Epigenetic analysis leads to identification of HNF1B as a subtype-specific susceptibility gene for ovarian cancer

Hui Shen, Brooke L. Fridley, Honglin Song et al.

HNF1B is overexpressed in clear cell epithelial ovarian cancer, and we observed epigenetic silencing in serous epithelial ovarian cancer, leading us to hypothesize that variation in this gene differentially associates with epithelial ovarian cancer risk according to histological subtype. Here we comprehensively map variation in HNF1B with respect to epithelial ovarian cancer risk and analyse DNA methylation and expression profiles across histological subtypes. Different single-nucleotide polymorphisms associate with invasive serous (rs7405776 odds ratio (OR) = 1.13, \( P = 3.1 \times 10^{-10} \)) and clear cell (rs11651755 OR = 0.77, \( P = 1.6 \times 10^{-8} \)) epithelial ovarian cancer. Risk alleles for the serous subtype associate with higher HNF1B-promoter methylation in these tumours. Unmethylated, expressed HNF1B, primarily present in clear cell tumours, coincides with a CpG island methylator phenotype affecting numerous other promoters throughout the genome. Different variants in HNF1B associate with risk of serous and clear cell epithelial ovarian cancer; DNA methylation and expression patterns are also notably distinct between these subtypes. These findings underscore distinct mechanisms driving different epithelial ovarian cancer histological subtypes.
Invasive epithelial ovarian cancer (EOC) has a strong heritable component, with an approximate three-fold increased risk associated with a first-degree family history. Much of the excess familial risk observed for EOC is unexplained, and efforts to identify common susceptibility genes have proven to be difficult. Seven regions harbouring susceptibility single-nucleotide polymorphisms (SNPs) for ovarian cancer have been identified through genome-wide association studies thus far, but candidate gene studies have been largely unsuccessful.

The Cancer Genome Atlas (TCGA) has fully characterized more than 500 serous EOC cases with respect to somatic mutation, DNA methylation, mRNA expression and germline genetic variants. These data are publicly available and can be analysed to identify candidate genes for association studies of the disease.

We conducted such an analysis of TCGA data and found a unique expression and methylation pattern of HNF1B characterized by downregulation of expression in most cases, with epigenetic silencing in about half of the cases, suggesting it might have a role in the serous subtype of ovarian cancer. In contrast, HNF1B overexpression is common in clear cell ovarian cancer.

The HNF1B gene (formerly known as TC2F2) encodes a POU-domain containing a tissue-specific transcription factor, and mutations in the gene cause maturity onset diabetes of the young type 5 (ref. 11). HNF1B is also a susceptibility gene for type II diabetes, prostate cancer and uterine cancer.

We report here on our comprehensive characterization of this gene in ovarian cancer and show evidence of a differential effect of HNF1B on the serous and clear cell subtypes of ovarian cancer. It appears that HNF1B has a loss-of-function role in serous and a gain-of-function role in clear cell ovarian cancers, and variants in this gene differentially affect genetic susceptibility to these subtypes.

Results

DNA methylation/expression analysis. From TCGA data (see Methods), HNF1B was observed to be epigenetically silenced in approximately half of the 576 primary serous ovarian tumours and downregulated by another mechanism in most of the other tumours, whereas no evidence of methylation was seen in the normal fallopian tube samples available from TCGA. We further assessed HNF1B-promoter methylation in an independent data set (OCRF panel; see Methods) and found the promoter region to be methylated in 42% of serous tumours and in none of the clear cell ovarian tumours. The pattern in serous tumours, in contrast to clear cell cancers, led to the evaluation of HNF1B as a candidate subtype-specific susceptibility gene for ovarian cancer.

DNA methylation and protein expression. The identification of HNF1B as a susceptibility gene for serous and clear cell ovarian cancer led us to further evaluate the relationship between HNF1B-promoter DNA methylation, protein expression and histological subtype. Immunohistochemistry (IHC) analysis for HNF1B protein expression in 1,149 ovarian cancers from the Ovarian Tumor Tissue Analysis Consortium, and DNA-methylation analysis on 269 of these tumours, revealed that the majority of clear cell tumours expressed the HNF1B protein and were unmethylated at the HNF1B promoter, whereas the majority of serous tumours lacked HNF1B protein expression and displayed frequent HNF1B-promoter methylation.

Although most clear cell tumours were devoid of HNF1B-promoter methylation, they revealed a surprisingly high frequency of CpG island hypermethylation at other sites across the genome, indicative of a CpG island methylator phenotype (CIMP). The few clear cell tumours lacking HNF1B expression exhibited HNF1B-promoter methylation, and a correspondingly low frequency of CpG island methylation throughout the genome, similar to the serous subtype.

Overexpression of HNF1B. Given the proposed role of HNF1B in clear cell tumorigenesis, we stably overexpressed the gene in immortalized endometriosis epithelial cells.

DNA methylation and genotype. We further investigated the relationship between risk allele genotypes and HNF1B DNA methylation in 231 serous ovarian cancers. The top serous risk SNP, rs7405776, showed only a borderline association with increased promoter methylation (P = 0.07; Fig. 3). Intriguingly, the association between SNPs in HNF1B and HNF1B-promoter DNA methylation strengthened as their location approached the promoter region, and the strongest signal came from a few SNPs, exemplified by rs11651755, overlapping with a polycomb repressive complex 2 (PRC2) mark in embryonic stem cells (P = 0.003; Fig. 3, Supplementary Fig. S8). We validated this SNP–methylation association in the TCGA data (Supplementary Fig. S9; see Methods). None of the probes used contained common SNPs in the sequence, excluding technical artifact as a confounder of this association.
hypothesized to be a cell of origin for clear cell ovarian cancers (Supplementary Fig. S10)\(^1\). EECs overexpressing HNF1B acquired an enlarged, flattened morphology and multi-nucleated cells accumulated in the cultures (Fig. 4a). Also, significant upregulation of HNF1B-associated genes SPP1, DPP4, and ACE2 was observed upon HNF1B overexpression in EECs (Fig. 4b).

**Discussion**

*HNF1B* appears to have a prominent role in ovarian cancer aetiology. It is the first clear cell ovarian cancer-susceptibility gene identified, and variation in the gene is also associated with risk of serous ovarian cancer at a genome-wide significance level. The gene is overexpressed in clear cell tumours and silenced in serous tumours. The strong association between HNF1B expression and CIMP methylation (\(P = 3 \times 10^{-10}\)), and the reciprocal nature of DNA methylation at the HNF1B-promoter CpG islands, versus other CpG islands across the genome, suggests that HNF1B-promoter methylation is not merely a CIMP passenger event; in fact, HNF1B expression may even contribute to the hypermethylation phenotype. Taken together, these data indicate differing roles for *HNF1B* in these invasive EOC.

---

**Figure 1** | Identification of *HNF1B* as a subtype-specific candidate gene for ovarian cancer and its establishment as a susceptibility gene. (a) The scatterplot compares the mRNA expression (y axis) versus DNA methylation (x axis) in serous ovarian tumours from TCGA (see Methods). Each blue dot is a serous tumour sample, whereas each pink dot is one of the ten normal fallopian tube samples. The HNF1B promoter is silenced in the majority of these tumours, either by an epigenetic (bottom right, high DNA methylation and low mRNA expression) or an unknown alternative mechanism. The mRNA expression data were integrated from three platforms (online Methods) and interpreted as log ratios, and we observe the same pattern with each individual expression platform (Supplementary Fig. S1). (b) HNF1B-promoter DNA methylation differs by histological subtype. Although unmethylated in the normal fallopian tissue, this locus is hypermethylated (beta value > 0.2) in approximately 50% of the TCGA (\(n = 576\); see Methods) serous cases as well as another independent set of 32 serous tumour samples (OCRF panel; see Methods), but remains unmethylated in clear cell tumours (OCRF panel; see Methods) (\(n = 4\)). These data are consistent with reported HNF1B expression in the clear cell tumours. (c) Genetic variants in the HNF1B locus are associated with risk of ovarian cancer histological subtypes. Plotted in each panel is the \(-\log_{10}(P\text{-value})\) from the SNP association with risk for each subtype (Manhattan plots) located in the 150-kb region described in the text. Imputed SNPs are indicated with a relatively lighter colour, whereas the genotyped SNPs are indicated with a darker colour. Dashed lines indicate the genome-wide significance threshold (5\(\times\)10\(^{-8}\)). The linkage disequilibrium plot on the bottom shows the \(r^2\) between the SNPs. Genomic coordinates are based on hg19 (Build37).
methylation is not a passenger event of global DNA-methylation changes. This suggests loci is highly correlated HNF1B expression. Also noteworthy is that the subtypes (panel B, ordered by average DNA methylation across the locus, either from the region flanking the HNF1B transcription start site normal ovarian tissue (4)

Table 1 | Association between invasive, serous and clear cell ovarian cancer for ten HNF1B SNPs that reached genome-wide significance in Whites.

| Reference/ alternate allele | Imputed $r^2$ | AAF | OR | 95% CI | $P$-value | OR | 95% CI | $P$-value | OR | 95% CI | $P$-value |
|-----------------------------|-------------|-----|-----|-------|---------|-----|-------|---------|-----|-------|---------|
| Univariate                  |             |     |     |       |         |     |       |         |     |       |         |
| rs3744763*                  | G/A         | 0.902 | 0.040 | 1.06 | 1.03 – 1.10 | 1.6 x 10^-4 | 1.13 | 1.09 – 1.17 | 4.0 x 10^-10 | 0.93 | 0.87 – 0.99 | 2.2 x 10^-7 |
| rs1736092641*               | G/CT        | 0.47 | 1.03 | 1.01 – 1.08 | 0.005 | 1.02 | 1.00 – 1.04 | 0.001 | 1.13 | 1.09 – 1.17 | 3.1 x 10^-9 | 0.85 | 0.79 – 0.88 | 3.8 x 10^-6 |
| rs7405776*                  | G/A         | 0.376 | 1.05 | 1.02 – 1.09 | 0.0001 | 1.13 | 1.09 – 1.17 | 3.1 x 10^-9 | 0.79 | 0.72 – 0.78 | 1.0 x 10^-5 | 0.80 | 0.73 – 0.82 | 4.1 x 10^-6 |
| rs757250*                   | C/T         | 0.372 | 1.05 | 1.02 – 1.09 | 9.5 x 10^-4 | 1.13 | 1.09 – 1.17 | 3.2 x 10^-10 | 0.80 | 0.73 – 0.82 | 4.1 x 10^-6 | 0.80 | 0.73 – 0.82 | 4.1 x 10^-6 |
| rs4239377*                  | A/G         | 0.402 | 1.04 | 1.01 – 1.07 | 0.018 | 1.11 | 1.07 – 1.16 | 2.6 x 10^-8 | 0.76 | 0.72 – 0.78 | 1.0 x 10^-5 | 0.78 | 0.72 – 0.82 | 1.0 x 10^-5 |
| rs1651755*                  | T/C         | 0.489 | 1.02 | 0.99 – 1.06 | 0.124 | 1.10 | 1.06 – 1.14 | 9.9 x 10^-7 | 0.77 | 0.70 – 0.82 | 1.6 x 10^-8 | 0.79 | 0.68 – 0.92 | 0.008 |
| rs61618281                  | G/A         | 0.0827 | 0.140 | 1.09 | 1.04 – 1.14 | 4.1 x 10^-4 | 1.19 | 1.13 – 1.24 | 1.1 x 10^-9 | 0.79 | 0.68 – 0.92 | 0.008 |
| rs31657964*                 | G/A         | 0.400 | 1.04 | 1.01 – 1.08 | 0.006 | 1.12 | 1.08 – 1.16 | 5.3 x 10^-9 | 0.80 | 0.73 – 0.82 | 4.6 x 10^-6 | 0.80 | 0.73 – 0.82 | 4.6 x 10^-6 |
| rs7507939*                  | C/T         | 0.400 | 1.04 | 1.01 – 1.08 | 0.006 | 1.12 | 1.08 – 1.16 | 5.3 x 10^-9 | 0.80 | 0.73 – 0.82 | 4.6 x 10^-6 | 0.80 | 0.73 – 0.82 | 4.6 x 10^-6 |
| rs16658063                  | G/C         | 0.963 | 0.398 | 1.05 | 1.02 – 1.08 | 0.003 | 1.12 | 1.08 – 1.17 | 1.8 x 10^-9 | 0.81 | 0.74 – 0.90 | 2.3 x 10^-5 |

AAF, alternative allele frequency; CI, confidence interval.

Different SNPs in the HNF1B gene regions explain the associations observed for serous and clear cell ovarian cancers. These different effects provide further support for the growing view that the histological subtypes of ovarian cancer represent distinct diseases, with endometriosis as a proposed cell of origin for clear cell disease and fallopian tube fimбриae as one for serous disease. Interestingly, no association was observed between HNF1B genotypes and endometrioid ovarian cancer despite the view that, like clear cell, endometriosis is also a cell of origin for this subtype. The lack of association may be due to a different transformation mechanism from endometriosis for the endometrioid subtype, given that although the HNF1B promoter remains unmethylated in the endometrioid subtype, the endometrioid subtype does not overexpress HNF1B. Alternatively, misclassification of high-grade serous EOC as high-grade endometrioid could result in a bias towards the null for the endometrioid subtype.

Figure 2 | HNF1B-promoter DNA methylation, protein expression and global DNA-methylation pattern by subtype. Each row is a tissue sample collected at the Mayo Clinic that belongs to one of the three categories: normal ovarian tissue (n = 7), clear cell ovarian tumours (n = 17) or serous ovarian tumours (n = 196). Endometrioid (n = 49) and mucinous (n = 7) tumours are not included in this figure. Each column represents a CpG locus, either from the region flanking the HNF1B transcription start site (panel A, ordered by genomic locations with an arrow indicating the transcription start site) or from a global panel of 1,003 CpG loci mapped to autosomal CpG island regions that distinguish clear cell and serous subtypes (panel B, ordered by average DNA methylation across the samples). For each horizontal panel group, the samples (rows) are ordered by HNF1B IHC status. The heatmap shows the DNA-methylation beta value, with blue indicating low DNA methylation and red indicating high methylation. Clear cell tumours showed less DNA methylation at the HNF1B-promoter region and correspondingly higher HNF1B protein expression. The clear cell tumours generally show a CIMP pattern where there is extensive gain of aberrant promoter methylation in a correlated manner. CIMP status (left side bar, defined as methylated at >80% of the 1,003 loci) is highly correlated HNF1B expression. Also noteworthy is that the HNF1B-promoter DNA methylation (panel a) is the opposite from the global pattern (panel b, Supplementary Fig. S8). This suggests HNF1B DNA methylation is not a passenger event of global DNA-methylation changes.

subtypes: a potential gain-of-function in clear cell ovarian cancer and loss-of-function in serous ovarian cancer, underscoring the heterogeneity of this disease.
DNA methyltransferase cross-talk has been proposed to be a mechanism of predisposition to cancer-specific hypermethylation. Our DNA-methylation data indicate that the causal risk alleles for the serous subtype may predispose the promoter to acquiring aberrant methylation, thereby promoting the development of serous but not clear cell tumours. This predisposition could be a direct functional effect of the SNP on the DNA-methylation machinery, or could act indirectly through differential binding affinity for PRC2 or one or more transcription factors. Given that we were able to fine-map the HNF1B region, it is unlikely that an unidentified common variant explains these associations. For serous ovarian cancer, the methylation signal suggests that the causal variant is most likely to be among those located within the region with the PRC2 mark for which we identified five SNPs with genome-wide significance.

This is the first study investigating the effects of overexpression of HNF1B in endometriosis, and the results support the hypothesis that HNF1B may have an oncogenic role in the initiation of clear cell ovarian cancers, as speculated by Gounaris et al. as a key step of endometriosis transformation. The observation in our data that HNF1B induces a polynucleated phenotype in EEC cells is intriguing, as clear cell ovarian cancers are often tetradiploid, more so than other ovarian cancer subtypes. The polynucleated phenotype may suggest that HNF1B overexpression in EECs perturbs cytokinesis, causing aneuploidy in some cells.

Histology re-review of the three clear cell tumours that do not express HNF1B revealed two scenarios: two samples with inconsistent evaluations between pathologists, and one consistently called clear cell. They might be cases that are especially difficult to classify, and therefore a molecular signature, for example, CIMP or HNF1B status, would be of great help in correctly classifying those tumours. The one sample that is called consistently clear cell tumour but does not express HNF1B might represent a rare subtype of clear cell carcinoma. With a larger cohort of clear cell ovarian cancers, these possibilities can be investigated.

To our knowledge, this is the first report of tumour DNA-methylation patterns leading to the identification of a germline susceptibility locus, underscoring the value of TCGA. Recent studies suggest a strong genetic component to inter-individual variation in tumour DNA methylation, and demonstrate both cis- and trans- associations between genotypes and DNA methylation. In addition, methylation quantitative trait loci were found to be enriched for expression quantitative trait loci. It has also been shown that epimutation is associated with genetic variation, for example, associations have been demonstrated between 5’ UTR MLH1 variants and MLH1 epigenetic silencing. Moreover, we have for the first time demonstrated the existence of a CIMP phenotype in ovarian cancer, highlighting the complicated nature of the disease.

In summary, variation in HNF1B is associated with serous and clear cell subtypes of ovarian cancer in opposite manner at genetic, epigenetic and protein expression levels. These observations are compatible with a tumour suppressor role in serous cancer and an oncogenic role in clear cell disease. Future efforts
should focus on understanding these mechanisms as they may have major clinical implications for ovarian cancer, based on better subtype stratification, potential novel treatment approaches and a better understanding of disease aetiology. Currently, effective chemotherapeutics for clear cell ovarian cancer is lacking, but our study reveals that HNF1B-expressing clear cell tumours have extensive epigenetic alterations that potentially make them good candidates for epigenetic therapies.

Methods

Molecular aspects

TCGA data access. We downloaded the TCGA serous ovarian cancer data packages from the TCGA public-access ftp (ftp://tcga-data.nci.nih.gov/tcGAfiles/ftp_auth/distro_ftpusers/anonymous/tumour/ov/). Data generated with the following platforms were used: Affymetrix HT Human Genome U133 Array Plate Set; Agilent 244K Custom Gene Expression G4502A-07-3; Affymetrix Human Exon 1.0 ST Array; and Illumina Infinium HumanMethylation27 Beadchip (a full list of the packages is provided in Supplementary Methods). The Illumina HumanMethylation1M-Duo DNA Analysis BeadChip Genotype data were downloaded from the controlled access data tier.

DNA methylation data production for the OCRF tumour panel. The Illumina Infinium HumanMethylation27 assay was performed as described on 32 serous and 4 clear cell ovarian tumours from USC Norris Comprehensive Cancer Center and Duke University (OCCRF tumour panel). The beta values for each sample and locus were calculated with mean non-background corrected methylated (M) and unmethylated (U) signal intensities with the formula M/(M+U), representing the percentage of methylated alleles. Detection P-values were calculated by comparing the set of analytical probe replicates for each locus to the set of 16 negative control probes. Data points with detection P-values > 0.05 were masked.

DNA methylation data production for the Mayo tumour panel. We also performed the Infinium HumanMethylation450 BeadChip assay on an independent set of tumour DNA in the Mayo Clinic Genotyping Shared Facility using recommended Illumina protocol. 1 μg of tumour DNA was bisulfite-converted using the Zymo EZ96 DNA Methylation Kit. Three samples failing quality control were removed, leaving DNA-methylation data on 333 ovarian cancer cases, including 254 serous and 17 clear cell tumours. Plate normalization was done with a linear model on the logit-transformed beta values, following back-transformation to the (0,1) range.

IHC assay. Previously built tissue microarrays, triplicate core, measuring 0.6 mm were cut at 4-μm thickness and mounted on superfrost slides. Slides were stained on a Ventana Benchmark XT using the manufacturer’s pretreat protocol CC1 standard (Supplementary Methods). A pathologist (MK) evaluated the IHC staining, and assigned the sample a score 0 in the absence of any nuclear staining, score 1 for any nuclear staining >1-50% or score 2 for >50% tumour cell nuclei-positive for HNF1B.

Genotype and DNA methylation association. We assessed the correlation of germline genotype at the nine genome-wide significant SNPs in serous cancer, with HNF1B DNA promoter methylation status using the Mayo Tumour Panel. Probe cg14487292 was used as it was most inversely correlated with mRNA expression. The nominal P-values are from two-sided tests for linear trend in the DNA-methylation beta values across the three genotypes for each locus. Bonferroni adjustment was not done for multiple comparisons as the SNPs are highly correlated. Validation was done with the TCGA data (Supplementary Appendix).

In vitro model of HNF1B overexpression. An immortalized EEC line was generated by lentiviral transduction of HRT (Addgene plasmid 12245) into primary EECs (Supplementary Fig. S10). HRT-immortalized EECs were transduced with lentiviral HNF1B-green fluorescent protein (GFP) or GFP (Genecopoeia) supernatants and positive cells selected with 400 ng ml⁻¹ puromycin (Sigma). GFP expression was confirmed by fluorescent microscopy; HNF1B expression was confirmed by real-time PCR (Supplementary Fig. S10).

For gene-expression studies, RNA was collected from cells using the Qiagen RNeasy kit with on-column DNase I digestion. An amount of 1 μg RNA was reverse transcribed using an MMLV reverse transcriptase enzyme (Promega), and relative mRNA level was assayed using the ABI 7900HT Fast Real-Time PCR.

Figure 4 | Phenotypic effects and downstream targets of HNF1B overexpression in immortalized EECs. (a) Morphological changes in EECs expressing a HNF1B GFP fusion protein (EEC(GFP)HNF1B). GFP-positive cells were sorted using flow cytometry. The arrows indicate five nuclei contained within a single EEC(GFP)HNF1B cell, showing the aberrant polynucleation that we observed in these cells. Using flow cytometry, we quantified the increase in polynucleation in EEC(GFP)HNF1B to be around eightfold compared with controls (data not shown). (b) Gene-expression analysis of HNF1B-target genes and clear cell ovarian cancer associated genes. *P < 0.01.
system utilizing the delta-delta Ct method. Statistical analyses were performed using Prism. Two-tailed paired t-tests with significance level of 0.05 were used.

Genetic association study

Study design. The genetic susceptibility aspect of this study was organized by the Collaborative Ovarian Gene-Environment Study, an ovarian, breast and prostate cancer consortium. The ovarian cancer part of this effort on which the current report is based is led by the Ovarian Cancer Association Consortium and included 43 studies (Supplementary Table S1). Following sample quality control, 44,308 subjects, including 16,111 patients with invasive EOC, 2,063 with low malignant potential (borderline) disease and 26,134 controls, were available for analysis; results presented here are restricted to invasive cancers. All studies obtained approval from their respective human research ethics committees, and all participants provided written informed consent.

Selection of SNPs. Data for 174 SNPs in this region were available from the Collaborative Ovarian Gene-Environment Study genotyping effort and provided full fine-mapping information in the 150-kb region surrounding HNF1B (hg18 coordinates 33,100,000–33,250,000). In addition, phase I haplotype data from the 1000 Genomes Project (January 2012) were used to implicate genotypes for SNPs across this region, resulting in available data on an additional 307 SNPs with MAF >0.02 in European Whites and imputation r2 >0.30 (IMPUTE 2.2).

SNP genotyping. The Ovarian Cancer Association Consortium genotyping was conducted by McGill University and Genome Quebec Innovation Centre (n = 19,806) and the Mayo Clinic Medical Genome Facility (n = 27,824) using an Illumina Infinium Omni 2.5 BeadChip. Genotypes were called using GenCall. Sample and SNP quality-control measures are described in the Supplementary Methods.

Statistical analysis. We used the program LAMP30 for principal components analysis to assign intercontinental ancestry based on the HapMap (release no. 22) genotype frequency data for European, African and Asian populations (Supplementary Methods). For LAMP-derived European ancestry groups for all patients of invasive cancer and for those with serious invasive cancer, we carried out unconditional logistic regression analyses within each study site, adjusted for the first five eigenvalues from the principal components analysis for European ancestry and then used a fixed-effects meta-analytic approach to obtain the summary OR estimate, 95% confidence interval and P-value. Details on analysis for the non-European groups are provided in the Supplementary Methods. Log-additive mode of inheritance was modelled (that is, co-dominant), treating each SNP as an ordinal variable.

For haplotype analysis, we used the tagSNPs program31 to obtain the haplotype dosage for each subject for the LAMP-derived European ancestry group for haplotypes with a frequency of ≥1%. The associations between haplotype and risks of serous and clear cell ovarian cancer were modelled by meta-analysis relative to the most common haplotype.

References

1. Lichtenstein, P. et al. Environmental and heritable factors in the causation of cancer—analyses of cohorts from twins of Sweden, Denmark, and Finland. N. Engl. J. Med. 343, 78–85 (2000).
2. Auranen, A. et al. Cancer incidence in the first-degree relatives of ovarian cancer patients. Br. J. Cancer 74, 280–284 (1996).
3. Antoniou, A. C. & Easton, D. F. Risk prediction models for familial breast cancer. Future Oncol. 2, 257–274 (2006).
4. Pharoah, P. D. P. et al. GWAS meta-analysis and replication identifies three novel common susceptibility loci for ovarian cancer. Nat. Genet. (e-pub ahead of print 27 March 2013; doi:10.1038/ng.2564) (2013).
5. Goede, L. E. et al. A genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31.1 and 8q24. Nat. Genet. 42, 874–879 (2010).
6. Bolton, K. L. et al. Common variants at 19p13 are associated with susceptibility to ovarian cancer. Nat. Genet. 42, 880–884 (2010).
7. Song, H. et al. A genome-wide association study identifies a new ovarian cancer susceptibility locus on 8p22.2. Nat. Genet. 41, 996–1000 (2009).
8. Bolton, K. L., Ganda, C., Berchuck, A., Pharoah, P. D. & Gayther, S. A. Roberts et al. common variants in ovarian cancer susceptibility and outcome: progress to date from the Ovarian Cancer Association Consortium (OCAC). J. Intern. Med. 271, 366–378 (2012).
9. Cancer Genome Atlas Network. Integrated genome analyses of ovarian carcinoma. Nature 474, 609–615 (2011).
10. Tsuchiya, A. et al. Expression profiling in ovarian clear cell carcinoma: identification of hepatocyte nuclear factor-1 beta as a molecular marker and a possible molecular target for therapy of ovarian clear cell carcinoma. Am. J. Pathol. 163, 2503–2512 (2003).
11. Horikawa, Y. et al. Mutation in hepatocyte nuclear factor-1 beta gene (TCF2) associated with MODY. Nat. Genet. 17, 384–385 (1997).
12. Gudmundsson, J. et al. Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. Nat. Genet. 39, 977–983 (2007).
13. Winckler, W. et al. Evaluation of common variants in the six known maturity-onset diabetes of the young (MODY) genes for association with type 2 diabetes. Diabetes 56, 685–693 (2007).
14. Berndt, S. I. et al. Large-scale fine mapping of the HNF1B locus and prostate cancer risk. Hum. Mol. Genet. 20, 3322–3329 (2011).
15. Sun, J. et al. Evidence for two independent prostate cancer risk-associated loci in the HNF1B gene at 17q12. Nat. Genet. 40, 1153–1155 (2008).
16. Thomas, G. et al. Multiple loci identified in a genome-wide association study of prostate cancer. Nat. Genet. 40, 310–315 (2008).
17. Spurle, A. B. et al. Genome-wide association study identifies a common variant closely linked with risk of endometrial cancer. Nat. Genet. 43, 451–454 (2011).
18. Pearce, C. L. et al. Association between endometriosis and risk of histological subtypes of ovarian cancer: a pooled analysis of case-control studies. Lancet Oncol. 13, 385–394 (2012).
19. Kurman, R. J. & Shah, J. L. The origin and pathogenesis of epithelial ovarian cancer: a proposed unifying theory. Am. J. Surg. Pathol. 34, 433–443 (2010).
20. Sankararaman, S., Sridhar, S., Kimmel, G. & Halperin, E. Clinical and biological characteristics of clear cell carcinomas of the ovary in FIGO stages I-II. Int. J. Oncol. 26, 177–183 (2005).
21. Bell, J. T. et al. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. Genome Biol. 12, R10 (2011).
22. Hitchins, M. P. et al. Dominantly inherited constitutional epigenetic silencing of MLH1 in a cancer-affected family is linked to a single nucleotide variation within the 5' UTR. Cancer Cell 20, 200–213 (2011).
23. Tsuchiya, A. et al. Cancer incidence in the first-degree relatives of ovarian cancer patients. Br. J. Cancer 74, 280–284 (1996).
24. Stram, D. O. et al. Modeling and E-M estimation of haplotype-specific relative risks from genotype data for a case-control study of unrelated individuals. Hum. Hered. 53, 179–190 (2003).
25. Ku, M. et al. Genome-wide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. PLoS Genet. 4, e1000242 (2008).
26. Ernst, J. et al. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature 473, 43–49 (2011).

Acknowledgements

We thank all the individuals who took part in this study and all the researchers, clinicians and administrative staff who have made possible the many studies contributing to this work. In particular, we thank: D. Bowtell, P.M. Webb, A. deFazio, D. Gertig, A. Green, P. Parsons, N. Hayward and D. Wildman (AUS); G. Peuteman, T. Van Brussel and D. Smeets (BEL); the staff of the genotyping unit, S LaBoissiere and F Robidoux (Genome Québec); U. Eibler and T. Koehler (GER); L. Gacucia (HMO); P. Schurrmair, F. Kramer, W. Zheng, T.-W. Park-Simon, K. Beer-Grondke and D. Schmidt (HJO); S. Windebank, C. Hilkier and J. Vollenweider (MAY); the state cancer registries of AL, AZ, CA, GA, CO, CT, DE, FL, GA, HI, 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 (NHS); L. Paddock, M. King, U. Chadrani, A. Samoila and Y. Bensam (NIO); M. Sherman, A. Hutchinson, N. Szenesia-Dabrowska, B. Peplonska, W. Zatonski, A. Soni, P. Chao and M. Stagner (POL); C. Lucarini, P. Harrington, the SEARCH team and ECRC (SEA); the Scottish Gynaecological Clinical Trials group and SCOCTRI investigators (SRO); W.-H. Chow and Y.-T. Gao (SWH); I. Jacobs, M. Widschwendter, N. Balogun, A. Ryan and J. Hutchinson (UK); and the Collaborative Ovarian Gene-Environment Study (COGS) project is funded through a European Commission’s Seventh Framework Programme grant (agreement number 221375–HEALTH-F2-2009-221375). 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 (PDP/RP10.07). The scientific development and funding of this project were (in part) supported by the Genetic Associations and Mechanisms in Oncology (GAME-ON) Network: a NCI Cancer Post-GWAS Initiative (U19-CA148112). This study made use of data generated by the Wellcome Trust Case Control consortium.
Hui Shen1,*, Brooke L. Fridley2,*, Honglin Song3,*, Kate Lawrenson4, Julie M. Cunningham5, Susan J. Ramus4, R01CA83918, U24 CA143882 and Intramural research funds); the US Army Medical Research Cen-
tre, the National Cancer Institute (R37-CA070867, R37-CA70867, U01-CA069417, U01-CA071966, P30-CA15083, R01-CA126841, R01-CA149429, R01-CA111415, R03-CA111433, R37-CA070867, R37-CA070867, U01-CN69419, U01-C011433, P30-CA15083, R01CA83918, U24 CA13882 and Intramural research funds); the US Army Medical Research and Material Command (DAMD17-98-1-6859, DAMD17-01-1-0729, DAMD17-02-1-0666, DAMD17-02-1-0669 and W81XWH-10-1-0280); the National

Author contributions

H. Shen, B.L.F., H. Song, K.L., M.K., G.C.T., S.A.G., P.D.P.P., P.W.L., E.L.G. and C.L.P. contributed to the preparation of the manuscript. All authors read and approved the final version. H. Shen, B.L.F., M.S.C., K.L., J.T., D.S., M.C.L., M.K., P.D.P.P., P.W.L., E.L.G. and C.L.P. carried out data analysis. S.J.R. and C.M.P. collated and prepared samples for genotyping. S.A.G. and K.L. performed functional analyses.

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: Shen, H., Fridley, B. L., Song, H. et al. Epigenetic analysis leads
to identification of HNFP1 as a subtype-specific susceptibility gene for ovarian cancer. Nat.
Commun. 4:1628 doi: 10.1038/ncomms2629 (2013).

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/3.0/

Hui Shen1,*, Brooke L. Fridley2,*, Honglin Song3,*, Kate Lawrenson4, Julie M. Cunningham5, Susan J. Ramus4, Mine S. Cicek6, Jonathan Tyrer3, Douglas Stram4, Melissa C. Larson7, Martin Köbel8, PRACTICAL Consortium9, Argyrios Ziogas10, Wei Zheng11, Hannah P. Yang12, Anna H. Wu4, Eva L. Wozniak13, Yin Ling Woo14, Boris Winterhoff15, Elisabeth Wik16,17, Alice S. Whittemore18, Nicolas Wentzensen12, Yin Ling Woo14, Boris Winterhoff15, Elisabeth Wik16,17, Alice S. Whittemore18, Nicolas Wentzensen12, Rachel Palmeri Weber19, Allison F. Vitonis20, Daniel Vincent21, Robert A. Vierkant7, Ignace Vergote22,23, David Van Den Berg4, Anne M. Van Altena24, Shelley S. Tworoger25,26, Pamela J. Thompson27, Daniel C. Tessier21, Kathryn L. Terry20,26, Soo-Hwong Teo28,29, Claire Templeman30, Daniel O. Stram4, Melissa C. Southey31, Weiva Sieh18, Nadeem Siddiqui32, Yuiri B. Shvetsov27, Xiao-Ou Shu10, Viji Shridhar5, Shan Wang-Gohrke33, Gianluca Severi34,35, Ira Schwaab36, Helga B. Salvesen16,17, Iwona K. Rzepecka37, Ingo B. Runnebaum38, Mary Anne Rossing39,40, Lorna Rodriguez-Rodriguez41, Harvey A. Risch42, Stefano P. Renner43, Elizabeth M. Poole25,26, Malcolm C. Pike4,44, Catherine M. Phelan45, Liisa M. Pelttari46, Tanja Pejovic47,48, James Paul49, Irene Orlove44, Siti Zawiah Omar14, Sara H. Olson44, Kunle Odunsi50, Stefan Nickels51, Heli Nevanlinna46, Roberta B. Ness52, Steven A. Narod53, Toru Nakashima54, Kirsten B. Moysich55, Alvaro N.A. Monteiro45, Joanna Moes-Sosnowska37, Francesmary Modugno56,57,58, Usha Menon13, John R. McLaughlin59,60, Valerie McGuire18, Keitaro Matsuo61, Noor Azmi Mat Adenan14, Leon F.A. G. Massuger24, Galina Lurie27, Lene Lundvall62, Jan Lubinski63, Jolanta Lisowski64, Douglas A. Levine55, Arto Leminen46, Alice W. Lee4, Nhu D. Le66, Sandrina Lambrechts22,23,67,
Diether Lambrechts22,67, Jolanta Kupryjanczyk37, Camilla Krakstad16,17, Gottfried E. Konecny68, Susanne Krüger-Kjaer62,69, Lambertus A. Klemmey70,71,72, Linda E. Kelemen73,74, Gary L. Keeney75, Beth Y. Karlan76, Rod Karevan4, Kimberly R. Kalli77, Hiroaki Kajiyama78, Bu-Tian Ji79, Allan Jensen69, Anna Jakubowska63, Edwin Iversen80, Satoyo Hosono61, Claus K. Høgdal62, Estrid Høgdal69,81, Maureen Hoatlin82, Peter Hillemanns83, Florian Heitz84,85, Rebecca Hein51,86, Philipp Harter84,85, Mari K. Halle16,17, Per Hall87, Jacek Gronwald63, Martin Gore88, Marc T. Goodman89, Graham G. Giles34,35,90, Aleksandra Gentry-Maharaj13, Montserrat Garcia-Closas91, James M. Flanagan92, Peter A. Fasching43,68, Arif B. Ekici93, Robert Edwards94, Diana Eccles95, Douglas F. Easton3, Matthias Dürst38, Andreas du Bois84,85, Thilo Dörfl96, Jennifer A. Doherty97, Evelyn Despierre22,23,67, Agnieszka Danko-Maściukowska37, Cezary Cybulski63, Daniel W. Cramer20,26, Linda S. Cook98, Xiaoping Chen99, Bridget Charbonneau98, Jenny Chang-Claude51, Ian Campbell100,101,102, Ralf Butzow46,103, Clareann H. Bunker57, Doerthe Bruéggemann30, Robert Brown92, Angela Brooks-Wilson104, Louise A. Brinton12, Natalia Bogdanova96, Matthew S. Block77, Elizabeth Benjamin105, Jonathan Beesley99, Matthias W. Beckmann43, Elisa V. Bandera41, Laura Baglietto34,35, François Bacot21, Sebastian M. Armasu7, Natalia Antonenkova106, Hoda Anton-Culver10, Katja K. Aben70,72, Dong Liang107, Xifeng Wu108, Karen Lu109, Michelle A.T. Hildebrandt108, Australian Ovarian Cancer Study Group110, Australian Cancer Study111, Joellen M. Schildkraut19,112, Thomas A. Sellers45, David Huntsman113, Andrew Berchuck114, Georgina Chenevix-Trench99, Simon A. Gayther4, Paul D.P. Pharoah3,115, Peter W. Laird1, Ellen L. Goode6 & Celeste Leigh Pearce4

1USC Epigenome Center, Keck School of Medicine, University of Southern California, Norris Comprehensive Cancer Center, Los Angeles, California 90033, USA. 2Department of Biostatistics, University of Kansas Medical Center, Kansas City, Kansas 66160, USA. 3Department of Oncology, University of Cambridge, Cambridge CB1 8RN, UK. 4Department of Preventive Medicine, Keck School of Medicine, University of Southern California Norris Comprehensive Cancer Center, Los Angeles, California 90033, USA. 5Division of Experimental Pathology, Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota 55905, USA. 6Division of Epidemiology, Department of Health Science Research, Mayo Clinic, Rochester, Minnesota 55905, USA. 7Division of Biomedical Statistics and Informatics, Department of Health Science Research, Mayo Clinic, Rochester, Minnesota 55905, USA. 8Department of Pathology and Laboratory Medicine, California Laboratory Services, University of California, California, Alberta, Canada T2N 2T9. 9A list of consortium members appears in Supplementary Note 1. 10Department of Epidemiology, Center for Cancer Genetics Research and Prevention, School of Medicine, University of California Irvine, Irvine, California 92697, USA. 11Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Vanderbilt University, Nashville, Tennessee 37232, USA. 12Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland 20892, USA. 13Gynaecological Cancer Research Centre, UCL EGA Institute for Women’s Health, London NW1 2BU, UK. 14Department of Obstetrics and Gynaecology, Faculty of Medicine, Affiliated to UM Cancer Research Institute, University of Malaya, Kuala Lumpur 59100, Malaysia. 15Department of Obstetrics and Gynecology, Mayo Clinic, Rochester, Minnesota 55905, USA. 16Department of Gynecology and Obstetrics, Haukeland University Hospital, Bergen 5006, Norway. 17Department of Clinical Science, University of Bergen, Bergen 5006, Norway. 18Department of Health Research and Policy, Stanford University School of Medicine, Stanford, California 94305, USA. 19Department of Community and Family Medicine, Duke University Medical Center, Durham, North Carolina 27708, USA. 20Obstetrics and Gynecology Epidemiology Center, Brigham and Women’s Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. 21Génome Québec, Montréal, Québec, Canada H3A OG1. 22Vesalius Research Center, VIB, Leuven 3000, Belgium. 23Division of Gynecologic Oncology, Department of Obstetrics and Gynecology, University Hospitals Leuven, Leuven 3000, Belgium. 24Department of Gynecology, Radboud University Medical Centre, Nijmegen HB 6500, The Netherlands. 25Channing Division of Network Medicine, Harvard Medical School, Brigham and Women’s Hospital, Boston, Massachusetts 02115, USA. 26Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts 02115, USA. 27Cancer Epidemiology Program, University of Hawaii Cancer Center, Honolulu, Hawaii 96813, USA. 28Cancer Research Initiatives Foundation, Sime Darby Medical Centre, Subang Jaya 47500, Malaysia. 29University Malaya Cancer Research Institute, University Malaya Medical Centre, University of Malaya, Kuala Lumpur 59100, Malaysia. 30Department of Obstetrics and Gynecology, Keck School of Medicine, University of Southern California, Los Angeles, California 90033, USA. 31Genetic Epidemiology Laboratory, Department of Pathology, University of Melbourne, Melbourne, Victoria VIC 3053, Australia. 32Department of Gynaecological Oncology, Glasgow Royal Infirmary, Glasgow G4 0SF, UK. 33Department of Obstetrics and Gynecology, University of Ulm, 89091 Ulm, Germany. 34Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Victoria VIC 3053, Australia. 35Centre for Molecular, Environmental, Genetic and Analytical Epidemiology, University of Melbourne, Melbourne, Victoria VIC 3053, Australia. 36Institut für Humangenetik Wiesbaden, Wiesbaden 65187, Germany. 37Department of Molecular Pathology, Maria Skłodowska-Curie Memorial Cancer Center, Institute of Oncology, Warsaw 02-781, Poland. 38Department of Gynecology and Obstetrics, Jena University Hospital, Jena 07743, Germany. 39Program in Epidemiology, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, USA. 40Department of Epidemiology, University of Washington, Seattle, Washington 98109, USA. 41Cancer Institute of New Jersey, Robert Wood Johnson Medical School, New Brunswick, New Jersey 08901, USA. 42Department of Chronic Disease Epidemiology, Yale School of Public Health, New Haven, Connecticut 06520, USA. 43Department of Gynecology and Obstetrics, University Hospital Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Comprehensive Cancer Center, Erlangen 91054, Germany. 44Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. 45Division of Population Sciences, Department of Cancer Epidemiology, Moffitt Cancer Center, Tampa, Florida 33612, USA. 46Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Central Hospital, Helsinki 00299 HUS, Finland. 47Department of Obstetrics and Gynecology, Oregon Health and Science University, Portland, Oregon 97229, USA. 48Knight Cancer Institute, Oregon Health and Science University, Portland, Oregon 97229, USA. 49Beaton West of Scotland Cancer Centre, Glasgow G12 OYN, UK. 50Department of Gynecologic Oncology, Roswell Park Cancer Institute, Buffalo, New York 14263, USA. 51Division of Cancer Epidemiology, German Cancer Research Center, Heidelberg 69120, Germany. 52School of Public Health,
