S, Ahmed AA, Aspuria P-J, Bast RC, Beral V, Berek JS, Ahmed S, Baynes C, Bolla MK, Wang Q, Dennis J, McGuffog L, Chen X, Weischer M, Nielsen SF, Maranian MJ, Ghoussaini M, Ahmed S, Baynes C, Bolla MK, Wang Q, Dennis J, McGuffog L, Barrowdale D, Lee A, Healey S, Lush M, Tessier DC, Vincent D, Baert F, Australian Cancer Study, Australian Ovarian Cancer Study, Kathleen Cunningham Foundation Consortium for Research into Familial Breast cancer (kConFab), Gene Environment Interaction and Breast Cancer (GENICA), Swedish Breast Cancer Study (SWE-BRCA), Hereditary Breast and Ovarian Cancer Research Group Netherlands (HEBON), Epidemiological Study of BRCA1 & BRCA2 Mutation Carriers (EMBRACE), Genetic Modifiers of Cancer Risk in BRCA1/2 Mutation Carriers (GEMO), SIBL, The Lyon, Albert Einstein, Boston, Cambridge, Harvard, Massachusetts, and World Health Organization Collaborative Group (2012) Inferring direct DNA binding from ChIP-seq. Nat Genet 44: 880–884.

Bowtell DD, Böhm S, Ahmed AA, Aspuria P-J, Bast RC, Beral V, Berek JS, Miller AB, Bo¨hm S, Ahmed AA, Aspuria P-J, Bast RC, Beral V, Berek JS, Ahmed S, Baynes C, Bolla MK, Wang Q, Dennis J, McGuffog L, Chen X, Weischer M, Nielsen SF, Maranian MJ, Ghoussaini M, Ahmed S, Baynes C, Bolla MK, Wang Q, Dennis J, McGuffog L, Barrowdale D, Lee A, Healey S, Lush M, Tessier DC, Vincent D, Baert F, Australian Cancer Study, Australian Ovarian Cancer Study, Kathleen Cunningham Foundation Consortium for Research into Familial Breast cancer (kConFab), Gene Environment Interaction and Breast Cancer (GENICA), Swedish Breast Cancer Study (SWE-BRCA), Hereditary Breast and Ovarian Cancer Research Group Netherlands (HEBON), Epidemiological Study of BRCA1 & BRCA2 Mutation Carriers (EMBRACE), Genetic Modifiers of Cancer Risk in BRCA1/2 Mutation Carriers (GEMO), SIBL, The Lyon, Albert Einstein, Boston, Cambridge, Harvard, Massachusetts, and World Health Organization Collaborative Group (2012) Inferring direct DNA binding from ChIP-seq. Nat Genet 44: 880–884.
regulates transcriptional changes between ovarian cancer and benign tumors. JCI Insight 1: e92388.

ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. Nature 489: 57–74.

Fletcher MNC, Castro MAA, Wang X, de Santiago I, O’Reilly M, Chinn S-F, Rueda OM, Caldas C, Ponder BAJ, Markowitz F, Meyer KB (2013) Master regulators of FGFR2 signalling and breast cancer risk. Nat Commun 4: 2464.

Freedman ML, Monteiro ANA, Gayther SA, Coetzee GA, Risch A, Plass C, Game D, Biase MA, Carlsson D, Gudagni M, James D, Liu P, Tichelaar JW, Vikis HG, You M, Mills IG (2011) Principles for the post-GWAS functional characterization of cancer risk loci. Nat Genet 43: 513–518.

Goel EL, Chenex-Vitch-Geng, Song H, Ramos SJ, Notariadou M, Lawson K, Widschwendter M, Vierkant RA, Larson MC, Kjaer SK, Birrer MJ, Berchuck A, Schildkraut J, Tomlinson I, Kiemeney LA, Cook LS, Gronwald J, Garcia-Closas M, Gore ME, Campbell I, Whittemore AS, Stuphen R, Phelan C, Anton-Culver H, Pearce CL, Lambrechts D, Ross MA, Chang-Claude J, Moysich KB, Goodman MT, Dörk T, Nevanlinna H, Ness RB, Rafnar T, Hogdall C, Hogdall E, Fridley BL, Cunningham JM, Sieh W, McGuire V, Godwin AK, Cramer DW, Hernandez D, Levine D, Lu K, Iversen ES, Palmiéri RT, Houlston R, van Altena AM, Aben KKH, Massuger LFAG, Brooks-Wilson A, Kelemen LE, ND, Jakubowska A, Lubinski J, Medrek K, Stafford A, Easton DF, Tyrer J, Bolton KL, Harrington P, Eccles D, Chen A, Molina AN, Davila BN, Arango H, Tsai Y-Y, Chen Z, Risck HA, McLauglin J, Narod SA, Ziogas A, Brewer W, Gentry-Maharaj A, Menon U, Wu AH, Stam DO, Pike MC, Wellcome Trust Case-Control Consortium, Beesley J, Webb FM, Australian Cancer Study (Ovarian Cancer), Australian Ovarian Cancer Study Group, Ovarian Cancer Association Consortium (OCAC), Chen X, Ekici AB, Thiel FC, Beckmann MW, Yang H, Wentzensen N, Lissowska J, Fasching PA, Despierre E, Amant F, Verhulst S, Do NT, Hein R, Wang-Gohrke S, Lurie G, Carney ME, Thompson PJ, Runnebaum I, Hillemanns P, Dürst M, Antonenkova N, Bogdanova N, Leminen A, Butzow R, Heikkinen T, Stefansson K, Sulem P, Besenbacher S, Sellers TA, Gayther SA, Pharaoh PDP, Ovarian Cancer Association Consortium (OCAC) (2010) A genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31 and 8q24. Nat Genet 42: 874–879.

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre CH, Ghish C, Clark CK (2010) Simple combinations of line-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38: 576–589.

Hung RJ, Ulrich CM, Goode EL, Braine H, Muir K, Chan AT, Marchand LL, Schildkraut J, Witte JS, Eeles R, Boffetta P, Spitz MR, Poirier JG, Rider DN, Fridley BL, Chen Z, Haiman C, Schumacher F, Easton DF, Landi MT, Brennan P, Houlston R, Christiansen DC, Field JK, Bikkebøller H, Risch A, Kote-Jarai Z, Wildfud L, Grönhag B, Chanoit S, Brittain SL, Kraw K, Lindström S, Al Omara AA, Song H, Phelan C, Wentzensen N, Peters U, Slattery ML, GECCO, Sellers TA, FOCCI, Casey G, Grub SR, CORECT, Hunter DJ, DRIVE, Amos CI, Henderson B. GAME-ON Network (2015) Cross cancer genomic investigation of inflammation pathway for five common cancer types: lung, ovary, prostate, breast, and colorectal cancer. J Natl Cancer Inst 107: dv4246.

Jiang J, Cui W, Vongsangnak W, Hu G, Shen B (2013) Post genome-wide association studies functional characterization of prostate cancer risk loci. BMC Genomics 14(Suppl 8): s9.

Kar SP, Tyrer JP, Li Q, Lawrenson K, Aben KKH, Anton-Culver H, Antonenkovna N, Chenex-Vitch-Geng T, Australian Cancer StudyAustralian Ovarian Cancer Study GroupBaker H, Bandera EV, Bean YT, Beckmann MW, Berchuck A, Bisogno B, Björge L, Bogdanova N, Brinton L, Brooks-Wilson A, Butzow R, Campbell I, Carty KD, Chang-Claude J, Chen YA, Chen Z, Cook LS, Cramer D, Cunningham JM, Cybulski C, Dansonka-Miszewska A, Dennis J, Dicks E, Doherty JA, Dörk T, du Bois A, Dörft M, Eccles D, Easton DF, Edwards RP, Ekici AB, Fasching PA, Fridley BL, Fradley GA, Gao Y-T, Gentry-Maharaj A, Giles GG, Glasspool R, Goode EL, Goodman MT, Gronwald J, Harrington P, Harrington P, Harter P, Hein R, Heitlau C, Houlston R, Christiansen DC, Field JK, Bikkebøller H, Risch A, Kote-Jarai Z, Wildfud L, Grönhag B, Chanoit S, Brittain SL, Kraw K, Lindström S, Al Omara AA, Song H, Phelan C, Wentzensen N, Peters U, Slattery ML, JECO, Sellers TA, FOCCI, Casey G, Grub SR, CORECT, Hunter DJ, DRIVE, Amos CI, Henderson B. GAME-ON Network (2015) Cross cancer genomic investigation of inflammation pathway for five common cancer types: lung, ovary, prostate, breast, and colorectal cancer. J Natl Cancer Inst 107: dv4246.
Kuchenbaecker KB, Ramus SJ, Tyrer J, Lee A, Shen HC, Beesley J, Lawrenson K, Li Q, Kar S, Seo J-H, Tyrer J, Spindler TJ, Lee J, Chen Y, Piedmonte M, O'Malley DM, de la Hoya M, Caldes T, Aittomäki K, Stoppa-Lyonnet D, Claes K, De Leeneer K, Kirk J, Rodriguez GC, EMBRACE Godwin AK, Schmutzler RK, Meindl A, Engel C, Sutter C, Bonanni B, Peissel B, Manoukian S, Radice P, Papi L, Ottini L, Fostira F, Komenaka I, Weitzel JN, Ganschow P, Peterlongo P, Bernard L, Viel A, Neuhausen SL, Ding YC, Hansen TVO, Jønson L, Gerdes A-M, Terry MB, Goldgar DE, Buys SS, Janavicius R, Dorfling CM, van Rensburg EJ, Sellers TA, Monteiro ANA, Freedman ML, Gayther SA, Pharoah PDP, Wilkens LR, Woo Y-L, Wu X, Wu AH, Yang H, Zheng W, Ziogas A, Wang-Gohrke S, Walsh C, Wentzensen N, Whittemore AS, Wicklund KG, Wilkens LR, Woo Y-L, Wu X, Wu A, Yang H, Zheng W, Zioagas A, Sellers TJA, Monteiro ANA, Freedman ML, Gayther SA, Pharoah PDP, Wilkens LR, Chang-Claude J, Wang-Gohrke S, Eilber U, Campbell K, Suh E, Duggan D, Carpten J, Trent JM, Gronberg H, Zheng SL, Isaacs WB, Xu J (2012) Association of prostate cancer risk with SNPs in regions containing androgen receptor binding sites captured by ChIP-On-chip analyses. The Prostate 72: 376–385.

Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, Reuter I, Chemennev D, Kruhl M, Hornischer K, Voss N, Stengmaier P, Lewicki-Potapov B, Saxel H, Kel AE, Wingender E (2006) TRANSFAC: a web-based transcriptional regulation database for eukaryotes. Nucleic Acids Res 34: D108–D110.

Neph S, Kuehn MS, Reynolds AP, Haugen E, Thurman RE, Johnson AK, Neph S, Kuehn MS, Reynolds AP, Haugen E, Thurman RE, Johnson AK, Rynes E, Maurano MT, Vierstra J, Thomas S, Sandstrom R, Humbert R, Dinger MK, Remm M, Tusnady GE, Iyer VS, Sandelin A, Liang D, Hildebrandt MAT, Wu X, Levine DA, Bisogna M, Berchuck A, Goode EL, Cunningham JM, Vierkant RA, Bruinsma F, Giles GG, Glasspool R, Goode EL, Cunningham JM, Vierkant RA, Wang-Gohrke S, Walsh C, Wentzensen N, Whittemore AS, Wicklund KG, Wilkens LR, Woo Y-L, Wu X, Wu AH, Yang H, Zheng W, Zioagas A, Monteiro A, Pharoah PD, Gayther SA, Freedman ML (2015) cis-eQTL analysis and functional validation of candidate susceptibility genes for high-grade serous ovarian cancer. Nat Commun 6: 8234.

Lee PH, O'Duccio J, Pu C, Thomas B, Purcell SM (2012) INRICH: interval-based enrichment analysis for genome-wide association studies. Bioinformatics 28: 1799–1797.

Lu Y, Sun J, Kader AK, Kim S-T, Kim J-W, Liu W, Sun J, Lu D, Feng J, Zha Y, Jin T, Zhang Z, Dimitrov L, Lowey J, Campbell K, Suh E, Duggan D, Carpten J, Trent JM, Gronberg H, Zheng SL, Isaacs WB, Xu J (2012) Association of prostate cancer risk with SNPs in regions containing androgen receptor binding sites captured by ChIP-On-chip analyses. The Prostate 72: 376–385.

Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, Reuter I, Chemennev D, Kruhl M, Hornischer K, Voss N, Stengmaier P, Lewicki-Potapov B, Saxel H, Kel AE, Wingender E (2006) TRANSFAC: a web-based transcriptional regulation database for eukaryotes. Nucleic Acids Res 34: D108–D110.

Neph S, Kuehn MS, Reynolds AP, Haugen E, Thurman RE, Johnson AK, Neph S, Kuehn MS, Reynolds AP, Haugen E, Thurman RE, Johnson AK, Rynes E, Maurano MT, Vierstra J, Thomas S, Sandstrom R, Humbert R, Dinger MK, Remm M, Tusnady GE, Iyer VS, Sandelin A, Liang D, Hildebrandt MAT, Wu X, Levine DA, Bisogna M, Berchuck A, Goode EL, Cunningham JM, Vierkant RA, Wang-Gohrke S, Walsh C, Wentzensen N, Whittemore AS, Wicklund KG, Wilkens LR, Woo Y-L, Wu X, Wu AH, Yang H, Zheng W, Zioagas A, Monteiro A, Pharoah PD, Gayther SA, Freedman ML (2015) cis-eQTL analysis and functional validation of candidate susceptibility genes for high-grade serous ovarian cancer. Nat Commun 6: 8234.

Lee PH, O'Duccio J, Pu C, Thomas B, Purcell SM (2012) INRICH: interval-based enrichment analysis for genome-wide association studies. Bioinformatics 28: 1799–1797.

Lu Y, Sun J, Kader AK, Kim S-T, Kim J-W, Liu W, Sun J, Lu D, Feng J, Zha Y, Jin T, Zhang Z, Dimitrov L, Lowey J, Campbell K, Suh E, Duggan D, Carpten J, Trent JM, Gronberg H, Zheng SL, Isaacs WB, Xu J (2012) Association of prostate cancer risk with SNPs in regions containing androgen receptor binding sites captured by ChIP-On-chip analyses. The Prostate 72: 376–385.

Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, Reuter I, Chemennev D, Kruhl M, Hornischer K, Voss N, Stengmaier P, Lewicki-Potapov B, Saxel H, Kel AE, Wingender E (2006) TRANSFAC: a web-based transcriptional regulation database for eukaryotes. Nucleic Acids Res 34: D108–D110.
Narod SA, Pharoah PDP, Phelan CM, Qu X, Ramus SJ, Risch H, Schildkraut JM, Song H, Stockwell H, Suphen R, Terry KL, Tyer J, Vierkant RA, Wentzensen N, Lancaster JM, Cheng JQ, Sellers TA, on behalf of the Ovarian Cancer Association Consortium (OCAC) (2011) LIN28B polymorphisms influence susceptibility to epithelial ovarian cancer. Cancer Res 71: 3896–3903.

Perrin-Wej Y, Lawrenson K, Shen HC, Velkova A, Tyer JP, Chen Z, Lin H-Y, Chen YA, Tsai Y-Y, Qu X, Ramus SJ, Karevan R, Lee J, Lee N, Larson MC, Aben KK, Anton-Culver H, Antonenko N, Antoniou AC, Armas SM. Australian Cancer StudyAustralian Ovarian Cancer Study, Bacot F, Baglietto L, Bandera EV, Barnholtz-Sloan JS, Beckmann MW, Birrer MJ, Bloom G, Bogdanova N, Brinton LA, Broaks-Arman SR, Brennan P, Brown R, Breslau N, Johnson EO, Hatsukami D, Pomerleau O, Swan GE, Goate AM, Rutter J, Bertelsen S, Fox L, Fugman D, Martin NG, Montgomery GW, Wang JC, Ballinger DG, Rice JP, Bierut LJ (2007) Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. Hum Mol Genet 16: 36–49.

Schmid R, Baumbach P, Ittrich C, Fundel-Clemens K, Huber W, Breslow N, Eils R, Weith A, Mennerich D, Quast K (2010) Comparison of normalization methods for Illumina BeadChipHumanHT-12 v3. BMC Genomics 11: 349.

Segre AV. DIAGRAM ConsortiumMAGIC investigatorsGroup L, Mootha VK, Daly MJ, Altschuler D (2010) Common inherited variation in mitochondrial genes is enriched for associations with type 2 diabetes or related glycemic traits. PLoS Genet 6: e1001058.

Song H, Ramus SJ, Tyer J, Bolton KL, Gentry-Maharaj A, Woiziek E, Anton-Culver H, Chang-Claude J, Cramer DW, DiCioccio R, Risch HA, Goode EL, Goodman MT, Schumacher FR, Eeles RA, Easton DF, Seminara D, Amos CI, Malhotra J, Risch A, Heinrich J, Hunter DJ, Henderson BE, Haiman CA, Modugno F, Moysich KB, Nakanishi T, Narod SA, Ness RB, McLaughlin JR, Bengtsson C, von Heyningen V, Bickmore W (2013) Regulation from a distance: long-range receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. Hum Mol Genet 16: 36–49.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 4.0 Unported License.
PAX8 and ovarian cancer susceptibility

BRITISH JOURNAL OF CANCER

www.bjcancer.com | DOI:10.1038/bjc.2016.426

535

Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)