nature COMMUNICATIONS

ARTICLE

Received 28 Nov 2012 | Accepted 21 Feb 2013 | Published 27 Mar 2013

DOI: 10.1038/ncomms2629 OPEN

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.

†A full list of authors and their affiliations appears at the end of the paper.

Invasive epithelial ovarian cancer (EOC) has a strong heritable component[1], with an approximate three-fold increased risk associated with a first-degree family history[2]. Much of the excess familial risk observed for EOC is unexplained[3], 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[4–7] thus far, but candidate gene studies have been largely unsuccessful[8].

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[9]. 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[10]. The *HNF1B* gene (formerly known as *TCF2*) 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[12,13], prostate cancer[12,14–16] and uterine cancer[17].

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 (Fig. 1a, Supplementary Fig. S1) 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 (Fig. 1b). 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.

SNP analysis. With all invasive cancer subtypes considered together, we found no genome-wide significant ($P<5\times10^{-8}$) *HNF1B* SNP associations among women of European ancestry (Table 1; Supplementary Data S1). However, when analyses were stratified by histological subtype, we observed genome-wide significant results for both serous and clear cell EOC subtypes, but with risk associations in opposite directions. The association was similar for high- and low-grade serous cancers. There was no evidence of association for mucinous or endometrioid subtypes (Fig. 1c). Associations in the non-European populations are shown in Supplementary Table S2.

Minor alleles at nine SNPs, six genotyped and three imputed, were associated with increased risk of invasive serous ovarian cancer at $P<5\times10^{-8}$ (Table 1). The risk signal spanned a 21.4-kb region from the 5′ untranslated region (UTR) through part of intron 4 of *HNF1B* (Fig. 1c). The most strongly associated SNP for invasive serous ovarian cancer (rs7405776, minor allele frequency (MAF) 36%) conferred a 13% increased risk per minor allele ($P=3.1\times10^{-10}$; Table 1, Supplementary Fig. S2A). The signals of this SNP and the eight other genome-wide significant SNPs were indistinguishable, given the linkage disequilibrium and resulting haplotype structure (Supplementary Figs S3, S4 and S5).

For the clear cell subtype, rs11651755 (MAF 45%) was associated with a 23% decreased risk of disease at a genome-wide significant level ($P=2\times10^{-8}$; Table 1, Supplementary Fig. S2B). This signal was distinct from the nine significant SNPs for invasive serous cancer (Table 1). The odds against the serous-associated SNP, rs7405776, as the true best hit for clear cell ovarian cancer were 244:1. Conversely, the odds against the clear cell SNP, rs11651755, as the true best hit for serous were 1808:1. Further, when rs11651755 and rs7405776 were jointly modelled, the signal for clear cell cancer was driven completely by rs11651755, whereas that for the serous disease was driven by rs7405776 (Table 1). The clear cell SNP (rs11651755) sits on five haplotypes, only three of which also contain the serous SNP (rs7405776; Supplementary Fig. S5). Thus, different SNPs in the *HNF1B* gene regions explain the associations observed for serous and clear cell 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 (Fig. 2, Supplementary Fig. S6).

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 (Fig. 2). HNF1B expression and CIMP methylation are strongly associated ($P=3\times10^{-16}$; Fig. 2). Further, minimal hypermethylation is observed in serous tumours overall, but HNF1B is one notable exception (Supplementary Fig. S7).

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 rs11658063, 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.

Overexpression of HNF1B. Given the proposed role of *HNF1B* in clear cell tumorigenesis, we stably overexpressed the gene in immortalized endometriosis epithelial cells (EECs), which are

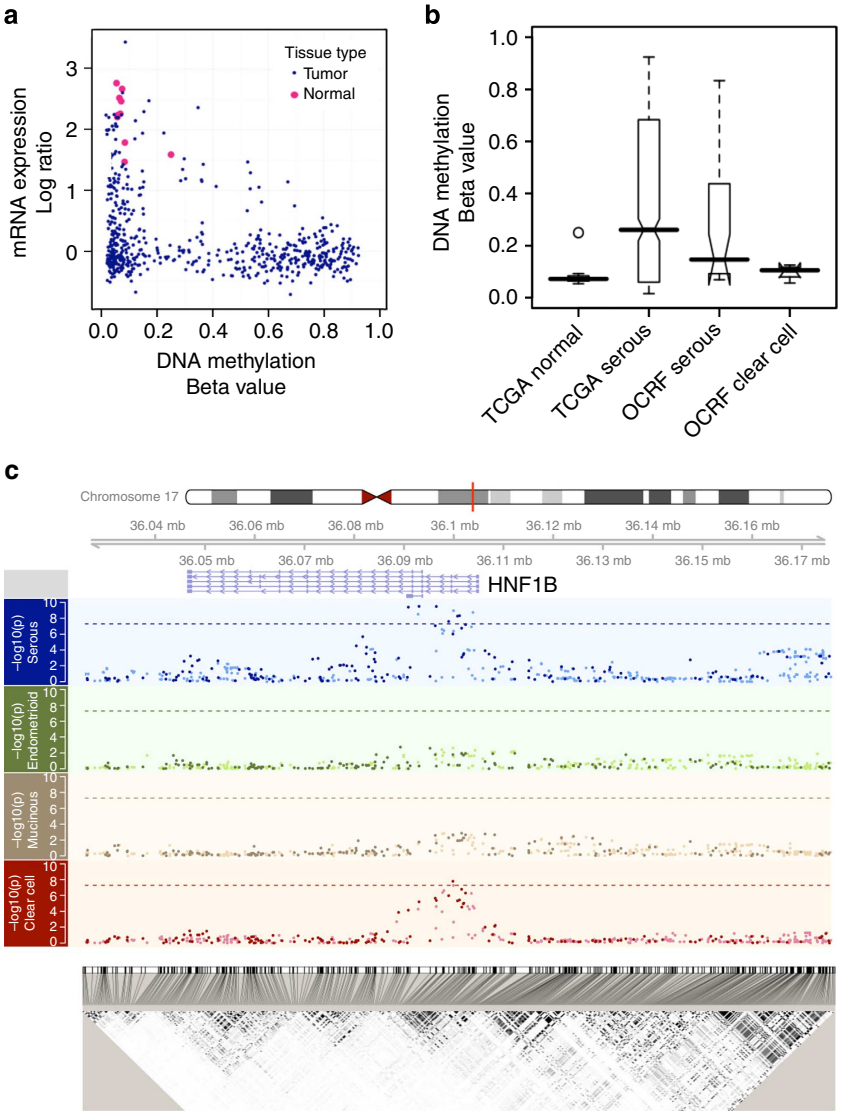

Figure 1 | Identification of *HNF1B* as a subtype-specific candidate gene for ovarian cancer and its establishment as a susceptibility gene. (**a**) The scatterplot compares the mRNA expression (y axis) versus DNA methylation (x axis) in serous ovarian tumours from TCGA (see Methods). Each blue dot is 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 $-\log10$ (*P*-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 (5×10^{-8}). The linkage disequilibrium plot on the bottom shows the r^2 between the SNPs. Genomic coordinates are based on hg19 (Build37).

hypothesized to be a cell of origin for clear cell ovarian cancers (Supplementary Fig. S10)[18]. 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 ($P=3\times10^{-16}$), 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

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	All invasive ($n=14,533$) OR	95% CI	*P*-value	Serous ($n=8,371$) OR	95% CI	*P*-value	Clear cell ($n=1,025$) OR	95% CI	*P*-value
Univariate												
rs3744763*	A/G		0.409	1.06	1.03 – 1.10	1.6×10^{-4}	1.13	1.09 – 1.17	4.0×10^{-10}	0.83	0.75 – 0.91	7.1×10^{-5}
17-36092841*,†	G/GT	0.902	0.403	1.05	1.01 – 1.08	0.005	1.13	1.08 – 1.17	3.1×10^{-9}	0.82	0.74 – 0.90	5.3×10^{-5}
rs7405776*	G/A		0.376	1.05	1.02 – 1.09	0.001	1.13	1.09 – 1.17	3.1×10^{-10}	0.80	0.73 – 0.88	6.2×10^{-6}
rs757210*	C/T		0.372	1.05	1.02 – 1.09	9.5×10^{-4}	1.13	1.09 – 1.17	3.2×10^{-10}	0.80	0.73 – 0.88	4.1×10^{-6}
rs4239217*	A/G		0.402	1.04	1.01 – 1.07	0.018	1.11	1.07 – 1.16	2.6×10^{-8}	0.79	0.72 – 0.87	1.0×10^{-6}
rs11651755*	T/C		0.489	1.02	0.99 – 1.06	0.124	1.10	1.06 – 1.14	9.9×10^{-7}	0.77	0.70 – 0.84	1.6×10^{-8}
rs61612821†	T/C	0.827	0.140	1.09	1.04 – 1.14	4.1×10^{-4}	1.19	1.13 – 1.26	1.1×10^{-9}	0.79	0.68 – 0.92	0.002
rs11657964*	G/A		0.400	1.04	1.01 – 1.08	0.006	1.12	1.08 – 1.16	5.3×10^{-9}	0.80	0.73 – 0.88	4.6×10^{-6}
rs7501939*	C/T		0.400	1.04	1.01 – 1.08	0.006	1.12	1.08 – 1.16	4.8×10^{-9}	0.80	0.73 – 0.88	3.7×10^{-6}
rs11658063‡	G/C	0.963	0.398	1.05	1.02 – 1.08	0.003	1.12	1.08 – 1.17	1.8×10^{-9}	0.81	0.74 – 0.90	2.3×10^{-5}
Bivariate												
rs7405776	G/A						1.11	1.06 – 1.18	9.8×10^{-5}	0.95	0.83 – 1.09	0.470
rs11651755	T/C						1.02	0.96 – 1.07	0.580	0.80	0.70 – 0.91	7.0×10^{-4}

AAF, alternate allele frequency; CI, confidence interval.
*Genotyped.
†Insertion.
‡Imputed.

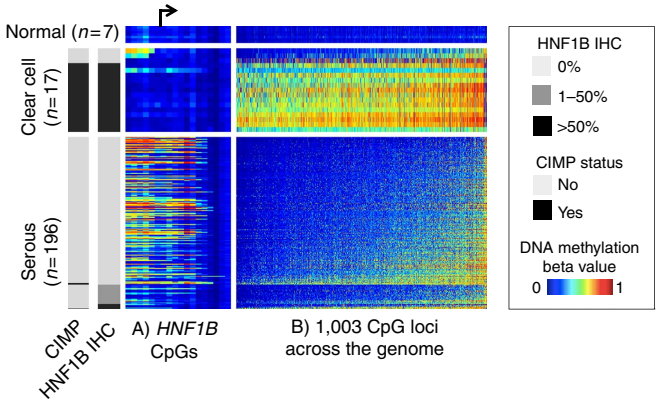

Figure 2 | *HNF1B*-promoter DNA methylation, protein expression and global DNA-methylation pattern by subtype. Each row is a tissue sample collected at the Mayo Clinic that belongs to one of the three categories: normal ovarian tissue ($n=7$), clear cell ovarian tumours ($n=17$) or serous ovarian tumours ($n=196$). 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 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.

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[18–24], with endometriosis as a proposed cell of origin for clear cell disease[18] and fallopian tube fimbriae as one for serous disease[22]. 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.

Variation in the 5′ UTR through the intron 4 region of *HNF1B* is also associated with susceptibility to prostate[12,14–16] and uterine cancer[17] (where minor alleles of certain SNPs are associated with decreased risk) and type II diabetes[12,13] (increased risk for the same or correlated SNP alleles; Supplementary Fig. S4). The opposing directions of these associations mirror the differential effects seen here in ovarian cancer. The most strongly associated SNP for both prostate[14] and uterine cancer[17] is rs4430796, correlated at $r^2=0.94$ with the top clear cell ovarian cancer SNP, rs11651755, suggesting a common risk variant. Although increased risk of type II diabetes has been reported with rs4430796 (ref. 12), Winckler *et al.*[13] have suggested that the best marker of diabetes risk is rs757210, which correlated at $r^2=0.97$ with our top serous SNP. Thus, the evidence suggests that a specific variant in *HNF1B* predisposes to clear cell ovarian, uterine and prostate cancers and that a different variants is associated with diabetes and serous ovarian cancer.

We were able to completely fine-map the *HNF1B* region, localize the signal and identify a handful of potentially causal SNPs. This is quite different from other regions of the genome where it is not uncommon to identify hundreds of candidate causal SNPs. Further, an important link, often missing when susceptibility loci are identified, is the functional role that the variant has in disease. In the case of serous ovarian cancer, the SNP–*HNF1B*-promoter DNA methylation association strengthens as it approaches the promoter region, particularly where it overlaps with a PRC2 mark. PRC2–

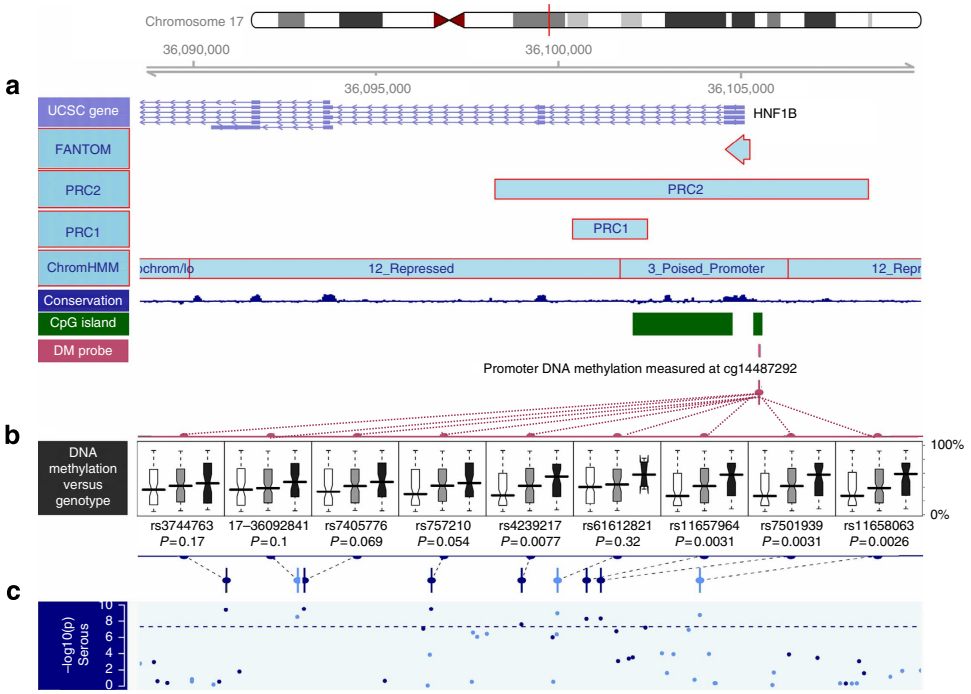

Figure 3 | Correlation of serous risk-associated SNPs with *HNF1B*-promoter DNA-methylation level. Plotted is the linkage disequilibrium region defined as $r^2 > 0.2$ with the top serous SNP rs7405776. (**a**) Annotation of the region in terms of (from top to bottom:) UCSC genes, FANTOM mark, PRC marks (PRC2 and PRC1)[32], the chromatin status determined in stem cells[33], the conservation score across this region and the CpG island information, on top of the location of the HM450 probe used in **b** boxplots of promoter DNA-methylation level of *HNF1B* (cg14487292) by SNP genotype with position indicated in **c**. This DNA-methylation probe was selected based on inverse association with mRNA expression for *HNF1B*, and does not contain any SNP with MAF >1% in its probe sequence. Each boxplot shows the distribution of DNA-methylation level by genotype (homozygous major—white; heterozygous—grey; and homozygous minor—black, where the minor alleles are the risk alleles). Two-sided *P*-values testing for trend are presented, and are computed for 231 Mayo Clinic high-grade, high-stage serous tumours to avoid confounding by histological subtypes, and also to be consistent with the TCGA data (primarily high-grade, high-stage serous). Results were similar with all subtypes combined. The risk alleles are associated with significantly increased DNA methylation. The association of rs11658063 genotype with promoter methylation is consistent across the entire region flanking *HNF1B* transcription start site, and stronger for the upstream promoter region (Supplementary Fig. S8).

DNA methyltransferase cross-talk has been proposed to be a mechanism of predisposition to cancer-specific hypermethylation[25]. 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.*[23] 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[26]. 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[27]. 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[28]. 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

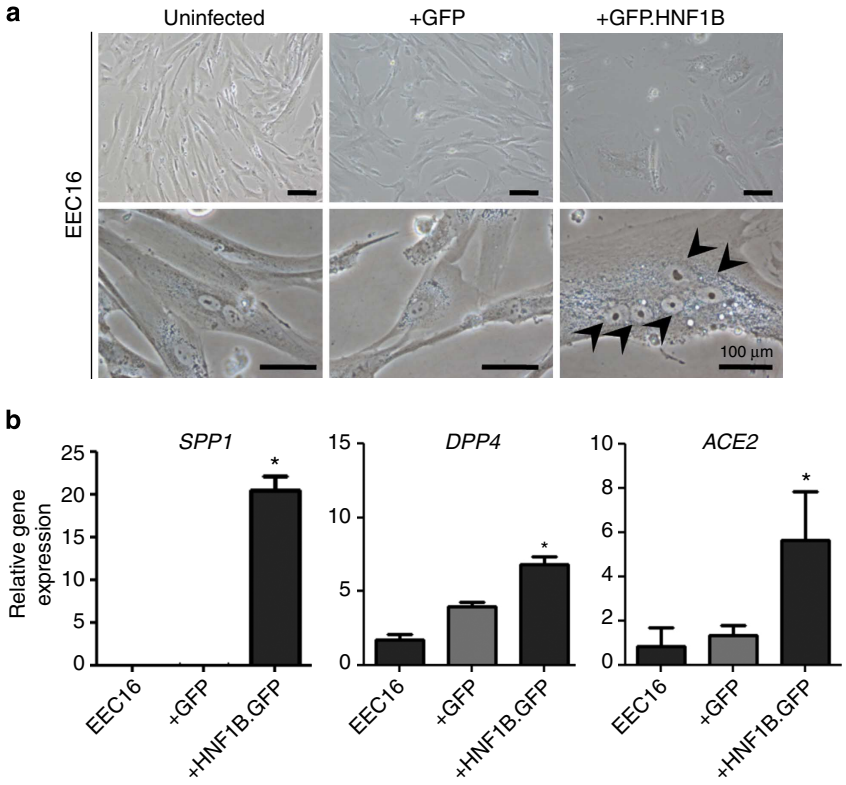

Figure 4 | Phenotypic effects and downstream targets of HNF1B overexpression in immortalized EECs. (**a**) Morphological changes in EECs expressing a HNF1B GFP fusion protein ($EEC^{GFP.HNF1B}$). GFP-positive cells were sorted using flow cytometry. The arrows indicate five nuclei contained within a single $EEC^{GFP.HNF1B}$ 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. $^*P > 0.01$.

should focus on understanding these mechanisms as they could 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 (<https://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 Human1M-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[9] on 32 serous and 4 clear cell ovarian tumours from USC Norris Comprehensive Cancer Center and Duke University ('OCRF 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[29]. 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 pretreatment 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 *hTERT* (Addgene plasmid 12245) into primary EECs (Supplementary Fig. S10). *TERT*-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

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 Oncological 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 Oncological 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 impute genotypes for SNPs across this region, resulting in available data on an additional 307 SNPs with MAF >0.02 in European Whites and imputation $r^2>0.30$ (IMPUTE 2.2).

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

Statistical analysis. We used the program LAMP[30] 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 serous 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 program[31] to obtain the haplotype dosage for each subject for the LAMP-derived European ancestry group for haplotypes with a frequency of $\geq 1\%$. 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 of twins from 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. Pharaoh, 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/ng2564) (2013).
5. Goode, E. L. *et al.* A genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31 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 9p22.2. *Nat. Genet.* **41,** 996–1000 (2009).
8. Bolton, K. L., Ganda, C., Berchuck, A., Pharaoh, P. D. & Gayther, S. A. Role of common genetic 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 genomic 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. Spurdle, A. B. *et al.* Genome-wide association study identifies a common variant associated 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. & Shih, IeM. The origin and pathogenesis of epithelial ovarian cancer: a proposed unifying theory. *Am. J. Surg. Pathol.* **34,** 433–443 (2010).
20. Gilks, C. B. Molecular abnormalities in ovarian cancer subtypes other than high-grade serous carcinoma. *J. Oncol.* **2010,** 740968 (2010).
21. Risch, H. A., Marrett, L. D., Jain, M. & Howe, G. R. Differences in risk factors for epithelial ovarian cancer by histologic type. Results of a case-control study. *Am. J. Epidemiol.* **144,** 363–372 (1996).
22. Crum, C. P. *et al.* The distal fallopian tube: a new model for pelvic serous carcinogenesis. *Curr. Opin. Obstet. Gynecol.* **19,** 3–9 (2007).
23. Gounaris, I., Charnock-Jones, D. S. & Brenton, J. D. Ovarian clear cell carcinoma—bad endometriosis or bad endometrium? *J. Pathol.* **225,** 157–160 (2011).
24. Kobel, M. *et al.* Ovarian carcinoma subtypes are different diseases: implications for biomarker studies. *PLoS Med.* **5,** e232 (2008).
25. Widschwendter, M. *et al.* Epigenetic stem cell signature in cancer. *Nat. Genet.* **39,** 157–158 (2007).
26. Skirnisdottir, I., Seidal, T., Karlsson, M. G. & Sorbe, B. Clinical and biological characteristics of clear cell carcinomas of the ovary in FIGO stages I-II. *Int. J. Oncol.* **26,** 177–183 (2005).
27. Bell, J. T. *et al.* DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biol.* **12,** R10 (2011).
28. Hitchins, M. P. *et al.* Dominantly inherited constitutional epigenetic silencing of MLH1 in a cancer-affected family is linked to a single nucleotide variant within the 5' UTR. *Cancer Cell* **20,** 200–213 (2011).
29. Bibikova, M. *et al.* High density DNA methylation array with single CpG site resolution. *Genomics* **98,** 288–295 (2011).
30. Sankararaman, S., Sridhar, S., Kimmel, G. & Halperin, E. Estimating local ancestry in admixed populations. *Am. J. Hum. Genet.* **82,** 290–303 (2008).
31. 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.* **55,** 179–190 (2003).
32. Ku, M. *et al.* Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet.* **4,** e1000242 (2008).
33. 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. Whiteman (AUS); G. Peuteman, T. Van Brussel and D. Smeets (BEL); the staff of the genotyping unit, S LaBoissière and F Robidoux (Génome Québec); U. Eilber and T. Koehler (GER); L. Gacucova (HMO); P. Schurmann, F. Kramer, W. Zheng, T.-W. Park-Simon, K. Beer-Grondke and D. Schmidt (HJO); S. Windebank, C. Hilker and J. Vollenweider (MAY); the state cancer registries of AL, AZ, AR, CA, 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. Chandran, A. Samoila and Y. Bensman (NJO); M. Sherman, A. Hutchinson, N. Szeszenia-Dabrowska, B. Peplonska, W. Zatonski, A. Soni, P. Chao and M. Stagner (POL); C. Luccarini, P. Harrington, the SEARCH team and ECRIC (SEA); the Scottish Gynaecological Clinical Trials group and SCOTROC1 investigators (SRO); W.-H. Chow and Y.-T. Gao (SWH); I. Jacobs, M. Widschwendter, N. Balogun, A. Ryan and J. Ford (UKO); and Carole Pye (UKR). The Collaborative Oncological Gene-Environment Study (COGS) project is funded through a European Commission's Seventh Framework Programme grant (agreement number 223175–HEALTH-F2-2009-223175). 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). 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.

A full list of the investigators who contributed to the generation of the data is available from http://www.wtccc.org.uk/. Funding for the project was provided by the Wellcome Trust under award 076113. The results published here are in part based upon data generated by The Cancer Genome Atlas Pilot Project established by the National Cancer Institute and National Human Genome Research Institute. Information about TCGA, and the investigators and institutions who constitute the TCGA research network, can be found at http://cancergenome.nih.gov/. G.C.T. is a Senior Principal Research Fellow of the National Health and Medical Research Council, Australia. D.F.E. is a Principal Research Fellow of Cancer Research UK. P.A.F. is supported by the Deutsche Krebshilfe. B.K. holds an American Cancer Society Early Detection Professorship (SIOP-06-258-01-COUN). L.E.K. is supported by a Canadian Institutes of Health Research Investigator award (MSH-87734). H.S.1 is supported by a National Institutes of Health training grant (T32GM067587), 'Training in Cellular, Biochemical and Molecular Sciences'. Funding of the constituent studies was provided by the American Cancer Society (CRTG-00-196-01-CCE); 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 and C536/A6689); the Celma Mastry Ovarian Cancer Foundation; the Danish Cancer Society (94-222-52); the Norwegian Cancer Society, Helse Vest, the Norwegian Research Council, ELAN Funds of the University of Erlangen-Nuremberg; the Eve Appeal; the Helsinki University Central Hospital Research Fund; Imperial Experimental Cancer Research Centre (C1312/A15589); the Ovarian Cancer Research Fund (PPD/USC.06); Nationaal 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 PO5C 028 14, 2 PO5A 068 27); the Roswell Park Cancer Institute Alliance Foundation; the US National Cancer Institute (K07-CA095666, K07-CA143047, K07-CA80668, K22-CA138563, N01-CN55424, N01-PC067001, N01-PC035137, P01-CA017054, P01-CA087696, P50-CA105009, P50-CA136393, R01-CA014089, R01-CA016056, R01-CA017054, R01-CA049449, R01-CA050385, R01-CA054419, R01-CA058598, R01-CA058860, R01-CA061107, R01-CA061132, R01-CA063682, R01-CA064277, R01-CA067262, R01-CA071766, R01-CA074850, R01-CA076016, R01-CA080742, R01-CA080978, R01-CA087538, R01-CA092044, R01-095023, R01-CA106414, R01-CA122443, R01-CA112523, R01-CA114343, R01-CA126841, R01-CA149429, R01-CA141154, R03-CA113148, R03-CA115195, R37-CA070867, R37-CA70867, U01-CA069417, U01-CA071966, P30-CA15083, R01CA83918, U24 CA143882 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 and W81XWH-10-1-0280); the National Health and Medical Research Council of Australia (199600 and 400281); the German Federal Ministry of Education and Research of Germany Programme of Clinical Biomedical Research (01 GB 9401); 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 Phi Beta Psi Sorority; 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, University College Hospital 'Womens Health Theme' and the Royal Marsden Hospital; and WorkSafeBC. We acknowledge the contributions of Kyriaki Michailidou, Ali Amin Al Olama and Karoline Kuchenbaecker to the iCOGS statistical analyses and Shahana Ahmed, Melanie J. Maranian and Catherine S Healey for their contributions to the iCOGS genotyping quality-control process. US National Health Institute/National Center for Research Resources/General Clinical Research Center M01- RR000056.

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 *HNF1B* as a subtype-specific susceptibility gene for ovarian cancer. *Nat. Commun.* 4:1628 doi: 10.1038/ncomms2629 (2013).

Hui Shen[1,*], Brooke L. Fridley[2,*], Honglin Song[3,*], Kate Lawrenson[4], Julie M. Cunningham[5], Susan J. Ramus[4], Mine S. Cicek[6], Jonathan Tyrer[3], Douglas Stram[4], Melissa C. Larson[7], Martin Köbel[8], PRACTICAL Consortium[9], Argyrios Ziogas[10], Wei Zheng[11], Hannah P. Yang[12], Anna H. Wu[4], Eva L. Wozniak[13], Yin Ling Woo[14], Boris Winterhoff[15], Elisabeth Wik[16,17], Alice S. Whittemore[18], Nicolas Wentzensen[12], Rachel Palmieri Weber[19], Allison F. Vitonis[20], Daniel Vincent[21], Robert A. Vierkant[7], Ignace Vergote[22,23], David Van Den Berg[4], Anne M. Van Altena[24], Shelley S. Tworoger[25,26], Pamela J. Thompson[27], Daniel C. Tessier[21], Kathryn L. Terry[20,26], Soo-Hwang Teo[28,29], Claire Templeman[30], Daniel O. Stram[4], Melissa C. Southey[31], Weiva Sieh[18], Nadeem Siddiqui[32], Yurii B. Shvetsov[27], Xiao-Ou Shu[10], Viji Shridhar[5], Shan Wang-Gohrke[33], Gianluca Severi[34,35], Ira Schwaab[36], Helga B. Salvesen[16,17], Iwona K. Rzepecka[37], Ingo B. Runnebaum[38], Mary Anne Rossing[39,40], Lorna Rodriguez-Rodriguez[41], Harvey A. Risch[42], Stefan P. Renner[43], Elizabeth M. Poole[25,26], Malcolm C. Pike[4,44], Catherine M. Phelan[45], Liisa M. Pelttari[46], Tanja Pejovic[47,48], James Paul[49], Irene Orlow[44], Siti Zawiah Omar[14], Sara H. Olson[44], Kunle Odunsi[50], Stefan Nickels[51], Heli Nevanlinna[46], Roberta B. Ness[52], Steven A. Narod[53], Toru Nakanishi[54], Kirsten B. Moysich[55], Alvaro N.A. Monteiro[45], Joanna Moes-Sosnowska[37], Francesmary Modugno[56,57,58], Usha Menon[13], John R. McLaughlin[59,60], Valerie McGuire[18], Keitaro Matsuo[61], Noor Azmi Mat Adenan[14], Leon F.A. G. Massuger[24], Galina Lurie[27], Lene Lundvall[62], Jan Lubiński[63], Jolanta Lissowska[64], Douglas A. Levine[65], Arto Leminen[46], Alice W. Lee[4], Nhu D. Le[66], Sandrina Lambrechts[22,23,67],

Diether Lambrechts[22,67], Jolanta Kupryjanczyk[37], Camilla Krakstad[16,17], Gottfried E. Konecny[68], Susanne Krüger Kjaer[62,69], Lambertus A. Kiemeney[70,71,72], Linda E. Kelemen[73,74], Gary L. Keeney[75], Beth Y. Karlan[76], Rod Karevan[4], Kimberly R. Kalli[77], Hiroaki Kajiyama[78], Bu-Tian Ji[79], Allan Jensen[69], Anna Jakubowska[63], Edwin Iversen[80], Satoyo Hosono[61], Claus K. Høgdall[62], Estrid Høgdall[69,81], Maureen Hoatlin[82], Peter Hillemanns[83], Florian Heitz[84,85], Rebecca Hein[51,86], Philipp Harter[84,85], Mari K. Halle[16,17], Per Hall[87], Jacek Gronwald[63], Martin Gore[88], Marc T. Goodman[89], Graham G. Giles[34,35,90], Aleksandra Gentry-Maharaj[13], Montserrat Garcia-Closas[91], James M. Flanagan[92], Peter A. Fasching[43,68], Arif B. Ekici[93], Robert Edwards[94], Diana Eccles[95], Douglas F. Easton[3], Matthias Dürst[38], Andreas du Bois[84,85], Thilo Dörk[96], Jennifer A. Doherty[97], Evelyn Despierre[22,23,67], Agnieszka Dansonka-Mieszkowska[37], Cezary Cybulski[63], Daniel W. Cramer[20,26], Linda S. Cook[98], Xiaoqing Chen[99], Bridget Charbonneau[98], Jenny Chang-Claude[51], Ian Campbell[100,101,102], Ralf Butzow[46,103], Clareann H. Bunker[57], Doerthe Brueggmann[30], Robert Brown[92], Angela Brooks-Wilson[104], Louise A. Brinton[12], Natalia Bogdanova[96], Matthew S. Block[77], Elizabeth Benjamin[105], Jonathan Beesley[99], Matthias W. Beckmann[43], Elisa V. Bandera[41], Laura Baglietto[34,35], François Bacot[21], Sebastian M. Armasu[7], Natalia Antonenkova[106], Hoda Anton-Culver[10], Katja K. Aben[70,72], Dong Liang[107], Xifeng Wu[108], Karen Lu[109], Michelle A.T. Hildebrandt[108], Australian Ovarian Cancer Study Group[110], Australian Cancer Study[111], Joellen M. Schildkraut[19,112], Thomas A. Sellers[45], David Huntsman[113], Andrew Berchuck[114], Georgia Chenevix-Trench[99], Simon A. Gayther[4], Paul D.P. Pharoah[3,115], Peter W. Laird[1], Ellen L. Goode[6] & Celeste Leigh Pearce[4]

[1] USC Epigenome Center, Keck School of Medicine, University of Southern California, Norris Comprehensive Cancer Center, Los Angeles, California 90033, USA. [2] Department of Biostatistics, University of Kansas Medical Center, Kansas City, Kansas 66160, USA. [3] Department of Oncology, University of Cambridge, Cambridge CB1 8RN, UK. [4] Department of Preventive Medicine, Keck School of Medicine, University of Southern California Norris Comprehensive Cancer Center, Los Angeles, California 90033, USA. [5] Division of Experimental Pathology, Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota 55905, USA. [6] Division of Epidemiology, Department of Health Science Research, Mayo Clinic, Rochester, Minnesota 55905, USA. [7] Division of Biomedical Statistics and Informatics, Department of Health Science Research, Mayo Clinic, Rochester, Minnesota 55905, USA. [8] Department of Pathology and Laboratory Medicine, Calgary Laboratory Services, University of Calgary, Calgary, Alberta, Canada T2N 2T9. [9] A list of consortium members appears in Supplementary Note 1. [10] Department of Epidemiology, Center for Cancer Genetics Research and Prevention, School of Medicine, University of California Irvine, Irvine, California 92697, USA. [11] Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Vanderbilt University, Nashville, Tennessee 37232, USA. [12] Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland 20892, USA. [13] Gynaecological Cancer Research Centre, UCL EGA Institute for Women's Health, London NW1 2BU, UK. [14] Department of Obstetrics and Gynaecology, Faculty of Medicine, Affiliated to UM Cancer Research Institute, University of Malaya, Kuala Lumpur 59100, Malaysia. [15] Department of Obstetrics and Gynecology, Mayo Clinic, Rochester, Minnesota 55905, USA. [16] Department of Gynecology and Obstetrics, Haukeland University Hospital, Bergen 5006, Norway. [17] Department of Clinical Science, University of Bergen, Bergen 5006, Norway. [18] Department of Health Research and Policy, Stanford University School of Medicine, Stanford, California 94305, USA. [19] Department of Community and Family Medicine, Duke University Medical Center, Durham, North Carolina 27708, USA. [20] Obstetrics and Gynecology Epidemiology Center, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. [21] Génome Québec, Montréal, Québec, Canada H3A 0G1. [22] Vesalius Research Center, VIB, Leuven 3000, Belgium. [23] Division of Gynecologic Oncology, Department of Obstetrics and Gynaecology and Leuven Cancer Institute, University Hospitals Leuven, Leuven 3000, Belgium. [24] Department of Gynaecology, Radboud University Medical Centre, Nijmegen HB 6500, The Netherlands. [25] Channing Division of Network Medicine, Harvard Medical School, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. [26] Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts 02115, USA. [27] Cancer Epidemiology Program, University of Hawaii Cancer Center, Hawaii 96813, USA. [28] Cancer Research Initiatives Foundation, Sime Darby Medical Centre, Subang Jaya 47500, Malaysia. [29] University Malaya Cancer Research Institute, University Malaya Medical Centre, University of Malaya, Kuala Lumpur 59100, Malaysia. [30] Department of Obstetrics and Gynecology, Keck School of Medicine, University of Southern California, Los Angeles, California 90033, USA. [31] Genetic Epidemiology Laboratory, Department of Pathology, University of Melbourne, Melbourne, Victoria VIC 3053, Australia. [32] Department of Gynaecological Oncology, Glasgow Royal Infirmary, Glasgow G4 0SF, UK. [33] Department of Obstetrics and Gynecology, University of Ulm, 89091 Ulm, Germany. [34] Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Victoria VIC 3053, Australia. [35] Centre for Molecular, Environmental, Genetic and Analytical Epidemiology, University of Melbourne, Melbourne, Victoria VIC 3010, Australia. [36] Institut für Humangenetik Wiesbaden, Wiesbaden 65187, Germany. [37] Department of Molecular Pathology, Maria Sklodowska-Curie Memorial Cancer Center, Institute of Oncology, Warsaw 02-781, Poland. [38] Department of Gynecology and Obstetrics, Jena University Hospital, Jena 07743, Germany. [39] Program in Epidemiology, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, USA. [40] Department of Epidemiology, University of Washington, Seattle, Washington 98109, USA. [41] Cancer Institute of New Jersey, Robert Wood Johnson Medical School, New Brunswick, New Jersey 08901, USA. [42] Department of Chronic Disease Epidemiology, Yale School of Public Health, New Haven, Connecticut 06520, USA. [43] Department of Gynecology and Obstetrics, University Hospital Erlangen, Friedrich-Alexander-University Erlangen-Nuremberg, Comprehensive Cancer Center, Erlangen 91054, Germany. [44] Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. [45] Division of Population Sciences, Department of Cancer Epidemiology, Moffitt Cancer Center, Tampa, Florida 33612, USA. [46] Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Central Hospital, Helsinki, 00029 HUS, Finland. [47] Department of Obstetrics and Gynecology, Oregon Health and Science University, Portland, Oregon 97239, USA. [48] Knight Cancer Institute, Oregon Health and Science University, Portland, Oregon 97239, USA. [49] Beatson West of Scotland Cancer Centre, Glasgow G12 0YN, UK. [50] Department of Gynecologic Oncology, Roswell Park Cancer Institute, Buffalo, New York 14263, USA. [51] Division of Cancer Epidemiology, German Cancer Research Center, Heidelberg 69120, Germany. [52] School of Public Health,

University of Texas, Houston, Texas 77030, USA. [53] Women's College Research Institute, University of Toronto, Toronto, Ontario, Canada M5G IN8. [54] Department of Gynecologic Oncology, Aichi Cancer Center Central Hospital, Nagoya 464-8681, Japan. [55] Department of Cancer Prevention and Control, Roswell Park Cancer Institute, Buffalo 14263, New York, USA. [56] Department of Obstetrics, Gynecology and Reproductive Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15213, USA. [57] Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania 15213, USA. [58] Women's Cancer Research Program, Magee-Womens Research Institute, University of Pittsburgh Cancer Institute, Pittsburgh, Pennsylvania 15213, USA. [59] Dalla Lana School of Public Health, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada M5T 3M7. [60] Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada M5G IX5. [61] Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Nagoya 464-8681, Japan. [62] Gynecologic Clinic, The Juliane Marie Center, Copenhagen University Hospital, DK-2100 Copenhagen, Denmark. [63] Department of Genetics and Pathology, International Hereditary Cancer Center, Pomeranian Medical University, Szczecin 70-115, Poland. [64] Department of Cancer Epidemiology and Prevention, Maria Sklodowska-Curie Memorial Cancer Center, Warsaw 02-781, Poland. [65] Department of Surgery, Gynecology Service, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA. [66] Cancer Control Research, BC Cancer Agency, Vancouver, British Columbia, Canada G12 0YN. [67] Laboratory for Translational Genetics, Department of Oncology, University of Leuven, Leuven 3000, Belgium. [68] Division of Hematology and Oncology, Department of Medicine, David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, California 90095, USA. [69] Virus, Lifestyle and Genes, Danish Cancer Society Research Center, DK-2100 Copenhagen, Denmark. [70] Department of Epidemiology, Biostatistics and HTA, Radboud University Medical Centre, Nijmegen, HB 6500, Netherlands. [71] Department of Urology, Radboud University Medical Centre, Nijmegen HB 6500, The Netherlands. [72] Comprehensive Cancer Center, Utrecht 1066CX, The Netherlands. [73] Department of Population Health Research, Alberta Health Services-Cancer Care, Calgary, Alberta, Canada T2N 2T9. [74] Department of Medical Genetics and Oncology, University of Calgary, Calgary, Alberta, Canada T2N 2T9. [75] Division of Anatomic Pathology, Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota 55905, USA. [76] Women's Cancer Program, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California 90048, USA. [77] Department of Medical Oncology, Mayo Clinic, Rochester, Minnesota 55905, USA. [78] Department of Obstetrics and Gynecology, Nagoya University Graduate School of Medicine, Nagoya 464-8601, Japan. [79] Division of Cancer Etiology and Genetics, National Cancer Institute, Bethesda, Maryland 20892, USA. [80] Department of Statistical Science, Duke University, Durham, North Carolina 27708, USA. [81] Molecular Unit, Department of Pathology, Herlev Hospital, University of Copenhagen, Copenhagen 2730, Denmark. [82] Department of Biochemistry and Molecular Biology, Oregon Health and Science University, Portland, Oregon 97239, USA. [83] Clinics of Obstetrics and Gynaecology, Hannover Medical School, Hannover 30625, Germany. [84] Department of Gynecology and Gynecologic Oncology, Kliniken Essen-Mitte, Essen 45136, Germany. [85] Department of Gynecology and Gynecologic Oncology, Dr. Horst Schmidt Klinik, Wiesbaden 65199, Germany. [86] PMV Research Group, Department of Child and Adolescent Psychiatry and Psychotherapy, University of Cologne, Cologne 50931, Germany. [87] Department of Epidemiology and Biostatistics, Karolinska Institutet, Stockholm 171-77, Sweden. [88] Gynecological Oncology Unit, Royal Marsden Hospital, London SW3 6JJ, UK. [89] Cedars Sinai Medical Center, Samuel Oschin Comprehensive Cancer Center Institute, Los Angeles, California 90048, USA. [90] Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Victoria VIC 3806, Australia. [91] Breakthrough Breast Cancer Research Centre, Division of Genetics and Epidemiology, Institute of Cancer Research, London SW7 3RP, UK. [92] Department of Surgery and Cancer, Imperial College London, London SW7 2AZ, UK. [93] Institute of Human Genetics, Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen 91054, Germany. [94] Maggee Women's Hospital, Pittsburg, Pennsylvania 15213, USA. [95] Faculty of Medicine, University of Southampton, University Hospital Southampton, Southampton S017 1BJ, UK. [96] Gynaecology Research Unit, Hannover Medical School, Hannover 30625, Germany. [97] Section of Biostatistics and Epidemiology, Geisel School of Medicine at Dartmouth, Lebanon, New Hampshire 03755, USA. [98] Division of Epidemiology and Biostatistics, Department of Internal Medicine, University of New Mexico, Albuquerque, New Mexico 87131, USA. [99] Department of Genetics, Queensland Institute of Medical Research, Herston, Queensland QLD 4006, Australia. [100] Peter MacCallum Cancer Centre, Research Division, Cancer Genetics Laboratory, Melbourne, Victoria VIC 3002, Australia. [101] Department of Pathology, University of Melbourne, Parkville, Victoria VIC 3053, Australia. [102] Sir Peter MacCallum Department of Oncology, University of Melbourne, Parkville, Victoria VIC 3002, Australia. [103] Department of Pathology, Helsinki University Central Hospital, Helsinki, 00029 HUS, Finland. [104] Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia, Canada V52 1L3. [105] Department of Pathology, Cancer Institute, University College London, London WC1E 6JJ, UK. [106] Belarusian Institute for Oncology and Medical Radiology Aleksandrov N.N., Minsk 223040, Belarus. [107] College of Pharmacy and Health Sciences, Texas Southern University, Houston, Texas 77044, USA. [108] Department of Epidemiology, University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. [109] Department of Gynecologic Oncology, University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. [110] A list of consortium members appears in Supplementary Note 2. [111] A list of consortium members appears in Supplementary Note 3. [112] Cancer Prevention, Detection and Control Research Program, Duke Cancer Institute, Durham, North Carolina 27708, USA. [113] Department of Pathology, Vancouver General Hospital, BC Cancer Agency, Vancouver, British Columbia, Canada V5Z 4E6. [114] Gynecologic Cancer Program, Duke Cancer Institute, Durham, North Carolina 27708, USA. [115] Department of Public Health and Primary Care, University of Cambridge, Cambridge CB1 8RN, UK. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to E.L.G. (email: egoode@mayo.edu).